Phylogenetic analysis of West Nile virus in Quebec, Canada, 2004–2016: Co-circulation of distinct variants harbouring conserved amino acid motifs in North America

Christian Therrien[a,*], Éric Fournier[a], Antoinette Ludwig[b], Joel Ménard[a], Hugues Charest[a], Christine Martineau[a]

[a] Laboratoire de santé publique du Québec, Institut national de santé publique du Québec. 20045 Chemin Sainte-Marie, Saint-Anne-de-Bellevue, H9X 3Y3, QC, Canada
[b] Laboratoire National de microbiologie, Agence de santé publique du Canada, 3200 Sicotte, Saint-Hyacinthe, QC, J2S 2M2, Canada

ABSTRACT

West Nile virus (WNV) was introduced for the first time in the western hemisphere in 1999 in New York City. In 2002, a phenotype-modifying mutation (Env-V159A) defined the first North American genotype WN02. So far, three genotypes has been described in North America but little is known about WNV evolution in Canada. We report the phylogenetic characterization of twenty-six WNV genomes isolated from mosquitoes in the province of Quebec. WNV strains found in Quebec are phylogenetically related to American strains collected in northern and southern regions. We also noted the presence of two robust monophyletic groups of isolates characterized by distinct conserved amino acid motifs. These emerging genotypes were detected for several years in different ecosystems. These results highlight the need for the maintenance of a nationwide surveillance to follow the dispersion of emergent WNV genotypes.

1. Introduction

West Nile virus (WNV) is an arthropod-borne flavivirus maintained in nature in an enzootic cycle between mosquito vectors and vertebrate hosts. Birds species of the Passeriformes order are the natural hosts and reservoirs for WNV (Pérez-Ramírez et al., 2014). Humans and other mammals are considered dead end hosts since they do not produce sufficiently high and prolonged viremia for efficient viral transmission during the blood meal of biting mosquitoes.

The majority of people infected by WNV are asymptomatic while 25% of them will experience febrile illnesses with non-specific flu-like symptoms (Petersen et al., 2013). Only a small fraction of susceptible patients (< 1%) will progress to severe neurological disorders which can conclude in death for ≈ 10% of the cases or leave patients with neurological sequelae (Petersen et al., 2013).

WNV is a single strand, positive-sense RNA enveloped virus harbouring a genome of 11,029 nucleotides (nt) (Grinev et al., 2014). The genome structure contains an open reading frame (ORF) of 10,302 nt flanked by a 5′ untranslated region (UTR) with a type I cap (m^{7}GpppAm) and a 3′UTR. The ORF codes for a unique polyprotein of 3433 amino acids which is cleaved into three structural (capsid, pre-membrane and envelope) and seven non-structural (NS1, NS2A, NS2B,

NS3, NS4A, NS4B and NS5) proteins by cellular and viral proteases (Assenberg et al., 2009).

Phylogenetic analysis of WNV genomes obtained from strains found in different parts of the world revealed the existence of several putative genetic lineages (Fall et al., 2017). The emergence of WNV in the Western Hemisphere was documented for the first time in New York City in the late summer of 1999 (Centers for Disease Control and Prevention (CDC), 1999). The complete genome of the New York strain (NY99) isolated from the brain of a dead Chilean flamingo was sequenced and clustered in WNV lineage 1a following phylogenetic analyses (Lanciotti et al., 1999). The newly introduced WNV strain was highly virulent to resident birds and caused several human cases of encephalitis and seven deaths (Centers for Disease Control and Prevention (CDC), 1999). The substitution of a threonine by a proline residue at position 249 of the viral helicase domain of NS3 was pointed out to play an important role in avian neurovirulence of the NY99 strain compared to other lineage 1 WNV strains of the old world (Brault et al., 2007). In 2002, a conserved amino acid substitution in the envelope protein (V159A) was described in several North American strains (Davis et al., 2005). Genetic and phenotypic studies of this novel variant showed a decreased extrinsic incubation period in *Culex tarsalis* and higher viral titers in *Culex pipiens* following experimental infection

---

(Moudy et al., 2007), the most important WNV vector species in North America. This adaptation to autochthonous vectors initiated the rapid dispersion of WNV throughout the continent. To date, several phylogenetic studies addressed the evolution of WNV in North America (Añez et al., 2013; Duggal et al., 2015; Grinev et al., 2016, 2008; Mann et al., 2013; McMullen et al., 2011). The analysis of hundreds of complete or partial genomic sequences of WNV strains, isolated from different hosts since 1999 in the United States, showed a relatively small but progressive genetic diversity and no significant geographical clustering of genotypes (Añez et al., 2013; Grinev et al., 2016; Mann et al., 2013). So far, three genotypes have been described in North America; NY99, NA/WN02 and SW/WN03. The NY99 genotype was mostly detected until 2004 and was gradually displaced by the NA/WN02 genotype (Giallonardo et al., 2016). The latter is defined in part by a conserved Env-V159A amino acid change, which is found in all North American WNV strains since 2002. The SW/WN03 genotype includes the Env-V159A change and an additional fixed non-synonymous mutation in the NS4A gene yielding A85T substitution (McMullen et al., 2011). Strains of the SW/WN03 genotype were mostly reported in the Central and Southwestern states but were also detected occasionally in New York State (Añez et al., 2013; Grinev et al., 2016; McMullen et al., 2011).

In Canada, the first incursion of WNV was documented in Windsor, Ontario in August 2001 by the report of a laboratory-confirmed infection of a dead bird (Giordano et al., 2017). The following year several human and equine encephalitis cases were reported in the provinces of Ontario and Quebec (Drebot et al., 2003). The phylogenetic analysis of 74 North American WNV strains revealed that strains of the NA/WN02 genotype were present in Canada as early as 2002 (Davis et al., 2005). Besides large WNV genomic studies performed by northern American states such as New York (Davis et al., 2005; Ebel et al., 2004), Illinois (Amore et al., 2010) and Connecticut (Armstrong et al., 2011), no studies addressed the evolution of WNV in Canada. In this study, we report the phylogenetic characterization of twenty-six WNV genomes isolated from *Culex pipiens* mosquitoes in the province of Quebec, Canada over a 7-year period.

2. Material and methods

2.1. Mosquito pools and WNV detection

The public health laboratory of the Province of Quebec (LSPQ) has a collection of over 44,000 mosquito pools. These pools were collected during the active surveillance programs held in Quebec from 2003 to 2006 and from 2013 to 2016, corresponding roughly to the two principal West Nile virus outbreaks recorded in Quebec. All RNA samples used in this study were extracted from *Cx. pipiens-restuans* pools, except for sample 21303VEX which was extracted from *Ae. vexans* pool. RNA extraction and real-time RT-PCR assays were used to detect WNV according to a previously published method (Lanciotti et al., 2000). WNV RNA positive samples were kept at −70 °C until further analyses. Twenty-nine WNV RNA positive samples were selected to cover the widest time and geographical distribution possible (Table 1). The selected isolates were geolocalized on a map of the south of Quebec using trapping site coordinates and identified by their genotype and sampling year.

2.2. Genomic sequencing

RNA-based genomic sequencing was attempted but failed to generate high levels of flavivirus-specific sequences due to high background interference from mosquito RNA (data not shown). A targeted approach was then put forward to specifically amplify WNV genomes by using four overlapping RT-PCR amplicons to cover most of the genome (nt 95 to 10,552), as reported previously (Bahuon et al., 2012). Nested RT-PCR was also performed when needed for low yield products using supplementary primers (Grinev et al., 2008). The RT-PCR

Table 1
List of selected WNV isolates collected in the province of Quebec, used for the genomic sequencing project (n = 29).

Strain ID	Trap ID	Year	City	GenBank accession no.
5237CPR1	OTW-305	2003	Ottawa	n/a
5889CPR1	LAV-050	2003	Laval	n/a
5326CPR1	MON-027	2003	Montréal	n/a
9120CPR2	HYA-008	2004	Saint-Hyacinthe	MH819468
10072CPR1	VAR-001	2004	Varennes	MH819461
10208CPR3	LAV-041	2004	Laval	MH819471
10704CPR1	RIC-003	2004	St-Jean-sur-Richelieu	MH819464
13217CPR4	DOR-001	2005	Dorval	MH819451
13452CPR3	SAP-001	2005	Sainte-Anne-des-Plaines	MH819457
13494CPR1	BLA-021	2005	Blainville	MH819458
13923CPR1	SMI-902	2005	Villeray-St-Michel	MH819448
21285CPR1	BOT-801	2006	Rosemont	MH819454
21303VEX	MON-832	2006	Côte St-Luc-Hamstead	MH819469
30712CPR1	MIR-005	2006	Mirabel	MH819455
VN1316CPR1	MON-026	2013	Montréal	MH819460
VN1329CPR1	POI-002	2013	Pointe-Claire	MH819467
VN1429CPR1	CHA-006	2013	Chateauguay	MH819447
VN1533CPR1	BOU-004	2013	Longueuil	MH819456
VN4365CPR1	MIC-101	2014	Montréal	MH819463
VN4794CPR1	LON-102	2014	Longueuil	MH819459
VN5038CPR1	AHU-118	2014	Montréal	MH819470
VN5229CPR1	HOM-114	2014	Montréal	MH819449
VN6319CPR1	BOU-008	2015	Boucherville	MH819462
VN6325CPR1	BRU-006	2015	Saint-Bruno	MH819465
VN6364CPR1	JUL-001	2015	Sainte-Julie	MH819452
VN6490CPR1	PAT-005	2015	Montréal	MH819446
VN6806CPR1	LAM-001	2016	Saint-Lambert	MH819453
VN6862CPR1	LAV-051	2016	Laval	MH819450
VN6864CPR2	MIC-101	2016	Montréal	MH819466

reactions were performed with the Superscript-III-One-Step RT-PCR System with Platinum Taq (Thermo Fisher Scientific, Mississauga, ON, Canada) and the nested RT-PCR were done with the Taq Expand High Fidelity PCR System (Roche Diagnostics, Laval, QC, Canada) following the manufacturer's instructions. All amplicons were purified using the QIAquick 96 PCR Purification Kit (QIAGEN), quantified with the Nanodrop 2000 (Thermo Fisher Scientific) and pooled at equimolar concentrations. The final DNA concentration was confirmed with the Qubit dsDNA HS assay (Thermo Fisher Scientific). DNA from each pool was used for library preparation using the Nextera XT DNA library preparation kit (Illumina, San Diego, CA, USA) following the manufacturer's instructions. Sequencing was carried out on the MiSeq instrument (Illumina) using a MiSeq v2 500-cycle kit. The quality of the raw sequence reads was verified with FastQC (Andrews, 2010). Adaptor and poor quality sequences (Q score < 20) were removed from reads with Trimmomatic v0.36 (Bolger et al., 2014). SMALT v0.7.6 (Ponstingl and Ning, 2010) was used to map sequencing reads on a WNV reference genome (Genbank accession number: AY646354). The consensus sequence from each sample was then recovered with SAMtools v0.1.19 (Li et al., 2009) and used in further analyses. Raw sequence data was deposited in NCBI under the BioProject accession PRJNA480662 and the SRA accession SRX4379698. Twenty-six assembled near-complete WNV whole genome sequences (GenBank accession nos. MH819446 -MH819471) were deposited in Genbank (Table 1). Genetic distances between the sequences and the prototype NY99 (AF196835) were determined with MEGA7 (Kumar et al., 2016).

2.3. WNV sequence dataset selection and phylogenetic analysis

In order to establish the genetic relationship of WNV Quebec isolates with North American isolates, 1090 WNV reference genomes were retrieved from the GenBank database. These sequences were chosen in order to span a representative range in terms of collection years, hosts

Table 2

Number of WNV sequences included in the dataset for phylogenetic studies according to years of collection, hosts and regions of isolation (n = 1116).

Years		Host	U.S. Regions	Countries
1999 (11)	2008 (74)	Mosquito (549)	Northeast (568)[a]	USA (1069)
2000 (18)	2009 (68)	Bird (498)	South (195)	Canada, Quebec (26)
2001 (86)	2010 (33)	Human (58)	Midwest (171)	Mexico (15)
2002 (148)	2011 (44)	Other (11)	West (160)	Columbia (3)
2003 (164)	2012 (105)		Other (22)	Argentina (2)
2004 (82)	2013 (31)			British Virgin Island (1)
2005 (76)	2014 (14)			
2006 (78)	2015 (13)			
2007 (64)	2016 (7)			

[a] Quebec isolates were included in the northeast region.

and regions (Table 2). U.S. regions were attributed to each sequence based upon the Census Regions and Divisions map of the United States (United States Census Bureau, Geography Division, 2013). WN02 and WN03 labels were given to each sequence based upon the presence of the ENV V159A and the combination of ENV-V159A and NS4A A85T substitutions, respectively. Duplicated sequences were removed from the dataset and sequences harbouring complete open reading frames (ORF, 10,299 nucleotides) were selected for further analyses. Multiple sequence alignment of the Quebec and North American ORF dataset was performed using MAFFT version v7.419 (Katoh and Standley, 2013). Selection of the best fitting model of molecular evolution was done with jModelTest v2.1.10 (Darriba et al., 2012; Guindon and Gascuel, 2003). According to the calculated Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC), the General Time Reversible (GTR) combined with a gamma (Γ) distribution of rate across site was selected. Bayesian Markov Chain Monte Carlo (MCMC) approach was used with the GTR + Γ model to infer the most probable phylogenetic tree from the posterior distribution. Exabayes mpi v1.5 (Aberer et al., 2014), which is well adapted for large-scale sequence analysis, was used for that purpose. Two independent runs were launched for 10,000,000 cycles with a sampling frequency of 100. Runs convergence in the parameters stationary distribution and sampling quality were evaluated by visualizing the evolution of the LnL in Tracer version v1.7.1 (Rambaut et al., 2018) and comparing the resulting Average Standard Deviation of Split Frequencies (ASDSF) diagnostic value and Effective Sampling Size (ESS) with the recommended ones (respectively 1–5% and > 100). Final consensus phylogenetic tree has been built with the 'sumt' command from the MrBayes package mpi v3.2.7a (Ronquist et al., 2012) after discarding the first 25% of sampled trees corresponding to the burn-in phase. Finally, consensus tree was plotted and branches coloured according to isolate genotype with R v3.5.2 using Ape library v5.3 (Paradis and Schliep, 2019).

3. Results

3.1. Mosquito surveillance and WNV screening

The mosquito surveillance program was initiated in the Province of Quebec soon after the first reported human cases in 2002 (Drebot et al., 2003). The active mosquito surveillance was discontinued for several years (2007–2012) following a decrease in the number of human cases. It was reinitiated in 2013 after the largest outbreak of human WNV disease cases (n = 134) in the summer of 2012 (http://www.msss.gouv. qc.ca/professionnels/zoonoses/virus-du-nil-occidental-vno/tableau-des-cas-humains-archives/). A total of 459 WNV positive mosquito pools (1%) were detected by RT-PCR out of 44,079 mosquito pools tested at the public health laboratory over an 8-year period. The tests were prioritized according to the mosquito species and its vector competence. The highest number of positive pools (≈ 90 %) was found among *Cx. pipiens-restuans*. Other WNV positive mosquitoes were found including *Ae. vexans* (10%), *Ochlerotatus sp* (< 3%) and *Anopheles sp*

(< 3 %).

3.2. Nucleotide variation and amino acid substitution

Nearly complete genomic sequences (≈ 94%) were obtained for 26 WNV Quebec isolates. The open reading frame (ORF) was completely sequenced, including parts of the non-coding regions. Unfortunately, isolates collected before 2004 could not be sequenced, potentially due to sample degradation. Sequence alignment comparison of the 26 nucleotide sequences with the prototype NY99 (AF196835) WNV strain identified 30–83 nt differences (0,29–0,81%) per 10,302 nt of the ORF. The number of amino acid substitutions ranged from two to seven amino acids, affecting mostly non-structural proteins (Supplementary Table 1). Conserved amino acid substitutions found in > 3 representative isolates are shown in Table 3. All isolates shared the conserved Env-V159A substitution, the typical genetic signature of the dominant NA/WN02 genotype firstly detected in Texas in 2002 and

Table 3

Conserved amino acid substitutions in WNV isolates, Quebec, Canada, 2004–2016 compared to NY99, WN02 and WN03 prototype strains[a].

Strain	Year	E	NS2A				NS3	NS4A	NS4B		
		159	52	95	188	208	334	85	14	104	240
NY99	1999	V	T	L	R	A	S	A	S	G	I
NA/WN02	2002	A	.	.	.	.	.	.	.	.	.
SW/WN03	2003	A	.	.	.	.	.	T	.	.	.
9120CPR2	2004	A	.	.	.	.	.	.	.	.	M
10072CPR1	2004	A	.	.	.	.	.	.	.	E	.
10208CPR3	2004	A	.	.	.	.	.	.	.	.	.
10704CPR1	2004	A	.	.	.	.	.	.	.	.	.
13217CPR4	2005	A	.	.	.	S	.	.	.	.	.
13452CPR3	2005	A	.	.	.	S	.	.	.	.	.
13494CPR1	2005	A	.	.	.	S	.	.	.	.	.
13923CPR1	2005	A	.	.	.	.	.	.	.	.	.
21285CPR1	2006	A	.	.	.	.	.	.	.	.	.
21303VEX2	2006	A	.	.	.	.	.	.	.	.	.
30712CPR1	2006	A	.	.	.	.	.	.	.	E	.
VN1316CPR1	2013	A	.	.	K	.	.	.	.	.	M
VN1329CPR1	2013	A	.	.	K	.	.	.	.	.	M
VN1429CPR1	2013	A	I	F	.	.	T	.	I	E	.
VN1533CPR1	2013	A	I	F	.	.	T	.	I	.	.
VN4365CPR1	2014	A	I	F	.	.	T	.	I	.	.
VN4794CPR1	2014	A	.	.	K	.	.	.	.	.	M
VN5038CPR1	2014	A	I	F	.	.	T	.	I	.	.
VN5229CPR1	2014	A	I	F	.	.	T	.	I	.	.
VN6319CPR1	2015	A	.	.	.	.	.	.	.	.	.
VN6325CPR1	2015	A	.	.	K	.	.	.	.	.	M
VN6364CPR1	2015	A	I	F	.	.	T	.	I	.	.
VN6490CPR1	2015	A	.	.	K	.	.	.	.	.	M
VN6806CPR1	2016	A	.	.	K	.	.	.	.	.	M
VN6862CPR1	2016	A	.	.	K	.	.	.	.	.	M
VN6864CPR2	2016	A	.	.	K	.	.	.	.	.	M

[a] For simplicity, only amino acid substitutions found three times or more are shown. Reference strains used: NY99 (AF196835), NA/WN02 (AY646354) and SW/WN03 (JF415921).
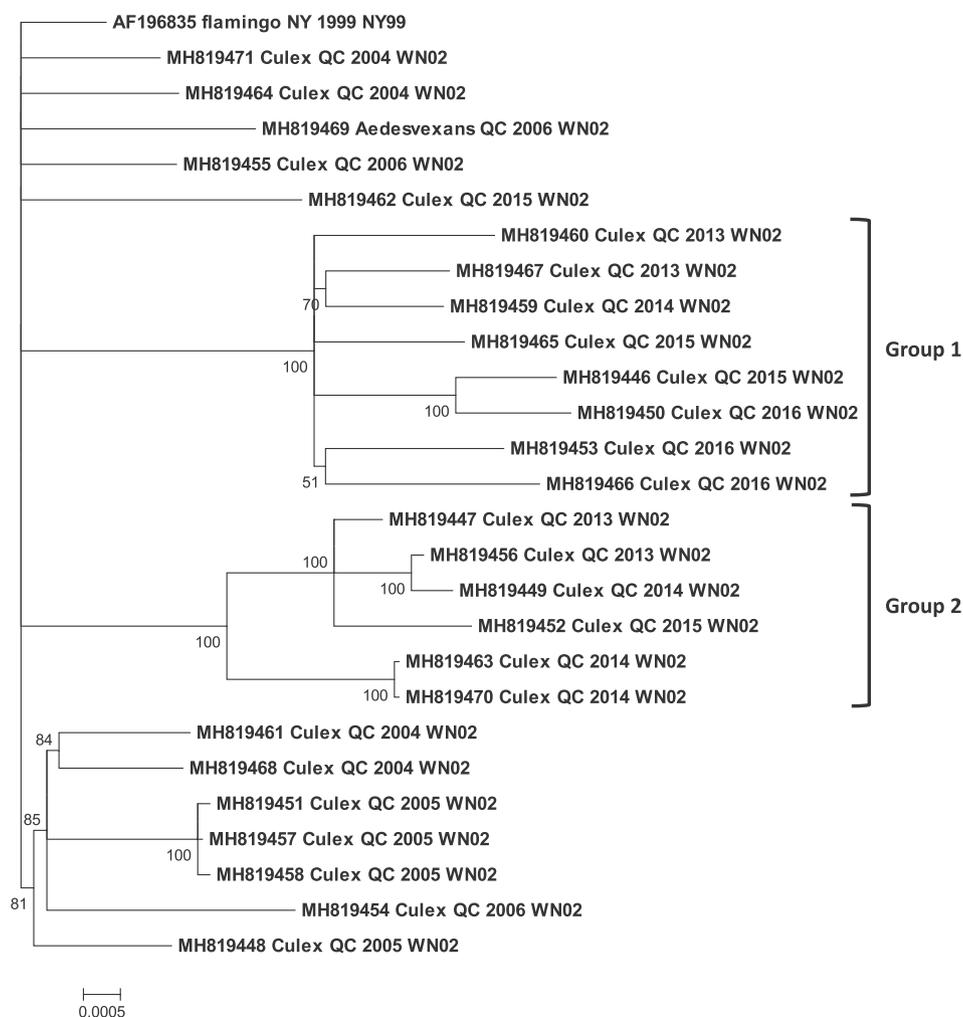
Fig. 1. The phylogenetic tree of Quebec West Nile virus isolates collected in 2004–2016. The tree was inferred using the 26 Quebec genomic sequences rooted with the prototype NY99 (AF196835).

now circulating in all American states (Beasley et al., 2003). Two distinct conserved amino acid motifs were observed in 14 isolates that were not found in the invading NY99 nor the dominant NA/WN02 strains. One of these motifs is characterized by the presence of two amino acids (termed KM for the remainder of the manuscript) changes in the non-structural proteins NS2A-R188K and NS4B-I240M. The other conserved motif (termed IFTI) is defined by four substitutions in three non-structural proteins, NS2A (-T52I and -L95F), NS3-S334T and NS4B-S14I. Three 2005 isolates shared the NS2A-A208S substitution, which was not detected in the following years. Finally, the NS4B-G104E substitution was detected in three isolates from 2004, 2006 and 2013. The remaining isolates harboured unique non-conserved amino acid substitutions. Interestingly, the genetic signature of the SW/WN03 genotype (NS4A-A85T) was not detected in Quebec isolates.

3.3. Phylogenetic analysis

The genetic relationship of the Quebec isolates was analyzed through Bayesian analysis. As shown in Fig. 1, the tree topology is characterized by the presence of two robust monophyletic clusters. These two groups of clustered sequences are distinguished by the presence of two specific amino acid motifs (KM or IFTI) identified by sequence alignment (Fig. 1). Group 1 is composed of eight isolates (VN6490, VN6862, VN1316, VN1329, VN4794, VN6325, VN6806 and VN6864) collected over a period of four years from 2013 to 2016 in three regional health units (Laval, Montreal and Monteregie). Several

group 1 isolates collected one year apart such as VN6490 and VN6862 are clustering together. Group 2 is composed of six isolates (VN1533, VN5229, VN1429, VN6364, VN4365 and VN5038) collected from 2013 to 2015 in two regional health units (Montreal and Monteregie). The average genetic distance of groups 1 and 2 sequences from NY99 (genbank No AF196835) is 0.68 % (min : 0.60 %; max 0.78 %) and 0.54 % (min : 0.49 %; max : 0.62 %), respectively. A third group composed of three 2005 isolates with the NS2A-A208S substitution (13452, 13217 and 13494) formed a well-resolved branch with isolates collected from 2004 to 2006. Besides the 2015 sequences, there was no shared specific amino acids motif among the other sequences of that node. Furthermore, the NS2A-A208S variants were not observed in the following years. Contrarily to the other isolates collected since 2013, the VN6319 isolate (2015) did not show any of the conserved motifs, suggesting that two subtypes were co-circulating in Quebec during that period.

The twenty-six sequenced isolates selected for this study were collected in four health units located in and around Montreal Island in the southern part of the province of Quebec (Fig. 2). Spatial distribution of isolates did not reveal any specific patterns of localization according to subtypes or years of collection.

An additional phylogenetic analysis was performed using the isolates from this study and 1090 published WNV genomes from North America retrieve from GenBank (Supplementary Fig. 1 for the completely resolved phylogenetic tree). As predicted by the amino acid substitution profiles, all Quebec sequences clustered within the NA/WN02 genotype (Fig. 3). The two well-resolved nodes (group 1 and 2)
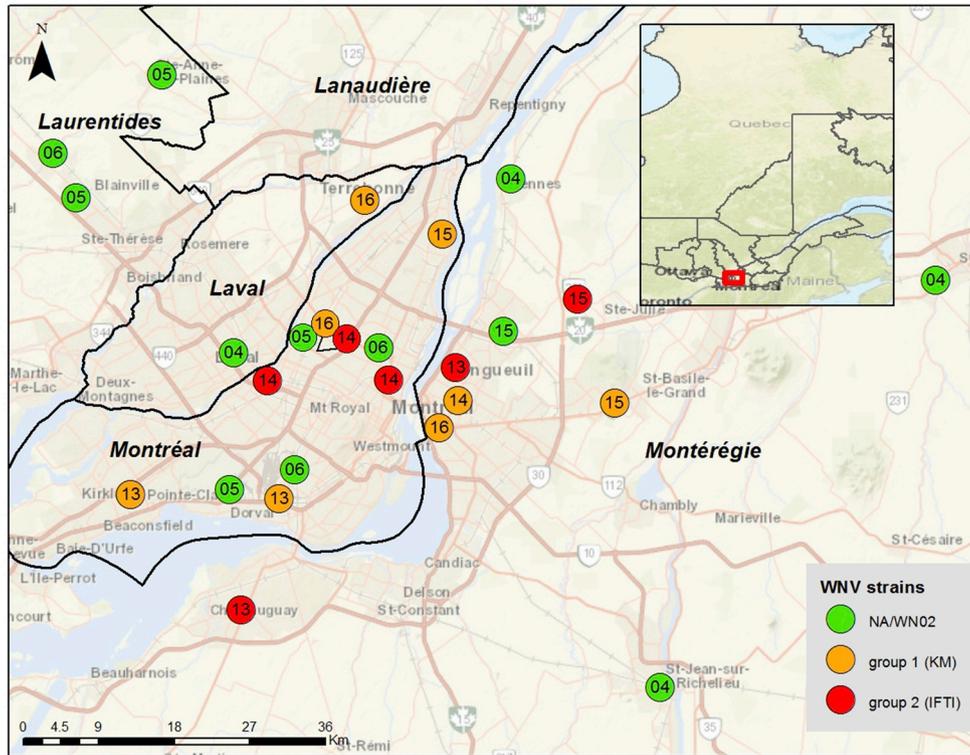
Fig. 2. Spatial distribution of WNV isolates collected in Quebec. The map is stratified by genotype or subtype and year of collection.

observed in Fig. 1 were also found in the North American phylogenic tree.

Group 1 includes sequences isolated from human, bird and

mosquito hosts. These isolates were collected between 2010 and 2016 in all regions of the US. Sequence alignment analysis revealed that the majority of group 1 sequences possess the NS2A-R188K and NS4B-
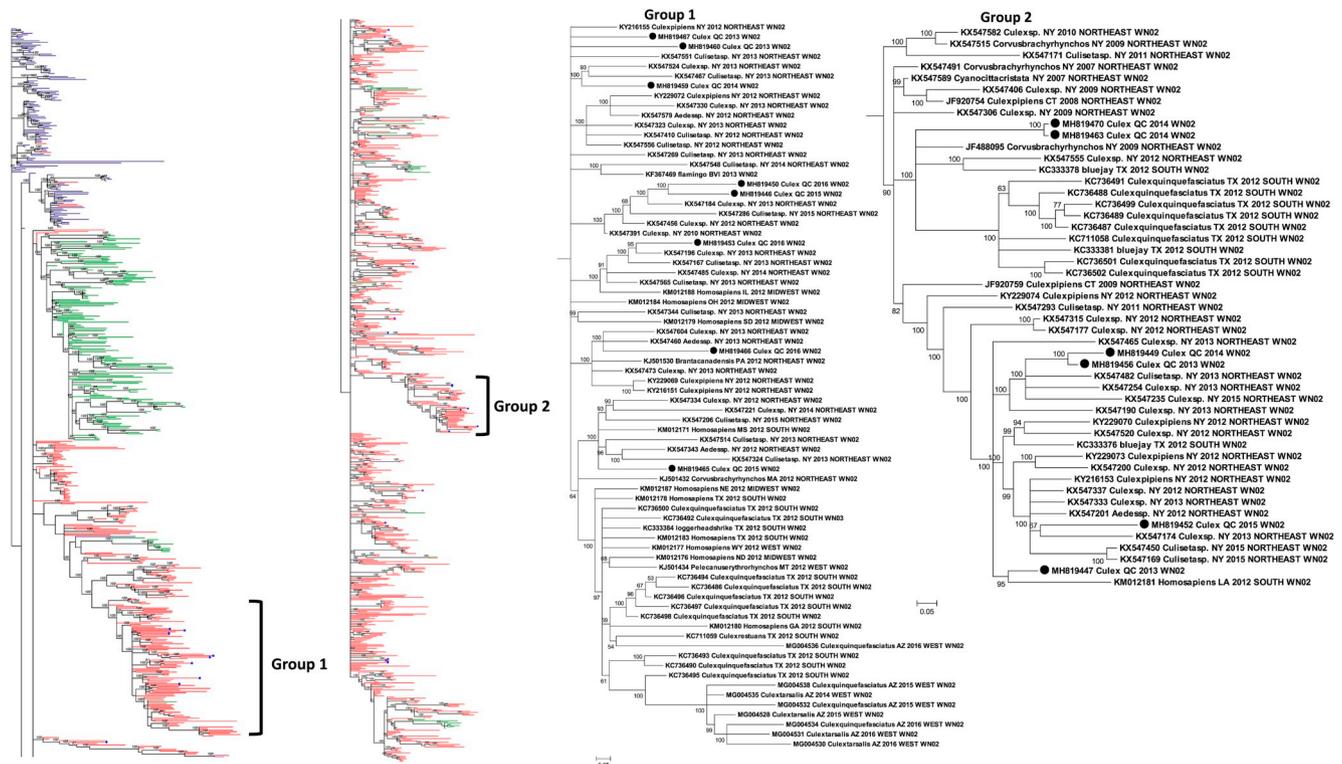

Fig. 3. Phylogenetic tree of 1090 published North American West Nile virus isolates and 26 Quebec isolates rooted with NY99 (AF196835) without tip labels. Tree branches are coloured according to genotype: NY99, blue; WN02, red and WN03, green. On the right, enlargement of two monophyletic groups clustering in the NA/WN02 genotype. Quebec isolates are identified with black dots. Sequence names include Genbank accession number, host species, 2-letter state abbreviation, year of isolation, U.S. region and genotype.

I240M (KM) motif. Eight mosquito isolates from Arizona collected in 2014–2016 reverted to arginine at position 188 of NS2A. The earliest detection of the KM motif was reported in a *Culex* isolate (KX547391) collected in Erie County, New York in 2010. The Quebec isolates are clustering closely with New York isolates and to a lesser extent to Midwestern and Southern isolates collected in 2012. Interestingly, one bird isolate collected in 2013 from the British Virgin Island in the Caribbean also presented the KM motif.

WNV sequences clustered in group 2 are geographically limited to northeastern and southern regions. All isolates of group 2 shared the NS2A (-T52I and -L95F), NS3-S334T and NS4B-S14I substitution motifs (IFTI), except for the human (KM012181) and mosquito (KC736499) isolates which have a threonine at position 52 of NS2A. Another mosquito isolate (KX547555) reverted to leucine at position 95 of NS2A. The IFTI motif was initially reported in WNV genome isolated from a crow and a blue jay (KX547491 and KX547589) recovered in New York State in 2007. The clustering pattern of Quebec "IFTI" isolates is similar to group 1 sequences with node specific association with New York strains collected between 2012 and 2013. Clustering of Quebec isolate MH819447 (2012) with isolate KM012181 collected in Louisiana in 2012 was also observed.

3.4. Conserved amino acid motifs and genotype

In order to identify amino acid motifs which could define new genotype, we determined the number of conserved non-silent mutations over time and their region of isolation across North America using our dataset (Table 4). Overall, 885 (26%) amino acid out of 3433 were substituted at least once or more in WNV isolates.

The most frequent and broadly distributed mutation detected in WNV isolates is the ENV V159A substitution (88 %). It was found either alone or in combination with other conserved amino acid motifs in all now established and emergent genetic groups derived from the dominant WN02. The NS4A A85T and NS5 K314R substitutions defining the WN03 genotype were found in 16 % and 13 % of isolates respectively. The NS4B I240M substitution is as frequent (15 %) as the NS4A A85T substitution (16%) and both mutations were detected for several years since 2002. Similarly to the ENV V159A substitution, we hypothesize that the conservation in time of the NS4A A85T or NS5 B I240M

substitutions might suggest an adaptation to hosts or environmental conditions specific to North America. Interestingly, the mutations found in the KM and IFTI isolates were also detected for several years with frequencies above or equal to 4 %.

We next determined the number and frequency of WNV genotypes according to the combination of fixed amino acid as a function of time and geographic localization (Table 5 and Fig. 4). Clearly, WN02 genotype is the most frequent variant in North America with 61 % of isolates harbouring this genetic signature. The WN03 genotype follows with 16 % of isolates. However, isolates of this genotype do not constitute a monophyletic group since 7 % of the isolates do not have the NS5 K314R substitution typical of WN03 (McMullen et al., 2011). It is noteworthy that WN03 isolates harbouring this substitution are mostly found in the Western region of the US. The KM (7%) and IFTI (4%) genetic groups are essentially located in the Northeast and Southern regions of the U.S. including Canada.

Fig. 4 shows the annual frequency of established and emerging WNV genotypes in North America. As previously reported the colonizing NY99 genotype was introduced in 1999 and gradually displaced by the WN02 genotype. To a lesser extent WN03 isolates were detected in 2003 and continued to slightly increase for several years and then decrease in numbers following 2011. The most recent genetic groups (IFTI and KM motif) were detected in 2007 and 2010 respectively. The frequency of these genotypes increased substantially with time concluding in a peak in 2013 which almost displaced circulating strains of the WN02 and WN03 genotypes.
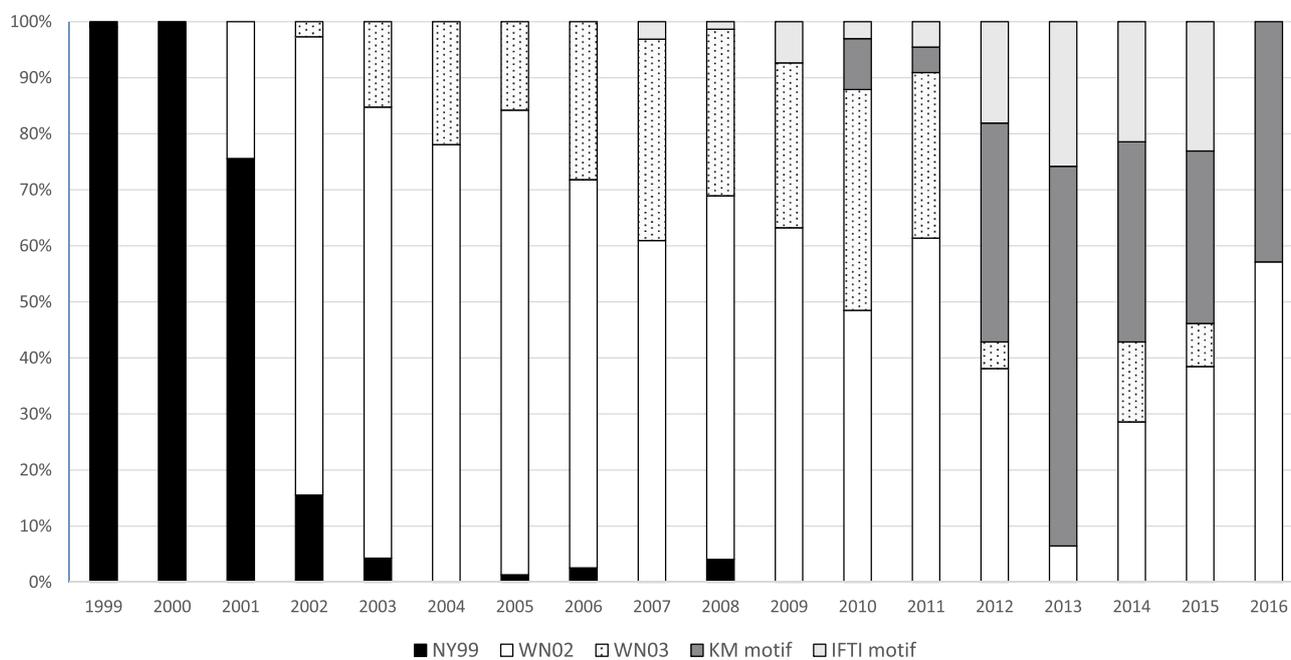
Table 5
Detection and distribution of genotypes per regions in North America.

Genotype [a]	West	Midwest	South	Northeast	Other	Total (%)
NY99	0	12	30	78	10	130 (12)
WN02	79	119	125	358	2	683 (61)
WN03 – NS5 K314R	14	15	7	37	2	75 (7)
WN03 + NS5 K314R	65	19	8	6	7	105 (9)
KM motif	2	6	15	55	1	79 (7)
IFTI motif	0	0	10	34	0	44 (4)

[a] WN02: ENV V159A; WN03: ENV V159A, NS4A A85T; KM: ENV V159A, NS4B I240M and NS2A R188K; IFTI: ENV V159A, NS2A T52I_ L95F, NS3 S334T and NS4B S14I. WN03 +/−, with or without the NS5 K314R substitution.

Table 4
Detection and distribution of conserved amino acid substitution in North American WNV isolates per year and region.

Protein	A.A. subs	Years	No Sequences[a]					
			West	Midwest	South	Northeast	Other	Total (%)
ENV	V159A	2001–2016	160	159	165	490	9	983 (88)
NS4A	A85T	2002–2015	79	34	15	43	9	180 (16)
NS4B	I240M	2002–2016	24	17	36	89	1	167 (15)
NS5	K314R	2002–2016	81	21	21	18	7	148 (13)
NS2A	R188K	2004–2016	4	8	20	81	2	115 (10)
NS2A	L95F	2003–2015	1	4	13	37	0	55 (5)
NS2A	T52I	2003–2015	1	1	12	37	0	51(5)
NS3	S334T	2007–2015	0	0	12	38	0	50 (4)
NS4B	S14I	2007–2015	0	0	12	37	0	49 (4)
NS4A	V135M	2001–2012	10	3	7	17	1	38 (3)
NS4B	G104E	2001–2015	0	2	1	32	0	35 (3)
NS2A	M90V	2001–2012	3	3	4	23	0	33 (3)
NS2A	H119Y	2002–2015	11	4	3	13	2	33 (3)
Env	V431I	2003–2011	31	0	0	0	0	31 (3)
NS4B	I245V	2000–2009	11	5	4	7	4	31 (3)
NS5	R546S	2001–2007	0	8	5	18	0	31 (3)
NS2A	A224T	2000–2007	0	1	2	27	0	30 (3)
NS4B	K24R	2003–2012	0	1	2	24	1	28 (3)
NS2A	V58I	2012–2016	10	2	15	0	0	27 (2)
C	V121A	2005–2013	12	1	3	10	0	26 (2)
NS4B	A83S	2001–2003	0	8	13	3	1	25 (2)
NS5	A860T	2003–2016	11	3	4	7	0	25 (2)

[a] Conserved amino acid substitutions found in ≥ 25 isolates are shown. Percentage of isolates calculated per total number of sequence in dataset (n = 1116).

Fig. 4. Annual proportion of established and emergent WNV genotypes.

4. Discussion

In this study, we examined the phylogeny of twenty-six WNV mosquito isolates collected over a seven-year period in the province of Quebec, Canada. All of the Quebec isolates clustered within the phylogenetic branch of the NA/WN02 genotype. The *Env* gene sequences obtained from dead bird isolates collected in 2002 in three Canadian provinces were also shown to cluster within the NA/WN02 clade (Davis et al., 2005). Isolates of the SW/WN03 genotype were not detected in this study. However, we cannot rule out the possibility that this genotype did circulate in Canada because of the limited sample size in our study. Indeed, our sampling method was biased both by the limited number of selected samples per year and by a lack of available isolates for six consecutive years (2007–2012). The SW/WN03 genotype was originally reported in the southwestern American states but was also detected in New York state repeatedly in 2005, 2007 and 2008 (McMullen et al., 2011).

The most significant finding of this study was the detection of two groups of genetically related sequences characterized by fixed amino acid motifs (KM or IFTI). Strains harboring these motifs were also reported by other phylogenetic studies (Grinev et al., 2016; Mann et al., 2013). Furthermore, the clustering of these North American sequences in two robust clades suggests the emergence of new WNV genotypes. Among the fifteen 2013–2016 Quebec isolates sequenced, only one clustered outside these two groups (VN6319).

The classification of North American WNV genotypes of lineage 1a was historically based upon the presence of fixed amino acid substitutions in isolates clustering in related genetic groups (Davis et al., 2005; McMullen et al., 2011). The finding of conserved motifs in Quebec isolates prompted us to examine the frequencies and the spatial distribution over time of these genetic signatures using a dataset of 1116 North American WNV isolates. WNV sequences harbouring the KM or IFTI motif were as frequent (7 and 4%, respectively) as WN03 isolates without the NS5 K314R substitution (7 %). WN02 genotype was the most prevalent genetic group for several years with frequencies over 60% (Fig. 4). However, emerging genotype such WN03 and the most recent KM and IFTI groups gradually displaced the WN02 genotype to levels below 50% since 2012. Considering their persistence over time and their successful establishment in different geographic regions, the KM and IFTI genetic groups should be considered as novel WNV

genotypes. Using the actual taxonomic denomination, we propose to name them according to the year of the first detection, WN10/KM and WN07/IFTI.

The first reports of sequences of the WN10 and WN07 genotypes were made in New York State. These strains dispersed rapidly to all regions of the United States and, as demonstrated here, to the north in the province of Quebec in the following years. One WN10 bird sequence was also reported as far as the British Virgin Island in the Caribbean in 2013. The dispersion of genetically related WNV sequences can be explained by the existence of several major bird flyways (Sorte et al., 2014). These migratory routes link the massive boreal forest found in Northern Canada to different U.S. Southern states and countries of the Americas. The majority of WN10 and WN07 sequences were probably initially dispersed via the Atlantic flyway linking the Northeast of Canada to the South of the U.S. Both genotypes were reported in Quebec, New York and Texas. A small number of WN10 isolates were also detected in the Western and Midwestern regions which could indicate the utilization of other terrestrial bird flyways. A recent study about the link between WNV phylogeography and terrestrial bird flyways suggested that WNV circulate via bird looped flight paths from North to South/South Central and then to the Midwest where WNV is distributed to all regions (Swetnam et al., 2018). Furthermore, the role of birds in WNV dispersal was strengthened by a field study about WNV prevalence in migrating birds along the Atlantic flyway during the fall migration (Dusek et al., 2009). The authors of the study identified several viremic bird species at specific trapping sites. For example, the Gray Catbird was reported several times in New York, New Jersey, Massachusetts, Virginia and Louisiana states (Dusek et al., 2009).

The successful establishment and dispersal of strains of the WN02 genotype were attributed phenotypically to a reduced extrinsic incubation period in *Culex* mosquitoes due to a conserved amino acid substitution in the envelope protein (ENV V159A). The functional relevance of amino acid substitutions found in the new genotypes are not known at this point. Several fixed non-synonymous mutations were observed in WNV strains isolated in Quebec. Up to four amino acid substitutions were identified in NS2A, one in NS3 and three in NS4B. These mutations were also detected as conserved motifs in many American isolates for several years. The transmembrane proteins NS2A and NS4B are localized in the endoplasmic reticulum of infected cells. They participate in virion assembly (Leung et al., 2008; Xie et al.,

2013), formation of the replication complex (Brinton, 2013) and inhibition of the α/β interferon signalling pathways (Liu et al., 2006; Muñoz-Jordán et al., 2005). NS3 is a multifunctional enzyme implicated in viral polypeptide cleavage and RNA processing (Assenberg et al., 2009). A WNV variant (B13) isolated from a persistently infected mice and harbouring the NS4A-V67I and NS4B-I240M substitutions was shown to up regulate the autophagic pathway in Vero cells (Blázquez et al., 2014). Using genetically engineered single or double NS4 mutants, it was shown that a single mutation was sufficient to increase the autophagic response in infected cells (Blázquez et al., 2014). The hijacking of the autophagy process in infected cells by viruses to promote replication is well known and was described for dengue and Zika viruses (Heaton et al., 2011; Liang et al., 2016). Several North American WNV isolates harbour the NS4B-I240M substitution, including nine Quebec isolates. Clearly, more genetic studies and functional assessment of laboratory mutants and natural variants are needed to unravel possible fine mechanistic interactions leading to phenotype variations.

In summary, our findings presented here indicate that WNV strains found in Quebec are phylogenetically related to American strains, sharing fixed amino acid motifs over time. These strains are most probably dispersed by migrating Passeriformes birds via the Atlantic flyway. Similarly to the well-characterized Env-V159A substitution, it is possible that the conserved amino acid motifs described in this study and others can provide variants with new or more competitive phenotypes. Genomic studies on larger and diversified cohorts (nationwide) will be necessary in the near future to detect the presence of WNV strains with fixed genetic signatures. Functional characterization of these strains is a prerequisite in order to better understand the relationship between genetic changes and viral protein functions. Furthermore, international surveillance cooperation would also be required to follow in real time the dispersion of emergent WNV genotypes having the potential of causing increased mortality and morbidity in human and wildlife populations.

Declaration of interest

The authors have no conflict of interest to declare.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.virol.2019.08.019.

References

Aberer, A.J., Kobert, K., Stamatakis, A., 2014. ExaBayes: massively parallel bayesian tree inference for the whole-genome era. Mol. Biol. Evol. 31, 2553–2556. https://doi.org/10.1093/molbev/msu236.

Amore, G., Bertolotti, L., Hamer, G.L., Kitron, U.D., Walker, E.D., Ruiz, M.O., Brawn, J.D., Goldberg, T.L., 2010. Multi-year evolutionary dynamics of West Nile virus in suburban Chicago, USA, 2005-2007. Philos. Trans. R. Soc. Lond. B Biol. Sci. 365, 1871–1878. https://doi.org/10.1098/rstb.2010.0054.

Andrews, S., 2010. FastQC: a Quality Control Tool for High Throughput Sequence Data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Añez, G., Grinev, A., Chancey, C., Ball, C., Akolkar, N., Land, K.J., Winkelman, V., Stramer, S.L., Kramer, L.D., Rios, M., 2013. Evolutionary dynamics of West Nile virus in the United States, 1999-2011: phylogeny, selection pressure and evolutionary time-scale analysis. PLoS Neglected Trop. Dis. 7, e2245. https://doi.org/10.1371/journal.pntd.0002245.

Armstrong, P.M., Vossbrinck, C.R., Andreadis, T.G., Anderson, J.F., Pesko, K.N., Newman, R.M., Lennon, N.J., Birren, B.W., Ebel, G.D., Henn, M.R., 2011. Molecular Evolution of West Nile Virus in a Northern Temperate Region: Connecticut, USA 1999-2008. Virology, vol. 417, 203–210. https://doi.org/10.1016/j.virol.2011.06.006.

Assenberg, R., Mastrangelo, E., Walter, T.S., Verma, A., Milani, M., Owens, R.J., Stuart, D.I., Grimes, J.M., Mancini, E.J., 2009. Crystal structure of a novel conformational state of the flavivirus NS3 protein: implications for polyprotein processing and viral replication. J. Virol. 83, 12895–12906. https://doi.org/10.1128/JVI.00942-09.

Bahuon, C., Desprès, P., Pardigon, N., Panthier, J.-J., Cordonnier, N., Lowenski, S., Richardson, J., Zientara, S., Lecollinet, S., 2012. IS-98-ST1 West Nile virus derived from an infectious cDNA clone retains neuroinvasiveness and neurovirulence properties of the original virus. PLoS One 7, e47666. https://doi.org/10.1371/journal.pone.0047666.

Beasley, D.W.C., Davis, C.T., Guzman, H., Vanlandingham, D.L., Travassos da Rosa, A.P.A., Parsons, R.E., Higgs, S., Tesh, R.B., Barrett, A.D.T., 2003. Limited evolution of West Nile virus has occurred during its southwesterly spread in the United States. Virology 309, 190–195.

Blázquez, A.-B., Martín-Acebes, M.A., Saiz, J.-C., 2014. Amino acid substitutions in the non-structural proteins 4A or 4B modulate the induction of autophagy in West Nile virus infected cells independently of the activation of the unfolded protein response. Front. Microbiol. 5, 797. https://doi.org/10.3389/fmicb.2014.00797.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinform. Oxf. Engl. 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Brault, A.C., Huang, C.Y., Langevin, S.A., Kinney, R.M., Bowen, R.A., Ramey, W.N., Panella, N.A., Holmes, E.C., Powers, A.M., Miller, B.R., 2007. A single positively selected West Nile viral mutation confers increased virogenesis in American crows. Nat. Genet. 39, 1162–1166.

Brinton, M.A., 2013. Replication cycle and molecular biology of the west nile virus. Viruses 6, 13–53. https://doi.org/10.3390/v6010013.

Centers for Disease Control and Prevention (CDC), 1999. Outbreak of West Nile-like viral encephalitis–New York, 1999. MMWR Morb. Mortal. Wkly. Rep. 48, 845–849.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772. https://doi.org/10.1038/nmeth.2109.

Davis, C.T., Ebel, G.D., Lanciotti, R.S., Brault, A.C., Guzman, H., Siirin, M., Lambert, A., Parsons, R.E., Beasley, D.W., Novak, R.J., Elizondo-Quiroga, D., Green, E.N., Young, D.S., Stark, L.M., Drebot, M.A., Artsob, H., Tesh, R.B., Kramer, L.D., Barrett, A.D., 2005. Phylogenetic analysis of North American West Nile virus isolates, 2001-2004: evidence for the emergence of a dominant genotype. Virology 342, 252–265.

Drebot, M.A., Lindsay, R., Barker, I.K., Buck, P.A., Fearon, M., Hunter, F., Sockett, P., Artsob, H., 2003. West Nile virus surveillance and diagnostics: a Canadian perspective. Can. J. Infect Dis. 14, 105–114.

Duggal, N.K., Reisen, W.K., Fang, Y., Newman, R.M., Yang, X., Ebel, G.D., Brault, A.C., 2015. Genotype-specific variation in West Nile virus dispersal in California. Virology 485, 79–85. https://doi.org/10.1016/j.virol.2015.07.004.

Dusek, R.J., McLean, R.G., Kramer, L.D., Ubico, S.R., Dupuis, A.P., Ebel, G.D., Guptill, S.C., 2009. Prevalence of West Nile virus in migratory birds during spring and fall migration. Am. J. Trop. Med. Hyg. 81, 1151–1158. https://doi.org/10.4269/ajtmh.2009.09-0106.

Ebel, G.D., Carricaburu, J., Young, D., Bernard, K.A., Kramer, L.D., 2004. Genetic and phenotypic variation of West Nile virus in New York, 2000-2003. Am. J. Trop. Med. Hyg. 71, 493–500.

Fall, G., Paola, N.D., Faye, M., Dia, M., Freire, C.C. de M., Loucoubar, C., Zanotto, P.M. de A., Faye, O., Sall, A.A., 2017. Biological and phylogenetic characteristics of west African lineages of West Nile virus. PLoS Neglected Trop. Dis. 11, e0006078. https://doi.org/10.1371/journal.pntd.0006078.

Giallonardo, F.D., Geoghegan, J.L., Docherty, D.E., McLean, R.G., Zody, M.C., Qu, J., Yang, X., Birren, B.W., Malboeuf, C.M., Newman, R.M., Ip, H.S., Holmes, E.C., 2016. Fluid spatial dynamics of West Nile virus in the United States: rapid spread in a permissive host environment. J. Virol. 90, 862–872. https://doi.org/10.1128/JVI.02305-15.

Giordano, B.V., Kaur, S., Hunter, F.F., 2017. West Nile virus in Ontario, Canada: a twelve-year analysis of human case prevalence, mosquito surveillance, and climate data. PLoS One 12, e0183568. https://doi.org/10.1371/journal.pone.0183568.

Grinev, A., Anez, G., Rios, M., 2014. Complete genome sequence of West Nile virus strains used for the formulation of CBER/FDA RNA reference reagents and lot release panels for nucleic acid testing. Genome Announc. 2 e00811-14-e00811-14. https://doi.org/10.1128/genomeA.00811-14.

Grinev, A., Chancey, C., Volkova, E., Añez, G., Heisey, D.A.R., Winkelman, V., Foster, G.A., Williamson, P., Stramer, S.L., Rios, M., 2016. Genetic variability of West Nile virus in U.S. Blood donors from the 2012 epidemic season. PLoS Neglected Trop. Dis. 10, e0004717. https://doi.org/10.1371/journal.pntd.0004717.

Grinev, A., Daniel, S., Stramer, S., Rossmann, S., Caglioti, S., Rios, M., 2008. Genetic variability of West Nile virus in US blood donors, 2002-2005. Emerg. Infect. Dis. 14, 436–444. https://doi.org/10.3201/eid1403.070463.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704. https://doi.org/10.1080/10635150390235520.

Heaton, N.S., Randall, G., Heaton, N.S., Randall, G., 2011. Dengue virus and autophagy. Viruses 3, 1332–1341. https://doi.org/10.3390/v3081332.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874. https://doi.org/10.1093/molbev/msw054.

Lanciotti, R.S., Kerst, A.J., Nasci, R.S., Godsey, M.S., Mitchell, C.J., Savage, H.M., Komar, N., Panella, N.A., Allen, B.C., Volpe, K.E., Davis, B.S., Roehrig, J.T., 2000. Rapid detection of West Nile virus from human clinical specimens, field-collected

mosquitoes, and avian samples by a TaqMan reverse transcriptase-PCR assay. J. Clin. Microbiol. 38, 4066–4071.

Lanciotti, R.S., Roehrig, J.T., Deubel, V., Smith, J., Parker, M., Steele, K., Crise, B., Volpe, K.E., Crabtree, M.B., Scherret, J.H., Hall, R.A., MacKenzie, J.S., Cropp, C.B., Panigrahy, B., Ostlund, E., Schmitt, B., Malkinson, M., Banet, C., Weissman, J., Komar, N., Savage, H.M., Stone, W., McNamara, T., Gubler, D.J., 1999. Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. Science 286, 2333–2337.

Leung, J.Y., Pijlman, G.P., Kondratieva, N., Hyde, J., Mackenzie, J.M., Khromykh, A.A., 2008. Role of nonstructural protein NS2A in flavivirus assembly. J. Virol. 82, 4731–4741. https://doi.org/10.1128/JVI.00002-08.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

Liang, Q., Luo, Z., Zeng, J., Chen, W., Foo, S.-S., Lee, S.-A., Ge, J., Wang, S., Goldman, S.A., Zlokovic, B.V., Zhao, Z., Jung, J.U., 2016. Zika virus NS4A and NS4B proteins deregulate Akt-mTOR signaling in human fetal neural stem cells to inhibit neurogenesis and induce autophagy. Cell Stem Cell 19, 663–671. https://doi.org/10.1016/j.stem.2016.07.019.

Liu, W.J., Wang, X.J., Clark, D.C., Lobigs, M., Hall, R.A., Khromykh, A.A., 2006. A single amino acid substitution in the west nile virus nonstructural protein NS2A disables its ability to inhibit alpha/beta interferon induction and attenuates virus virulence in mice. J. Virol. 80, 2396–2404. https://doi.org/10.1128/JVI.80.5.2396-2404.2006.

Mann, B.R., McMullen, A.R., Swetnam, D.M., Salvato, V., Reyna, M., Guzman, H., Bueno Jr., R., Dennett, J.A., Tesh, R.B., Barrett, A.D., 2013. Continued evolution of West Nile virus, Houston, Texas, USA, 2002-2012. Emerg. Infect. Dis. 19, 1418–1427.

McMullen, A.R., May, F.J., Li, L., Guzman, H., Bueno, R., Dennett, J.A., Tesh, R.B., Barrett, A.D.T., 2011. Evolution of new genotype of West Nile virus in north America. Emerg. Infect. Dis. 17, 785–793. https://doi.org/10.3201/eid1705.101707.

Moudy, R.M., Meola, M.A., Morin, L.L., Ebel, G.D., Kramer, L.D., 2007. A newly emergent genotype of West Nile virus is transmitted earlier and more efficiently by Culex mosquitoes. Am. J. Trop. Med. Hyg. 77, 365–370.

Muñoz-Jordán, J.L., Laurent-Rolle, M., Ashour, J., Martínez-Sobrido, L., Ashok, M., Lipkin, W.I., García-Sastre, A., 2005. Inhibition of alpha/beta interferon signaling by the NS4B protein of flaviviruses. J. Virol. 79, 8004–8013. https://doi.org/10.1128/JVI.79.13.8004-8013.2005.

Paradis, E., Schliep, K., 2019. Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinform. Oxf. Engl. 35, 526–528. https://doi.org/10.1093/bioinformatics/bty633.

Pérez-Ramírez, E., Llorente, F., Jiménez-Clavero, M.Á., 2014. Experimental infections of wild birds with West Nile virus. Viruses 6, 752–781. https://doi.org/10.3390/v6020752.

Petersen, L.R., Brault, A.C., Nasci, R.S., 2013. West Nile virus: review of the literature. J. Am. Med. Assoc. 310, 308–315. https://doi.org/10.1001/jama.2013.8042.

Ponstingl, H., Ning, Z., 2010. SMALT – A New Mapper for DNA Sequencing Reads. F1000Research 1. https://doi.org/10.7490/f1000research.327.1.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in bayesian phylogenetics using tracer 1.7. Syst. Biol. 67, 901–904. https://doi.org/10.1093/sysbio/syy032.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542. https://doi.org/10.1093/sysbio/sys029.

Sorte, F.A.L., Fink, D., Hochachka, W.M., Farnsworth, A., Rodewald, A.D., Rosenberg, K.V., Sullivan, B.L., Winkler, D.W., Wood, C., Kelling, S., 2014. The role of atmospheric conditions in the seasonal dynamics of North American migration flyways. J. Biogeogr. 41, 1685–1696. https://doi.org/10.1111/jbi.12328.

Swetnam, D., Widen, S.G., Wood, T.G., Reyna, M., Wilkerson, L., Debboun, M., Symonds, D.A., Mead, D.G., Beaty, B.J., Guzman, H., Tesh, R.B., Barrett, A.D.T., 2018. Terrestrial bird migration and West Nile virus circulation, United States. Emerg. Infect. Dis. 24, 2184–2194. https://doi.org/10.3201/eid2412.180382.

United States Census Bureau, Geography Division, 2013. Census Regions and Divisions of the United States.

Xie, X., Gayen, S., Kang, C., Yuan, Z., Shi, P.-Y., 2013. Membrane topology and function of dengue virus NS2A protein. J. Virol. 87, 4609–4622. https://doi.org/10.1128/JVI.02424-12.