npg

ARTICLE

Dengue hemorrhagic fever is associated with polymorphisms in JAK1

Luciano K Silva[1,7], Ronald E Blanton*[,2,7], Antonio R Parrado[3], Paulo S Melo[1], Vanessa G Morato[4], Eliana AG Reis[1], Juarez P Dias[5], Jesuina M Castro[5], Pedro FC Vasconcelos[6], Katrina AB Goddard[3], Maurício L Barreto[4], Mitermayer G Reis[1] and M Glória Teixeira[4]

To identify genes associated with the clinical presentation of dengue, 50 cases of probable or possible dengue hemorrhagic fever (DHF), 236 dengue fever (DF), and 236 asymptomatic infections were genotyped for 593 single-nucleotide polymorphisms (SNPs) in 56 genes across the type 1 interferon (IFN) response pathway as well as other important candidate genes. By single locus analysis comparing DHF with DF, 11 of the 51 markers with $P < 0.05$ were in the JAK1 gene. Five markers were significantly associated by false discovery rate criteria ($q < 0.20$ when $P < 6 \times 10^{-4}$). The JAK1 SNPs showed differential distribution by ethnicity and ancestry consistent with epidemiologic observations in the Americas. The association remained significant after controlling for ancestry and income. No association was observed with markers in the gene encoding CD209 (DC-SIGN). An association between DHF and JAK1 polymorphisms is in agreement with expression profiles showing generalized decreased type 1 IFN-stimulated gene expression in these patients.
European Journal of Human Genetics (2010) 18, 1221–1227; doi:10.1038/ejhg.2010.98; published online 30 June 2010

INTRODUCTION

Dengue is one of the most important arboviral infections because of its global distribution, its prevalence, and the gravity of its most severe presentation, dengue hemorrhagic fever (DHF). Most infections are asymptomatic,[1,2] but 100 million each year have dengue fever (DF), which frequently presents as an acute, weeklong febrile illness with severe joint pains that is followed by a month of generalized weakness and lack of well being. An additional 1% of those symptomatic with dengue virus (DENV) infection will develop DHF, where there is spontaneous bleeding and at times shock and death. Most cases of DHF occur in individuals who have had earlier infection with one or more of the four viral serotypes or who have non-neutralizing maternal antibodies.[1] In many endemic areas, such as Brazil, the majority of individuals have had multiple infections,[2] yet, only 2–4% of those with secondary infections will develop DHF.[3] Additional host or viral factors, therefore, must be involved.

Although there may be additional epidemiological associations between DHF, such as age, some chronic diseases, and nutritional status[4], a contribution from host genetics to the risk for DHF is suggested by ethnicity-based differences in disease rates.[5,6] Populations of African origin[7,8] as well as other ethnic groups[5,7–9] seem to be relatively resistant. Genetic association studies have also pointed associations with multiple HLA loci[10] and, most recently, polymorphisms in the gene encoding CD209 (DC-SIGN).[11]

In 2002, the murine *flav* locus, a locus responsible for resistance to all flaviviruses in mice, was identified as the OAS1b gene.[12,13] As this gene in mice differs substantially from its homolog in human beings, we generalized this to be the evidence that there are elements in the type 1 interferon (IFN) response pathway that can significantly influence the clinical presentation of dengue. We, therefore, chose to study a range of genetic candidates, in particular those of the IFNα response pathway, while controlling for important non-genetic variables also associated with DHF.

MATERIALS AND METHODS

Recruitment

During Brazil's most active dengue transmission year to date (2002–2003), the State Secretariat of Health of Bahia identified 91 potential cases of DHF by active surveillance in the hospitals of the city of Salvador. Of these, 65 met the initial study criteria for suspected cases, that is a dengue-like syndrome, spontaneous bleeding, and thrombocytopenia. Using World Health Organization, population-based criteria[6] 31 had evidence of a capillary leak syndrome as defined by Brazilian Ministry of Health norms for hemoconcentration (http://dtr2001.saude.gov.br/svs/pub/GBDIP/guia_bolso_4ed.pdf) and a hemagglutination inhibition (HI) serology consistent with dengue infection. HI is not specific among flaviviruses, and the serology was not conducted in the acute phase, thus these cases represent hospitalized probable DHF cases. An additional 34 met all criteria, but did not have data for determination of hemoconcentration. These were classified as hospitalized possible DHF. All of the cases were

[1]Fundação Oswaldo Cruz, Centro de Pesquisas Gonçalo Moniz, Laboratory of Pathology and Molecular Biology, Rua Waldemar Falcão, Candeal, Salvador, Bahia, Brazil; [2]Center for Global Health and Disease, Case University, Wolstein Research Building, Cleveland, OH, USA; [3]Division of Genetic Epidemiology, Department of Epidemiology and Biostatistics, Case Western Reserve University, Wolstein Research Building, Cleveland, OH, USA; [4]Universidade Federal da Bahia, Instituto de Saúde Coletiva, Rua Basílio da Gama s/n, Canela, Salvador, Bahia, Brazil; [5]Directoria de Vigilância Epidemiológica, Secretaria da Saúde do Estado da Bahia, 4a Avenida do CAB, Lado B, CAB Salvador, Bahia, Brazil; [6]Department of arbovirology and hemorrhagic fever, Instituto Evandro Chagas, Belém, Pará, Brazil
*Correspondence: Dr R Blanton, Center for Global Health and Diseases, Case Western Reserve University, 2103 Cornell Road, Wolstein Research Building, Cleveland, OH 44106-7286, USA. Tel: +1 216 368 4814; Fax: +1 216 368 4825; E-mail: reb6@case.edu
[7]These authors contributed equally to this work.
These results were presented in part at the American Society of Tropical Medicine Annual Meeting in 2006.
Received 30 September 2009; revised 4 May 2010; accepted 7 May 2010; published online 30 June 2010
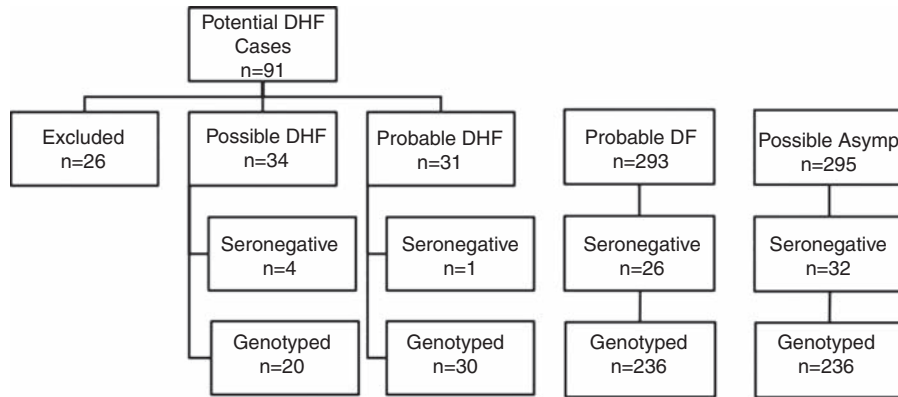
Figure 1 Recruitment. Potential DHF cases – identified by the Bahia State Secretariate of Health from febrile hospitalized patients during the 2002 and 2003 dengue transmission season based on hospitalization and severity of disease. Excluded – failed to show documented clinical criteria: spontaneous bleeding and thrombocytopenia (platelets <100 000) plus one or more of the following: headache, retro-orbital pain, muscle pain, joint pain, rash, and leukopenia. Possible DHF – hospitalized patients with all of the clinical criteria, but no documentation of hemoconcentration by Brazilian Ministry of Health hematocrit criteria: children up to 12 years >38%, women >40%, men >45%. Probable DHF – all of the WHO criteria for DHF documented. Probable DF – neighbors of suspected or confirmed DHF who claimed to have symptoms of dengue (see excluded) during January–May of 2002 or 2003. Possible Asymp (asymptomatic) – neighbors of suspected or confirmed DHF who claimed never to have had symptoms consistent with DF or DHF. Seronegative – IgG hemagglutination inhibition titer <1:20. Genotyped – all suspected and confirmed DHF cases enrolled were genotyped. Some enrolled probable DF and Asymp controls selected at random were not genotype because of limitations in genotyping capacity.

consistent with DHF grade II with spontaneous bleeding, but no evidence of shock or circulatory failure (Figure 1). There were no reported deaths in Salvador because of dengue in 2002. In 2004, 55 of the 65 probable or possible of DHF cases were contacted at home and agreed to enter the study. Five of these were excluded, as they were seronegative for DENV, and this left 20 possible and 30 probable hospitalized cases of DHF for analysis. For each DHF case, at least four neighbors who had symptoms of dengue in 2002 or 2003 and four who claimed never to have had symptoms were also enrolled. Symptomatic cases conformed to the 1997 WHO classification of probable dengue febrile illness (DF). These controls were all identified within 1 km of their index case and generally were within 100 m.

Relatives and household members of cases or controls were excluded. The controls were matched for age either less than or greater than 15, but not for sex, as there is usually little or no difference for DHF.[5,14] In addition to the interview for demographic, socioeconomic, and environmental data, blood was collected for genomic DNA and dengue serology. Written informed consent was obtained from all participants and/or their guardians. The study was approved by the ethical review boards of University Hospitals of Cleveland, the Oswaldo Cruz Foundation, Bahia, and the National Commission on Ethics in Research (CONEP), Brazilian Ministry of Health, and followed the guidelines of the US Department of Health and Human Services.

Socioeconomic and environmental variables

During a home visit, trained interviewers recorded answers on a standardized form to questions about schooling, household income, household appliances, water sources, water storage, sewage disposal, and home environment. Self-reported ethnicity was solicited according to the classification (white, mixed, black, Indian/Asian) recognized by the Brazilian Census (Instituto Brasileiro de Geografia e Estatística-IBGE). Unlinked markers were used to estimate percent genetic ancestry with the software program Structure 2.1[15] under the admixture model.[16] 'Income' was reported in units of monthly minimum salaries (1 MMS ≈ US$97.03 in 2004) as set by the Brazilian government. Income was significantly correlated with variables related to formal education and disposable income by linear regression analysis (in decreasing order of significance – computer, water filter, years of schooling, internet, open water containers, washing machine, freezer, household planters, planters observed in the household, cell phone). The reported income by ethnic categories also paralleled that reported to the Brazilian census for the city of Salvador (Supplemental Figure 1). Income was not recorded for 13% ($n=83$) of those interviewed, so an income index was created using the sum of coefficients from linear regression analysis of demographic and environmental variables with income as the dependent variable.[16]

Serology, genomic DNA isolation, and quantitation

All plasma samples were screened at 1:20 (initial dilution) at the Evandro Chagas Institute (a WHO Collaborating Center for Arbovirus Reference and Research) in Belém, Pará, Brazil using the HI test as modified by Shope and Sather.[17] When the response to a dengue serotype was four times greater than other flaviviruses or ≥1280, the HI response was considered positive. All serum samples were also tested by IgG ELISA for dengue, yellow fever, and other flaviviruses circulating in Brazil.

Genomic DNA was extracted from 2 ml of buffy coat by QIAamp Flexigen kit according to the manufacturer's instructions and quantified by spectrophotometer. All samples were diluted to a final concentration of 100 ng/μl before genotyping.

Single-nucleotide polymorphisms selection and genotyping

We genotyped 728 single-nucleotide polymorphisms (SNPs) in 56 important genes of the type 1 IFN response pathway or with potential for association with dengue severity (Supplemental Table 1). Genes were selected for their function in viral sensing, signal receptors, signal transduction, effector or suppressor activity, or involvement suggested by the literature. We sought to identify ~20 markers in each gene based on markers at the SNPper website (http://snpper.chip.org/) in 2006. Markers separated by at least 60 bases and within a region from 500 bases upstream of the translational start to the polyadenylation signal were targeted. SNPs with earlier validation on dbSNP at the NCBI website (http://www.ncbi.nlm.nih.gov/) were preferred. HapMap data were not available at this time. If the minor allele frequency was available, we considered those with minor allele frequency >0.05. We also included SNPs located in the promoter, intron–exon boundaries, published regulatory motifs in the promoter, and 3′ UTR,[10] as well as SNPs with known functional significance or genetic association. To aid ancestry estimates, 40 ancestry informative markers[18] were selected for genotyping.

For genotyping, 250 ng of DNA was arrayed in 96-well plates and processed using Illumina's (Illumina, CA, USA) microbead array technology.[19] Genotypes of six replicates and two trios were used to guide the clustering and calculate error rates. The trios consisted of multi-ethnic families from Salvador not identified through their dengue status and composed of father, mother, and son. Two analysts interpreted the genotyping results using GenCall software v.

6.2.0.4 (Illumina). As recommended by Illumina, the limit for low GenCall Score was set at 0.25, and loci with lower scores were excluded.

Data analysis

Cases of hospitalized possible DHF and hospitalized probable DHF were analyzed separately and combined. Although those reporting never experiencing dengue are likely to have had milder symptoms, some may have been merely stoic or not very knowledgeable about the disease. Therefore, the asymptomatic group was also analyzed separately. Genotyping error rates were calculated on inconsistencies for six replicates per plate, on male hemizygosity for X-linked markers, and on offspring genotypes inconsistent with parental genotypes for the trios. Linkage disequilibrium (LD) was calculated with the program Haploview.[20] The Bonferroni's correction for multiple tests was calculated for all independent SNPs. Independent SNPs were considered as those not located within haplotype blocks as defined in Haploview along with SNPs selected at random from within the haplotype blocks (n=305). In addition, the q-value, a measure of statistical significance for the false discovery rate (FDR),[21] was calculated using the program Qvalue.[22]

A sliding window approach, implemented in the DECIPHER package of S.A.G.E. 5.4.2,[23] was used to analyze the association with contiguous haplotypes. Haplotype windows or sets were formed from the first SNP in a selected region and the next $n-1$ SNPs, where n indicates the size of the combined SNPs or window size. The analysis was repeated with a haplotype set consisting of the second SNP and the next $n-1$ SNPs in the region. The process was continued until the window consisted of the last n markers in the region. Haplotype frequencies were estimated using the expectation-maximization algorithm, and a likelihood ratio test (LRT) was used to compare the distribution of haplotypes between groups. The likelihood ratio asymptotically follows a χ^2 distribution, and this distribution is conservative when there are rare haplotypes, thus an empirical P-value for the LRT statistic was also computed. Alleles in a single gene showing significant association with DHF were also grouped as a haplotype and analyzed using logistic regression. Genotypic and haplotypic analyses were controlled for sex, age, ancestry, and income.

The association between disease severity and other variables was tested using multivariable logistic regression with the program SPSS v. 10 (SPSS Inc., IL, USA). Each marker in the JAK 1 gene was tested for associations with DHF controlling for sex, age, proportion of African genetic ancestry, and an index of income. The percent African ancestry was earlier estimated for each individual[16] using the program Structure,[15] and this proportion was used as an independent variable in the regression analysis. Genotypes for the markers associated with DHF were tested using dominant, recessive, and codominant models.

RESULTS

Population and genotyping

From active hospital surveillance in Salvador, Bahia, Brazil, 55 of those with hospitalized/possible DHF or hospitalized/probable DHF in 2002 were contacted and agreed to participate in the study. Five (8%) of those with hospitalized/possible DHF or hospitalized/probable DHF were eliminated because they were seronegative for dengue as compared with 9 and 12% for the outpatient DF and asymptomatic groups, respectively. Although it was not possible to determine the infecting serotype, 80% of the viral recoveries at the State's Central Laboratory were DENV-3 in 2002 (P Melo, personal communication, master's thesis). The DF households had significantly lower income than any of the other groups. Otherwise, there were no statistically significant differences for age or sex between the clinical categories of dengue. Although the hospitalized/possible DHF group were younger, they did not otherwise differ from the confirmed cases for income, ancestry (Table 1) or allele frequencies for associated markers (not shown).

Genotyping was successful for 93.5% of the markers and 98.8% of the samples. The two analysts' genotype calls were highly correlated ($r=0.99 \pm 0.03$). For successful markers, 1% of genotypes failed on

Table 1 Epidemiologic associations with dengue clinical presentation

| | Clinical presentation[a] | | | |
Characteristics	Asymp (n=236)	DF (n=236)	DHF possible (n=20)	DHF probable (n=30)
Age[b]	32 ± 18	30 ± 15	24 ± 15[c]	33 ± 15
Sex (% male)	42	39	55	33
Ethnicity				
% White	16	13	30	20
% Mixed	52	50	55	63
% Black	30	35	15	17
% Indian	2	2	0	0
Income[d]	3.7 ± 1.9	3.0 ± 1.7[e]	4.0 ± 1.7	4.9 ± 2.3

[a]Asymp, possible asymptomatic; DF, probable dengue fever; DHF, dengue hemorrhagic fever.
[b]Mean age in years.
[c]Significantly different from the asymptomatic and probable DHF presentations.
[d]Geometric mean family income ± geometric SD, income in monthly minimum salaries (~US$97.03 in 2004).
[e]Significantly different from all other presentations, $P<0.05$.

average, and 83% were polymorphic (minor allele frequency >0.05). In all, 77.6% (593/768) of the markers were genotyped successfully and were polymorphic. The mean and median minor allele frequencies were 0.269 and 0.265, respectively. HW proportions were observed for >98% of the polymorphic markers. Those that failed were examined for potential genotyping errors and did not show a significant association with DHF.

The genotyping error rate was estimated from analysis of six replicates and two trios. In addition, 25 X-linked markers were used to check hemizygosity in males, and the one heterozygous male sample was excluded. The error rate by replicate inconsistencies was 0.13%, by sexual inconsistencies 0.01, and by Mendelian inconsistencies 0.07%. Seven polymorphic SNPs were also genotyped using a primer extension method (DNAprint, Sarasota, FL, USA) for 65 samples and no discordant genotypes were identified.

Single locus and genotype analysis

Comparing the hospitalized possible/probable DHF cases with the outpatient DF cases, there were 58 markers with $P<0.05$ out of 593 informative markers. The gene with the largest percentage of significant markers was JAK1. There were 11 out of 18 SNPs in JAK1 with $P<0.05$, and 6 of these had the most significant values overall (Table 2). Five markers in JAK1 (rs11208534, rs2780831, rs310196, rs310222, and rs310216) were significantly associated with DHF (Table 2) using a conservative cutoff of $q<0.2$ for the FDR and three markers nearly reached significance using the Bonferroni's correction (1.6×10^{-4}).

The effect size of JAK1 polymorphisms on DHF risk was assessed by logistic regression of each locus for DHF compared with DF and controlled for age, sex, income, and ancestry. Increased risk was associated with homozygous genotypes for rs11208534 (TT) and rs2780831 (GG), whereas the TT genotype was protective for rs310196 (Table 3). Risk for the combined hospitalized/probable and possible DHF compared with outpatient DF was associated only with the TT genotype for marker rs11208534 (odds ratio (OR): 5.19, confidence interval (CI): 2.13–12.66), GG for rs2780831 (OR: 2.61, CI: 1.40–4.90), whereas the homozygous TT genotype was protective for rs310196 (OR: 0.30, CI: 0.15–0.57). In the stratified analysis, the hospitalized possible DHF cases had ORs similar to the probable cases. The same was true for the comparison of DHF with asymptomatic

controls; however, the strength of association was weaker for this comparison and only reached significance when the probable and possible DHF cases were combined (not shown). The comparison of the combined possible asymptomatic and DF individuals with the combined DHF cases produced no significant associations.

The P-values and ORs for the association DHF with DF were mapped to the JAK1 intron–exon structure (Figure 2). The most significant SNPs as well as those with the greatest impact on risk for DHF were located toward the 5′ end of the JAK1 gene. Although the P-values for the three principal SNPs in JAK1 were similar, the OR for rs11208534 (3.9) was the largest.

The LD structure of JAK1 for this sample and marker set was composed of four LD blocks and four independent markers. The SNPs rs11208534 and rs2780831 were located in the same linkage block and rs310196 was an independent marker (D′ rs2780831/rs11208534=1, rs2780831/rs310196=0.924, and rs11208534/rs310196=0.856).

Haplotype analysis

Haplotype analysis was performed using a sliding window approach in which SNPs were grouped 2, 3, 4, and 5 at a time. We found that the P-value of the JAK1 markers declined with windows >2. However, associations for markers in RNaseL and IKBKE increased with increasing window size compared with SNP analysis (Figure 3).

In another approach, the high-risk homozygotes for the three most significant SNPs (rs11208534, rs2780831, and rs310196) were combined into a single haplotype and compared for association with DHF (Table 4). As rs11208534 and rs2780831 were in strong LD, we also evaluated the two independent SNPs separately. The haplotypes produced smaller P-values and more narrow CIs (Table 3). When controlled for age, sex, ancestry, and income, logistic regression for combined DHF cases gave P=0.0025, OR: 2.7 (95% CI: 1.4–5.3) for the haplotype consisting of three SNPs, and P=0.0023, OR: 2.9 (95% CI: 1.5 – 5.8) for the two SNP haplotype.
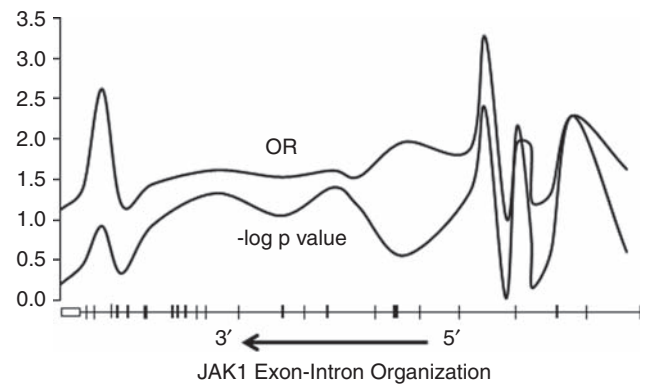
Table 2 Rank order of single-nucleotide polymorphism associations chromosome 1

Gene	SNP	Location	P-value[a]	q-value[b]
JAK1	rs11208534	Intron 3	4×10^{-4}	0.1272
JAK1	rs310196	Intron 1	6×10^{-4}	0.1272
JAK1	rs2780831	Intron 2	1×10^{-3}	0.1413
JAK1	rs310222	Intron 7	0.0026	0.2459
JAK1	rs310216	Intron 10/exon 11 boundary	0.0029	0.2459
IKBKE	rs1059704	Intron	0.0048	0.2591
JAK1	rs2256298	Intron 6/exon 7 boundary	0.0055	0.25911
JAK1	rs2991269	Exon 15	0.0062	0.2629
IKBKE	rs11118132	Intron	0.0203	0.3896
IKBKE	rs11117909	Intron	0.0219	0.3896

[a]χ^2 for genotype associations.
[b]False discovery rate significance (cutoff <0.2).

Figure 2 Distribution of JAK1 P-values and ORs. The P-values were converted to −log10 p and OR values were transformed by inversion in which necessary to ≥1. Exon–intron organization based on map structure in Genbank (Build 36.3) is shown below.
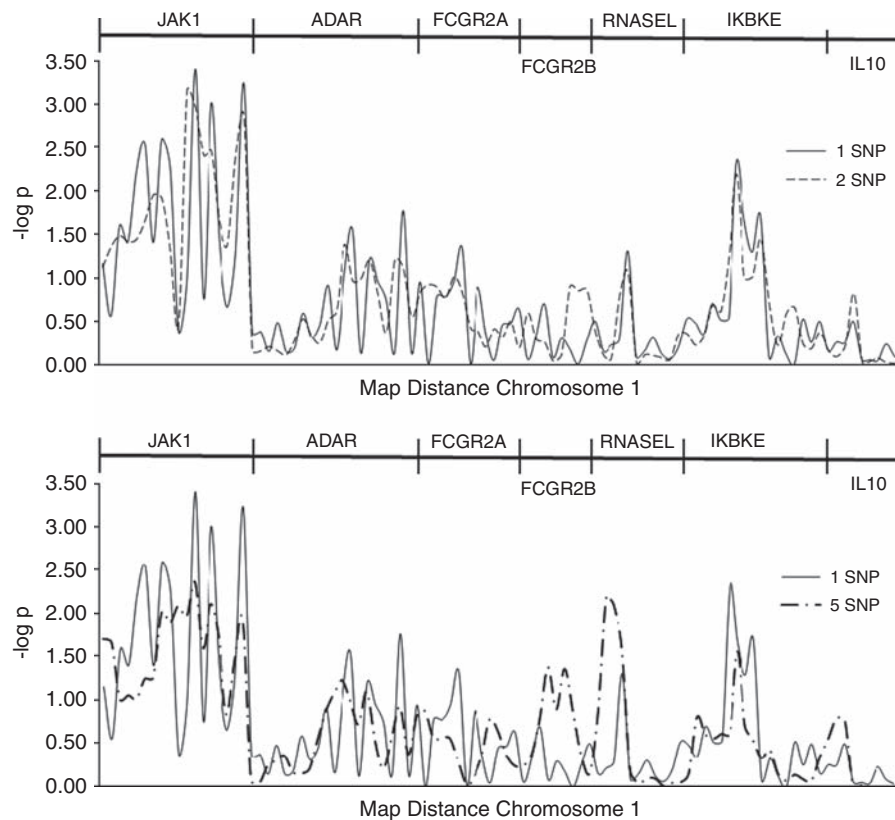
Table 3 Genotype risk for DHF[a]

DHF	SNP	Genotype	Without covariates		With covariates[b]	
			P	OR (CI 95%)	P	OR (95% CI)
Probable[c]	rs11208534	TT	5.5×10^{-3}	4.7 (1.6–13.8)	1.4×10^{-2}	3.9 (1.3–11.8)
	rs2780831	GG	5.7×10^{-2}	2.1 (<1.0–4.6)	1.5×10^{-1}	1.8 (0.8–4.0)
	rs310196	TT	2.2×10^{-2}	0.4 (0.2–0.9)	6.8×10^{-2}	0.5 (0.2–1.1)
	Haplotype 1[d]	3 SNPs	1.3×10^{-2}	2.7 (1.2–5.8)	4.4×10^{-2}	2.3 (1.0–5.2)
	Haplotype 2	2 SNPs	1.2×10^{-2}	2.7 (1.2–6.0)	3.7×10^{-2}	2.3 (1.0–5.2)
Possible	rs11208534	TT	1.4×10^{-2}	6.4 (1.5–28.3)	4.0×10^{-2}	4.9 (1.1–22.5)
	rs2780831	GG	7.1×10^{-3}	3.6 (1.4–9.3)	3.9×10^{-2}	2.9 (1.1–7.9)
	rs310196	TT	2.2×10^{-3}	0.2 (0.06–0.5)	1.0×10^{-2}	0.2 (0.1–0.7)
	Haplotype 1	3 SNPs	1.3×10^{-3}	6.3 (2.1–19.5)	7.9×10^{-3}	4.9 (1.5–16.0)
	Haplotype 2	2 SNPs	1.6×10^{-3}	4.6 (1.8–11.7)	1.3×10^{-2}	3.6 (1.3–9.6)
Combined	rs11208534	TT	2.9×10^{-4}	5.2 (2.1–12.7)	2.0×10^{-3}	4.2 (1.7–10.4)
	rs2780831	GG	2.7×10^{-3}	2.6 (1.4–4.9)	2.3×10^{-2}	2.1 (1.1–4.1)
	rs310196	TT	3.0×10^{-4}	0.3 (0.2–0.6)	4.3×10^{-3}	0.4 (0.2–0.7)
	Haplotype 1	3 SNPs	2.1×10^{-4}	3.3 (1.8–6.2)	2.5×10^{-3}	2.7 (1.4–5.3)
	Haplotype 2	2 SNPs	1.1×10^{-4}	3.7 (1.9–7.1)	2.3×10^{-3}	2.9 (1.5–5.8)

CI, confidence interval; DHF, dengue hemorrhagic fever; OR, odds ratio; SNP, single-nucleotide polymorphism.
[a]Genotype models tested were recessive/dominant and codominant. Risk was associated with the homozygotes TT and GG, respectively, for the markers rs11208534 and rs2780831. The homozygote TT was protective for rs310196. ORs compare risk for DHF (n=50) with DF (n=236) given the genotype.
[b]Covariates modeled: age, sex, genetic ancestry, and income index.
[c]Probable DHF (n=30) cases fit criteria for possible DHF, but also have documented evidence of capillary leak. Possible DHF (n=20) cases have a positive dengue serology, clinical syndrome consistent with dengue, spontaneous bleeding, and low platelet counts. Combined (n=50) represents probable and possible cases.
[d]Haplotype 1 – rs11208534, rs2780831, and rs310196. Haplotype 2 – rs11208534 and rs310196.

Figure 3 Sliding window haplotype analysis on chromosome 1. SNPs were sequentially grouped in windows of 2, 3, 4, and 5 and tested for association with DHF. The results for 1, 2, and 5 SNP windows are shown. The *y* axis shows *P*-values transformed to $-\log_{10} p$. The *x* axis provides the relative map order and distance along chromosome 1 for each marker. The order of each candidate gene is shown at the top of the graph.

Table 4 Genotypic associations with ancestry

		rs11208534				rs2780831				rs310196	
		Mean ± SD				Mean ± SD				Mean ± SD[a]	
Genotype	N	Ancestry	Income index	Genotype	N	Ancestry	Income index	Genotype	N	Ancestry	Income index
TT	341	0.42 ± 0.22	0.51	GG	172	0.39 ± 0.21	0.55	TT	256	0.40 ± 0.22	0.52
TC	157	0.50 ± 0.22	0.47	AG	222	0.47 ± 0.23	0.50	TG	212	0.49 ± 0.21	0.47
CC	23	0.61 ± 0.18	0.45	AA	94	0.52 ± 0.21	0.55	GG	52	0.57 ± 0.20	0.48
P value[b]		3.95×10^{-6}	4.7×10^{-2}			9.00×10^{-6}	4.4×10^{-2}			1.30×10^{-8}	3.24×10^{-3}

[a]Proportion African ancestry in those with this genotype.
[b]One-way ANOVA.

Effect of population stratification and covariates
Of all the demographic variables tested, only African ancestry and income entered the regression model. Therefore, both the genotypic and haplotypic regression analyses were controlled for these variables as well as for age and sex. The three markers most associated with DHF (rs310196, rs2780831, and rs11208534) each showed significant associations with African ancestry and income in this population (Supplemental Table 2). Only rs11208534 and rs310196 were also genotyped in the HapMap populations, and these also showed significant differences between the European and African populations for these markers (minor allele frequencies 0.12 *vs* 0.23 and 0.12 *vs* 0.50, respectively, *P* < 0.01). In the Salvador, for rs11208534, self-identified whites had a minor allele frequency of 0.13 compared with 0.22 for blacks. For rs310196, the frequencies were 0.20 and

0.37, respectively, *P* < 0.05. The proportion of African ancestry was significantly lower for the susceptible genotypes, consistent with the association of African ancestry with protection from DHF in this population.[16] There was a weaker association between income and each of the genotypes. There were no statistically significant differences for sex, income, or genotypes in JAK1 between the probable and possible cases, except for one SNP not associated with DHF. This indicates that these two groups are epidemiologically and genetically similar for these important variables and justifies combining them.

CD209 and DHF in Salvador, Brazil
We genotyped 11 markers in the CD209 gene including the promoter polymorphism most strongly implicated in an earlier study

(rs4804803, a.k.a. DC-SIGN1 −336[18]), but we found no association with DHF in this Brazilian population (Supplemental Table 3).

DISCUSSION

The study of dengue in the Americas is a unique opportunity to witness the sequential introduction of a 'new' pathogen into a naive population. DENV-2 was introduced to the city of Salvador in 1995. This was followed by the introduction of DENV-1 and the co-circulation of these viruses until by 2002 was 70% of the population was seropositive for at least one serotype.[2] The year 2002 saw the introduction of DENV-3 and the city's first cases of DHF. These were the cases collected for this study. Typical for most endemic areas, the majority of those infected did not experience DHF despite a high percentage of secondary infections. So, individual host genetics may influence the risk for developing DHF.

In this study, the strongest genetic association with DHF was found for polymorphisms located at the 5′ end of the JAK1 gene. Identification of JAK1 as an important gene in the clinical presentation of dengue is consistent with an emerging picture from diverse lines of investigation. Mouse studies have shown that some components of the type 1 IFN response pathway can control the response to all flaviviruses.[12,13] In addition, flaviviruses have developed specific strategies to neutralize signaling downstream of the receptor to circumvent this response.[24,25] Supportive evidence is also provided by a global analysis of expression profiles in individuals with dengue shock syndrome compared with DF, which indicated that multiple IFNα-regulated genes were under-expressed.[26] The effect on multiple genes at the end of a pathway may suggest suppression of an early step in signaling. As JAK1 is one of the two signaling proteins associated with the type I IFN receptor (the other is TKY2), regulatory polymorphisms in this gene could produce the pattern of downstream under-expression seen in severe dengue. JAK1 is also a component of many other signaling pathways, but the accumulation of evidence from other studies[12,13,24–26] suggests that its importance here is as a component of the IFNα/β response pathway.

The association of the identified JAK1 SNPs with DHF is also consistent with the epidemiology of this condition in the western hemisphere. The decrease in the susceptibility of allele frequency for those with greater proportions of African ancestry follows the epidemiologic and immunologic differences observed in the distribution of DHF.[3,27–29]

We did not observe any association between DHF and the CD209 locus in this population (Supplemental Table 3), despite an earlier study showing ORs >14 for an SNP (rs4804803) in this gene.[11] The failure to show an association with this gene and SNP may be due to population-based genetic heterogeneity between Brazilians and Thais. There are also important differences in the clinical presentation of the DF patients for the two studies. All of the Thais with clinically apparent dengue were hospitalized and were classified as severe dengue, whereas the Brazilians with probable DF in this study were outpatients and unlikely to be classified as severe. As the study was conducted retrospectively, acute sera were not available for most of those enrolled. Thus, our DF patients correspond to probable cases by WHO criteria, unlike the hospitalized DF cases in the Thai study in which acute serologic profiles were obtained. Finally, the study in Thailand also enrolled children exclusively, whereas we primarily enrolled adults, thus, age-specific differences and disease severity may also explain the failure to see an association with CD209 in our population. These enrollment characteristics for the two studies were driven by the differential age distribution of DHF between these two regions. Although DHF cases in Latin America

occurred predominantly in adults in 2004, in Southeast Asia, they occurred most frequently in children.[30]

As this was a retrospective study, we were unable to distinguish primary and secondary infections. Whether this poses a problem in classification for genetic studies is unclear. In the small number of studies in which the analysis was stratified in this way, most found no difference for the loci examined, except in the largest HLA studies.[11,31–34] There may be genetic differences based on earlier immune status, but apparently not for all associations. A retrospective study may also suffer more from misclassification, as the IgG serology is less specific than the IgM serology obtained in the convalescent period. The clinical classifications should thus be considered provisional.

The strongest association between JAK1 and DHF was observed when those cases with DHF were contrasted with the DF cases, rather than the asymptomatic, but seropositive controls. A similar pattern of association was observed in the study in Thailand for CD209. This may suggest additional genetic loci or environmental factors, such as viral inoculum or nutrition might have a function in addition to the risk genotypes. The strong influence of socioeconomic factors also suggests a function for environmental factors or may simply correlate with the usage of or access to medical services.

The location of the associated polymorphism indicates either that there is a regulatory element in one of the first three introns or that these SNPs are in LD with an as yet unidentified polymorphism in an exon. At present there has been only one SNP identified in an exon of this region of JAK1, rs35237903, which is 3730 bp downstream of the SNP with the strongest signal, rs11208534. Clearly, resequencing efforts should be directed to the 5′ region of the gene. Studies to replicate the findings reported here are underway in a region of Brazil with a different population profile as well as laboratory-based investigations of JAK1 and DENV interactions.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

1 Halstead SB: The pathogenesis of dengue. Molecular epidemiology in infectious disease. *Am J Epidemiol* 1981; **114**: 632–648.
2 Teixeira Md G, Barreto ML, Costa MdC N, Ferreira LDA, Vasconcelos PF, Cairncross S: Dynamics of dengue virus circulation: a silent epidemic in a complex urban area. *Trop Med Int Health* 2002; **7**: 757–762.
3 Guzman MG, Kouri G, Valdes L *et al*: Epidemiologic studies on dengue in Santiago de Cuba, 1997. *Am J Epidemiol* 2000; **152**: 793–799;discussion 804.
4 Guzman MG, Kouri G: Dengue: an update. *Lancet Infect Dis* 2002; **2**: 33–42.
5 Guha-Sapir D, Schimmer B: Dengue fever: new paradigms for a changing epidemiology. *Emerg Themes Epidemiol* 2005; **2**: 1.
6 World Health Organization: Dengue Haemorrhagic Fever: Diagnosis, Treatment, Prevention and Control. *Geneva, WHO* 1997.
7 Goh KT, Yamazaki S: Serological survey on dengue virus infection in Singapore. *Trans R Soc Trop Med Hyg* 1987; **81**: 687–689.
8 Shekhar KC, Huat OL: Epidemiology of dengue/dengue hemorrhagic fever in Malaysia – a retrospective epidemiological study. 1973–1987. Part II: dengue fever (DF). *Asia Pac J Public Health* 1992; **6**: 126–133.
9 Wallace HG, Lim TW, Rudnick A, Knudsen AB, Cheong WH, Chew V: Dengue hemorrhagic fever in Malaysia: the 1973 epidemic. *Southeast Asian J Trop Med Public Health* 1980; **11**: 1–13.
10 Xie X, Lu J, Kulbokas EJ *et al*: Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals. *Nature* 2005; **434**: 338–345.

11 Sakuntabhai A, Turbpaiboon C, Casademont I *et al*: A variant in the CD209 promoter is associated with severity of dengue disease. *Nat Genet* 2005; **37**: 507–513.

12 Perelygin AA, Scherbik SV, Zhulin IB, Stockman BM, Li Y, Brinton MA: Positional cloning of the murine flavivirus resistance gene. *Proc Natl Acad Sci USA* 2002; **99**: 9322–9327.

13 Mashimo T, Lucas M, Simon-Chazottes D *et al*: A nonsense mutation in the gene encoding 2′-5′-oligoadenylate synthetase/L1 isoform is associated with West Nile virus susceptibility in laboratory mice. *Proc Natl Acad Sci USA* 2002; **99**: 11311–11316.

14 Siqueira Jr JB, Martelli CM, Coelho GE, Simplicio AC, Hatch DL: Dengue and dengue hemorrhagic fever, Brazil, 1981–2002. *Emerg Infect Dis* 2005; **11**: 48–53.

15 Pritchard JK, Stephens M, Donnelly P: Inference of population structure using multilocus genotype data. *Genetics* 2000; **155**: 945–959.

16 Blanton RE, Silva LK, Morato VG *et al*: Genetic ancestry and income are associated with dengue hemorrhagic fever in a highly admixed population. *Eur J Hum Genet* 2008; **16**: 762–765.

17 Shope R, Sather G Arboviruses; in: Lennette E, Schmidt N (eds): *Diagnostic Procedures for Viral, Rickettsial and Chlamydial Infections*. American Public Health Association: Washington, 1979, pp 167–171.

18 Parra EJ, Kittles RA, Shriver MD: Implications of correlations between skin color and genetic ancestry for biomedical research. *Nat Genet* 2004; **36**: S54–S60.

19 Oliphant A, Barker DL, Stuelpnagel JR, Chee MS: BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. *Biotechniques* 2002; **Suppl**: 56–58, 60–51.

20 Barrett JC, Fry B, Maller J, Daly MJ: Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263–265.

21 Benjamini Y, Hochberg Y: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B* 1995; **57**: 289–300.

22 Storey J: A direct approach to false discovery rates. *J R Stat Soc Series B* 2002; **64**: 479–498.

23 S.A.G.E.: *S.A.G.E. (Statistical Analysis for Genetic Epidemiology)* 2006.

24 Jones M, Davidson A, Hibbert L *et al*: Dengue virus inhibits alpha interferon signaling by reducing STAT2 expression. *J Virol* 2005; **79**: 5414–5420.

25 Munoz-Jordan JL, Sanchez-Burgos GG, Laurent-Rolle M, Garcia-Sastre A: Inhibition of interferon signaling by dengue virus. *Proc Natl Acad Sci USA* 2003; **100**: 14333–14338.

26 Simmons CP, Popper S, Dolocek C *et al*: Patterns of host genome-wide gene transcript abundance in the peripheral blood of patients with acute dengue hemorrhagic fever. *J Infect Dis* 2007; **195**: 1097–1107.

27 Bravo JR, Guzman MG, Kouri GP: Why dengue haemorrhagic fever in Cuba? 1. Individual risk factors for dengue haemorrhagic fever/dengue shock syndrome (DHF/DSS). *Trans R Soc Trop Med Hyg* 1987; **81**: 816–820.

28 de la C, Sierra B, Garcia G *et al*: Ethnicity and difference in dengue virus-specific memory T cell responses in Cuban individuals. *Viral Immunol* 2006; **19**: 662–668.

29 Halstead SB, Streit TG, Lafontant JG *et al*: Haiti: absence of dengue hemorrhagic fever despite hyperendemic dengue virus transmission. *Am J Trop Med Hyg* 2001; **65**: 180–183.

30 Halstead SB: Dengue in the Americas and Southeast Asia: do they differ? *Rev Panam Salud Publica* 2006; **20**: 407–415.

31 Nguyen TP, Kikuchi M, Vu TQ *et al*: Protective and enhancing HLA alleles, HLA-DRB1*0901 and HLA-A*24, for severe forms of dengue virus infection, dengue hemorrhagic fever and dengue shock syndrome. *PLoS Negl Trop Dis* 2008; **2**: e304.

32 Sierra B, Alegre R, Perez AB *et al*: HLA-A, -B, -C, and -DRB1 allele frequencies in Cuban individuals with antecedents of dengue 2 disease: advantages of the Cuban population for HLA studies of dengue virus infection. *Hum Immunol* 2007; **68**: 531–540.

33 Stephens HA, Klaythong R, Sirikong M *et al*: HLA-A and -B allele associations with secondary dengue virus infections correlate with disease severity and the infecting viral serotype in ethnic Thais. *Tissue Antigens* 2002; **60**: 309–318.

34 Vejbaesya S, Luangtrakool P, Luangtrakool K *et al*: TNF and LTA gene, allele, and extended HLA haplotype associations with severe dengue virus infection in ethnic Thais. *J Infect Dis* 2009; **199**: 1442–1448.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)