Quantifying genetic and nongenetic contributions to malarial infection in a Sri Lankan population

Margaret J. Mackinnon*†, D. M. Gunawardena‡§, Jagath Rajakaruna§, Sudath Weerasingha§, Kamini N. Mendis§, and Richard Carter*§

*Institute of Cell, Animal, and Population Biology, University of Edinburgh, Edinburgh, EH9 3JT, United Kingdom; ‡Anti-Malaria Campaign, Provincial Ministry of Health, Badulla, Sri Lanka; and §Malaria Research Unit, Department of Parasitology, Faculty of Medicine, University of Colombo, Colombo 8, Sri Lanka

Edited by David Weatherall, University of Oxford, Oxford, United Kingdom, and approved August 23, 2000 (received for review June 12, 2000)

Explaining the causes of variation in the severity of malarial disease remains a major challenge in the treatment and control of malaria. Many factors are known to contribute to this variation, including parasite genetics, host genetics, acquired immunity, and exposure levels. However, the relative importance of each of these to the overall burden of malarial disease in human populations has not been assessed. Here, we have partitioned variation in the incidence of malarial infection and the clinical intensity of malarial disease in a rural population in Sri Lanka into its component causes by pedigree analysis of longitudinal data. We found that human genetics, housing, and predisposing systematic effects (e.g., sex, age, occupation, history of infections, village) each explained approximately 15% of the variation in the frequency of malarial infection. For clinical intensity of illness, 20% of the variation was explained by repeatable differences between patients, about half of which was attributable to host genetics. The other half was attributable to semipermanent differences among patients, most of which could be explained by known predisposing factors. Three percent of variation in clinical intensity was explained by housing, and an additional 7% was explained by current influences relating to infection status (e.g., parasitemia, parasite species). Genetic control of *Plasmodium falciparum* infections appeared to modulate the frequency and intensity of infections, whereas genetic control of *Plasmodium vivax* infections appeared to confer absolute susceptibility or refractoriness but not intensity of disease. Overall, the data show consistent, repeatable differences among hosts in their susceptibility to clinical disease, about half of which are attributable to host genes.

Humans infected with malaria parasites experience a remarkable range of disease severity, from very mild symptoms to rapid death (1–3). Understanding the causes of this variation would be useful to clinicians and disease control experts alike. For example, does a child become ill with malaria because of its genetic make-up, its history of previous exposure, its current level of exposure, the genotype of its infecting parasite, its physiological status (e.g., nutritional) at the time of infection, or chance? How will malaria control programs that reduce transmission or boost acquired immunity change the population's mean and variation in morbidity, and what will happen if control is relaxed? To understand how to combat malaria effectively, it is necessary to know the relative importance of factors that contribute to the striking variation among humans in their experience of malarial disease.

One of the major classes of factors known to affect susceptibility to malaria and its clinical outcome is the genetics of the host. Early studies based on across-population associations indicated that mutations in the hemoglobin genes and other blood-related disorders had large protective effects on malaria (reviewed in refs. 4 and 5), although the mechanisms by which they protect are still not understood. Later studies based on within-population analyses in humans and experimental crosses in mice revealed more genes conferring resistance to malaria (6–17). It is clear that the trait is genetically complex.

Understandably, the focus in these studies has been on the genes themselves, rather than on their contribution to disease burdens in the overall population. The latter, therefore, remains largely undefined. A further feature of these genetic studies is that they have used a broad spectrum of possible measures of resistance, including parasite density (8–10, 18–21), severity of symptoms (7, 15, 22–27), and immune responsiveness (16, 17, 28). As there is no real consensus as to which measures are good indicators of protection from disease, it can be argued that the importance of host genetics is most meaningfully assessed by observing the frequency and severity of clinical disease, as reported by the patient, in the field.

In this study, we stand back from the issue of single genes and mechanisms and ask the question, "What is the total contribution of host genetic factors to disease burdens in the field relative to environmental influences and nongenetic personal factors?" We were able to address this question for a pedigree-known rural population of Sri Lanka by statistical analysis of longitudinal data on the frequency of becoming clinically ill and the severity of symptoms. By incorporating pedigree information alongside other information on the individual into the analyses, e.g., on house, sex, age, and previous exposure, we were able to separate out genetic factors from nongenetic factors and thereby make an assessment of the relative importance of these influences on the risk of malarial infection and its outcome in a typical population of a malaria-endemic area.

Materials and Methods

Study Area and Population. The study was conducted in a farming district in the dry lowland coastal plains of southeast Sri Lanka. The incidence of malaria is moderately seasonal, and the area is considered to be one of "unstable" but endemic malaria. The number of infectious bites per person per night is of the order of 0.001–0.01 (29), which compares to values of 0.1–1 in many areas of stable malaria in Africa (30). Most people in our study population have about one clinical attack per year, and generally 1–2% of the population are infected (at detectable levels) with either *Plasmodium vivax* or *Plasmodium falciparum* at any point in time (29). Epidemics of *P. falciparum* malaria used to occur at 7- to 10-year intervals (31), but since 1986, *P. falciparum* has persisted (32, 33). Further details of the study site are given elsewhere (29, 33).

Data Collection. The study population comprised 1,771 individuals of all ages living in eight contiguous villages in a total of 410 houses. This population was part of a larger one in which demographic

This paper was submitted directly (Track II) to the PNAS office.

Abbreviation: h^2, heritability.

†To whom reprint requests should be addressed. E-mail: M.Mackinnon@ed.ac.uk.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.220267997. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.220267997

information on age, sex, previous malarial infections, and various predisposing factors were recorded on the entire population since 1986 (33). The current study began in January 1992. For 19 months, the incidence of malaria episodes was monitored by passive case detection, i.e., individuals showing symptoms voluntarily presented themselves at the research center or occasionally at the local hospital. Use of bednets was negligible during this study period. Upon positive diagnosis of malaria based on symptoms and presence of parasites upon microscopic examination of thick blood films, a record of severity of 11 symptom scores, parasitemia, temperature and time since onset of symptoms, and related variables was made (see below) (3). Patients were treated with curative doses of chloroquine and primaquine at this time.

Pedigrees for most of the individuals in this study population were ascertained by interviewing members of each household and recording the relationships among all household members (e.g., parent, child, full-sib, half-sib, grandparent, unrelated). Records were also made of the house identities of any other first- or second-degree relatives living in other households in the study to establish genetic ties between households. The total pedigree comprised 2,692 individuals, including absent or dead relatives. This included 441 nuclear families (mother–father couples with at least one child) with an average of 2.2 children each. Sixty percent of the monitored population of 1,771 people shared a house with first-degree relatives (full-sibs or parents) who were also in the study, and 11% had second-degree relatives (cousins, uncles, aunts, or grandparents), usually living in separate houses. For the remainder of people, house, genetic, and nongenetic effects could not be disentangled because there was only one person per house/family with data.

The traits under analysis fell into two groups. Group A traits related to the frequency of becoming clinically ill and included the number of clinical attacks per person with either species, and the number of attacks due to each species separately. Each of these traits was analyzed with and without records from people who had no attacks during the study period. Group B traits related to the severity of the illness. They included the following: scores for 11 symptoms as reported by the patient in integer units on a scale of 0 to 2 or 3 [backache, arthralgia, myalgia, headache, hypochondrial (spleen-associated) pain, nausea, vomiting, anorexia, cold, shivering, sweating] (3), the total of the 11 scores, the first principal component of the following groups of related symptoms: aches (backache, arthralgia, myalgia), fever (cold, shivering, sweating), and gastrointestinal traits (nausea, vomiting, anorexia), the number of days since the onset of symptoms (obtained by interview), the temperature at time of presentation, the species of infecting parasite (obtained by morphology from blood films), parasitemia at the time of presentation (obtained from blood films) for each species separately, and gametocyte prevalence, i.e., whether gametocytes were detected in thick blood films or not (also split by species).

Before statistical analysis, records were disregarded when consecutive attacks of the same species occurred within 42 days of each other because it is probable that most such occurrences were attributable to the failure of drug treatment rather than to a new infection.

Statistical Analysis. The objective of the analysis was to estimate the relative contribution of the following five sources of variation to the total variation observed for the trait: (*i*) systematic environmental effects ("fixed effects"), (*ii*) additive genetic effects, (*iii*) other "person" effects, (*iv*) house effects, and (*v*) unexplained residual variation. To achieve this, a mixed model with fixed and random effects to estimate, respectively, systematic differences among groups of individuals (e.g., male vs. female) and sources of random variation among individuals (e.g., additive genes) was fitted. By incorporating pedigree information into the model via a genetic relationship matrix that encompasses all known genetic relationships among all individuals in the population, an estimate of the additive genetic variance, and hence the heritability, can be obtained (34). The model fitted to the data was $y = b + a + m + c + e$, where y is the vector of observations on individual people, b represents a series of fixed effects, a, m, c, and e are the random effects for additive genetic merit, "other" personal merit, house, and residual error, respectively. These random components are assumed to be normally distributed with variances and covariances of $A\sigma_a^2$, $I\sigma^2_m$, $I\sigma^2_c$, and $I\sigma^2_e$, where A is the additive genetic relationship matrix and I is an identity matrix. The order of A is the number of people in the entire pedigree (including those without records), the order of I for $I\sigma^2_m$, and $I\sigma^2_e$ is the number of people with records, and the order of I in $I\sigma^2_c$ is the number of houses. The proportions of the observed phenotypic variance (σ^2_p) due to all of the fixed effects combined and to each random effect were calculated to yield estimates of heritability ($h^2 = \sigma^2_a/\sigma^2_p$), "other" personal effects ($m^2 = \sigma^2_m/\sigma^2_p$), and house ($c^2 = \sigma^2_c/\sigma^2_p$). m^2 includes all effects attributable to a person other than the additive genetic effects, e.g., "permanent environmental" effects due to, say, acquired immunity, as well as nonadditive genetic effects due to dominance and epistasis. The house component (c^2) represents effects common to a house other than additive genetic effects and "other" personal effects. The term for m was excluded from the model for analysis of group A traits because there was only one record per person, and repeated records on the same person are required to separate m^2 from a^2. In this case, any permanent environmental and nonadditive genetic effects (σ^2_m) are included in the residual (σ^2_e).

The model was fitted to the data using the method of restricted maximum likelihood to find the best estimates of the parameters. The computer program used was DFREML (35). As this, and all other available genetic analysis programs, does not allow for non-normal distributions of the trait, the data were also analyzed under the same model, but without genetic relationships, using the PROC MIXED procedure of SAS (36) in combination with macros to perform the analysis on the logit-transformed scale for categorical traits (37) or on the logit-transformed scale for Poisson-distributed traits (group A), allowing for over- or underdispersion (38). Based on comparisons of the results from the SAS vs. DFREML analyses (i.e., variance component estimates, fixed effect estimates, distributions of residuals), it was decided that for group A traits the genetic (DFREML) analysis should be performed on the residuals from a fixed effects analysis on the transformed scale. For categorical data, DFREML analyses were performed on the untransformed scale. Parasitemia data, which were highly skewed, were log-transformed before analysis.

A range of models differing in their fixed effects was fitted to explore the overlap between fixed and random effects, e.g., parasitemia at presentation may have had a genetic basis, but in some models was included as a fixed effect. Fixed effects were grouped into two classes: those that might predispose the patient to a clinical attack (e.g., age, occupation) and those immediate factors that were likely to affect the level of illness, (e.g., parasitemia, parasite species). Zero, one, or both sets of these factors were fitted in the model (see footnote in Table 1). For group B traits, predisposing effects included sex, age, occupation, village, whether born in an endemic area or not, the number of previous attacks, and the number of months since a previous attack occurred (the latter two factors being fitted as both linear and quadratic covariates, and the remainder fitted as class variables), whereas immediate effects included temperature and parasitemia at the time of presentation, parasite species, whether gametocyte positive or not, the number of days symptomatic, and the number of hours since the last paroxysm occurred (the latter two being fitted as linear covariates). An additional "nuisance" fixed effect for interviewer was also fitted for group B traits to

Table 1. Summary statistics and proportions of variance explained by genetics and other sources for the number of clinical attacks per person and the severity of symptoms in a population in Sri Lanka as estimated by different models

Trait	No. of records	Mean*	Phenotypic variance,* σ_P^2	Percentage of phenotypic variance explained†: Random effects (SE)‡: Genetic h^2	Other personal m^2	Total personal $h^2 + m^2$	House c^2	Fixed effects§: Model 1A	2A	3A	
Number of infections (group A traits)											
With zeroes¶											
Both species	1771	0.71	1.16	14–18 (7)			12–21 (4)	0	10	18	
P. falciparum	1769	0.23	0.24	8–21 (7)			4–13 (3)	0	9	16	
P. vivax	1771	0.48	0.85	10–17 (7)			11–17 (4)	0	9	16	
Without zeroes¶											
Both species	739	1.69	1.11	15–21 (12)			3–7 (6)	0	5	14	
P. falciparum	737	0.55	0.41	10–25 (13)			5–10 (6)	0	7	9	
P. vivax	739	1.14	1.28	2–5 (11)			8–11 (6)	0	2	9	
							Model	1B	2B	3B	4B
Sickness scores (group B traits)											
Total score	1246	11.0	17.3	8–10 (7)	1–14 (6)	11–22	3–4 (4)	6	16	10	20
Aches‖	1076	3.90	7.09	4–6 (9)	9–24 (8)	13–29	0–2 (5)	9	28	10	29
Gastrointestinal‖	1076	1.82	2.07	10–13 (9)	2–4 (8)	14–17	3–4 (4)	1	4	6	8
Fever‖	1076	3.09	2.78	12–15 (9)	0–5 (9)	14–17	5–7 (4)	2	3	6	8
Headache	1144	2.11	0.87	3–5 (7)	17–22 (8)	20–27	0–7 (4)	7	9	7	10
Hypo. pain	1093	0.35	0.28	6–13 (9)	0–1 (9)	6–13	1–5 (5)	2	3	6	8
Other measures (group B traits)											
Temperature	1201	38.03	1.14	0–0 (7)	0–9 (6)	0–9	3–5 (9)	1	8	23	26
Parasitemia (%)											
P. falciparum	406	0.063	0.102	16–24 (22)	1–12 (21)	16–35	0–1 (11)	0	11	2	11
P. vivax	839	0.094	0.030	0–6 (15)	6–19 (15)	12–19	4–10 (7)	2	12	21	21
Gametocytemic											
P. falciparum	406	0.042	0.040	0–0 (24)	0–1 (21)	0–1	0–0 (5)	1	0	0	0
P. vivax	839	0.277	0.200	0–1 (11)	2–5 (11)	4–6	0–0 (3)	5	6	14	14
No. days symp.	1246	2.81	2.76	3–5 (8)	5–7 (7)	8–15	0–3 (5)	0	4	2	5

*On the untransformed scale, unadjusted for fixed effects.
†Range of estimates from models with different fixed effects.
‡Averaged over all models, where estimable.
§Fixed effects in each model are as follows. Group A traits: model 1A, none; model 2A, village; model 3A, village and other predisposing (sex, age, occupation); group B traits: model 1B, nuisance effect (interviewer); model 2B, nuisance and predisposing effects; model 3B, nuisance and immediate effects; model 4B, nuisance, predisposing and immediate effects. See text.
¶"With zeroes" and "Without zeroes" includes and excludes, respectively, records from people who had no clinical attacks during the study period.
‖Trait analyzed was the first principal component of a group of related traits.

account for interviewer bias in the recording process. For group A traits, the predisposing effects fitted were village, sex, age, and occupation; immediate effects were not fitted. Significance tests of whether variance components were non-zero were performed using likelihood ratio tests and t tests based on the sampling errors from the information matrix (39). As between fixed and random effects, there was potential confounding among random effects, especially between genetic and house effects because members of the same family usually lived in the same house. An indication of the level of confounding among random effects (and hence the reliability of individual estimates) was obtained by setting h^2, m^2, or c^2 to zero and then observing the change in other estimates, and from the sampling correlations among estimates, derived from the information matrix (39).

Results

Summary of Data. Summary statistics of the data (raw means and variances) are given in Table 1. Of the 1,771 people recruited into the study, 42% experienced at least one clinical attack during the 19-month period. After discarding 322 of the 1,564 (21%) recorded attacks that seemed to be due to recrudescences resulting from drug failure (i.e., within 42 days of the previous attack), of those people who had an attack, 58% had only one attack, 25% had two attacks, and the remaining 17% had three or more attacks, giving an average number of attacks per person of 0.7 within 19 months. There was a different age-prevalence pattern for the two species: whereas the number of *P. vivax* infections decreased in a linear fashion after the peak rate at age 3–6 years, the rate of *P. falciparum* attacks did not reach a peak until the ages of 31–35, after which it declined linearly. Before this study, the estimated average number of previous attacks experienced to date (obtained by interview) was 8.4, which came to 0.3 per year of age (average age of 25). Thirty-eight percent of the population reported to the clinic within 2 days of the onset of symptoms. Most clinical symptoms were mild to moderate (score of 1), and when totaled over 11 symptoms, they gave an average of 11.0 out of a possible total of 26. Accompanying parasitemias were low (average of 0.08%), and in 80% of cases, temperatures were not higher than 39°C.

After removing the recrudescent infections, the majority (74%) of which were *P. falciparum*, the proportion of infections due to *P. falciparum* infections was 36%. Only two cases of mixed-species infections were recorded. Of those people who had more than one attack, the frequencies of consecutive

infections were as follows: 11% were *P. falciparum* infections followed by *P. falciparum* infections, 21% were *P. falciparum* infections followed by *P. vivax* infections, 14% were *P. vivax* infections followed by *P. falciparum* infections, and 54% were *P. vivax* infections followed by *P. vivax*. This observed pattern of consecutive infections deviated significantly from that expected ($\chi_1^2 = 12.0, P < 0.001$); most of this deviation was due to an excess of consecutive *P. falciparum* infections.

Sources of Variation. The proportion of total variation explained by additive genetic effects, by other personal effects, by house effects, and by fixed effects are presented in Table 1. Individual estimates from each model are not given; instead, the range of estimates from the three or four models fitted is shown. Estimates of variance components on the transformed scale from SAS analyses were similar to those from DFREML analyses on the untransformed scale and so are not presented. Results for individual symptom scores are not presented for traits that were included in groups of related symptoms (e.g., body aches) analyzed as principal components. For fixed effects, only the proportion of variation explained by a whole group of fixed effects (e.g., predisposing factors as a group, or immediate effects as a group) is shown.

Statistical tests based on likelihood ratios of the hypothesis that each estimate was different from zero did not give different results to t tests based on the standard error and so are not shown. In general, estimates of h^2 and m^2 were not significantly different ($P > 0.05$) from zero if below 15–20%, whereas estimates of c^2 were not significant if below 10%. Estimates of the variation accounted for by the fixed effects were not significant ($P > 0.05$) when greater than 2–3%. Differences between estimates of h^2, m^2, and c^2 for the two species separately were never statistically significant ($P > 0.05$). Correlations between estimates of h^2 and m^2 were around -0.6 and between h^2 and c^2 were around -0.8, reflecting the predominantly nested structure of the data (e.g., families tended to live in the same house). However, this confounding is not expected to lead to bias in the estimates.

Number of Infections (Group A Traits). When individuals who remained uninfected with malaria during the period of the study were included in the analysis (Table 1), heritabilities (h^2, the genetic component) of the number of infections of either *P. falciparum* or *P. vivax* were around 15–20%. When those who did not become infected during the study were excluded from the data, the heritable component of the number of infections with *P. vivax* was reduced from around 15% to less than 5%, whereas for *P. falciparum* it remained around 15–20%. Although these different estimates for the two species could not be distinguished statistically, this might suggest that the genetic effect protecting against *P. vivax* was acting as an absolute determinant of whether the infection occurred or not, compared with a more graded protection upon challenge from *P. falciparum* parasites.

House (c^2) accounted for a further 15% of the variation in susceptibility to infection and/or clinical disease with either *P. falciparum* or *P. vivax* when people with zero infections were included, but fell to about 5–10% when these were excluded. This, combined with the fact that village accounted for 5–10% of the total variation (model 2A vs. model 1A), indicates that there was a significant element of spatial variation in exposure levels, i.e., a non-uniform distribution of risk of infection between different houses and different villages. This "clustering" was also reflected in overdispersion of the distribution of the number of infections (variance to mean ratio ranging from 1.3 to 1.6) when only the genetic contribution was considered; this disappeared (i.e., followed a Poisson distribution) when village and house were included in the model (data not shown). Village effects seemed to be more influential on *P. falciparum* than *P. vivax*, as inclusion of village in the model reduced the estimates of h^2 by around 10% and 3% in the two species, respectively, and the estimates of c^2 by around 7% and 2%, respectively. Thus, there seemed to be some confounding between incidence of *P. falciparum* and village, and this may have led to overestimation of the genetic and housing influences on infection rates with this species. Even so, the estimates of h^2 for *P. falciparum* were always higher than for *P. vivax* when analyzed under the same model, and the reverse was true for estimates of c^2. Other predisposing factors (excluding village), i.e., sex, age, occupation, endemic-born, etc., accounted for a further 5–10% of variation in the number of clinical episodes of malaria (model 3A vs. model 2A), which probably reflects the effects of acquired immunity on protection against infection and/or clinical disease. Inclusion of these predisposing effects in the model changed the estimates of h^2 and c^2 by less than 2%.

Sickness Scores and Other Measures of Infection (Group B Traits). Heritabilities (h^2) for intensity of clinical disease traits as measured by symptom scores were around 10–15% for each of fever symptoms, gastrointestinal symptoms, and hypochondrial pain (Table 1). Heritabilities for body ache and headache symptoms were lower (5%). The opposite pattern was observed for "other" personal effects (m^2), which were stronger for body aches and headache (15–20%) than for other symptoms (<5%). Similarly, predisposing effects (age, sex, occupation, village, endemic born; model 2B vs. model 1B) also explained more variation in body aches and headache (28% and 9%, respectively) than in other symptoms that were largely unaffected by these factors. Thus, body aches and headache seemed to be more influenced by factors related to previous exposure (as reflected in m^2 and predisposing effects), whereas fever, gastrointestinal symptoms, and hypochondrial pain seemed to be more controlled by host genetic influences. Altogether, between-person variation, both genetic and nongenetic (h^2+m^2), accounted for around 20% of the variation for most symptoms. Thus, people tended to be consistent in the degree of sickness that they reported; this was also reflected in the 8–15% repeatability of the number of days that a person was symptomatic before reporting to the clinic. Many of these repeatable differences in symptoms were able to be explained by known predisposing factors since the estimates of m^2 were, on average, reduced by 7% when these factors were included in the model as fixed effects. Thus, m^2 and predisposing effects were somewhat interchangeable in the model. In contrast, estimates of h^2 and c^2 changed by less than 2% when predisposing effects were fitted; similarly, estimates of h^2, m^2, and c^2 changed little when current influences on the infection were included in the model.

Variation between houses (c^2) in intensity of clinical disease was low (around 3%) and not significantly different from zero.

For traits relating to the status of the infection (temperature, parasitemia, and gametocyte prevalence at presentation), estimates of heritability were generally low (<5%). A notable exception was the moderate heritability (around 18–24%) of *P. falciparum* parasitemia. In contrast, the heritability of *P. vivax* parasitemia was low (0–7%). These results parallel the higher heritability of the number of attacks of *P. falciparum* than of the number of attacks of *P. vivax* (see group A traits, above). Conversely, the "other" personal component (m^2) was moderately high for *P. vivax* (7–20%) and rather lower for *P. falciparum* (0–13%).

To appreciate the magnitude of the contribution of host genetics to the overall variation in the frequency and clinical intensity of malarial infections among members of this population, consider the following. The mean symptom scores of the top and bottom 10% of the population for intensity of clinical symptoms were 16.3 and 5.7, respectively, giving a difference between these groups of 10.6 (from a population with a mean of 11.0 and range of 0 to 25). If it were possible to assign people into the top and bottom 10% of clinical responders according to their genetic, rather than their phenotypic characteristics, the mean

symptom scores of these two groups would be 11.9 and 10.0, respectively, a difference of 1.9. That is, genetic effects accounted for roughly one-fifth of the observed spread of symptom scores between the most susceptible and the most protected in the population. Similarly, the 10% of people who had the most and the 10% with the fewest infections during the period of the study had means of 3.2 and 0 infections, respectively (population mean of 0.71, range of 0 to 7). If the population could be assigned to these groups on the basis of their genetic characteristics, the mean numbers of infections (predicted from the distribution on the transformed scale) would be 1.1 and 0.5, respectively, i.e., an approximately 2-fold difference. Put another way, the least protected 10% of the population were twice as vulnerable to malarial infection as the most protected 10% because of their genetic characteristics. The equivalent ratios if the population was divided into the top and bottom 50%, or the top and bottom 5%, would be 1.4- and 3-fold, respectively. Approximately the same ratios would be found if the population were to be similarly divided into low-risk vs. high-risk houses or villages.

Discussion

This study was designed to partition the remarkable amount of variation in the frequency and intensity of malarial disease observed in the field (1–3) into its major sources. These sources may be broadly categorized as intrinsic factors due to the host (genetics, acquired immunity, physiological factors), extrinsic factors of the environment (exposure-related effects, e.g., house location, occupation), and temporary infection-status effects (e.g., parasitemia and temperature at presentation, duration of symptoms). We were able to attribute around 15–20% of the variation in the frequency and intensity of clinical episodes to stable between-person differences, about one-third of which could be attributed to predisposing, semipermanent variation among people due to previous exposure, and about half of which was due to permanent, heritable differences. Thus, there are genes segregating in this rural Sri Lankan population that help or hinder the host's ability to control malarial infection. These combined genetic effects are sufficiently large to be detectable against a background of stochastic and heterogeneous malaria transmission. They are responsible for an approximately 2-fold difference in the rate at which the genetically most susceptible compared with the genetically least susceptible people experience clinical infections. The contribution of host genetics found in this study is likely to be an underestimate of the importance of host genes on a global scale as malaria-related mortality in this population is virtually zero, largely due to good primary health care. We may also have underestimated the importance of host genetics because of inaccuracies in the pedigree due to misreporting of paternity, the subjective nature of reporting severity of symptoms, and the tendency to not report very mild infections.

An interesting finding of this study is that the genetic effects on disease susceptibility appeared to be fundamentally different between *P. falciparum* and *P. vivax* infections. Whereas the h^2 of the frequency of infections with *P. falciparum* did not change when those with zero infections were excluded from the analysis, the h^2 of *P. vivax* infections was much reduced when these were excluded. Also, the h^2 of parasite density was higher in *P. falciparum* than in *P. vivax* infections. Although these species differences in h^2 were not statistically significant and were somewhat dependent on whether spatial variation was taken into account in the model, the consistent nature of the species pattern across different traits is suggestive that real differences exist in the genetic mechanisms that control the two species. The pattern suggests that the genetic control of *P. falciparum* infection does not confer absolute susceptibility or refractoriness to infection but rather modulates the probability of infection and the density of parasites during infection. Such effects would be consistent with the known presence in South Asian populations of inherited blood disorders that give partial protection against clinical illness from *P. falciparum* such as sickle cell trait (hemoglobin S), hemoglobin E, the thalassemias, and glucose 6-phosphate dehydrogenase deficiency (40). Degrees of genetic protection against *P. falciparum* infection could also occur through immunological mechanisms such as the postulated HLA-controlled blocking of parasite development in the liver (7, 41) or through other immune response genes (16, 17, 28). The effects are also consistent with the existence of genes affecting parasite density during blood infections of *P. falciparum* such as the single locus genetic effect on parasite density during *P. falciparum* infection, which has been reported in populations of West Africa (8, 11–14). In contrast, the genetic effects on *P. vivax* seemed to determine in an absolute way whether or not an individual became infected with *P. vivax* malaria rather than modulating the response to infections already acquired. This type of genetic effect is consistent with the known effects of the Duffy blood group antigen (18), which is, to our knowledge, the only known human genetic influence on resistance to *P. vivax* infection. This variant confers absolute refractoriness to *P. vivax* infection because of the inability of *P. vivax* merozoites to invade Duffy negative red blood cells (18). Unlike *P. falciparum* (42), *P. vivax* does not appear to have alternative invasion pathways by which *P. vivax* can circumvent this obstacle. However, the Duffy negative blood group is rare outside West African populations (43–46), although it may be emerging in Papua New Guinea (47). Our study thus raises the question of whether this or other genes conferring absolute refractoriness to *P. vivax* infection are segregating in this Sri Lankan population. Alternative explanations for the apparent species difference in the partitioning of genetic and environmental effects is that the species differ in their ecology and hence environmental dependence, or differ in the way they generate acquired immunity, or that the genetic control of *P. vivax* is predominantly non-additive (dominant and/or epistatic) in nature. All of these are interesting possibilities that deserve deeper investigation because of their implications for disease control in areas where these species coexist.

In addition to the genetic effects on susceptibility to, and frequency of, malarial infection and parasite density, we also found evidence of significant genetic effects upon the clinical intensity of the infections. Genetic effects were strongest for gastrointestinal symptoms, fever, and hypochondrial (splenic) pain. What might be the basis of the genetic control on clinical disease observed here? The host may be protected from disease through some of the genetic mechanisms implicated above that limit the growth rate of the asexual parasite population (8, 10–14, 21), perhaps through cytokine-mediated fever responses (15, 48, 49), or some physiological alteration of the red cell environment that is unfavorable to parasite growth (50, 51), or through better immune response to parasite antigens (16, 17). Disease severity may also be directly limited by host genetic control of the pathological effects induced by the parasites such as the cytokines tumor necrosis factor-α and tumor necrosis factor-β (23, 27, 52, 53). It is not possible from our analysis to say which of these mechanisms were involved because the heritability estimates include all of the additive genetic variation, not just that from individual genes. Thus, the precise genetic mechanisms offering protection from malaria in this population will remain speculative until linkage studies involving candidate genes, as revealed by other studies (10–14, 21), lead to identification of the genes involved.

Equally important as genetics in determining disease burden were nongenetic intrinsic factors (acquired immunity) and environmental factors (e.g., housing); these effects are well known and have been discussed widely in the malaria literature (32, 33, 54–58), although rarely with respect to their contribution to

overall variation in disease. Both of these influences exhibited different patterns for the two parasite species, indicating that different immune mechanisms and ecology are operating in *P. falciparum* and *P. vivax*, as is well established (54, 58). Moreover, there seemed to be a strong interaction between these two species as evident by a marked deficit in mixed-species infections in this study; such deficits have been reported in other areas (59–61) but are poorly understood (62). The suggestion from this study that there are genes that protect against *P. falciparum* and not against *P. vivax,* and vice versa, and the possible interactions between host genetics and disease ecology (26) add further insights into this issue.

We are grateful to the staff of the Kataragama Malaria Research Station, the Malaria Research Unit, Faculty of Medicine, Colombo, and to the Ministry of Education, Moneragalla District, Sri Lanka, for assistance in the collection of the data. We thank W. G. Hill and A. F. Read for comments on the manuscript. M.J.M. was supported by the Leverhulme Trust, and the work in Sri Lanka was supported by the United Nations Development Program/World Bank/World Health Organization Special Program for Research and Training in Tropical Diseases.

1. Greenwood, B. M., Marsh, K. & Snow, R. (1991) *Parasitol. Today* **7,** 277–281.
2. Marsh, K. (1992) *Parasitology* **104,** 53–69.
3. Karunaweera, N. D., Carter, R., Grau, G. E. & Mendis, K. (1998) *Am. J. Trop. Med. Hyg.* **58,** 204–210.
4. Livingstone, F. B. (1971) *Annu. Rev. Genet.* **5,** 33–64.
5. Weatherall, D. J. (1987) *Ann. Trop. Med. Parasitol.* **81,** 539–548.
6. Stevenson, M. M., Nesbitt, M. N. & Skamene, E. (1988) in *Current Topics in Microbiology and Immunology* (Springer, Heidelberg), pp. 325–328.
7. Hill, A. V. S., Allsopp, C. E. M., Kwiatkowski, D., Anstey, N. M., Twumasi, P., Rowe, P. A., Bennett, S., Brewster, D., McMichael, A. J. & Greenwood, B. M. (1991) *Nature (London)* **352,** 595–600.
8. Abel, L., Cot, M., Mulder, L., Carnevale, P. & Feingold, J. (1992) *Am. J. Hum. Genet.* **50,** 1308–1317.
9. Stevenson, M. M., Lyange, J. J. & Skamene, E. (1982) *Infect. Immun.* **38,** 80–88.
10. Foote, S. J., Burt, R. A., Baldwin, T. M., Presente, A., Roberts, A. W., Laural, Y. L., Lew, A. M. & Marshall, V. M. (1997) *Nat. Genet.* **17,** 380–381.
11. Garcia, A., Cot, M., Chippaux, J. P., Ranque, S., Feingold, J., Demenais, F. & Abel, L. (1998) *Am. J. Trop. Med. Hyg.* **58,** 480–488.
12. Garcia, A., Marquet, S., Bucheton, B., Hillaire, D., Cot, M., Fievet, N., Dessein, A. J. & Abel, L. (1998) *Am. J. Trop. Med. Hyg.* **58,** 705–709.
13. Rihet, P., Abel, L., Traore, Y., Aucan, C. & Fumoux, F. (1998) *Genet. Epidemiol.* **15,** 435–450.
14. Rihet, P., Traore, Y., Abel, L., Aucan, C., Traore-Leroux, T. & Fumoux, F. (1998) *Am. J. Hum. Genet.* **63,** 498–505.
15. Jepson, A., Banya, W., Sisay-Joof, F., Hassan-King, M., Bennett, S. & Whittle, H. C. (1995) *J. Infect. Dis.* **172,** 316–319.
16. Jepson, A., Sisay-Joof, F., Banya, W., Hassan-King, M., Frodsham, A., Bennett, S., Hill, A. V. S. & Whittle, H. C. (1997) *Br. Med. J.* **315,** 96–97.
17. Stirnadel, H. A., Beck, H.-P., Alpers, M. P. & Smith, T. A. (1999) *Genet. Epidemiol.* **17,** 16–34.
18. Miller, L. H., Mason, S. J., Dvorak, J. A., McGinniss, M. H. & Rothman, I. K. (1975) *Science* **189,** 561–563.
19. Cattani, J. A., Gibson, F. D., Alpers, M. P. & Crane, G. G. (1987) *Trans. R. Soc. Trop. Med. Hyg.* **81,** 705–709.
20. Cot, M., Abel, L., Roisin, A., Barro, D., Yada, A., Carnevale, P. & Feingold, J. (1993) *Am. J. Trop. Med. Hyg.* **48,** 358–364.
21. Fortin, A., Belouchi, A., Tam, M. F., Cardon, L., Skamene, E., Stevenson, M. M. & Gros, P. (1997) *Nat. Genet.* **17,** 382–383.
22. Stevenson, M. M., Lyanga, J. J. & Skamene, E. (1982) *Infect. Immun.* **38,** 80–88.
23. McGuire, W., Hill, A. V. S., Allsopp, C. E. M., Greenwood, B. M. & Kwiatkowski, D. (1994) *Nature (London)* **371,** 508–511.
24. Genton, B., Al-Yaman, F., Mgone, C., Alexander, N., Paniu, M. M. & Alpers, M. P. (1995) *Nature (London)* **378,** 564–565.
25. Ruwende, C., Khoo, S. C., Snow, R. W., Yates, S. N. R., Kwiatkowski, D., Gupta, S., Warn, P., Allsopp, C. E. M., Gilbert, S. C., Peschu, N., *et al.* (1995) *Nature (London)* **376,** 246–249.
26. Williams, T. N., Maitland, K., Bennett, S., Ganczakowski, M., Peto, T. E. A., Newbold, C. I., Bowden, D. K., Weatherall, D. J. & Clegg, J. B. (1996) *Nature (London)* **383,** 522–525.
27. Wattavidanage, J., Carter, R., Perera, K. L. R. L., Munasingha, A., Bandara, S., McGuiness, D., Wickramasinghe, A. R., Alles, H. K., Mendis, K. N. & Premawansa, S. (1999) *Clin. Exp. Immunol.* **115,** 350–355.
28. Sjöberg, K., Lepers, J. P., Raharimalala, L., Larsson, A., Olerup, A., Marbiah, N. T., Troye-Blomberg, M. & Perlmann, P. (2000) *Proc. Natl. Acad. Sci. USA* **89,** 2101–2104.
29. Mendis, C., Gamage-Mendis, A. C., Dezoysa, A. P. K., Abhayawardena, T. A., Carter, R., Herath, P. R. J. & Mendis, K. N. (1990) *Am. J. Trop. Med. Hyg.* **42,** 298–308.
30. Hay, S. I., Rogers, D. J., Toomer, J. F. & Snow, R. W. (2000) *Trans. R. Soc. Trop. Med. Hyg.* **94,** 113–127.
31. Rajendram, S. & Jayewickreme, S. H. (1951) *Indian J. Malariol.* **5,** 1–89.
32. Gamage-Mendis, A. C., Carter, R., Mendis, C., Dezoysa, A. P. K., Herath, P. R. J. & Mendis, K. N. (1991) *Am. J. Trop. Med. Hyg.* **45,** 77–85.
33. Gunawardena, D. M., Wickramasinghe, A. R., Muthuwatta, L., Weerasinghe, S., Rajakaruna, J., Senanayaka, T., Kotta, P. K., Attanayake, N., Carter, R. & Mendis, K. N. (1998) *Am. J. Trop. Med. Hyg.* **58,** 533–542.
34. Meyer, K. (1989) in *Estimation of Genetic Parameters*. eds. Hill, W. G. & Mackay, T. F. C. (CAB International, Wallingford), pp. 161–167.
35. Meyer, K. (1991) DFREML User Notes (University of New England, Armidale, Australia), Version 2.0.
36. SAS Institute, Inc. (1990) SAS/STAT *User's Guide* (SAS Institute, Inc., Cary, NC).
37. Lipsitz, S. R., Kyungmann, K. & Zhao, L. (1994) *Stat. Med.* **13,** 1149–1163.
38. Wolfinger, R. (1997) GLIMIX: A SAS macro for fitting generalized linear mixed models using PROC MIXED and the Output Delivery System (ODS) (SAS Institute, Inc., Cary, North Carolina).
39. Kendall, M. & Stuart, A. (1979) *The Advanced Theory of Statistics* (MacMillan, New York).
40. Livingstone, F. B. (1967) *Abnormal Hemoglobins in Human Populations: A Summary and Interpretation* (Aldine Publishing Company, Chicago).
41. Hill, A. V. S., Elvin, J., Willis, A. C., Aidoo, M., Allsopp, C. E. M., Gotch, F. M., Gao, X. M., Takiguchi, M., Greenwood, B. M., Townsend, A. R. M., *et al.* (1992) *Nature (London)* **360,** 434–439.
42. Dolan, S. A., Proctor, J. L., Alling, D. W., Okubo, Y., Wellems, T. E. & Miller, L. H. (1994) *Mol. Biochem. Parasitol.* **64,** 55–63.
43. Sanger, R., Race, R. R. & Jack, J. (1955) *Br. J. Haematol.* **1,** 370–374.
44. Mourant, A. E. (1974) *Philos. Trans. R. Soc. London* **268,** 251–255.
45. Lewis, M., Kaita, H. & Choun, B. (1972) *Vox Sanguinis* **23,** 523–527.
46. Ward, R. H., Gershowitz, H., Layrisse, M. & Neel, J. V. (1975) *Am. J. Hum. Genet.* **27,** 1–30.
47. Zimmerman, P. A., Woolley, I., Masinde, G. L., Miller, S. M., McNamara, D. T., Hazlett, F., Mgone, C., Alpers, M. P., Genton, B., Boatin, B. A., *et al.* (1999) *Proc. Natl. Acad. Sci. USA* **96,** 13973–13977.
48. Kwiatkowski, D., Cannon, J. G., Manogue, K. R., Cerami, A., Dinarello, C. A. & Greenwood, B. M. (1989) *Clin. Exp. Immunol.* **77,** 361–366.
49. Kwiatkowski, D. (1995) *Parasitol. Today* **11,** 206–212.
50. Pasvol, G., Weatherall, D. J. & Wilson, R. J. M. (1977) *Nature (London)* **270,** 171–173.
51. Pasvol, G., Weatherall, D. J. & Wilson, R. J. M. (1978) *Nature (London)* **274,** 701–703.
52. Kwiatkowski, D., Hill, A. V. S., Sambou, I., Twumasi, P., Castracane, J., Manogue, K. R., Cerami, A., Brewster, D. R. & Greenwood, B. M. (1990) *Lancet* **336,** 1201–1204.
53. Karunaweera, N. D., Grau, G. E., Gamage, P., Carter, R. & Mendis, K. N. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 3200–3203.
54. Taliaferro, W. H. (1949) in *Immunity to the Malaria Infections*, ed. Boyd, M. F. (Saunders, London), pp. 935–965.
55. Cohen, S. (1979) *Proc. R. Soc. London Ser. B.* **203,** 323–345.
56. van der Hoek, W., Konradsen, F., Dijkstra, D. S., Amerasinghe, P. H. & Amerasinghe, F. P. (1998) *Trans. R. Soc. Trop. Med. Hyg.* **92,** 265–269.
57. Ghebreyesus, T. A., Haile, M., Witten, K. H., Getachew, A., Yohannes, M., Lindsay, S. W. & Byass, P. (2000) *Trans. R. Soc. Trop. Med. Hyg.* **94,** 17–21.
58. Boyd, M. F. (1949) *Malariology* (Saunders, London).
59. Cohen, J. E. (1973) *Q. Rev. Biol.* **48,** 467.
60. Maitland, K., Williams, T. N., Bennett, S., Newbold, C. I., Peto, T. E. A., Viji, J., Timothy, R., Clegg, J. B., Weatherall, D. J. & Bowden, D. K. (1996) *Trans. R. Soc. Trop. Med. Hyg.* **90,** 614–620.
61. Bruce, M. C., Donnelly, C. A., Alpers, M. P., Galinski, M. R., Barnwell, J. W., Walliker, D. & Day, K. P. (2000) *Science* **287,** 845–848.
62. Maitland, K., Williams, T. N. & Newbold, C. I. (1997) *Parasitol. Today* **13,** 227–230.