

Genetic Variants and Susceptibility to Neurological Complications Following West Nile Virus Infection

Mark Loeb,¹ Sasha Eskandarian,² Mark Rupp,⁴ Neil Fishman,⁵ Leanne Gasink,⁵ Jan Patterson,⁶ Jonathan Bramson,² Thomas J Hudson,³ and Mathieu Lemire³

¹Department of Medicine, Clinical Epidemiology, and Biostatistics, and ²Department of Pathology and Molecular Medicine, Michael G DeGroot Institute for Infectious Disease Research, McMaster University, Hamilton, Ontario; ³Ontario Institute for Cancer Research, Toronto, Canada; ⁴Department of Internal Medicine, University of Nebraska Medical Center, Omaha; ⁵Department of Medicine, University of Pennsylvania Medical Center, Hospital of the University of Pennsylvania, Philadelphia; and ⁶Department of Medicine, University of Texas Health Center and South Texas VA, San Antonio

To determine genetic factors predisposing to neurological complications following West Nile virus infection, we analyzed a cohort of 560 neuroinvasive case patients and 950 control patients for 13 371 mostly nonsynonymous single-nucleotide polymorphisms (SNPs). The top 3 SNPs on the basis of statistical significance were also in genes of biological plausibility: rs2066786 in *RFC1* (replication factor C1) ($P = 1.88 \times 10^{-5}$; odds ratio [OR], 0.68 [95% confidence interval {CI}, .56–.81]); rs2298771 in *SCN1A* (sodium channel, neuronal type I α subunit) ($P = 5.87 \times 10^{-5}$; OR, 1.47 [95% CI, 1.21–1.77]); and rs25651 in *ANPEP* (ananyl aminopeptidase) ($P = 1.44 \times 10^{-4}$; OR, 0.69 [95% CI, .56–.83]). Additional genotyping of these SNPs in a separate sample of 264 case patients and 296 control patients resulted in a lack of significance in the replication cohort; joint significance was as follows: rs2066786, $P = .0022$; rs2298771, $P = .005$; rs25651, $P = .042$. Using mostly nonsynonymous variants, we therefore did not identify genetic variants associated with neuroinvasive disease.

Over the past decade, West Nile virus (WNV) has emerged as an important human pathogen in North America, where it has been reported in a majority of states and provinces in the United States and Canada [1–3]. Of those infected, approximately 20% develop mild symptoms (“West Nile fever,” eg, fever, malaise, headache, myalgia, rash) [4] and about 1 in 150 develop meningitis or encephalitis [4, 5]. The incidence of severe neurological syndromes increases with age and includes encephalitis, meningitis, acute flaccid paralysis, peripheral neuropathy, polyradiculopathy,

optic neuritis, and acute demyelinating encephalitis [5–9]. Pathological changes within the central nervous system appear to be due to several factors, including the direct result of viral proliferation within neuronal and glial cells, cytotoxic immune response to infected cells, diffuse perivascular inflammation, and microglial nodule formation [10–12].

Given the fact that only a small minority of those infected develop severe disease and the fact that risk factors, apart from older age and immunosuppression [1, 13–15], are not well defined, there is a strong rationale to suspect a genetic predisposition to WNV neurological complications. Several putative susceptibility genes have been reported in single candidate gene studies. An association between symptomatic WNV disease and homozygosity for the *CCR5* Δ 32 mutation in the chemokine receptor gene *CCR5* was initially reported [16, 17]. More recently, this association was not replicated, but a study suggestive of a link to clinical manifestations of infection with *CCR5* Δ 32 mutation was reported [18]. The plausibility of a susceptibility

Received 13 March 2011; accepted 2 May 2011.

Potential of conflicts of interest: none reported.

Correspondence: Mark Loeb, MD, McMaster University, 1200 Main St W, Hamilton, ON L8N 3Z5, Canada (loebm@mcmaster.ca).

The Journal of Infectious Diseases 2011;204:1031–7

© The Author 2011. Published by Oxford University Press on behalf of the Infectious Diseases Society of America. All rights reserved. For Permissions, please e-mail: journals.permissions@oup.com

0022-1899 (print)/1537-6613 (online)/2011/2047-0010\$14.00

DOI: 10.1093/infdis/jir493

locus in the oligoadenylate synthetase (OAS) gene cluster was reported [19, 20]. More recently, a candidate gene study examining the OAS gene cluster alone suggested a predisposition to WNV infection [21].

Because exposure to WNV is difficult to control for, we sought to examine genetic susceptibility to complications given documented WNV infection. We hypothesized that the use of coding nonsynonymous markers would result in the discovery of ≥ 1 genetic variants associated with severe WNV disease that may be directly likely to lead to a causal mechanism. We report an association between the replication factor C1 gene, *RFC1*, and severe WNV disease in the first phase of association testing. We found no association between *CCR5* Δ 32 mutation and a susceptibility to severe complications to WNV infection. Similarly, we found no association between the OAS gene cluster and severe WNV disease. Our results indicate that *RFC1* may potentially play a role in the etiology of severe WNV illness.

PARTICIPANTS AND METHODS

Study Participants

During the period September 2004 through December 2009, we enrolled participants from states and provinces in the United States (Nebraska, Pennsylvania, Texas, and Colorado) and Canada (Alberta and Saskatchewan) that were among those with the highest number of individuals with severe WNV disease. Individuals were eligible if they were aged ≥ 18 years and had evidence of WNV infection on the basis of Centers for Disease Control and Prevention diagnostic laboratory criteria [22]. Case patients were defined as patients who met criteria for WNV infection along with modified case definitions for meningitis, encephalitis, or acute flaccid paralysis [9]. Control participants were individuals who met criteria for infection but who did not meet clinical case definitions for meningitis or encephalitis. Because the majority of testing for WNV is performed centrally by state or provincial laboratories, the state or provincial health departments contacted persons with laboratory-confirmed WNV infection by letter and by telephone to see if they were willing to be approached about the study. Clinical information about WNV infection was obtained by contacting patients' physicians and through review of medical records. The study received institutional review board approval from McMaster University, McGill University, Nebraska Medical Center, University of Pennsylvania, and University of Texas at San Antonio.

Genotyping

Blood samples were collected at the time of enrollment and genomic DNA obtained with the use of a DNA isolation kit (Qiagen). For the first stage of the study, genotyping was conducted using the Illumina HumanNS-12 BeadChip, based largely on single-nucleotide polymorphisms (SNPs) selected by the Wellcome Trust Case Control Consortium that included all

known nsSNPs with $>1\%$ minor allele frequency in European populations at the time of study design [23]. The Illumina HumanNS-12 chip included 13 371 SNPs, of which the majority were nonsynonymous variants but also included synonymous, untranslated, and tagSNPs. For the validation and replication stage, 2 panels of 34 SNPs each were designed using Sequenom MassARRAY IPLEX Gold; these included 47 SNPs with the highest association in the derivation cohort at 1 of 2 time points; 15 tagSNPs to fine map the gene that showed the highest association in the derivation cohort; 5 coding nonsynonymous SNPs uncovered in genes that were sequenced; and *CCR5* Δ 32 (rs333), a 32-base pair (bp) deletion in *CCR5* that was shown to be associated with clinical symptoms of WNV infection [18]. Duplicate samples were included, and laboratory technicians were blind to the case status of samples.

Sample Quality Control

Illumina HumanNS-12 Array. There were 1677 unique samples that were genotyped. Twenty samples were present in duplicate, and 1 was in triplicate. Four Centre d'Etude du Polymorphisme Humain controls were genotyped multiple times. Three duplicate samples that did not correspond to the same DNA were removed from further analysis. From the remaining samples, we calculated a pairwise reproducibility rate of 99.9926% before applying any filters on the markers. We computed the mean number of alleles shared identical by state in all possible pairs of individuals to detect population structure and cryptic relatedness in the sample. The set of SNPs used to compute this distribution was restricted to the autosomes and was selected in such a way that no 2 SNPs were allowed to be in linkage disequilibrium with each at $r^2 > 0.2$, in any window of 50 consecutive SNPs. To minimize admixture and other population effects, we restricted the genetic studies to samples of Caucasian origin. We excluded samples obtained from individuals who were not of self-declared non-Hispanic white ancestry, and we conducted a multidimensional scaling analysis comparing samples from the present study to HapMap samples derived from 3 populations [24] in order to identify samples from individuals whose ancestry was likely of non-Caucasian origin. Pairs of individuals who were identified as being likely related were broken up: we kept the case patient if both members of the pair were phenotypically discordant; otherwise, we kept the first participant enrolled. Seemingly related pairs involving 2 participants from distinct collection centers were discarded, on the basis that this was caused by either a sample switch or a contamination. Using this strategy, 29 individuals (17 control patients, 3 case-control pairs due to switches, and 3 case-control pairs with shared DNA) were excluded. Using all the SNPs on the X chromosome, we found inconsistencies regarding the sex of 6 individuals, who were excluded from the analysis. Samples that did not achieve a 90% call rate were excluded from the analysis.

Sequenom Assays. In a second stage, 2 panels of 34 SNPs were each genotyped in a total of 2350 unique DNA samples, a number that includes the 1677 samples genotyped on the Illumina HumanNS-12 array for the primary analysis. The samples that were removed in the primary analysis (except those that were removed for reason of low call rate) were also removed in the second stage. Samples that were not obtained from participants of self-declared non-Hispanic white ancestry were excluded. Samples that did not reach an 85% call rate in 1 panel were excluded from the analysis of that panel.

SNP Quality Control

To be included in the analysis, an SNP had to have a call rate $\geq 95\%$. Genotypic frequencies had to be consistent with the rules of Hardy-Weinberg equilibrium among control patients, at a significance level above a threshold that was chosen so that the distribution of Hardy-Weinberg significance levels among the remaining SNPs did not deviate from what was expected by chance alone ($P > .0005$ for the Illumina HumanNS-12 assays; $P > .001$ for the Sequenom assays).

Association Analysis

We used logistic regression to test for association between single SNPs and case-control status, assuming a log-additive effect of the alleles on the risk. For SNPs on the X chromosome, males were treated as homozygous females [25]. We adjusted for the collection center (ie, used as covariate) to account for possible local variations in allele frequencies or local selection. SNPs were deemed to be significant if their P values were below the 5% level, subject to a Bonferroni correction for the number of SNPs tested ($P < 4.7 \times 10^{-6}$, accounting for 10 625 tests). Power to detect an association at this significance level was calculated using CaTS [26].

Forty-eight samples were randomly selected among the cases for resequencing with a view to fine-mapping the *RFC1*, *SCN1A*, and *ANPEP* exons and promoter regions. Primers were designed for amplification of the coding regions, as well as 1kb regions upstream and downstream of the genes.

RESULTS

Nonsynonymous SNP Testing of WNV Severe Disease

We initially genotyped the Illumina HumanNS-12 array in 1677 participants: 608 case patients with neuroinvasive disease (112 meningitis, 72 encephalitis, 195 meningo-encephalitis, and 229 acute flaccid paralysis), 994 control patients, and 75 of equivocal or unknown status. There were 63 samples excluded (44 case patients, 16 control patients, and 3 of equivocal or unknown status) because of non-European ancestry (Supplementary figure 1), 26 samples excluded (2 case patients, 21 control patients, and 1 of unknown status) because of cryptic relatedness, 2 samples (1 case patient and 1 of unknown status) removed

because of low call rates ($< 95\%$), 6 samples (1 case patient and 5 control patients) removed for inconsistent sex information, and 3 samples genotyped in duplicate (3 control patients) but that corresponded to different DNAs. There were 307 SNPs excluded from analysis for not reaching a sufficient call rate; 52 SNPs failed Hardy-Weinberg equilibrium at a level $P < .0005$; and 2387 SNPs with minor allele frequency less than 1% (including 1246 monomorphic SNPs) were excluded. The first-stage analysis was performed on 10 625 SNPs in 560 case patients and 950 control patients. A summary of their characteristics is shown in Table 1.

The results from this analysis are summarized in Table 2 for all SNPs selected to undergo replication (see below). The 3 results with the highest levels of significance were from genes potentially biologically relevant: rs2066786 ($P = 1.88 \times 10^{-5}$; odds ratio [OR], 0.68 for the major allele [95% confidence interval {CI}, .57–.81]), an exonic synonymous SNP (Pro847Pro) found in the replication factor C1 gene (*RFC1* [MIM 102579]); rs2298771 ($P = 5.87 \times 10^{-5}$; OR, 1.47 [95% CI, 1.22–1.77]), a missense SNP (Ala1056Thr) in the sodium channel, voltage-gated, type I, α subunit gene (*SCN1A* [MIM 182389]); and rs25651 ($P = 1.44 \times 10^{-4}$; OR, 0.69 [95% CI, .56–.83]), a missense SNP (Ser752Asn) found in the alanyl (membrane) aminopeptidase gene (*ANPEP* [MIM 151530]). As described in detail below, each of these genes has biological plausibility, with *RFC1* potentially having a role in viral replication [27], *SCN1A* having a role in seizure disorders [28], and *ANPEP* a receptor for human coronavirus [29].

Novel Coding Nonsynonymous Variants in *RFC1*, *SCN1A*, and *ANPEP* Regions

The 3 genes were sequenced (exons and promoter regions only) in a set of 48 randomly chosen cases (totaling 96 chromosomes), in order to find potential novel candidate risk factors. A total of 5 coding nonsynonymous variants not already genotyped were uncovered in the 3 genes (1 in *RFC1*, 2 in *SCN1A*, and 2 in *ANPEP*). All these novel variants showed at most 2 instances of the nonreference allele out of the 96 chromosomes, making them rare variants, each with estimated frequency $< 2\%$ in the case patients. In addition, we confirmed a number of known and rare coding nonsynonymous SNPs that were already indexed in the dbSNP database at the time of the sequencing experiment.

Validation of Significant Markers

To validate findings from the first stage of the analysis, we genotyped 68 SNPs, including 47 SNPs taken on the basis of the results of the analysis of the Illumina HumanNS-12 array, at 2 different time points during the course of the project (the top 15 SNPs on the basis of a preliminary analysis of the first 445 case patients and 813 control patients collected, and the top 32 SNPs not already selected on the basis of the analysis of the complete sample, after pruning SNPs found to be in high linkage disequilibrium [$r^2 > 0.8$] with better ranked SNPs); 15 tagSNPs to try to refine the association seen in *RFC1* with rs2066786,

Table 1. Characteristics of Study Participants

Characteristic	Study Center									Total
	Alberta	Colorado	Manitoba	Nebraska	Ohio	Ontario	Pennsylvania	Saskatchewan	Texas	
Primary samples										
Control patients	75	73	0	357	0	36	41	340	28	950
Female sex, %	53.3	67.1	...	51	...	58.3	61	63.2	57.1	57.7
Mean age, years (SD)	49.6 (12.5)	50.8 (11)	...	54.6 (13.3)	...	52.7 (11.2)	54.7 (13.3)	52.5 (13.6)	52.5 (14.1)	53 (13.2)
Case patients	17	38	5	199	0	25	81	31	164	560
Acute flaccid paralysis	3	4	1	48	...	3	32	2	116	209
Encephalitis	6	8	0	33	...	7	3	0	11	68
Meningitis	3	5	0	41	...	3	28	15	7	102
Meningo-Encephalitis	5	21	4	77	...	12	18	14	30	181
Female sex, %	23.5	42.1	40.0	52.3	...	52.0	51.9	64.0	42.0	48.2
Mean age, years (SD)	56.8 (13.8)	56.6 (13)	60.2 (12.9)	55.6 (15.2)	...	54.7 (14.7)	62.9 (15.1)	59.4 (14.5)	61.3 (14.8)	58.7 (15)
Replication samples										
Control patients	3	24	0	182	6	0	21	0	60	296
Female sex, %	66.7	66.7	...	57.7	66.7	...	47.6	...	66.7	59.8
Mean age, years (SD)	54.7 (1.53)	54.3 (11.6)	...	54.8 (13.8)	58.3 (11.3)	...	55.2 (13.9)	...	56.6 (12.2)	55.2 (13.1)
Case patients	0	20	0	36	17	0	101	0	90	264
Acute flaccid paralysis	...	13	...	11	3	...	41	...	52	120
Encephalitis	...	0	...	7	0	...	5	...	2	14
Meningitis	...	3	...	12	5	...	25	...	13	58
Meningo-Encephalitis	...	4	...	6	9	...	30	...	23	72
Female sex, %	...	45	...	52.8	64.7	...	49.5	...	45.6	49.2
Mean age, years (SD)	...	59.2 (14.3)	...	56.9 (17.1)	70 (15.7)	...	64.6 (15.0)	...	61.5 (13.8)	62.4 (15.1)

Abbreviation: SD, standard deviation.

a variant that does not change the amino acid in the amino acid sequence of the gene; and the 5 coding nonsynonymous SNPs uncovered after sequencing the *RFC1*, *SCN1A*, and *ANPEP* genes. In addition, we included in a panel rs333, a 32-bp deletion in *CCR5* shown to be associated with increased risk of WNV infection, for which a design on the Illumina HumanNS-12 array was not available. Two control patients had insufficient call rates on both Sequenom panels and were excluded from the analysis; an additional 25 samples (6 case patients, 16 control patients, and 1 equivocal) failed only 1 Sequenom panel and were excluded only for SNPs on that panel. After excluding samples with low call rates, the SNPs were tested in the original set of (up to, depending on panel failure) 560 case patients and 950 control patients, in an additional set of 560 samples (264 case patients and 296 control patients), and in all samples jointly considered. The results are summarized in Supplementary table 1. For rs2066786 in *RFC1*, $P = .58$ in the replication set and a joint $P = .0022$. For rs2298771 in *SCN1A*, $P = .58$ in the replication cohort, with joint $P = .0050$. For rs25651 in *ANPEP*, the replication P value is .037, but the risk allele (the one that is more frequent among the case patients) is

different from the risk allele in the original sample set, and as a result the joint P value increases to .042. None of the following 44 ranked SNPs replicated in the additional samples genotyped.

Fine mapping of *RFC1* did not provide additional insights: significance levels are comparable to the original significance observed for rs2066786 and can be explained by the linkage disequilibrium structure in the region (Supplementary figure 2).

Effect of Rare Coding Nonsynonymous Variants

The rare coding nonsynonymous variants uncovered by sequencing *RFC1*, *SCN1A*, and *ANPEP* were genotyped. No significant associations (Fisher exact test) were found between these rare mutations and the case/control status (data not shown). Collapsing the counts of carriers of rare variants across SNPs within genes did not provide additional insights.

Lack of Replication of Associations Reported in the Literature

An association between symptomatic WNV disease and homozygosity for the *CCR5*Δ32 mutation in the chemokine receptor gene *CCR5* was initially reported [16, 17]. More recently, this association was not replicated, but results suggestive of a link to clinical manifestations of infection with *CCR5*Δ32

Table 2. Results From the Primary Association Analyses Between Single-Nucleotide Polymorphisms (SNPs) and Risk of Neuroinvasive Disease in Patients Infected With the West Nile Virus

SNP	Rank	Chromosome	Position	Allele		Frequency		HWE	OR (95% CI)	P
				1	2	Cases	Controls			
rs2066786 ^a	1	4	38978423	T	C	0.3869	0.4639	0.2952	0.6781 (.5675–.8101)	.0000188
rs2298771 ^a	2	2	166601033	C	T	0.3571	0.3014	0.1659	1.469 (1.218–1.772)	.0000587
rs25651 ^a	3	15	88136791	T	C	0.2857	0.3423	0.003926	0.685 (.5636–.8326)	.0001444
rs11575302 ^b	5	7	50575187	A	G	0.01964	0.01105	1	3.825 (1.867–7.838)	.0002469
rs2177336 ^b	8	3	197001272	T	C	0.1464	0.1989	0.1278	0.6659 (.5282–.8395)	.0005802
rs7163367 ^a	10	15	88061148	A	T	0.3964	0.4584	0.7443	0.7349 (.6154–.8777)	.0006718
rs10839601 ^b	12	11	6696711	A	G	0.1821	0.1484	0.6071	1.457 (1.163–1.824)	.001052
rs3738573 ^a	17	1	84636843	C	G	0.3222	0.368	0.5295	0.7398 (.6170–.8869)	.001129
rs3733890 ^b	18	5	78457714	A	G	0.2631	0.3193	0.5488	0.7293 (.6024–.8830)	.001214
rs323347 ^a	19	8	30825765	G	A	0.1571	0.1974	0.2596	0.6853 (.5438–.8637)	.001363
rs11652709 ^c	22	17	53626092	C	G	0.3527	0.2979	0.07428	1.335 (1.118–1.595)	.00145
rs1805073 ^b	23	5	78362505	C	G	0.2562	0.3047	0.3995	0.7346 (.6071–.8889)	.001519
rs2824721 ^b	25	21	18591988	G	A	0.2071	0.2484	0.02988	0.7123 (.5767–.8797)	.001638
rs1058587 ^b	26	19	18360421	G	C	0.2816	0.2405	0.1822	1.368 (1.126–1.664)	.001657
rs10778292 ^a	27	12	102784416	C	T	0.1179	0.1526	0.2576	0.6536 (.5010–.8527)	.001722
rs5370 ^b	29	6	12404240	T	G	0.2429	0.1979	0.3086	1.387 (1.127–1.707)	.002003
rs3816988 ^a	30	15	60898791	C	T	0.1986	0.1575	0.2685	1.423 (1.136–1.783)	.002161
rs12371985 ^b	32	12	93466790	A	G	0.01518	0.03316	1	0.3779 (.2021–.7064)	.002295
rs2288777 ^b	33	5	150900300	G	A	0.05446	0.07474	0.6419	0.5692 (.3958–.8186)	.002369
rs794999 ^b	34	3	12204014	A	G	0.2839	0.2479	0.6635	1.351 (1.113–1.640)	.002371
rs16826069 ^b	35	1	39569641	G	A	0.258	0.21	0.5573	1.369 (1.115–1.680)	.002647
rs2278428 ^b	36	19	60109865	C	A	0.06071	0.08105	0.2701	0.5904 (.4183–.8333)	.002725
rs2161468 ^b	38	19	9949270	G	C	0.4482	0.3868	0.632	1.301 (1.093–1.548)	.003051
rs2270962 ^a	39	10	102006033	T	C	0.05714	0.07526	0.6462	0.5893 (.4152–.8363)	.003071
rs2270915 ^b	40	5	32822145	G	A	0.2392	0.1986	0.2619	1.365 (1.111–1.678)	.003095
rs4432013 ^b	41	11	4687755	G	A	0.1911	0.2174	0.634	0.7146 (.5718–.8931)	.003141
rs1527014 ^b	42	12	16288943	G	A	0.03839	0.02789	0.03282	2.024 (1.267–3.235)	.0032
rs13095016 ^b	43	3	196999988	C	G	0.1437	0.1811	0.8272	0.6977 (.5492–.8865)	.003223
rs11079339 ^b	44	17	53625440	G	A	0.1607	0.1326	0.1565	1.423 (1.124–1.803)	.003431
rs2236358 ^b	47	1	226002384	A	G	0.05446	0.03789	1	1.836 (1.219–2.764)	.00361
rs6263 ^b	48	7	50563393	C	T	0.01071	0.02158	0.3566	0.3336 (.1592–.6989)	.003623
rs2523421 ^b	51	6	29498418	G	A	0.1146	0.09259	0.07721	1.496 (1.136–1.969)	.004082
rs11076256 ^a	53	16	57309966	T	C	0.09732	0.06789	0.795	1.585 (1.155–2.174)	.004313
rs1128349 ^b	54	7	72735589	T	C	0.4304	0.4716	0.3625	0.7757 (.6515–.9235)	.004326
rs7574414 ^b	55	2	227952148	A	G	0.1518	0.1289	0.4732	1.426 (1.116–1.822)	.004548
rs2287939 ^b	57	5	34034639	A	G	0.2571	0.3011	0.316	0.7519 (.6172–.9158)	.004606
rs2290182 ^b	59	16	65171508	C	T	0.09643	0.1279	0.243	0.6592 (.4931–.8813)	.004909
rs7143633 ^b	61	14	20619732	G	C	0.1616	0.1328	0.6737	1.418 (1.111–1.810)	.00505
rs3754274 ^b	63	1	84640197	A	G	0.271	0.3027	1	0.7602 (.6275–.9211)	.005131
rs7591849 ^b	64	2	159821126	G	T	0.467	0.4168	0.9468	1.276 (1.076–1.514)	.00514
rs11540407 ^b	65	12	74186648	T	C	0.2308	0.26	0.4003	0.7472 (.6088–.9171)	.005299
rs9380006 ^a	66	6	27764477	C	A	0.1482	0.1747	0.1773	0.7209 (.5711–.9099)	.005871
rs1047881 ^b	75	1	211097780	C	T	0.4768	0.4289	0.3896	1.271 (1.068–1.513)	.006946
rs2241988 ^a	89	3	57517212	T	C	0.4786	0.4184	0.5941	1.261 (1.061–1.499)	.008448
rs6046 ^a	132	13	112821159	A	G	0.1027	0.1221	0.8795	0.708 (.5396–.9291)	.01275
rs2240154 ^a	135	19	954171	T	C	0.1759	0.1553	0.9019	1.337 (1.063–1.681)	.01302
rs5748648 ^a	259	22	15660821	A	G	0.04286	0.06632	0.599	0.6385 (.4315–.9450)	.02491

SNPs were selected to undergo replication as detailed below. Abbreviations: A, adenine; C, cytosine; CI, confidence interval; G, guanine; HWE, Hardy-Weinberg equilibrium; OR, odds ratio; T, thymine.

^a Selected for replication according to results from 445 case patients and 813 control patients.

^b Selected for replication according to results from the complete sample of case patients and control patients.

^c Selected for replication according to results from the complete sample of case patients and control patients but replaced with rs10853004 in the replication panels. Frequency estimates and ORs (along with their 95% CIs) are calculated with respect to allele 1. Also included are significance levels for tests of HWE.

mutation were reported [18]. The plausibility of a susceptibility locus in the OAS gene cluster was reported [19, 20]. More recently, a candidate gene study examining the OAS gene cluster alone suggested a predisposition to WNV infection [21]. There were 13 (1.58%) of 821 case patients who were homozygous for the *CCR5*Δ32 deletion, compared with 13 (1.05%) of 1233 control patients (OR, 1.51 [95% CI, .70–3.27]; $P = .30$). For genes in the OAS cluster, ORs ranged from 1.1 to 1.13 for 4 OAS variants with all P values $>.05$ (Supplementary table 2).

Power to Detect an Association

We calculated the power that a sample consisting of 824 case patients and 1246 control patients (the sample used in the joint analysis) has to detect an association with a polymorphism whose risk is assumed to act additively with the number of risk alleles an individual possesses assuming homogeneity across collection centers. Supplementary figure 3 shows that the sample has 80% power to detect an association at $P < 4.7 \times 10^{-6}$ (a level deemed to be significant even after a Bonferroni correction for all SNPs tested) for relative risks as low as 1.5 (depending on the risk allele frequency).

DISCUSSION

We sought to test for common genetic polymorphisms associated with severe WNV complications in a North American population using mainly nonsynonymous SNPs. We identified a variant in each of 3 genes (*RFC1*, *SCN1A*, *ANPEP*) potentially implicated in susceptibility to such neurological complications. Joint analysis showed for *RFC1*-rs2066786, $P = .0022$ *SCN1A*-rs2298771, $P = .005$ and *ANPEP*-rs25651, $P = .042$. SNPs in other candidate genes (OAS, *TLR3* [Toll-like receptor 3 gene]) were not significant. Our findings reveal preliminary information that genetic variants may play a role in susceptibility to severe WNV disease.

Human replication factor C (RFC), also called activator-1, is a multimeric primer-recognition protein consisting of 5 distinct subunits [27]. Human RFC was purified from extracts of HeLa cells as a host factor essential for the *in vitro* replication of simian virus 40 DNA. RFC allows for efficient elongation of DNA in the presence of human single-stranded DNA binding protein. Using interaction cloning, Uchiyama et al found that the large subunit of RFC interacts with the DNA sequence repeats of telomeres [30]. They found that RFC recognizes the 5'-phosphate termini of double-stranded telomeric repeats. The authors suggested that RFC may be involved in telomere stability or turnover. Presumably, if *RFC1* plays a pathogenic role in WNV, it could be with respect to facilitating viral replication.

To explain the lack of replication of rs2066786, we assessed the effect of this variant on the 3 subcategories of neuroinvasive disease (meningitis, encephalitis, or acute flaccid paralysis). As shown in Table 2, there was no difference in effects. Population structure in the replication cohort is another possible

explanation for lack of replication. Indeed, continental differences in European ancestry have been described [31]. However, we did not identify any differences even adjusting for ancestry informative markers. We reviewed phenotype data and could not identify any systematic differences in phenotyping.

An association between the 32-bp deletion in *CCR5* and West Nile clinical symptoms has been reported [18]. We found no evidence of any effect. One possible reason for the discrepancy is that we compared case patients with control patients, all of whom were symptomatic when infected with WNV. In contrast, the published reports compared case patients with control patients with no symptoms. Notably, we found no association between any of the OAS polymorphisms and WNV. We believe that this highlights the importance of conducting human population-based studies. Although TLR was also reported to be a mechanism in a murine model, we saw no significant effect in this population-based study [32].

Strengths of this study include the rigorous definitions and procedures used to confirm case and control status. All clinical and laboratory data had to have supportive documentation, including laboratory records and clinical information. Selection of case patients and control patients was as unbiased as could be achieved. In the states and provinces that we selected, all of those who were infected were invited to participate through notification from public health laboratories where the great majority of testing occurred. We acknowledge that sample size for such an association study was limited, but this is a function of the number of WNV neuroinvasive cases seen in North America.

Funding

This work was supported by the National Institute of Allergy and Infectious Diseases, National Institutes of Health (contract HHSN266200400066C).

References

- Petersen LR, Hayes EB. West Nile virus in the Americas. *Med Clin North Am* **2008**; 92:1307–22.
- Gubler DJ. The continuing spread of West Nile virus in North America. *Clin Infect Dis* **2007**; 45:1039–46.
- Kramer LD, Styer LM, Ebel GD. A global perspective on the epidemiology of West Nile virus. *Annu Rev Entomol* **2008**; 53:61–81.
- Mostashari F, Bunning ML, Kitsutani PT, et al. Epidemic West Nile encephalitis, New York, 1999: results of a household-based seroepidemiological survey. *Lancet* **2001**; 358:261–4.
- Brown JA, Factor DL, Tkachenko N, et al. West Nile viremic blood donors and risk factors for subsequent West Nile fever. *Vector Borne Zoonotic Dis* **2007**; 7:479–88.
- O'Leary DR, Marfin AA, Montgomery SP, et al. The epidemic of West Nile virus in the United States, 2002. *Vector Borne Zoonotic Dis* **2004**; 4:61–70.
- Davis LE, DeBiasi R, Goade DE, et al. West Nile virus neuroinvasive disease. *Ann Neurol* **2006**; 60:286–300.
- Pepperell C, Rau N, Kraiden S, et al. West Nile virus infection in 2002: morbidity and mortality among patients admitted to hospital in south central Ontario. *CMAJ* **2003**; 168:1399–405.
- Sejvar JJ, Haddad MB, Tierney BC, et al. Neurologic manifestations and outcome of West Nile virus infection. *JAMA* **2003**; 290:511–5.

10. Smith RD, Konoplev S, DeCourten-Myers G, et al. West Nile virus encephalitis with myositis and orchitis. *Hum Pathol* **2004**; 35:254–8.
11. Agamanolis DP, Leslie MJ, Caveny EA, et al. Neuropathological findings in West Nile virus encephalitis: a case report. *Ann Neurol* **2003**; 54:547–51.
12. Gyure KA. West Nile virus infections. *J Neuropathol Exp Neurol* **2009**; 68:1053–60.
13. Iwamoto M, Jernigan DB, Guasch A, et al. Transmission of West Nile virus from an organ donor to four transplant recipients. *N Engl J Med* **2003**; 348:2196–203.
14. Kumar D, Drebot MA, Wong SJ, et al. A seroprevalence study of West Nile virus infection in solid organ transplant recipients. *Am J Transplant* **2004**; 4:1883–8.
15. Kumar D, Prasad GV, Zaltzman J, et al. Community-acquired West Nile virus infection in solid-organ transplant recipients. *Transplantation* **2004**; 77:399–402.
16. Glass WG, McDermott DH, Lim JK, et al. CCR5 deficiency increases risk of symptomatic West Nile virus infection. *J Exp Med* **2006**; 203:35–40.
17. Lim JK, Louie CY, Glaser C, et al. Genetic deficiency of chemokine receptor CCR5 is a strong risk factor for symptomatic West Nile virus infection: a meta-analysis of 4 cohorts in the US epidemic. *J Infect Dis* **2008**; 197:262–5.
18. Lim JK, McDermott DH, Lisco A, et al. CCR5 deficiency is a risk factor for early clinical manifestations of West Nile virus infection but not for viral transmission. *J Infect Dis* **2010**; 201:178–85.
19. Bonnevie-Nielsen V, Field LL, Lu S, et al. Variation in antiviral 2',5'-oligoadenylate synthetase (2'5'AS) enzyme activity is controlled by a single-nucleotide polymorphism at a splice-acceptor site in the OAS1 gene. *Am J Hum Genet* **2005**; 76:623–33.
20. 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.
21. Lim JK, Lisco A, McDermott DH, et al. Genetic variation in OAS1 is a risk factor for initial infection with West Nile virus in man. *PLoS Pathog* **2009**; 5:e1000321.
22. Centers for Disease Control and Prevention. Encephalitis or meningitis, arboviral (includes California serogroup, Eastern equine, St. Louis, Western equine, West Nile, Powassan). 2001. Case Definition. http://www.cdc.gov/osels/ph_surveillance/nndss/casedef/encephalitiscurrent.htm Accessed 17 August 2011.
23. The Wellcome Trust Case Control Consortium. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* **2007**; 39:1329–37.
24. The International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature* **2007**; 449:851–61.
25. Clayton DG. Sex chromosomes and genetic association studies. *Genome Med* **2009**; 1:110.
26. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet* **2006**; 38:209–13.
27. Okumura K, Nogami M, Taguchi H, et al. Assignment of the 36.5-kDa (RFC5), 37-kDa (RFC4), 38-kDa (RFC3), and 40-kDa (RFC2) subunit genes of human replication factor C to chromosome bands 12q24.2-q24.3, 3q27, 13q12.3-q13, and 7q11.23. *Genomics* **1995**; 25:274–8.
28. Baulac S, Gourfinkel-An I, Picard F, et al. A second locus for familial generalized epilepsy with febrile seizures plus maps to chromosome 2q21-q33. *Am J Hum Genet* **1999**; 65:1078–85.
29. Yeager CL, Ashmun RA, Williams RK, et al. Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature* **1992**; 357:420–2.
30. Uchiyama F, Ohta T, Tanuma S. Replication factor C recognizes 5-prime-phosphate ends of telomeres. *Biochem Biophys Res Commun* **1996**; 229:310–5.
31. Heath SC, Gut IG, Brennan P, et al. Investigation of the fine structure of European populations with applications to disease association studies. *Eur J Hum Genet* **2008**; 16:1413–29.
32. Wang T, Town T, Alexopoulou L, Anderson JF, Fikrig E, Flavell RA. Toll-like receptor 3 mediates West Nile virus entry into the brain causing lethal encephalitis. *Nat Med* **2004**; 10:1366–73.