MAJOR ARTICLE

A Major Gene Controls Leprosy Susceptibility in a Hyperendemic Isolated Population from North of Brazil

Fernando P. Lázaro,[1,a] Renata I. Werneck,[1,a] Ciane C. O. Mackert,[1] Aurélie Cobat,[3,4] Flávia C. Prevedello,[1] Raphaela P. Pimentel,[1] Geraldo M. M. Macedo,[2] Marco A. M. Eleutério,[1] Guilherme Vilar,[1] Laurent Abel,[3,4,5] Marília B. Xavier,[2] Alexandre Alcaïs,[3,4,5,b] and Marcelo T. Mira[1,b]

[1]Advanced Molecular Investigation Core, Graduate Program in Health Sciences, Pontifical Catholic University of Paraná, Curitiba, Paraná, and [2]Tropical Medicine Core, Federal University of Pará, Belém, Brazil; [3]Laboratoire de Génétique Humaine des Maladies Infectieuses, Institut National de la Santé et de la Recherche Médicale U550, and [4]Université Paris René Descartes, Faculté Médecine Necker, Paris, France; and [5]Laboratory of Human Genetics of Infectious Diseases, Rockefeller Branch, Rockefeller University, New York, New York

Background. Leprosy is a chronic infectious disease that affects 250,000 new individuals/year worldwide. Genetic analysis has been successfully applied to the identification of host genetic factors affecting susceptibility to leprosy; however, a consensus regarding its mode of inheritance is yet to be achieved.

Methods. We conducted a complex segregation analysis (CSA) on leprosy using data from the Prata Colony, an isolated, highly endemic former leprosy community located at the outskirts of the Brazilian Amazon. The colony offers large multiplex, multigenerational pedigrees composed mainly by descendents of a small number of original leprosy-affected families. Our enrollment strategy was complete ascertainment leading to the inclusion of the whole colony (2005 individuals, 225 of whom were affected) distributed in 112 pedigrees. CSA was performed using REGRESS software.

Results. CSA identified a best-fit codominant model, with a major gene accounting for the entire familial effect observed. The frequency of predisposing allele was estimated at 0.22. Penetrance for homozygous individuals for the predisposing allele >30 years old ranged from 56% to 85%, depending on sex.

Conclusions. A strong major gene effect in the isolated, hyperendemic Prata Colony indicates enrichment of genetic risk factors, suggesting a population particularly suitable for leprosy gene identification studies.

Leprosy is a chronic infectious disease caused by *Mycobacterium leprae* that affects 250,000 new individuals worldwide every year, with the majority of cases concentrated in India and Brazil [1]. On exposure, most individuals develop efficient immunity against *M. leprae* with no signs of clinical disease. However, in a small proportion of exposed individuals, leprosy manifests in a spectrum of clinical forms, ranging from the localized, tuberculoid to the systemic, lepromatous disease [2], associated with a Th1 or Th2 type of immune response presented against the pathogen, respectively [3]. Despite a global leprosy elimination effort coordinated by the World Health Organization since 1991, the disease persists in 118 countries, and Brazil, Nepal, and Timor-Leste have still not achieved the elimination goal of a prevalence rate <1 case per 10,000 persons [1].

The clinical outcome of infection is the result of interaction between variables related to the host, the path-

Received 21 August 2009; accepted 3 December 2009; electronically published 13 April 2010.

[a] F.P.L. and R.I.W. contributed equally to this work.

[b] A.A. and M.T.M. share senior authorship.

Reprints or correspondence: Dr Marcelo T. Mira, Advanced Molecular Investigation Core, Pontifical Catholic University of Paraná, Imaculada Conceição 1155, Prado Velho, Curitiba, Paraná, CEP 80215–901, Brazil (m.mira@pucpr.br).

The Journal of Infectious Diseases 2010; 201(10):1598–1605

0022-1899/2010/20110-0021$15.00
DOI: 10.1086/652007

Potential conflicts of interest: none reported.

Financial support: UNICEF/UNDP/World Bank/WHO Special Program for Research and Training in Tropical Diseases; Brazilian Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (projects 410470/2006-6 [MCT/CNPq/MS-SCTIE-DECIT 25/2006, Estudo de Doenças Negligenciadas] and 401057/2005-4 [MCT/CNPq/MS-SCTIE-DECIT 35/2005, Estudo da Hanseníase]); Order of Malta Grants for Leprosy Research—2008; INSERM-CNPq joint program (project 490844/2008-1 [ASCIN/CNPq 61/2008, Convênios Bilaterais Europa]); Rockefeller University Center for Clinical and Translational Science (grant 5UL1RR024143-03 to Laboratory of Human Genetics of Infectious Disease) and Rockefeller University (Laboratory of Human Genetics of Infectious Disease); Agence Nationale de la Recherche, Ministère Français de l'Éducation Nationale de la Recherche et de la Technologie (A.A., L.A.); CAPES/PROSUP (scholarships to F.P.L. and R.I.W.); and CAPES/PDEE (scholarship to R.I.W.).

ogen, and the environment [4]. In this complex scenario, an important role for host genetic risk factors controlling susceptibility to disease has become increasingly evident. For example, the following have been implicated with leprosy phenotypes: variants of HLA class I and II; HLA-linked genes *MICA* and *MICB* [5, 6], *TNFA* [7, 8], and *KIR* [9], among others; and non-HLA genes *IL-10* [10], *VDR* [11], and *SLC11A1* (formerly *NRAMP1*) [12, 13]. In addition, 3 model-free genome-wide linkage studies involving different leprosy phenotypes have been performed, localizing loci harboring susceptibility genes on chromosomal regions 10p13 [14], 17q11–q21 [15], 6q25–q26, and 6p21 [16]. However, new leprosy candidate genes have emerged only from chromosomes 6q25–q27 and 6p21: high-density association mapping demonstrated susceptibility variants for leprosy per se (ie, the disease regardless its clinical form) located on the shared regulatory region of the *PARK2* and *PACRG* genes on chromosome 6q25–q27 [17, 18] and a functional regulatory site of the *LTA* gene located on 6p21 [19].

In addition to molecular, DNA-based studies, classic observational genetic epidemiology tools, such as complex segregation analysis (CSA), have been used to advance our understanding of the genetic basis of complex diseases. Several CSA studies have been performed in leprosy-affected population samples of different ethnic background using a variety of analytical methods, with the objective to identify the best-fit model of inheritance for leprosy phenotypes. Although a few studies could not distinguish between environmental and genetic effects or suggested a predominant environmental effect controlling susceptibility to disease [20, 21], more indicate the existence of a recessive major gene (MG) controlling susceptibility to lepromatous disease [22], nonlepromatous disease [23–25], and leprosy per se [25, 26].

Of note, the term "major gene" means that its effect is important enough to be distinguished from other genes effects but does not assume it is the only gene involved. In contrast, 2 studies concluded in favor of the existence of a dominant or codominant MG related to susceptibility to leprosy per se in families from Thailand [27] and Vietnam [28]. More recently, Shaw et al [7] described a best-fit model for susceptibility to leprosy per se that included 2 loci, one a recessive MG and the other a recessive modifier. The model was then used in parametric linkage followed by association analysis, using the same family sample that confirmed the role of HLA class II and III alleles, located on chromosome 6p21, in leprosy control. Such a strategy (ie, CSA followed by model-based linkage analysis) was also successful in identifying human genetic factors controlling other infectious diseases, as in the seminal works performed in schistosomiasis [29–31] and tuberculosis [32], highlighting the power of this approach.

Importantly, none of the other leprosy studies have been performed using an isolated, hyperendemic population homogeneously exposed to the disease. In addition, enrollment strategies are always prone to ascertainment bias that is completely controlled only if the entire target population is included. For example, complete ascertainment was applied in 2 CSA studies performed for susceptibility to human T lymphotropic virus type 1 and human herpes virus type 8 in an isolated hyperendemic population from French Guiana, with notable results [33, 34]. In leprosy, the CSA conducted in 1988 in the Desirade Island involved 1600 individuals recruited from an area with disease frequency of ~30 cases per 1000 persons, the highest at the time [25]. To date, the Desirade study is the closest to a theoretical ideal design for a CSA on leprosy, in which the entire population from an isolated hyperendemic area would be recruited.

Here we present the results of a CSA performed in a unique collection of leprosy multiplex, multigenerational families corresponding to the complete population of the Colony of Santo Antônio do Prata (the Prata Colony), located in the Amazonic state of Pará, in the north of Brazil. The Prata is a former leprosy colony created in the early 1920s to isolate leprosy-affected individuals. Isolation was compulsory until 1962; however, the population of the colony remains isolated, probably owing to the strong stigma still associated with the disease. Preliminary assessment indicated the highest disease prevalence ever reported worldwide, homogenous environmental and socioeconomic variables, and a predominance of a mixed ethnic group.

MATERIALS AND METHODS

Population recruitment. The entire population of the Prata Colony was contacted by the research team over a period from April 2006 to December 2007. The investigation was approved by the Research Ethics Committee of the Pontifical Catholic University of Paraná, the Brazilian National Board for Ethics in Research, and the Ethics Research Committee of the World Health Organization.

To assure that all households were contacted, visits were planned according to a previously existing subdivision of the colony into 6 sectors. All households in 1 sector were visited before the entire research team moved on to the next sector. The procedure was repeated until all sectors were systematically included. All adult individuals were independently interviewed by trained personnel. Information regarding individuals <18 years old was provided by the parents.

Epidemiological data collection and phenotype definition. Demographic data (sex, age, ethnicity, and location of household), and affection status was collected for all individuals, who were classified as affected or not by leprosy per se. Ethnicity was defined by trained personnel based on the following phenotypic characteristics: skin color, hair type, and conformation of the nose and lips [35]. Age and disease status were self-

Table 1. Segregation Analysis of Leprosy Per Se in the Prata Population

Model	Q_a	α_{AA}	α_{Aa}	α_{aa}	γPO		γSS		β sex	β log age	τAAa	τAaa	τaaa	−2lnL + C
					Unaffected	Affected	Unaffected	Affected						
I. Sporadic	(0)	−3.67	[αAA]	[αAA]	(0)	(0)	(0)	(0)	−0.52	2.27	…	…	…	55
II. FD														
a. PO	(0)	−3.78	[αAA]	[αAA]	−0.44	0.62	(0)	(0)	−0.51	2.33	…	…	…	42
b. SS	(0)	−3.85	[αAA]	[αAA]	(0)	(0)	−0.06	0.61	−0.51	2.24	…	…	…	46
c. PO + SS	(0)	−3.90	[αAA]	[αAA]	−0.23	0.56	−0.08	0.49	−0.48	2.31	…	…	…	33
III. MG and FD (PO + SS)	0.21	−7.62	−5.38	0.29	−0.23	0.28	−0.06	0.06	−0.66	4.50	(0)	(0.5)	(1)	7
IV. MG														
a. Codominant	0.22	−7.98	−5.94	−0.68	(0)	(0)	(0)	(0)	−0.75	4.26	(0)	(0.5)	(1)	5
b. Recessive	0.27	−6.49	[−6.49]	−0.97	(0)	(0)	(0)	(0)	−0.72	3.82	(0)	(0.5)	(1)	14
V. Absence of transmission	0.06	−6.76	−6.75	−0.87	(0)	(0)	(0)	(0)	−0.69	4.32	0.68	[τAAA]	[τAAA]	27
VI. General transmission	0.15	−7.59	−5.60	−1.05	(0)	(0)	(0)	(0)	−0.64	4.13	0.00	0.31	1.00	0

NOTE. $C = -1015$, corresponding to twice the logarithm of the likelihood (2lnL) of the best-fitting model (model VI); FD, familial dependency; MG, major gene; PO, parent-offspring; Q frequency of leprosy predisposing allele *a*; SS, sibling-sibling; α, baseline risk of being affected on a logit scale corresponding to 3 genotypes: *AA* (αAA), *Aa* (αAa), and *aa* (αaa); β, covariable regression coefficients; γPO and γSS, regression coefficients associated with familial dependencies PO and SS, respectively; τAAa, τAaa, and τaaa, probabilities of transmitting *a* for individuals *AA, Aa,* and *aa,* respectively. Terms in brackets represent parameters fixed to the same value as the preceding estimated parameter; terms in parentheses, fixed parameters for hypothesis; ellipses, irrelevant parameters in the model.

reported at the interview. For self-reported affected individuals, both affected status and age were confirmed on cross-checking using 3 independent sources available at the local health care center: the patient's medical record, a copy of the compulsory notification form, and a registry book used for treatment follow-up.

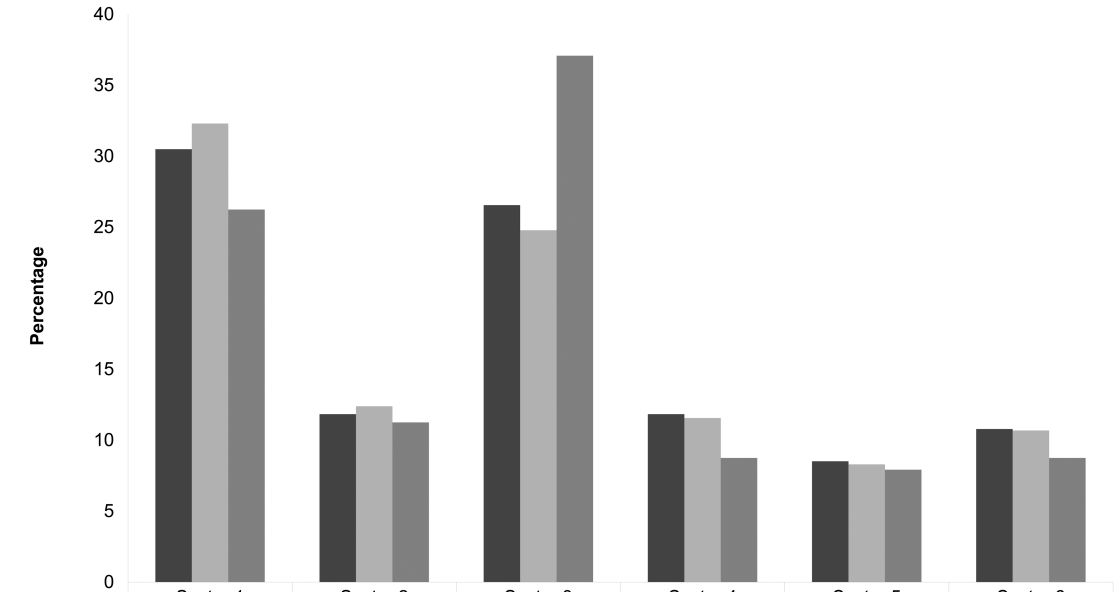
Pedigree reconstruction. To allow for pedigree reconstruction, parental information was also obtained for all individuals. Pedigree drawing was performed using Cranefoot software (version 3.2.2) [36]. Information regarding members of informative pedigrees who were not living in the colony (moved or deceased) and therefore not personally interviewed was used on confirmation of epidemiological and clinical status, as described above. If disease status could not be checked, these individuals were coded as "unknown." One very large pedigree was decomposed, due to computational limitations, according to systematic criteria that reduced the loss of information [37], before being included in the CSA. If necessary to link nuclear pedigrees, individuals were duplicated or created ("dummies"), a procedure that did not affect the outcome of either the epidemiological description or the CSA.

Statistical methods. The phenotype of interest for the CSA was a binary trait—that is, affected or not affected by leprosy. Age was coded as the natural logarithm of age in years, which was the best-fitting age function based on Akaike's information criterion [38]. Ethnicity was coded as a categorical variable with 3 classes: white (reference class), black, and mixed race. The analyses were performed using logistic regression analysis as implemented in the LOGISTIC procedure of SAS software, version 9.1 (SAS Institute).

Segregation analysis was done by using the regressive logistic model [39], which specified a regression relationship between the probability that a person would be affected and a set of explanatory variables, including MG, phenotype of preceding relatives, and other covariates. The regressive model allowed analysis of large families as a whole, simultaneous estimation of genetic and risk factor effects, and consideration of different patterns of familial correlations for leprosy status. Sporadic transmission (Table 1, model I) includes only the nongenetic covariates with a significant effect on disease susceptibility (for example, the question whether the phenotype of an individual is influenced by those of other family members is not addressed).

In addition to the significant covariates, familial correlations (Table 1, model II) are included in the model, using the class D pattern of familial dependencies (FD) [39]. Four types of phenotypic FD were considered: father-mother (FM), father-offspring (FO), mother-offspring (MO), and sibling-sibling (SS), with corresponding regression coefficients denoted as Γ'FM, Γ'FO, Γ'MO, and Γ'SS, respectively. To account for the phenotype "unaffected," each Γ' parameter is a vector of 2 coefficients [39]: Γ'FM = γFMunaf and γFMaff, Γ'FO = γFOunaf and γFOaff, Γ'MO = γMOunaf and γMOaff, and Γ'SS = γSSunaf and γSSaff. γXunaf and γXaff can be interpreted as the impact on the risk that an individual will develop leprosy conditional on the observation that he or she has a relative of type X affected and unaffected, respectively. The higher the parameter, the higher the increase in the risk of developing leprosy and symmetrically.

The presence of significant FDs can be due to any source of unmeasured shared environmental factor (eg, higher exposure to the microbe due to the geographic location of the family). In particular, this should be suspected in the presence of sig-



	Sector 1	Sector 2	Sector 3	Sector 4	Sector 5	Sector 6
Households	30.5	11.83	26.56	11.83	8.51	10.79
Individuals	32.3	12.39	24.78	11.56	8.29	10.68
Affecteds	26.25	11.25	37.08	8.75	7.92	8.75

Figure 1. Distribution of households, all individuals, and leprosy-affected individuals across the 6 sectors of the colony. Values inside the cells correspond to the exact percentages in each sector of the colony. An apparent increase in leprosy cases in sector 3 is due to the location in this sector of the permanent shelter for elderly patients with leprosy who have no relatives living in the colony.

nificant spouse-spouse dependencies. To rule out this possibility, an MG effect is included in the model (Table 1, model III). The estimated parameters of the MG are Q (the frequency of allele *a* predisposing someone to be affected by leprosy) and α_{AA}, α_{Aa}, and α_{aa} (the 3 baseline risks of being affected on the logit scale for the 3 genotypes *AA, Aa,* and *aa,* respectively).

Several studies [40–42] have shown that the identification of an MG effect (ie, a mix of 3 distributions, 1 per genotype, explains the data better than a single distribution) was not specific enough to demonstrate the existence of an MG. They have suggested that 2 additional tests were needed. These 2 tests rely on the parent-offspring (PO) transmission pattern of the major effect. The PO transmission is parameterized in terms of the 3 classic transmission probabilities, as defined by Elston and Stewart [43]: τAAa, τAaa, and τaaa, which denote the probabilities of transmitting *a* for individuals *AA, Aa,* and *aa,* respectively [43]. Mendelian transmission is obtained by setting $\tau aaa = 1$, $\tau Aaa = 0.5$, and $\tau AAa = 0$; in this case the major effect is actually an MG.

Two additional models including a major effect were considered: (1) an "absence of transmission" model (model IV) in which 3 types of individuals (*aa, Aa,* and *AA*) are specified but in which absence of PO transmission is obtained by setting $\tau aaa = \tau Aaa = \tau AAa$; and (2) a more general transmission model (model V) in which the 3 τ values are estimated. Segregation of an MG can be inferred if we fail to reject mendelian transmission of the major effect in comparison with the general transmission model and we reject the nontransmission hypothesis, also in comparison with the general transmission model (this latter test rules out the possibility that the failure to reject mendelian transmission of the major effect was due to lack of power). Parameter estimation and hypothesis testing were performed using classic likelihood strategies. Because of the study design (ie, exhaustive collection of the whole population), there was no need for ascertainment correction. CSA was performed as implemented in REGRESS software, version 4.8 [44], which incorporates the regressive approach into the LINKAGE package [45].

RESULTS

Population description. Of the 2007 individuals invited, only 2 declined to participate; therefore, a total of 2005 individuals (1012 male and 993 female) signed an informed consent and were enrolled in the study. Leprosy prevalence was 12.82% (257 confirmed cases among 2005 individuals), distributed equally throughout the colony (Figure 1). For the CSA, after removal of single subjects and uninformative families, a total of 1867 individuals (225 affected and 1642 unaffected) were included, distributed in 112 pedigrees. Therefore, the observed prevalence in the CSA families was 12.05%, very close to the global prevalence. All clinical forms of disease were represented, with 104

(40.5%) of the cases being lepromatous, 53 (20.6%) tuberculoid, 53 (20.6%) borderline, and 47 (18.3%) of indeterminate clinical form. The mean and median ages at diagnosis were 27.2 and 25 years, respectively.

Analysis of covariates. In the univariate analysis, age had a strong effect on disease status (Table 2). The distribution of disease according to age class shows an increase in the proportion of leprosy-affected individuals in each age class as age increases up to >60 years old. Males were affected more often than females (13.8% vs 10.2%, respectively). The distribution of leprosy across ethnicity shows a higher proportion of affected among black subjects (19.2%) than among those who were white (10.6%) or mixed race (11%) ($P = .002$). Multivariate logistic regression analysis confirmed the strong effect of age ($P<10^{-4}$) and sex ($P = .002$); however, the ethnicity effect became not significant, because it was due to an age-confounding effect.

CSA results. Results of the CSA are shown in Table 1. Because we never observed significant FM correlation, these results were not included in the table. Moreover, in the different analyses we never observed a significant difference between FO and MO dependency; therefore, only a global PO dependency (model IIa) ($\gamma_{FOunaf/aff} = \gamma_{MOunaf/aff} = \gamma_{POunaf/aff}$) was considered.

There was evidence of a strong FD, because the sporadic model without FD was rejected against the model that included PO plus SS correlation (model I vs IIc, $\chi^2(2df) = 22$; $P = 10^{-5}$). When the model with SS correlation was compared with the model with PO plus SS correlation, the first model was rejected (model IIb vs IIc, $\chi^2(2df) = 13$; $P = .0015$). Of note, the absence of FM correlation is a good indicator that the significant SS and PO correlations are not caused by an unmeasured environmental factor shared by the family members. The inclusion of a codominant major effect to PO plus SS FDs resulted in a highly significantly better fit (model IIc vs III, $\chi^2(3df) = 26$; $P = 10^{-5}$). Interestingly, removal of residual PO plus SS dependency did not significantly affect the fitness (model III vs IVa, $\chi^2(4df) = 2$; $P = .73$). The codominant model was significantly better than a recessive model (IVb vs IVa, $\chi^2(1df) = 9$; $P = .0027$). Finally, the transmission of the codominant major effect was compatible with the mendelian hypothesis (IVa vs VI; $\chi^2(3df) = 5$; $P = .17$), and the hypothesis of no transmission was rejected (V vs VI, $\chi^2(2df) = 27$; $P = 10^{-6}$); In summary, the CSA favored a codominant MG influencing leprosy per se. Under this codominant MG model ($\alpha_{AA} \neq \alpha_{Aa} \neq \alpha_{aa}$), the frequency of the predisposing allele *a* was estimated at 0.22. The same analysis was performed using the mixed population only, with similar results (data not shown).

Table 2. Factors Influencing the Onset of Leprosy in Univariate and Multivariate Logistic Regression Analysis

Covariate	Sample size, no.[a]	No. (%) of affected subjects	*P*	
			Univariate analysis	Multivariate analysis
Age, years			$<10^{-4}$	$<10^{-4}$
0–20	1019	28 (2.7)		
21–40	520	54 (10.38)		
41–60	233	86 (36.9)		
>60	95	57 (60)		
Sex			.01	.002
Male	933	129 (13.8)		
Female	934	96 (10.2)		
Ethnicity			.002	.365
White	187	20 (10.6)		
Mixed race	1451	161 (11)	.08[b]	.57
Black	229	44 (19.2)	.01[c]	.67

[a] Samples represent only informative individuals included in the complex segregation analysis.
[b] Comparison between mixed-race and white subjects.
[c] Comparison between black and white subjects.

DISCUSSION

Even though several initiatives have successfully identified genes associated with host susceptibility to leprosy phenotypes [5, 46], basic answers regarding the genetic model involved are still not precisely known, as reflected by discordant results of several CSA studies [7, 20–28]. These discrepancies may be due to the tremendously difficult task of accounting for all environmental, socioeconomic and cultural variables involved, as well as limitations regarding population enrollment. For example, the large number of individuals required for a powerful CSA often imply population sample heterogeneity and ascertainment bias, unless the population is exposed to very high disease risk and/or is entirely included in the analysis.

The objective of this study was to conduct a CSA using data obtained from the entire population of the Colony of Santo Antônio do Prata, a former leprosy colony with unique characteristics, including very high disease frequency with equal distribution across the community (indicating homogenous exposure), a high degree of isolation, and homogenous socioeconomic, environmental, and ethnic backgrounds. We hypothesize that, because the colony was founded almost exclusively by leprosy-affected individuals, genetic risk factors are enriched within the community, composed today of a large number of extended, multiplex, multigenerational families. In this context, the Prata CSA detected the existence of an MG controlling susceptibility to leprosy per se, inherited after a codominant model with the frequency of the predisposing allele *a* estimated at 0.22.

Our results are not far from a recessive model, as observed in a previous CSA using a population from the Desirade Islands [25]—small discrepancies may be due to ascertainment correction procedures adopted in the Desirade and not necessary in the Prata study—and such recessive-like models may be specific to isolated population that descended from inhabitants

isolated because they had leprosy. Under this codominant, recessive-like model, ~5% of the population is *aa* homozygous and therefore highly predisposed to leprosy. As shown in Figure 2, in addition to the strong effect of the underlying genotype, disease penetrance was also influenced by sex and age; as an example, at age 30 years, penetrance ranged from 0.85 for *aa* men to 0.0006 for *AA* women.

A previous, classic study of a population from Malawi demonstrated an increased risk for leprosy among dwelling contacts of a leprosy case patient. Disease risk decreased for household contacts and was the lowest for no contacts, leading the authors to conclude that there was an environmental or behavioral cause for familial aggregation [47]. In the Prata colony, the MG effect was detected in the presence of FD; however, when FD was removed from the model, the MG effect became better noticed, indicating that the observed familial aggregation was entirely explained by genetics, as first suggested by the absence of FM correlation. Interestingly, the median age at diagnosis of leprosy in the Prata (25 years) is lower than the Brazilian national median of 39 years [48]. Although early diagnosis of cases in a hyperendemic population under constant monitoring is somewhat expected, this observation is compatible with recent findings suggesting that early leprosy cases are more likely to have a genetic basis [19]. Unfortunately, more precise inference about the nature of the MG effect observed is limited by the inability of CSA to distinguish between 1 single, codominant gene with a very strong effect and several codominant genes with milder effects that play additively on the risk. For example, an exciting possibility is that the MG effect observed is largely due to the *PARK2/PACRG* [17, 18] and *LTA* [19] effects, described elsewhere in different populations. Alternatively—and not less exciting—new major susceptibility genes are yet to be described. For further investigation, molecular, DNA-based studies are necessary.

It is widely accepted that genetic control of the pathogenesis of leprosy follows a 2-step model, in which different genes affect susceptibility to leprosy per se and clinical manifestation of disease [49]. Our study focused only on the phenotype for leprosy per se, because of 2 limitations: (1) the distribution of clinical forms of disease must be interpreted as an approximate estimate, given the impossibility of confirming this very complex phenotype retrospectively, based exclusively on heterogeneous clinical information recorded by different physicians over a time period that spans almost 100 years; and (2) the smaller number of cases in each clinical class significantly decreases the power of the analysis. Therefore, the MG effect detected here is likely to modify the onset of leprosy independent of its clinical form [5].

The results of the Prata CSA support the hypothesis of an enrichment of genetic risk factors in this particular population, making it ideal for future gene mapping analysis. Parameters of the genetic model generated by the CSA may be used for

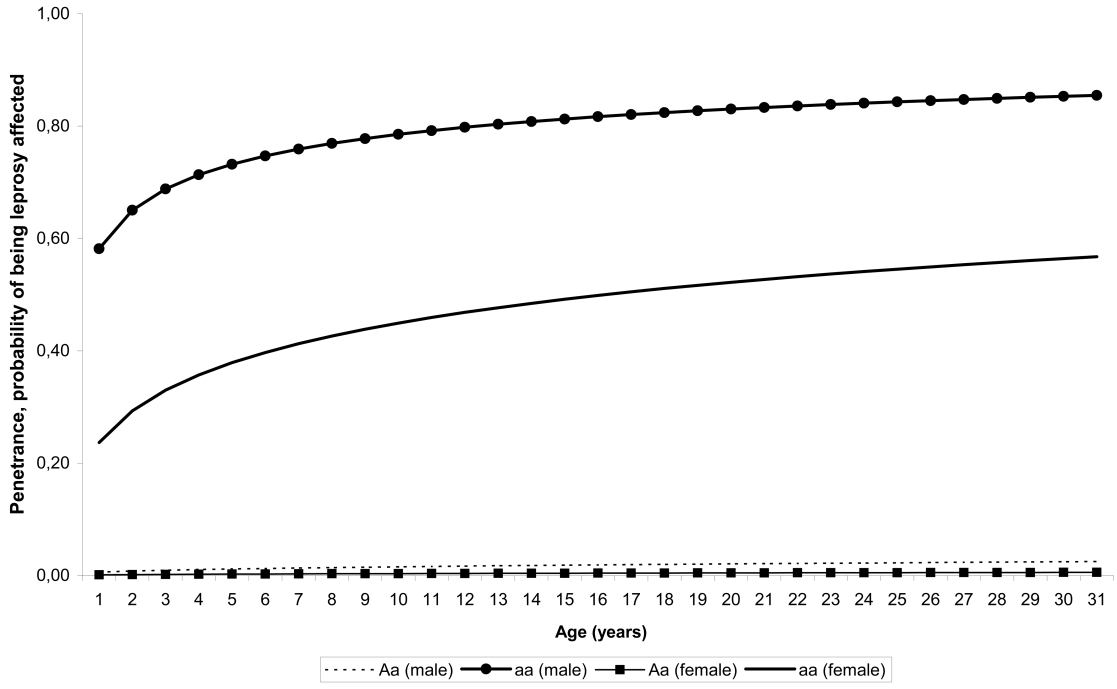


Figure 2. Penetrance (ie, probability of being leprosy affected) according to age, genotype, and sex, as predicted in model IVa (Table 2); *a* is the leprosy-predisposing allele. Exact penetrance values at age 30 years are as follows: females, 0.56 for *aa* and 0.005 for *Aa*; males, 0.85 for *aa* and 0.024 for *Aa*. Penetrance for homozygous *AA* individuals at age 30 years was 0.003 for males and 0.0006 for females (not shown).

model-based linkage analysis followed by high-density association mapping using the already-mapped Prata families. Specifying the nature of the genetic component controlling susceptibility to leprosy may have a profound impact on the development of new strategies for prevention, diagnosis, and treatment of the disease. For example, individuals genetically at risk and living under high exposure would be candidates for chemoprophylaxis, a highly debated issue today in Brazil and other countries of endemicity [50]. Genetic susceptibility could also help in the understanding of issues such as disease relapse and drug resistance, the mechanisms of which are still widely unknown in leprosy. Finally, a deeper insight into the molecular basis of leprosy pathogenesis would certainly improve the understanding not only of this particular disease but of infection in general.

Acknowledgments

We thank all affected individuals and their families for agreeing to participate in the study. We also thank the local team from the Federal University of Pará, as well as the leaders and health workers of the Prata Colony for their help in guiding our team through the community, socially and geographically. A special acknowledgment is addressed to Dr Erwin Schurr for his critical reading and valuable suggestions during the development of the study and preparation of the manuscript.

References

1. World Health Organization. Global leprosy situation. Wkly Epidemiol Rec **2008**; 83:293–300.
2. Ridley DS, Jopling WH. Classification of leprosy according to immunity: a five-group system. Int J Lepr Other Mycobact Dis **1966**; 34: 255–273.
3. Scollard DM, Adams LB, Gillis TP, Krahenbuhl JL, Truman RW, Williams DL. The continuing challenges of leprosy. Clin Microbiol Rev **2006**; 19:338–381.
4. Casanova JL, Abel L. The human model: a genetic dissection of immunity to infection in natural conditions. Nat Rev Immunol **2004**; 4: 55–66.
5. Mira MT. Genetic host resistance and susceptibility to leprosy. Microbes Infect **2006**; 8:1124–1131.
6. Tosh K, Ravikumar M, Bell JT, Meisner S, Hill AV, Pitchappan R. Variation in MICA and MICB genes and enhanced susceptibility to paucibacillary leprosy in South India. Hum Mol Genet **2006**; 15:2880–2887.
7. Shaw MA, Donaldson IJ, Collins A, et al. Association and linkage of leprosy phenotypes with HLA class II and tumour necrosis factor genes. Genes Immun **2001**; 2:196–204.
8. Santos AR, Suffys PN, Vanderborght PR, et al. Role of tumor necrosis factor-alpha and interleukin-10 promoter gene polymorphisms in leprosy. J Infect Dis **2002**; 186:1687–1691.
9. Franceschi DS, Mazini PS, Rudnick CC, et al. Association between killer-cell immunoglobulin-like receptor genotypes and leprosy in Brazil. Tissue Antigens **2008**; 72:478–482.
10. Pereira AC, Brito-de-Souza VN, Cardoso CC, et al. Genetic, epidemiological and biological analysis of interleukin-10 promoter single-nucleotide polymorphisms suggests a definitive role for −819C/T in leprosy susceptibility. Genes Immun **2009**; 10:174–180.
11. Fitness J, Floyd S, Warndorff DK, et al. Large-scale candidate gene study of leprosy susceptibility in the Karonga district of northern Malawi. Am J Trop Med Hyg **2004**; 71:330–340.
12. Abel L, Sánchez FO, Oberti J, et al. Susceptibility to leprosy is linked to the human NRAMP1 gene. J Infect Dis **1998**; 177:133–145.
13. Alcaïs A, Sanchez FO, Thuc NV, et al. Granulomatous reaction to intradermal injection of lepromin (Mitsuda reaction) is linked to the human NRAMP1 gene in Vietnamese leprosy sibships. J Infect Dis **2000**; 181:302–308.
14. Siddiqui MR, Meisner S, Tosh K, et al. A major susceptibility locus for leprosy in India maps to chromosome 10p13. Nat Genet **2001**; 27:439–441.
15. Miller EN, Jamieson SE, Joberty C, et al. Genome-wide scans for leprosy and tuberculosis susceptibility genes in Brazilians. Genes Immun **2004**; 5:63–67.
16. Mira MT, Alcaïs A, Van Thuc N, et al. Chromosome 6q25 is linked to susceptibility to leprosy in a Vietnamese population. Nat Genet **2003**; 33:412–415.
17. Mira MT, Alcaïs A, Nguyen VT, et al. Susceptibility to leprosy is associated with PARK2 and PACRG. Nature **2004**; 427:636–640.
18. Alcaïs A, Mira M, Casanova JL, Schurr E, Abel L. Genetic dissection of immunity in leprosy. Curr Opin Immunol **2005**; 17:44–48.
19. Alcaïs A, Alter A, Antoni G, et al. Stepwise replication identifies a low-producing lymphotoxin-alpha allele as a major risk factor for early-onset leprosy. Nat Genet **2007**; 39:517–522.
20. Shields ED, Russell DA, Pericak-Vance MA. Genetic epidemiology of the susceptibility to leprosy. J Clin Invest **1987**; 79:1139–1143.
21. Serjeantson S, Wilson SR, Keats BJ. The genetics of leprosy. Ann Hum Biol **1979**; 6:375–393.
22. Smith DG. The genetic hypothesis for susceptibility to lepromatous leprosy. Hum Genet **1979**; 50:163–177.
23. Haile RW, Iselius L, Fine PE, Morton NE. Segregation and linkage analysis of 72 leprosy pedigrees. Hum Hered **1985**; 35:43–52.
24. Demenais F, Goulet V, Feingold N, et al. Genetic study of leprosy in a Caribbean island: Desirade. In: Skamene E, ed. Genetic control of host resistance to infection and malignancy. New York: Alan R. Liss, **1985**:319–324.
25. Abel L, Demenais F. Detection of MGs for susceptibility to leprosy and its subtypes in a Caribbean island: Desirade island. Am J Hum Genet **1988**; 42:256–266.
26. Feitosa MF, Borecki I, Krieger H, Beiguelman B, Rao DC. The genetic epidemiology of leprosy in a Brazilian population. Am J Hum Genet **1995**; 56:1179–1185.
27. Wagener DK, Schauf V, Nelson KE, Scollard D, Brown A, Smith T. Segregation analysis of leprosy in families of northern Thailand. Genet Epidemiol **1988**; 5:95–105.
28. Abel L, Vu DL, Oberti J, et al. Complex segregation analysis of leprosy in southern Vietnam. Genet Epidemiol **1995**; 12:63–82.
29. Abel L, Demenais F, Prata A, Souza AE, Dessein A. Evidence for the segregation of a major gene in human susceptibility/resistance to infection by *Schistosoma mansoni*. Am J Hum Genet **1991**; 48:959–970.
30. Marquet S, Abel L, Hillaire D, et al. Genetic localization of a locus controlling the intensity of infection by *Schistosoma mansoni* on chromosome 5q31-q33. Nat Genet **1996**; 14:181–184.
31. Dessein AJ, Hillaire D, Elwali NE, et al. Severe hepatic fibrosis in *Schistosoma mansoni* infection is controlled by a major locus that is closely linked to the interferon-gamma receptor gene. Am J Hum Genet **1999**; 65:709–721.
32. Greenwood CM, Fujiwara TM, Boothroyd LJ, et al. Linkage of tuberculosis to chromosome 2q35 loci, including NRAMP1, in a large aboriginal Canadian family. Am J Hum Genet **2000**; 67:405–416.
33. Plancoulaine S, Gessain A, Joubert M, et al. Detection of a major gene predisposing to human T lymphotropic virus type I infection in children among an endemic population of African origin. J Infect Dis **2000**; 182:405–412.
34. Plancoulaine S, Gessain A, van Beveren M, Tortevoye P, Abel L. Evidence for a recessive major gene predisposing to human herpesvirus

8 (HHV-8) infection in a population in which HHV-8 is endemic. J Infect Dis **2003**; 187:1944–1950.
35. Krieger H, Morton NE, Mi MP, Azevêdo E, Freire-Maia A, Yasuda N. Racial admixture in north-eastern Brazil. Ann Hum Genet **1965**; 29: 113–125.
36. Mäkinen VP, Parkkonen M, Wessman M, Groop PH, Kanninen T, Kaski K. High-throughput pedigree drawing. Eur J Hum Genet **2005**; 13:987–989.
37. Liu F, Kirichenko A, Axenovich TI, van Duijn CM, Aulchenko YS. An approach for cutting large and complex pedigrees for linkage analysis. Eur J Hum Genet **2008**; 16:854–860.
38. Akaike H. A new look at statistical model identification. IEEE Trans Autom Control **1974**; 19:716–723.
39. Bonney GE. Regressive logistic models for familial disease and other binary traits. Biometrics **1986**; 42:611–625.
40. Demenais F, Lathrop M, Lalouel JM. Robustness and power of the unified model in the analysis of quantitative measurements. Am J Hum Genet **1986**; 38:228–234.
41. Demenais F, Abel L. Robustness of the unified model to shared environmental effects in the analysis of dichotomous traits. Genet Epidemiol **1989**; 6:229–234.
42. Demenais FM, Laing AE, Bonney GE. Numerical comparisons of two formulations of the logistic regressive models with the mixed model in segregation analysis of discrete traits. Genet Epidemiol **1992**; 9:419–435.
43. Elston RC, Stewart J. A general model for the genetic analysis of pedigree data. Hum Hered **1971**; 21:523–542.
44. Demenais F, Lathrop GM. REGRESS: a computer program including the regressive approach into the LINKAGE programs. Genet Epidemiol **1994**; 11:291.
45. Lathrop GM, Lalouel JM, Julier C, Ott J. Strategies for multilocus linkage analysis in humans. Proc Natl Acad Sci USA **1984**; 81:3443–3446.
46. Alter A, Alcaïs A, Abel L, Schurr E. Leprosy as a genetic model for susceptibility to common infectious diseases. Hum Genet **2008**; 123:227–235.
47. Fine PE, Sterne JA, Pönnighaus JM, et al. Household and dwelling contact as risk factors for leprosy in northern Malawi. Am J Epidemiol **1997**; 146:91–102.
48. Ministério da Saúde. Hanseníase. http://portal.saude.gov.br/portal/saude/Gestor/visualizar_texto.cfm?idtxt = 27445. Accessed 21 August 2009.
49. de Vries RR, Fat RF, Nijenhuis LE, van Rood JJ. HLA-linked genetic control of host response to *Mycobacterium leprae*. Lancet **1976**; 2:1328–1330.
50. Goulart IM, Goulart LR. Leprosy: diagnostic and control challenges for a worldwide disease. Arch Dermatol Res **2008**; 300:269–290.