

Differential Frequencies of *HLA-DRB1*, *DQA1*, and *DQB1* Alleles and Haplotypes Are Observed in the Arbovirus-Related Neurological Syndromes

Paulin Sonon,^{1,9} Maria Lúcia Brito Ferreira,² Renata Santos Almeida,¹ Neifi Hassan Saloum Deghaide,³ Glauco Henrique Willcox,⁴ Elizabeth Lima Guimarães,⁴ Antônio Fernando da Purificação Júnior,⁵ Marli Tenório Cordeiro,⁵ Carlos Alexandre Antunes de Brito,⁶ Maria de Fátima Militão de Albuquerque,⁷ Roberto D. Lins,⁵ Eduardo A. Donadi,³ and Norma Lucena-Silva¹

¹Immunology Department, Aggeu Magalhães Institute, Oswaldo Cruz Foundation, Recife, Pernambuco, Brazil, ²Hospital da Restauração Governador Paulo Guerra, Recife, Pernambuco, Brazil, ³Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, Sao Paulo, Brazil, ⁴Laboratório HLA Diagnóstico, Recife, Pernambuco, Brazil, ⁵Virology Department, Aggeu Magalhães Institute, Oswaldo Cruz Foundation, Recife, Pernambuco, Brazil, ⁶Internal Medicine Department, Federal University of Pernambuco, Recife, Pernambuco, Brazil, and ⁷Public Health Department, Aggeu Magalhães Institute, Oswaldo Cruz Foundation, Recife, Pernambuco, Brazil

Background. We took advantage of the 2015–2016 Brazilian arbovirus outbreak (Zika [ZIKV]/dengue/chikungunya viruses) associated with neurological complications to type *HLA-DRB1/DQA1/DQB1* variants in patients exhibiting neurological complications and in bone marrow donors from the same endemic geographical region.

Methods. *DRB1/DQA1/DQB1* loci were typed using sequence-specific oligonucleotides. In silico studies were performed using X-ray resolved dimer constructions.

Results. The *DQA1*01*, *DQA1*05*, *DQB1*02*, or *DQB1*06* genotypes/haplotypes and *DQA1/DQB1* haplotypes that encode the putative *DQA1/DQB1* dimers were overrepresented in the whole group of patients and in patients exhibiting peripheral neurological spectrum disorders (PSD) or encephalitis spectrum disorders (ESD). The *DRB1*04*, *DRB1*13*, and *DQA1*03* allele groups protected against arbovirus neurological manifestation, being underrepresented in whole group of patients and ESD and PSD groups. Genetic and in silico studies revealed that *DQA1/DQB1* dimers (1) were primarily associated with susceptibility to arbovirus infections; (2) can bind to a broad range of ZIKV peptides (235 of 1878 peptides, primarily prM and NS2A); and (3) exhibited hydrophilic and highly positively charged grooves when compared to the *DRA1/DRB1* cleft. The protective dimer (*DRA1/DRB1*04*) bound a limited number of ZIKV peptides (40 of 1878 peptides, primarily prM).

Conclusion. Protective haplotypes may recognize arbovirus peptides more specifically than susceptible haplotypes.

Keywords. *HLA-DRB1*; *HLA-DQA1*; *HLA-DQB1*; ZIKV; DENV; CHIKV; Brazil.

Arbovirus infections are primarily transmitted by mosquito-borne viruses. In Brazil, the *Aedes aegypti* mosquito is responsible for the transmission of dengue (DENV), Zika (ZIKV) and chikungunya (CHIKV) viruses [1]. In 2015–2016, a unique outbreak encompassing ZIKV, CHIKV, and the endemic DENV caused severe neurological complications primarily observed in infants producing an encephaly [2, 3], and among adults causing several types of neurological disorders, particularly peripheral neurological spectrum disorders (PSD) and encephalitis spectrum disorders (ESD) [3, 4].

Arbovirus gene variability may contribute to mosquito virus adaptation, enhancing infectivity, pathogenicity, fitness, and host transmissibility. In particular, residue exchange in structural or nonstructural (NS) viral proteins [5], including mutations of the NS1 (ZIKV, DENV) and envelope E-1/E-2 (CHIKV) proteins, contribute to virus adaptation to the vectors, facilitating rapid transmission [6]. In addition, a subgenomic flaviviral RNA fragment of DENV-2, which is important for viral replication, may accumulate in the mosquito salivary gland, enhancing virus transmission to the host [7].

Single nucleotide variations in human genes responsible for innate immunity have been associated with susceptibility to or severity of arboviral infections. For instance, mutations at *IFNA* [8], *TLR* (Toll-like receptors) [9], and other genes important for immune response [5] have been associated with severe arbovirus infections. Regarding adaptive immunity, histocompatibility genes are remarkably variable, encoding molecules that play a major role in virus peptide presentation as well as in virus elimination. HLA class I and II genes have been evaluated in dengue and chikungunya infections [10–13] but have not been evaluated in Zika infection.

Received 3 September 2020; editorial decision 8 December 2020; accepted 11 December 2020; published online December 15, 2020.

Correspondence: Norma Lucena-Silva, PhD, MD, Immunogenetic Laboratory, Immunology Department, Aggeu Magalhães Institute, Oswaldo Cruz Foundation, Av. Moraes rego, s/n, Campus da UFPE, Cidade Universitária, 50670420, Recife, PE, Brazil (nlucena@cpqam.fiocruz.br).

The Journal of Infectious Diseases® 2021;224:517–25

© The Author(s) 2020. Published by Oxford University Press for the Infectious Diseases Society of America. All rights reserved. For permissions, e-mail: journals.permissions@oup.com. DOI: 10.1093/infdis/jiaa764

The heterodimers encoded by HLA class II genes (*DPA1/B1*, *DQA1/B1*, and *DRA1/B1*) are primarily involved in antigen presentation to CD4 T cells and exhibit folded α and β chains that form the peptide binding groove. In contrast to the *DRA1* chain that presents negligible residue variation, the other HLA class II α and β chains are highly variable, primarily at the peptide binding groove, as an adaptation to recognize pathogen-derived antigens to be presented to T-cell receptors. *HLA-DQA1* alleles have been associated with susceptibility to flavivirus infections in Greek and Brazilian patients [12, 14]. The *HLA-DQB1*03:02* and *DQB1*02:02* genotypes are associated with dengue hemorrhagic fever and dengue fever, respectively [11]. *HLA-A*31* and *HLA-DRB1*08* allele groups conferred risk to dengue shock syndrome during the secondary infection, while patients with the *HLA-A*24* and *HLA-DRB1*12* allele groups are susceptible to dengue hemorrhagic fever during the primary dengue infection [15].

We took advantage of the 2015–2016 arbovirus outbreak that occurred in the northeastern state of Pernambuco to (1) follow up patients exhibiting neurological complications for at least 2 years; (2) type *HLA-DQA1/DQB1/DRB1* allele groups, genotypes, and haplotypes; and (3) study in silico the interaction of arbovirus peptides with the putative *DQA1/DQB1* and *DRA1/DRB1* dimers formed by some of the differentially observed *DRA1/DRB1* and *DQA1/DQB1* haplotypes.

METHODS

Patients and Controls

Patients ($n = 90$, 43 men) with average age 45.75 years (SD ± 16.90), presenting with neurological symptoms during the arbovirus outbreak, were recruited at Hospital da Restauração of Recife, State of Pernambuco, Brazil. Patient sera or plasma and cerebrospinal fluid were used for etiological infection diagnosis using enzyme-linked immunosorbent assay (ELISA; immunoglobulin M and immunoglobulin G), plaque reduction neutralization test, and quantitative real-time polymerase chain reaction (qRT-PCR) for ZIKV, CHIKV, and DENV [16, 17].

At the time of diagnosis, patients exhibited laboratory evidence for flavivirus (ZIKV or DENV) infections (13 patients), or togavirus (CHIKV) infection (23 patients), or flavivirus plus togavirus infections (28 patients). In the remaining 26 patients it was not possible to discriminate ZIKV from DENV infection due to the well-known serologic cross-reactivity between flaviviruses (ZIKV and DENV); however, they exhibited similar neurological manifestation as seen in confirmed ZIKV-infected patients. All patients were also categorized according to the major neurological syndrome into ESD (29 patients) and PSD (46 patients). Few cases (15 patients) exhibited other neurological complications, that is optic neuritis, peripheral facial paralysis, and brain stroke. These patients were included in the analysis of the association between HLA class II alleles with the arbovirus-induced neurological complications; however, they

were excluded when we investigated the allele frequencies between the major clinical presentation forms (ESD/PSD). In parallel, we studied 200 healthy bone marrow donors (112 men) with average age 35.74 years (SD ± 7.75) from the same geographical and endemic area as the patients (Table 1).

The project protocol was approved by the ethics committee of the Hospital da Restauração (protocol No. CAAE63883517.4.3001.5198), and all the participants signed the informed consent forms.

DNA Extraction and HLA-II-DRB1/DQA1/DQB1 Typing

Genomic DNA was obtained from peripheral blood mononuclear cells using the DNAzol reagent (Thermo Fisher Scientific, Invitrogen), following the supplier's instructions. The *HLA-DRB1/DQA1/DQB1* loci were typed using sequence-specific oligonucleotides (LABType SSO kit, One Lambda) and named in accordance with the National Marrow Donor Program code (<https://hml.nmdp.org/MacUI/>).

Statistical Analysis

Genetic Association Studies

Allele, genotype, haplotype frequencies, and Hardy-Weinberg equilibrium (HWE) were performed using ARLEQUIN version 3.5.2 software [18]. HWE was tested by the exact test of Guo and Thompson [19]. Haplotypes were inferred with the same ARLEQUIN software using the expectation-maximization algorithm. The univariate analyses were performed by 2×2 contingency tables (binary variables), using R software (R version 3.4.2) [20]. The strength of the association was evaluated by the odds ratio (OR and 95% confidence interval [CI]), and *P* values were obtained using the 2-tailed Fisher exact test.

Association analyses were performed using only variants that exhibited a frequency $> 1\%$ (Supplementary Table 1A–1D). The following analyses were performed: (1) whole group of arbovirus infected patients versus controls, (2) ESD plus PSD patients versus control, (3) ESD syndrome versus controls, and (4) PSD syndrome versus controls. To limit the false positivity and enhance the accuracy of the results, we adopted a *P* value significance threshold of .005 [21] for univariate analysis followed by the Bonferroni correction (*P_c*). We included in multivariable analysis only univariate associations (Supplementary Tables 2 and 3) that reached the stringent $P \leq .005$ and/or $P_c \leq .05$ threshold. The multivariable analyses were done using the multiple logistic regression model, including the dichotomous dependent variables (infected/controls, ESD plus PSD/controls, ESD/controls, or PSD/controls) and the independent variables (*HLA-DR/DQ* variants), adjusted for age and sex (covariables; Supplementary Table 4). The model was determined by using the step procedure and the choice of the final model was made on the basis of the lowest Akaike information criterion [22]. Multivariate adjusted *P* values $\leq .05$ were considered significant.

Binding Affinity of Virus Peptides to HLA-II Dimers

Because ZIKV is the prototype arbovirus associated with neurological complications, we used the complete ZIKV polyprotein, divided into 15-residue peptides (same size as the native peptides in the experimentally devised structures) with an overlap of 12 residues between adjacent peptides, for in silico studies. Variations among the 15 ZIKV virus strains were also considered, according to sequences deposited at National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). The binding free energy ($\Delta\Delta G$) of the coupled peptide into the resolved native X-ray crystallographic structure (PDB ID: 5KSV; 5NI9) [23, 24] was compared to the values obtained for the virus peptides when coupled (replacing the native γ -chain) into the α - and β -beta chains of HLA-II dimers. Once the structure of each 15-mer peptide was computed, the binding free energy was compared to the value of the native peptide. The $\Delta\Delta G$ was calculated using Rosetta version 3.10.2 software [25, 26]. Geometry optimizations were performed using the Broyden-Fletcher-Goldfarb-Shanno algorithm, complying with the Armijo-Goldstein condition and the binding free energy was computed using the atomistic energy function REF2015 [27]. Sampling of the peptide backbone was enhanced by the backrub protocol. REF2015 has been parameterized against high-resolution structural data and its energy unit was expressed in kcal/mol [27, 28]. Electrostatic potential surfaces were calculated using the adaptive Poisson-Boltzmann solver [29] and plotted onto the HLA van der Waals surface, as rendered by visual molecular dynamics [30]. To estimate the cross-reactivity among virus peptides and HLA dimers, peptide homology between DENV1-4, CHIKV, and ZIKV was evaluated applying BLASTp [31], and using polyproteins retrieved from UNIPROT [32].

RESULTS

HLA-DRB1/DQA1/DQB1 Alleles, Genotypes, and Haplotypes

Besides comparing the frequencies of the *DRB1/DQA1/DQB1* alleles, genotypes, and haplotypes between patients and controls, we also evaluated some of the putative dimers formed by HLA class II haplotypes that exhibited differential frequencies between patients and controls. To perform dimer analyses, we considered (1) only *DRB1* allele groups for the *DRA1/DRB1* haplotypes because the *DRA1* locus has a negligible polymorphism, and (2) both *cis* and *trans* combinations for *DQA1/DQB1* haplotypes. The terms *DRA1/DRB1* and *DQA1/DQB1* haplotypes refer to genetic studies and the putative molecules encoded by these haplotypes are *DRA1/DRB1* and *DQA1/DQB1* dimers. For genetic analyses, only *cis* haplotypes (*cis* dimers) were computed. The *HLA-DRB1/DQA1/DQB1* loci fit the HWE for patients and controls, exhibiting *P* values > .05 (Supplementary Table 1A).

Several comparisons were performed exhibiting significant *P* values; however, in the results section, only comparisons that

Table 1. Demographic and Phenotypic Features of Control Population and Patients Infected by Arbovirus in the 2015–2016 Epidemics in the Brazilian Northeastern State of Pernambuco

Demographic and Phenotypic Features	Patients, No. (%) ^a	Controls, No. (%)
Total participants	90 (31.03)	200 (68.97)
Sex		
Female	47 (16.21)	88 (30.34)
Male	43 (14.83)	112 (38.62)
Age		
< Median, No.	43	97
Median, γ (No.)	46 (2)	34 (11)
> Median, No.	43	92
Patient phenotypes		
Infected by flavivirus, ZIKV or DENV	39 (43.33)	...
Infected by togavirus, CHIKV	23 (25.56)	...
Infected by togavirus + flavivirus	28 (31.11)	...
Syndromes caused by arbovirus ^b		
ESD	29 (32.22)	...
PSD	46 (51.11)	...
OND ^c	15 (16.67)	...

Abbreviations: CHIKV, chikungunya virus; DENV, dengue virus; ESD, encephalitis spectrum disorders; OND, other neurological disorders; PSD, peripheral spectrum disorders; ZIKV, Zika virus.

^aPatients infected by arboviruses ZIKV, DENV, and/or CHIKV.

^bESD: cerebellitis, encephalitis, meningoencephalitis, rhombencephalitis, myelitis, and myeloradiculitis; and PSD: polyneuropathy, polyradiculoneuropathy, radiculoneuropathy, and Guillain-Barré syndrome.

^cA few patients were also diagnosed with OND, including optic neuritis, peripheral facial paralysis, and brain stroke.

reached significance of $P \leq .005$ for univariate analysis and $P \leq .05$ for multivariable analysis adjusted for sex and age are shown (Table 2 and Table 3).

HLA-DRB1/DQA1/DQB1 Variants Are Associated With Susceptibility/Protection to Arbovirus Neurological Complications

Compared to controls, the *DRB1*04* allele group or the *DRB1*04* dimer was overrepresented in controls compared to whole group of patients (univariate analysis $P = .0003$; multivariable $P = .044$; Table 2), while the *DRB1*03/DQA1*01/DQB1*02* extended haplotype was overrepresented in the whole group of patients (univariate analysis $P = .0002$; multivariable $P = .011$; Table 2).

Considering the stratified ESD plus PSD or ESD or PSD patient groups in comparison to controls: (1) the *DQA1*03* allele group exhibited higher frequency in controls than in both the ESD plus PSD patient group (univariate analysis $P = .001$; multivariable $P = .033$; Table 3) and the ESD patient group (univariate analysis $P = .004$; multivariable $P = .044$; Table 3); (2) the *DRB1*03/DQA1*01/DQB1*02* extended haplotype was overrepresented in the ESD plus PSD patient group (univariate analysis $P = .0002$; multivariable $P = .005$; Table 3); and (3) the *DRB1*03* allele group or dimer (univariate analysis $P = .0005$;

Table 2. Multivariate Analyses for Comparisons of HLA-DRB1, DQA1, and DOB1 Alleles, Genotypes, Dimers, and Extended Haplotypes, Exhibiting P Value ≤ .005 in Univariate Analysis Between the Whole Group of Patients With Arbovirus (ZIKV, DENV, and CHIKV) Infections Versus Controls

Variables	Univariate (Unadjusted)					Multivariate (Adjusted)				
	Patient, %	Control, %	OR	95% CI	P	OR	95% CI	Z Value	Model Dispersion Coefficient ^a	P
DRB1*04 allele/dimer	15.70	36.50	0.33	.16–.63	.0003	0.49	.23–.96	-2.010	1.06	.044
DRB1*15 allele/dimer	36.00	18.50	2.47	1.35–4.49	.002	1.66	.88–3.11	1.570	1.06	.116
DRB1*04/DRB1*13	0.00	9.00	0.00	.00–.46	.001	0.00	0–1 × 10 ⁺²⁰	-0.018	1.06	.986
DRB1*03/DOA1*01/DOB1*02	11.10	1.00	12.26	2.53–117.64	.0002	8.51	1.91–60.75	2.540	1.06	.011

Significant multivariate P values adjusted by sex and age are shown in bold.

Abbreviations: CHIKV, chikungunya virus; CI, confidence interval; DENV, dengue virus; OR, odds ratio; ZIKV, Zika virus.

^aModel dispersion is the summary(model.final\$deviance/summary(model.final)\$df.residual) ratio. The model is dispersed when this ratio exceeds 1.5.

Table 3. Multivariate Analyses for Comparisons of HLA-DRB1, DQA1, and DOB1 Alleles, Genotypes, Dimers, and Extended Haplotypes, Exhibiting P Value ≤ .005 in Univariate Analysis Between Patients Presenting With ESD Plus PSD, Only ESD, and Only PSD Versus Controls

Variables	Univariate (Unadjusted)					Multivariate (Adjusted)				
	ESD + PSD, ESD, or PSD, %	Control, %	OR	95% CI	P	OR	95% CI	Z Value	Model Dispersion Coefficient ^a	P
Patients presenting with ESD plus PSD										
DQA1*03	20.00	41.50	0.35	.17–.68	.001	0.48	.24–.92	-2.129	1.03	.033
DRB1*04/DRB1*13	0.00	9.00	0.00	.00–.57	.005	0.00	0–3 × 10 ⁺¹⁹	-0.017	1.03	.986
DRB1*03/DOA1*01/DOB1*02	12.00	1.00	13.35	2.67–130.25	.0002	10.30	2.38–72.23	2.809	1.03	.005
Patients presenting with only ESD										
DQA1*03	13.80	41.50	0.23	.06–.69	.004	0.30	.08–.89	-2.016	0.69	.044
DQA1*01/DOB1*06	41.40	17.00	3.42	1.36–8.43	.005	2.17	.88–5.30	1.709	0.69	.087
Patients presenting with only PSD										
DRB1*03 allele/dimer	28.30	8.00	4.49	1.81–11.05	.0005	3.34	1.27–8.69	2.473	0.81	.013
DRB1*13 allele/dimer	10.90	33.00	0.25	.07–.67	.002	0.25	.08–.67	-2.566	0.81	.010
DQA1*01/DOA1*05	34.80	12.50	3.71	1.65–8.24	.0007	3.56	1.48–8.53	2.863	0.81	.004

Significant multivariate P values adjusted by sex and age are shown in bold.

Abbreviations: CI, confidence interval; ESD, neurological encephalitis spectrum; OR, odds ratio; PSD, peripheral spectrum disorders.

^aModel dispersion is the summary(model.final\$deviance / summary(model.final)\$df.residual) ratio. The model is dispersed when this ratio exceeds 1.5.

multivariable $P = .013$; Table 3) and the DQA1*01/DQA1*05 genotype (univariate analysis $P = .0007$; multivariable $P = .004$; Table 3) were found in higher frequency in PSD patients. In addition, the DRB1*13 allele group or dimer was overrepresented in controls compared to PSD patients (univariate analysis $P = .002$; multivariable $P = .010$; Table 3).

Binding Affinity of Arbovirus-Derived Peptides to DRA1/DRB1 and DQA1/DQB1 Dimers

Considering that (1) some protein structures of the DRA1/DRB1 and DQA1/DQB1 dimers coupled with 15-residue peptides have been resolved by X-ray crystallography [23, 24], (2) some dimers are available for in silico studies [23, 33], and (3) ZIKV is the major agent associated with neurological manifestations, we used a library of ZIKV polyproteome peptides to evaluate their interaction with dimers. We selected available dimer constructions associated with protection (DRA1/DRB1*04) or susceptibility (DQA1*05/DQB1*02) to arbovirus-related neurological complications. Both dimers were evaluated for their binding affinities for 1878 15-mer peptides covering the entire of ZIKV polyprotein of the 15 strains. Binding free energy ($\Delta\Delta G$) values for the native structures of the DQA1*05/DQB1*02 and DRA1/DRB1*04:01 complexes were -69 and -54 kcal/mol, respectively. Accordingly, peptides that presented

binding free energy values lower than the values for the native peptides would have higher affinity and, therefore, were able to bind to HLA dimers. The DQA1*05/DQB1*02 dimer can bind with high affinity to a remarkable number of hypothetical ZIKV peptides. A total of 235 15-mer ZIKV polyprotein-derived peptides could be presented by the DQA1*05/DQB1*02 dimer (Supplementary Table 5), whereas only 40 peptides were bound to DRA1/DRB1*04:01 dimer with higher affinity than the native counterpart (Supplementary Table 6). The distribution of the high-affinity peptides across the ZIKV proteins that can bind to the dimers is shown in Figure 1A. To assess statistical relevance regarding the ZIKV proteins presented by HLA dimers, protein length should be considered. The relative coverage (percentage of ZIKV 15-mer peptides that can bind to a dimer with higher affinity than the dimer's native peptide) of the high-affinity peptides on each ZIKV protein for both HLA molecules is presented in Figure 1B. Notably, among peptides exhibiting highest affinities to bind to the DQA1*05/DQB1*02 dimer, the ZIKV proteins prM and NS2A were highly represented, whereas no peptide from NS2B protein was able to bind to the DRA1/DRB1*04:01 dimer (Figure 1A and 1B). There was little overlap between the peptides that bind to each HLA dimer (Supplementary Tables 5 and 6). The discrepancy in the ZIKV peptide repertoire that can bind to each HLA can be

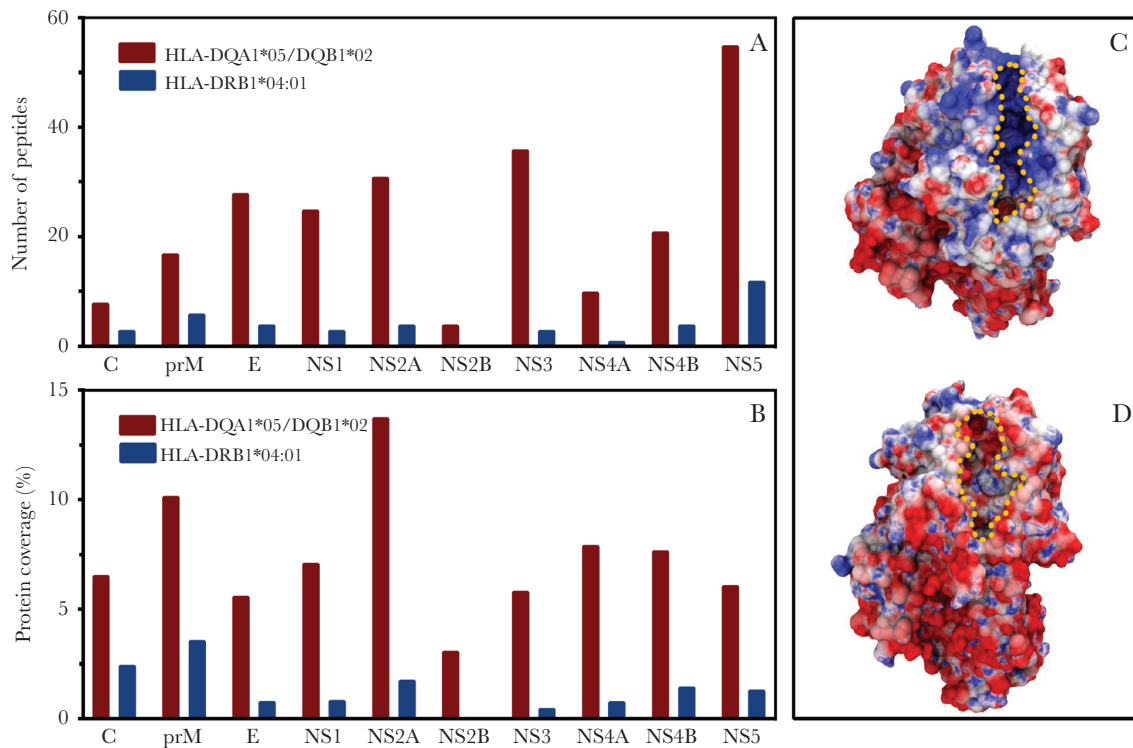


Figure 1. Distribution of 15-mer peptides derived from the Zika virus polyprotein capable of binding with high affinity to HLA-DQA1*05/DQB1*02 and HLA-DRA1/DRB1*04:01 dimers and HLA electrostatic surface potential. *A*, Number of peptides from each protein. *B*, Protein coverage, defined as the ratio of the number of peptides over the length of each protein. *C* and *D*, Electrostatic potential (from -6 to $+6$ kJ/mol/e) plotted onto the van der Waals surface of the DQA1*05/DQB1*02 dimer (*C*) and DRA1/DRB1*04:01 dimer (*D*). Negatively charged areas are represented in red, positively charged areas are represented in blue, and neutral areas are represented in white; peptide binding clefts are highlighted by a dotted yellow line.

understood by inspecting the differences in charge surface of each binding site. Plotting of the electrostatic potential onto the protein van der Waals surface revealed that DQA1*05/DQB1*02 dimer has a highly positively charged groove (Figure 1C), while HLA/DRB1*04:01 dimer displays a more amphipathic and yet less hydrophilic site (Figure 1D). In addition, 24, 27, 5, and 38 ZIKV peptides predicted to bind to the DQA1*05/DQB1*02 dimer were homologous to DENV1–4 peptides, respectively (Supplementary Tables 7). Interestingly, only a few ZIKV peptides, which were similar to each other and homologous to DENV1 (6 peptides), DENV2 (6 peptides), and DENV4 (8 peptides), were predicted to bind to DQA1/DRB1*04 dimer (Supplementary Tables 8).

DISCUSSION

HLA-I molecules have a pivotal role in viral infection control [34]; however, the class II presentation of viral peptides to T-cell CD4⁺ is the first step of the adaptive immunity that may impact on viral disease outcome [35], and HLA-II gene variants are expected to modulate host immune response against viruses, depending on residue features of the binding groove. Mutations in HLA-II genes have been associated with susceptibility/protection for many arbovirus infections, including DENV, West Nile virus, and other flaviviruses [11, 12, 14, 15, 36]. We evaluated the association between HLA-II alleles, genotypes, haplotypes, and their encoded dimers on the susceptibility to arbovirus-related neurological complications.

Although *DRB1* alleles (*DRB1*15* and *DRB1*03* considered as alleles and haplotypes) were associated with susceptibility to arbovirus infections (Supplementary Tables 2 and 3), the most striking associations were observed with *DQA1*05* (genotype and haplotype), *DQA1*01* (genotype and haplotype), *DQB1*02* (genotype and haplotype), and *DQB1*06* (genotype and haplotype), which were overrepresented in arbovirus-infected patients, considered as the whole group and as major neurological syndromes (ESD and PSD). Accordingly, the haplotypes that encoded the *DQA1*01/DQB1*02* and *DQA1*01/DQB1*06* dimers showed strong associations with susceptibility to arbovirus infections and ESD or PSD. Notably, a strong linkage disequilibrium (LD) has been reported for *DRB1* and *DQB1* alleles, and in the Brazilian population the *DRB1*15* allele group is primarily in LD with the *DQB1*06* group, whereas the *DRB1*03* group is primarily associated with the *DQB1*02*, *03*, and *04* groups [37, 38]. In addition to *DRB1*15*, the *DRB1*11* group, which exhibited a frequency closely similar to controls, also exhibit strong LD with *DQB1*06* alleles. Therefore, LD between *DRB1/DQB1* alleles may also contribute to the increased frequency of *DQB1*06* alleles, and the participation of *DRB1* alleles (particularly *DRB1*15* and *DRB1*03*) may be due to a hitchhiking LD with *DQB1* alleles. Besides *DRB1/DQB1* LD, some *DRB1/DQA1/DQB1* extended haplotypes were also associated with susceptibility to arbovirus neurological infections.

The increased frequency of the haplotypes *DRB1*03/DQA1*01/DQB1*02* and *DRB1*15/DQA1*05/DQB1*06* (Supplementary Tables 2 and 3) may yield several *cis* dimers (*DQA1*01/DQB1*02* and *DQA1*05/DQB1*06*). In parallel, several *trans* dimers may also be formed, including several combinations of *DQA1*01* and *DQA1*05* at one haplotype with *DQB1*02* and *DQB*06* at the other haplotype. To further understand the role of *DQA1/DQB1* dimers we performed an *in silico* study to evaluate the peptide profile that can be presented by such dimers, and we illustrated these studies by coupling ZIKV peptides to dimers. Surprisingly, a single dimer (*DQA1*05/DQB1*02*) presented hundreds of virus peptides. Considering that a great number of *cis* and *trans* *DQA1/DQB1* dimers can be generated by the susceptibility haplotypes, the number of virus peptides to be presented to T cells may drastically increase. If, on the one hand, the increased number of peptides may result in a diverse immune response against virus peptides, but on the other hand, the association of several haplotypes with susceptibility to arbovirus may indicate that peptides may be redundant and/or cross-reacting, impairing the effective immune response against the virus.

In contrast to susceptibility alleles, the most striking association with protection was conferred by the *DRB1*04* allele group, which was underrepresented in the whole group of patients. The *in silico* studies showed that only dozens of Zika virus peptides were presented by the *DQA1/DRB1*04* dimer, contrasting to hundreds of peptides that can be presented by *DQA1/DQB1* dimers. The reduced number of virus peptides may indicate that a more robust and more specific immune response may be mounted by *DRB1* molecules.

The chemical nature of the clefts formed by *DQA1*05/DQB1*02* or *DRB1/DRB1*04* are quite distinct. While the former is mainly hydrophilic and positively charged (Figure 1C), the latter is less hydrophilic and exhibits an amphipathic character (Figure 1D). In addition, little overlap is seen for the repertoire of ZIKV peptides that can bind to each dimer. As shown in Figure 1, the *DQA1*05/DQB1*02* dimer is highly promiscuous to ZIKV hypothetical peptides. Among those, it is reported that prM and NS2A proteins share a role in the modulation of virion assembly [39], and prM has been associated with neurovirulence [39]. In contrast, HLA-*DRB1*04* is more specific (binds to fewer peptides) and prM was the most represented protein (Figure 1B). Notably, immune response against prM is likely to be essential for protection because DNA-based Zika vaccines that did not code for prM did not achieve appropriate level of protection [40].

Peptide promiscuity among HLA class I and II molecules is a widely discussed subject; however, molecules that present larger number of peptides promote immune response against a broad range of pathogens [41, 42], but may not be able to respond to new virulent pathogens [41]. In contrast, particular dimers may provide protection against highly virulent pathogens, such

as human immunodeficiency virus [43], hepatitis C virus [44, 45], and *Leishmania* species [42]. These observations are consistent with our findings, in which some DQA1/DQB1 dimers (for instance, DQA1*05/DQB1*02) act as generalist molecules, while others (DRA1/DRB1*04) act as specialist molecules towards ZIKV peptides. It has also been proposed that pathogen diversity drives the evolution of promiscuous HLA-II alleles in regions of high pathogen diversity [46]. Indeed, a higher prevalence of the 4 DENV serotypes in the northeastern region of Brazil [47] may account for the higher number of ESD and PSD cases compared to other regions in the country. Interestingly, our results show that ZIKV NS2A protein is highly covered by the DQA1*05/DQB1*02 dimer, and the cross-reactivity between ZIKV and DENV T-cell responses [48, 49] may account for the difficulty in discriminating between these infections.

One of the limitations of this exploratory report is the size of our cohort (90 patients). Few studies have addressed the association of HLA-DR-DQ genes with arbovirus infections, particularly focusing on dengue [11, 12, 50] and chikungunya [10, 13], encompassing sample sizes ranging from 39 to 110 patients; however, no studies have been conducted to evaluate the effect of these genes in neurological complications. In the present study, all 90 cases analyzed exhibited neurological complication of arbovirus disorders, diagnosed in a defined time in a restricted geographical area of Brazil, possibly presenting with the same virus strains. Despite the sample size, in this series we detected several HLA-II associations that (1) reached the stringent P values ($P \leq .005$ and/or Bonferroni corrected $P_c \leq .05$) threshold fixed for univariate analyses, (2) reached the stringent the P value .05 threshold for the multivariable analysis following correction for sex and age, and (3) exhibited a biological relevance in *in silico* studies, explaining at least in part the increased frequency of neurological complications following the arbovirus outbreak that occurred in the state of Pernambuco, Brazil. Another limitation was the controls. The ideal control group would be healthy individuals who presented positive serology or previous history of dengue, chikungunya, and Zika infections; however, the patients exhibited these infections in several combinations, impairing the availability of controls with the same viral serologic profile but without neurological manifestations. To circumvent these limitations, we chose healthy bone marrow donors because (1) they are from the same geographical and endemic region as the patients, (2) they represent the allele frequency of the studied population, and (3) the frequency deviation from that of the patients gives a clue to the differential alleles implicated in disease pathogenesis, satisfying the objectives of the study.

In conclusion, the ensemble of genetic and *in silico* studies indicate that *DQA1* and *DQB1* alleles primarily considered as haplotypes are associated with susceptibility to arbovirus infection, possibly due to (1) the large number of DQA1/DQB1 molecules that can be generated by the haplotypes exhibiting differential frequencies between patients and controls, (2)

broad peptide specificity, as shown by the large number of arbovirus peptides that these dimers can present, and (3) the distinct chemical nature of the DQA1/DQB1 dimer cleft when compared to that formed by DRA1/DRB1 molecules. In contrast, the protection conferred by the specific DRA1/DRB1 dimer may associated with the presentation of a limited number of arbovirus peptides, suggesting that protective molecules may recognize arbovirus peptides more specifically and more efficiently.

Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Acknowledgments. The authors thank patients and relatives for their free consent, and the health professionals of the neurology service in the Hospital da Restauração for their great collaboration.

Author contributions. E. A. D. and N. L.-S. conceived, designed, and applied for grants to perform the study. M. L. B. F., C. A. A. B., M. F. M. A., M. T. C., and R. S. A. collected samples and provided clinical and laboratory information. N. H. S. D., G. H. W., and E. L. G. performed HLA-II gene typing. P. S. processed sequenced data and wrote the original manuscript draft. P. S. and R. S. A. performed statistical analyses. A. F. P. J. and R. D. L. performed the computer simulations. P. S., R. S. A., M. T. C., C. A. A. B., M. F. M., R. D. L., E. A. D., and N. L.-S. performed data curation. All authors have critically read and edited the manuscript.

Financial support. This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq; grant numbers 440760/2016-0 and 302060/2019-7); Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (grant number 88881.130769/2016-01); and the Aggeu Magalhães Institute/Fundação Oswaldo Cruz. P. S. was supported by Fundação Oswaldo Cruz (grant number INOVA FIOCRUZ/02/2019 PDJ); N. L.-S. and R. D. L. are CNPq scholarship holders (grant numbers 310364/2015-9 and 310892/2019-8); and A. F. P. J. was sponsored by Fundação de Amparo a Ciência e Tecnologia do Estado de Pernambuco. Computer allocation was provided by Laboratório Nacional de Computação Científica.

Potential conflicts of interest. All authors: No reported conflict of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Wu P, Yu X, Wang P, Cheng G. Arbovirus lifecycle in mosquito: acquisition, propagation and transmission. *Expert Rev Mol Med* 2019; 21:e1.

2. Keesen TSL, de Almeida RP, Gois BM, et al. Guillain-Barré syndrome and arboviral infection in Brazil. *Lancet Infect Dis* **2017**; 17:693–4.
3. Wilder-Smith A, Gubler DJ, Weaver SC, Monath TP, Heymann DL, Scott TW. Epidemic arboviral diseases: priorities for research and public health. *Lancet Infect Dis* **2017**; 17:e101–6.
4. Araujo LM, Ferreira ML, Nascimento OJ. Guillain-Barré syndrome associated with the Zika virus outbreak in Brazil. *Arq Neuropsiquiatr* **2016**; 74:253–5.
5. Ketkar H, Herman D, Wang P. Genetic determinants of the re-emergence of arboviral diseases. *Viruses* **2019**; 11:150.
6. Liu Y, Liu J, Du S, et al. Evolutionary enhancement of Zika virus infectivity in *Aedes aegypti* mosquitoes. *Nature* **2017**; 545:482–6.
7. Pompon J, Manuel M, Ng GK, et al. Dengue subgenomic flaviviral RNA disrupts immunity in mosquito salivary glands to increase virus transmission. *PLoS Pathog* **2017**; 13:e1006535.
8. Yakub I, Lillibridge KM, Moran A, et al. Single nucleotide polymorphisms in genes for 2'-5'-oligoadenylate synthetase and RNase L in patients hospitalized with West Nile virus infection. *J Infect Dis* **2005**; 192:1741–8.
9. Dutta SK, Tripathi A. Association of Toll-like receptor polymorphisms with susceptibility to chikungunya virus infection. *Virology* **2017**; 511:207–13.
10. Chaaithanya IK, Muruganandam N, Anwesh M, et al. HLA class II allele polymorphism in an outbreak of chikungunya fever in Middle Andaman, India. *Immunology* **2013**; 140:202–10.
11. Falcón-Lezama JA, Ramos C, Zuñiga J, et al. HLA class I and II polymorphisms in Mexican Mestizo patients with dengue fever. *Acta Trop* **2009**; 112:193–7.
12. Polizel JR, Bueno D, Visentainer JE, et al. Association of human leukocyte antigen DQ1 and dengue fever in a white Southern Brazilian population. *Mem Inst Oswaldo Cruz* **2004**; 99:559–62.
13. Thanapati S, Hande A, Das R, Gurav Y, Tripathy AS. Association of human leukocyte antigen class II allele and haplotypes in chikungunya viral infection in a western Indian population. *Trans R Soc Trop Med Hyg* **2014**; 108:277–82.
14. Sarri CA, Markantoni M, Stamatis C, et al; MALWEST Project. Genetic contribution of MHC class II genes in susceptibility to West Nile virus infection. *PLoS One* **2016**; 11:e0165952.
15. Malavige GN, Rostron T, Rohanachandra LT, et al. HLA class I and class II associations in dengue viral infections in a Sri Lankan population. *PLoS One* **2011**; 6:e20581.
16. Brito Ferreira ML, Antunes de Brito CA, Moreira ÁJP, et al. Guillain-Barré syndrome, acute disseminated encephalomyelitis and encephalitis associated with Zika virus infection in Brazil: detection of viral RNA and isolation of virus during late infection. *Am J Trop Med Hyg* **2017**; 97:1405–9.
17. Pessôa R, Patriota JV, Lourdes de Souza MD, Felix AC, Mamede N, Sanabani SS. Investigation into an outbreak of dengue-like illness in Pernambuco, Brazil, revealed a cocirculation of Zika, chikungunya, and dengue virus type 1. *Medicine (Baltimore)* **2016**; 95:e3201.
18. Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online* **2007**; 1:47–50.
19. Guo SW, Thompson EA. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* **1992**; 48:361–72.
20. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, **2017**. <http://www.R-project.org/>. Accessed 21 December 2020.
21. Benjamin DJ, Berger JO, Johannesson M, et al. Redefine statistical significance. *Nat Hum Behav* **2018**; 2:6–10.
22. Akaike H. A new look at statistical model identification. *IEEE Trans Autom Control* **1974**; 19:716–23.
23. Nguyen TB, Jayaraman P, Bergseng E, Madhusudhan MS, Kim CY, Sollid LM. Unraveling the structural basis for the unusually rich association of human leukocyte antigen DQ2.5 with class-II-associated invariant chain peptides. *J Biol Chem* **2017**; 292:9218–28.
24. Pieper J, Dubnovitsky A, Gerstner C, et al. Memory T cells specific to citrullinated α -enolase are enriched in the rheumatic joint. *J Autoimmun* **2018**; 92:47–56.
25. Kellogg EH, Leaver-Fay A, Baker D. Role of conformational sampling in computing mutation-induced changes in protein structure and stability. *Proteins* **2011**; 79:830–8.
26. Mandell DJ, Kortemme T. Computer-aided design of functional protein interactions. *Nat Chem Biol* **2009**; 5:797–807.
27. Alford RF, Leaver-Fay A, Jeliakov JR, et al. The Rosetta all-atom energy function for macromolecular modeling and design. *J Chem Theory Comput* **2017**; 13:3031–48.
28. Park H, Bradley P, Greisen P Jr, et al. Simultaneous optimization of biomolecular energy functions on features from small molecules and macromolecules. *J Chem Theory Comput* **2016**; 12:6201–12.
29. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* **2001**; 98:10037–41.
30. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph* **1996**; 14:33–8, 27–8.
31. Altschul SF, Gish W, Miller W, et al. Basic local alignment search tool. *J Mol Biol* **1990**; 215:403–410.
32. The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* **2019**; 47:D506–15.

33. Li Z, Zhao Y, Pan G, et al. A novel peptide binding prediction approach for HLA-DR molecule based on sequence and structural information. *BioMed Res Int* **2016**; 2016:1–10.
34. Maucourant C, Petitdemange C, Yssel H, et al. Control of acute arboviral infection by natural killer cells. *Viruses* **2019**; 11:131.
35. Reynolds CJ, Suleyman OM, Ortega-Prieto AM, et al. T cell immunity to Zika virus targets immunodominant epitopes that show cross-reactivity with other flaviviruses. *Sci Rep* **2018**; 8:672.
36. Lanteri MC, Kaidarova Z, Peterson T, et al. Association between HLA class I and class II alleles and the outcome of West Nile virus infection: an exploratory study. *PLoS One* **2011**; 6:e22948.
37. Reis PG, Sell AM, Sakita KM, et al. HLA-DRB1, DQA1 and DQB1 diversity in a mixed population of Paraná, Southern Brazil. *Hum Immunol* **2015**; 76:153.
38. Sotomaior VS, Fauz FR, Schafhauser C, et al. HLA-DQA1 and HLA-DQB1 alleles and haplotypes in two Brazilian Indian tribes: evidence of conservative evolution of HLA-DQ. *Hum Biol* **1998**; 70:789–97.
39. Nambala P, Su WC. Role of Zika virus prM protein in viral pathogenicity and use in vaccine development. *Front Microbiol* **2018**; 9:1797.
40. López-Camacho C, Abbink P, Larocca RA, et al. Rational Zika vaccine design via the modulation of antigen membrane anchors in chimpanzee adenoviral vectors. *Nat Commun* **2018**; 9:2441.
41. Kaufman J. Generalists and specialists: a new view of how MHC class I molecules fight infectious pathogens. *Trends Immunol* **2018**; 39:367–79.
42. Singh G, Pritam M, Banerjee M, Singh AK, Singh SP. Genome based screening of epitope ensemble vaccine candidates against dreadful visceral leishmaniasis using immunoinformatics approach. *Microb Pathog* **2019**; 136:103704.
43. Goulder PJ, Walker BD. HIV and HLA class I: an evolving relationship. *Immunity* **2012**; 37:426–40.
44. Pybus OG, Thézé J. Hepacivirus cross-species transmission and the origins of the hepatitis C virus. *Curr Opin Virol* **2016**; 16:1–7.
45. Dazert E, Neumann-Haefelin C, Bressanelli S, et al. Loss of viral fitness and cross-recognition by CD8⁺ T cells limit HCV escape from a protective HLA-B27-restricted human immune response. *J Clin Invest* **2009**; 119:376–86.
46. Manczinger M, Boross G, Kemény L, et al. Pathogen diversity drives the evolution of generalist MHC-II alleles in human populations. *PLoS Biol* **2019**; 17:e3000131.
47. Castanha PM, Cordeiro MT, Martelli CM, Souza WV, Marques ET Jr, Braga C. Force of infection of dengue serotypes in a population-based study in the northeast of Brazil. *Epidemiol Infect* **2013**; 141:1080–8.
48. Steinhagen K, Probst C, Radzimski C, et al. Serodiagnosis of Zika virus (ZIKV) infections by a novel NS1-based ELISA devoid of cross-reactivity with dengue virus antibodies: a multicohort study of assay performance, 2015 to 2016. *Euro Surveill* **2016**; 21:30426.
49. Wong SJ, Furuya A, Zou J, et al. A multiplex microsphere immunoassay for Zika virus diagnosis. *EBioMedicine* **2017**; 16:136–40.
50. Chen Y, Liao Y, Yuan K, Wu A, Liu L. HLA-A, -B, -DRB1 alleles as genetic predictive factors for dengue disease: a systematic review and meta-analysis. *Viral Immunol* **2019**; 32:121–30.