



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Evolution of infectious bronchitis virus in Taiwan: Positively selected sites in the nucleocapsid protein and their effects on RNA-binding activity

Shu-Ming Kuo^{a,1}, Hsiao-Wei Kao^{a,1}, Ming-Hon Hou^{b,1}, Ching-Ho Wang^c,
Siou-Hong Lin^a, Hong-Lin Su^{a,d,*}

^a Department of Life Sciences, Agricultural Biotechnology Center, National Chung Hsing University, Taichung 40227, Taiwan

^b Institute of Genomics and Bioinformatics, Agricultural Biotechnology Center, National Chung Hsing University, Taichung 40227, Taiwan

^c School of Veterinary Medicine, National Taiwan University, Taipei 10617, Taiwan

^d Department of Physical Therapy, Graduate Institute of Rehabilitation Science, China Medical University, Taichung, Taiwan

ARTICLE INFO

Article history:

Received 27 April 2012

Received in revised form 15 October 2012

Accepted 17 October 2012

Keywords:

Infectious bronchitis virus

RNA recombination

Coronavirus

Nucleocapsid

Positive selection

ABSTRACT

RNA recombination has been shown to underlie the sporadic emergence of new variants of coronavirus, including the infectious bronchitis virus (IBV), a highly contagious avian pathogen. We have demonstrated that RNA recombination can give rise to a new viral population, supported by the finding that most isolated Taiwanese (TW) IBVs, similar to Chinese (CH) IBVs, exhibit a genetic rearrangement with the American (US) IBV at the 5' end of the nucleocapsid (*N*) gene. Here, we further show that positive selection has occurred at two sites within the putative crossover region of the N-terminal domain (NTD) of the TW IBV N protein. Based on the crystal structure of the NTD, the stereographic positions of both predicted selected sites do not fall close to the RNA-binding groove. Surprisingly, converting either of the two residues to the amino acid present in most CH IBVs resulted in significantly reduced affinity of the N protein for the synthetic RNA repeats of the viral transcriptional regulatory sequence. These results suggest that modulating the amino acid residue at either selected site may alter the conformation of the N protein and affect the viral RNA–N interaction. This study illustrates that the N protein of the current TW IBV variant has been shaped by both RNA recombination and positive selection and that the latter may promote viral survival and fitness, potentially by increasing the RNA-binding capacity of the N protein.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Among RNA viruses, the coronavirus has the largest genome, consisting of a 27.7 kb, single-stranded, positive-sense RNA. The structural genes of the coronavirus genome encode spike (*S*), membrane (*M*), envelope (*E*) and nucleocapsid (*N*) proteins. Frequent point mutations in the hypervariable regions of the *spike 1* (*S1*) gene

contribute to most of the antigenic determinants of IBV. The N protein participates in binding to the viral RNA genome and forms a ribonucleoprotein (RNP) complex. The N- and C-terminal domains (NTD and CTD) of the N protein are mainly involved in RNA binding and oligomerization, respectively (Jayaram et al., 2006). Recent evidence indicates that the NTD shows strong affinity for the transcriptional regulatory sequence (TRS), which lies within the viral leader sequence and also precedes each viral open reading frame (Spencer and Hiscox, 2006). This NTD–TRS binding is critical for the control of viral life cycle, including viral packaging, genomic replication and viral transcription (Grossoehme et al., 2009; Hurst et al., 2010).

* Corresponding author at: 250 Kuo Kuang Road, Taichung 40227, Taiwan. Tel.: +886 4 22840416; fax: +886 4 22854391.

E-mail address: suhonglin@nchu.edu.tw (H.-L. Su).

¹ These authors contributed equally.

In addition to maintaining the viral RNA in an ordered conformation for replication and transcription, the N protein is also involved in the regulation of cellular transcription, actin reorganization and apoptosis (Kopecky-Bromberg et al., 2007; Surjit et al., 2006). Recently, an antibody against the N protein was used as a diagnostic marker of coronaviral infection (Mourez et al., 2007). Moreover, immunization of the N protein alone can elicit sufficiently protective cellular immunity against lethal IBV challenges, indicating that the N protein is an immune-dominant antigen (Cowley et al., 2010).

Both genetic mutation and recombination contribute to the natural evolution of RNA viruses. Genetic variants attributed to RNA recombination are often found in coronaviruses such as mouse hepatitis virus (MHV) (Keck et al., 1987) and infectious bronchitis virus (IBV) (Cavanagh, 2007; Dolz et al., 2008). IBV is an important chicken pathogen that causes severe economic losses for global poultry industry (Cavanagh, 2007). Genetic recombination in the IBV *S1* and *N* genes has been documented in different field isolates, based on sequence comparisons and phylogenetic incongruence (Cavanagh, 2007). The high frequency of RNA recombination in coronaviruses is likely caused by their unique mechanism of RNA synthesis, which involves discontinuous transcription and polymerase jumping (Makino et al., 1986).

In addition to the sporadic generation of new variants, RNA recombination may also give rise to new viral populations, as suggested by our study of Taiwanese (TW) IBVs (Kuo et al., 2010). In general, TW IBVs are genetically closer to Chinese (CH) isolates than to American (US) strains. The only exception is the *N* gene of the TW IBV, which shows higher similarity to that of US strains. In particular, recombination in the NTD of the *N* gene has been detected in most analyzed TW IBVs. This phylogenetic incongruence suggests that RNA recombination can drive viral evolution at a population level.

However, TW IBVs that carry the recombinant NTD may present the same antigenic epitopes, at least in part, as the US Massachusetts- or Connecticut-serotype vaccine strains, which are routinely administered in chicken to control IBV in Taiwan. This sequence rearrangement may therefore pose a survival disadvantage for the TW recombinants by increasing their susceptibility to vaccine-induced immune surveillance (Chua et al., 2004). Whether this recombination event in the TW IBV enables the offspring of the recombinant founder to become dominant and evolutionally preserved is rather unclear. In this study, we show that TW IBV recombinants have undergone positive selection at two amino acids within the putative crossover region of the NTD. Replacing either of these two residues with the amino acid present in most CH IBVs significantly attenuated the binding capacity of the N protein to synthetic RNA repeats of the viral TRS. These findings suggest that the fitness of recombinant TW IBVs has been increased by positive selection conferring a replication/transcription advantage, even though these variants have been exposed to immune pressure in fowl vaccinated against US-like strains.

2. Materials and methods

2.1. IBV strains

Detail information of the recruited IBV strains is provided in supplementary Table 1 (Table S1), including the Genbank accession number and the place where the strain was isolated. We analyzed 72 IBV strains, including 26 TW strains, 15 US strains and 31 CH strains. The criteria of recruiting the IBV strains based on the decoded both full-length *S1* and *N* genes. China SD0611 strain is the only exception, whose *S1* gene is still not sequenced.

Among the IBV strains, 8 IBVs in the same clade of ME tree in Fig. 1B were chosen from each TW, CH and US groups for the BI analysis and positive selection. TW2992/02 and 3374/05 were excluded because of their phylogenetic discordance in *S1* and 3382/06 was excluded due to the discordance in *N*. The length of the tested genes is all the same. Notably, TW2575/98 is completely sequenced and is the most studied strain in Taiwan (Kuo et al., 2010).

2.2. Phylogenetic analysis

Phylogenetic trees based on ME algorithms on the full-length *S1* and full-length *N* genes were analysis by MEGA 4.0. Bootstrap values, estimated from 1000 replicates of the ME analysis, are given.

For BI analysis, we compiled the *N* genes of various IBV strains, including a Taiwanese group (TW2575/98, TW1171/92, TW2296/95, TW3374/05, TW2992/02, TW3071/03, TW97-4 and TP/64), an American group (Mass 41, H120, Cal99, CU-T2, Beaudette, Gray, Connecticut and ArkDPI) and a Chinese group (LDT3, S14, BJ, LX4, CK/CH/LTJ/95I, CK/CH/LHB/96I, SH and SD0611). Multiple sequence alignment was performed using the ClustalW program. The DNA sequences were translated into amino acid sequences using the software DAMBE 4.5.20.

Phylogenetic trees were constructed with the 24 amino acid sequences of the *N* genes by BI analysis. The best-fit models and parameters for initial settings of the phylogenetic programs were selected by ProtTest 1.2.6 for BI algorithms on the basis of the Bayesian information criterion. For the IBV N 1-124 data set, the best-fit model of substitution was the JTT model with a gamma substitution parameter of 0.76. For the IBV N 125-409 data set, the best-fit model was the JTT model with a gamma substitution parameter of 0.96. MrBayes 3.1.2 was used for BI analysis. Random starting trees were used. A total of 10 million generations of Markov chains were run. Trees were saved every 100 generations, resulting in 100,000 trees in the initial samples. The burn-in (number of initial trees that were discarded) was set to 25,000. A majority-rule consensus tree was generated from the remaining samples (75,000 trees), and the percentage of samples recovering any particular clade represented the clade's posterior probability.

The best-fit model of GTR+I+G was selected for Bayesian analysis by the program of Modeltest 3.7 using Akaike information criterion. In the analyses, 2,000,000 generations of Markov chains were run. Trees were saved every 100 generations, yielding initial samples with a total

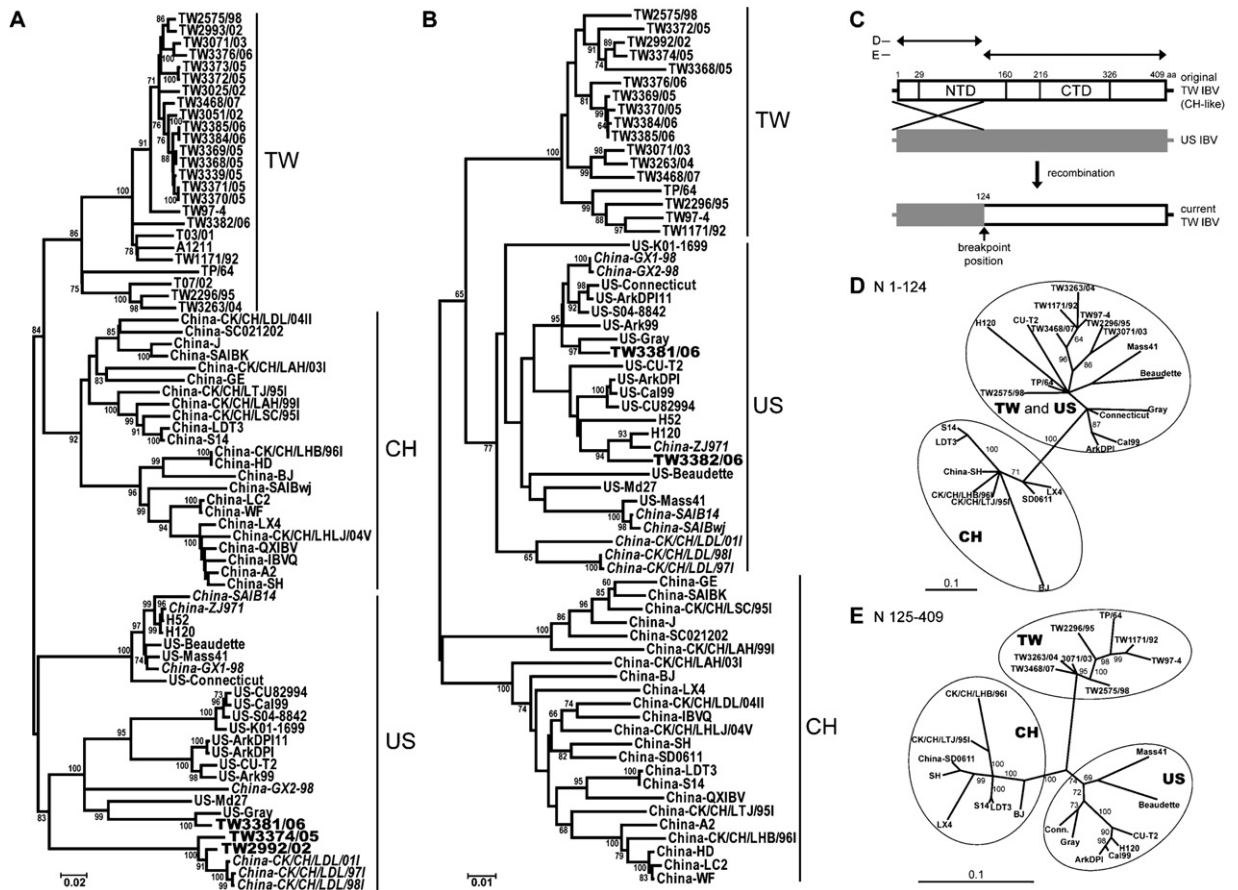


Fig. 1. Phylogenetic relationships among the IBVs. (A and B) Phylogenetic trees based on ME algorithms on the full-length *S1* (A) and full-length *N* genes (B) show the relationships among the three major IBV groups, Taiwanese (TW), Chinese (CH) and American (US). Bootstrap values, estimated from 1000 replicates of the ME analysis in MEGA 4.0, are given in percentage (%). The default setting of the MEGA program chooses the longest branch as the root. Scale bar indicates the number of nucleotide replacements per site. Exceptional TW strains are marked in bold. (C) Diagram illustrating RNA recombination in TW IBVs. NTD, N-terminal domain; CTD, C-terminal domain. (D and E) The phylogenetic analyses of the crossover (aa 1–124) (D) and non-crossover regions (aa 125–409) (E) of the *N* protein were based on BI analysis in MrBayes 3.1.2. The scale bars correspond to the number of amino acid replacements per site.

size of 20,000. The stationary phase of log-likelihood was reached within 500,000 generations. Thus, burn-in (the numbers of the initial trees were discarded) was set to 5000. Majority rule consensus trees, constructed from the 15,000 remaining trees, were used to determine the posterior probabilities of each node.

2.3. Recombination detection

Genetic recombination was evaluated by the recombination detection program (RDP) 3.0 and the genetic algorithms for recombination detection (GARD) program to predict the putative cross-over region and single breakpoint position. Detailed information was provided in our previous study (Kuo et al., 2010).

2.4. Positive selection analysis

Based on BI analysis of the 24 IBVs, variable selective pressures were evaluated at individual codon positions of the TW *N* protein. We applied paired models of variable d_N/d_S distribution among amino acid sites, including M3

(discrete) versus M0 (one ratio), M2a (positive selection) versus M1a (nearly neutral) and M8 (beta and d_N/d_S) versus M7 (beta), in the codon-based phylogenetic models (CODEML) program within phylogenetic analysis by maximum likelihood (PAML) 4. The likelihood ratio test statistic was calculated as twice the log likelihood (L) difference between the two models and labeled as 2 ($L_1 - L_0$) in Table 1. The value was compared with a chi-square test with one degree of freedom (d.f.), which is equal to the difference in the number of free parameters between the compared models.

2.5. Homology modeling of three-dimensional protein structure

A homology model of the NTD of the IBV *N* protein (TP/64) was built using the automated mode of the SWISS-MODEL program (Arnold et al., 2006). We used the crystal structure of the NTD of the IBV Beaudette strain (PDB No. 2BTL) as a template to construct a plausible conformation. The deduced models and template were superimposed, showing the overall structure with root mean square

Table 1

Amino acid sites under positive selection ($d_N/d_S > 1$) and Likelihood ratio tests of IBV N gene between models that allow positive selection (M2a, M3, M8) and neutral models (M0, M1a, M7).

Parameter	M0	M3	M1a	M2a	M7	M8
$-L$	6185.85	6095.77	6114.53661	6097.91	6119.18	6098.87
$2(L_1 - L_0)$	180.16 (comparison between M0 and M3)		33.25322 (comparison between M1a and M2a)		40.62 (comparison between M7 and M8)	
chi-square test d.f. (between 2 models)	$p < 0.001$ 4		$p < 0.001$ 2		$p < 0.001$ 2	
Parameter estimates	$\omega = 0.10026$	$\omega_0 = 0.02378$, $p_0 = 0.76904$; $\omega_1 = 0.34981$, $p_1 = 0.22444$; $\omega_2 = 1.84622$, $p_2 = 0.00652$	Not allow	$\omega_0 = 0.02948$, $p_0 = 0.80713$; $\omega_1 = 1$, $p_1 = 0.01564$; $\omega_2 = 0.4$, $p_2 = 0.17724$	$\omega_0 = 0$, $p_0 = 0.5$; $\omega_1 = 0.24135$, $p_1 = 0.5$	$\omega_0 = 0$, $p_0 = 0.47944$; $\omega_1 = 0.15589$, $p_1 = 0.47944$; $\omega_2 = 1$, $p_2 = 0.04111$
Positive selected sites	Not allow	64 Q , $p^a = 0.997$; 123 V , $p^a = 0.995$	Not allow	64 Q, $p^b = 0.846$; 123 V, $p^b = 0.808$	Not allow	64 Q , $p^b = 0.981$; 123 V , $p^b = 0.972$; 223 M, $p^b = 0.635$; 350 G, $p^b = 0.572$; 360 P, $p^b = 0.532$

a: probability in naive empirical bayes (NEB) analysis; b: probability in bayes empirical bayes (BEB) analysis; d.f: degree of freedom; ω : d_N/d_S ratio; p: proportion; L: log likelihood; M0: one ratio model; M1a: nearly neutral model; M2a: positive selection model; M3: discrete model; M7: beta model; M8: beta and ω model.

deviations of ~ 1.35 Å. Model fidelity was confirmed using the PROCHECK and Adaptive Poisson–Boltzmann Solver modules to examine the main-chain torsion angles and electrostatic distribution, respectively (Laskowski et al., 1993).

2.6. Site-directed mutagenesis

The IBV N gene was cloned into pET-28a(+) (Merck-Novagen, USA) via *Bam*H I and *Xho* I sites to synthesize a recombinant IBV N protein with a histidine tag. Mutagenesis of N gene was performed according to the instructions provided with the QuikChange site-directed mutagenesis kit (Stratagene, USA). The sense oligonucleotides used for mutagenesis are as follows: 5'-GAT AAT GAA AAT CTT AAA CCA AGC CAG CAG CAT GG-3', Thr to Pro at aa 64; 5'-CCT GAT AAT GAA AAT CTT AAA AAT AGC CAG CAG CTT GGA TAC TGG-3', Thr to Asn at aa 64; 5'-GCT GCA AAG GGT GCT GAT GTT AAA TCT AGA TCT AAT C-3', Thr to Val at aa 123. These mutants were confirmed by nucleotide sequencing.

2.7. Protein expression and purification

The pET-28 plasmid containing the full-length N protein was transferred into *E. coli* BL21(DE3) competent cells (Merck-Novagen), and the transformed bacteria were selected and cultured at 37 °C in LB broth containing 25 µg/ml kanamycin. N protein expression was induced by adding 1 mM IPTG for 4 h when the OD₆₀₀ value reached 0.4–0.6. To prevent protein degradation, the induction condition was carried out at 25 °C and supplemented with cocktail protease inhibitors (Sigma–Aldrich). The cells were collected, suspended in lysis buffer (50 mM sodium monobasic phosphate, 300 mM sodium chloride, 1 mM imidazole, 1 mg/ml lysozyme) with protease inhibitor and then sonicated. The crude cell lysates were applied to resin

affinity columns through a fast protein liquid chromatography (FPLC) system (AKTA prime plus, Amersham Pharmacia) with buffer containing 8 M urea. Resin-packed columns (Amersham Pharmacia) were equilibrated with a buffer consisting of 0.1 M NaH₂PO₄, 0.01 M Tris–HCl, 8 M urea, and 1 mM imidazole (pH 8.0) and then washed with a buffer consisting of 0.1 M NaH₂PO₄, 0.01 M Tris–HCl, 8 M urea, and 10 mM imidazole (pH 6.3). N protein fractions were eluted with a buffer containing 0.1 M NaH₂PO₄, 0.01 M Tris–HCl, 8 M urea, and 250 mM imidazole (pH 4.5) at a flow rate of 2.0 ml/min. Purified proteins were further dialyzed and refolded with refolding buffer (50 mM Tris–HCl, 50 mM NaCl, 0.4 M L-Arginine, 1 mM EDTA, 0.2 mM phenylmethanesulfonyl fluoride, 0.5 mM oxidized glutathione, 5 mM reduced glutathione, pH 8.0). Samples were slowly refolded at 4 °C for four days as the urea concentration was lowered from 4 M to 0 M using a dialysis membrane with a molecular cutoff 30–60 kDa (Millipore). Concentrations of the purified N proteins were determined by the Bradford assay (Bio-Rad, USA). In addition, the N proteins were examined by 10% SDS-PAGE and stained with Coomassie brilliant blue R250 (USB–Affymetrix).

2.8. Surface plasmon resonance (SPR) analysis

The binding capacity of N proteins for viral RNA was performed on a BIAcore 3000A SPR instrument (Amersham Pharmacia) equipped with a research-grade SensorChip SA5. This apparatus measures binding capacity by monitoring changes in the refractive index of the sensor chip surface. These changes, recorded in resonance units (RU), are assumed to be proportional to the mass of the molecules bound to the chip. Oligomers of the repeated TRS sequence, 5'-(CUUACAA)₄-3', were synthesized using an automated RNA synthesizer, labeled with biotin and

purified by gel electrophoresis. The oligomer probes were manually immobilized to the streptavidin-coated biosensor chip. The purified N proteins were dissolved in a solution consisting of 50 mM Tris–HCl, 50 mM NaCl, 1 mM EDTA, 0.5 mM oxidized glutathione, and 5 mM reduced glutathione at pH 7.3. The protein was applied to the chip surface at a flow rate of 30 μ l/min for 140 s to reach equilibrium. Before fitting to the 1:1 Langmuir model, binding data were corrected by subtracting the control to account for simple refractive index differences. Sensorgrams depicting interactions between RNA and N proteins were obtained using BIA evaluation 3 software (version 3).

2.9. Statistical analysis

One-way analysis of variance (ANOVA) was performed to determine significant differences in N protein RNA-binding capacity in three independent experiments. The Tukey's post hoc test was further used for multiple comparisons among the data collected for the wild type and variant N proteins. Statistical significance was set at $p < 0.05$.

3. Results

3.1. Recombination in the N gene of TW IBVs

We analyzed 72 IBV strains from Taiwan, China and US in this study. The available full-length *S1* and *N* gene sequences were the criteria for recruiting these strains (Genbank, accessed on April 13th, 2011). Based on a minimum evolution (ME) analysis using MEGA 4.0, the phylogenetic topology of the full-length *S1* gene revealed that the most tested TW IBVs are similar to the CH IBVs (Fig. 1A). The TW3374/05 and TW2992/02 strains were the only exceptions, showing greater similarity to the US group. In contrast, analysis of the full-length *N* genes of the IBV strains by ME (Fig. 1A and B) and Bayesian analyses (supplementary Fig. S1) demonstrated that the TW IBVs are phylogenetically closer to the US IBVs than to the CH IBVs (Fig. 1B), indicating phylogenetic incongruence between the results for the *S1* and *N* genes.

Twenty-four strains with complete sequences for the *N* gene were randomly chosen from the three TW, US and CH viral pools and subjected to phylogenetic analysis on the N protein. Putative sporadic recombinant strains, such as TW3374/05 and TW2992/02, are excluded to avoid misinterpretation of further analytic results. One candidate crossover region between the TW and US IBV strains is located between amino acid (aa) 1 and 124 of the N protein (Kuo et al., 2010), as predicted by the recombination detection program (RDP) (Martin and Rybicki, 2000) used in our previous study (Fig. 1C) (Kuo et al., 2010). The p values of the RDP and Bootscan algorithms for this recombination were 5.1×10^{-4} and 1.0×10^{-5} , respectively (Kuo et al., 2010).

Because aligning the N genes of severe acute respiratory syndrome (SARS) virus, human coronavirus-OC43 (HCoV-OC43) and MHV with IBV N sequences gives many gaps across these sequences, an ancestor strain as an outlier sequence was absent in Fig. 1A and B (data not shown). To

further confirm our hypothesis of RNA recombination, unrooted Bayesian inference (BI) algorithm was applied in this study (Fig. 1D and E). BI analysis revealed that the putative crossover region (Fig. 1D) in the TW N protein, but not the non-recombinant region (Fig. 1E), belongs to the monophyletic group of US N proteins. Moreover, to demonstrate that the trees of Fig. 1D and E are representative, we included all the strains in Fig. 1B and constructed their Bayesian trees (Fig. S2) according to the two segments of N-terminal and C-terminal amino acid sequences of IBV N protein (Fig. 1C). The results in Fig. S2 illustrated similar topologies as those in Fig. 1D and E, strongly supporting the presence of RNA recombination between the TW and US IBVs in the *N* gene sequence (Fig. 1D and E and Fig. S2).

To further clarify whether the phylogenetic incongruence is caused by different evolutionary rates in different parts of the *N* gene, a ML test that is not biased by evolutionary rate variation was applied to recheck the phylogenetic relationships (Holmes and Rambaut, 2004). No significant difference was observed between the results from the BI and those from the ML test (supplementary Fig. S3), confirming that RNA recombination probably occurred between TW and US IBVs in the 5'-terminal region of the *N* gene and that the phylogenetic incongruence was not caused by point mutation or variation in local evolutionary rates.

3.2. Positive selection of the N gene in TW IBVs

Although the N protein is evolutionally conserved among IBVs, positive selection may occur at individual amino acid (aa) residues. To investigate this possibility, full-length *N* genes of the 24 IBVs (shown in Fig. 1D and E) were subjected to codon-based phylogenetic models (CODEML) within the phylogenetic analysis by maximum likelihood (PAML) programs (Yang, 2007). Based on Bayesian analyses, neutral and positive selection models were compared using likelihood ratio tests.

The neutral models (M0, M1a and M7), selection models with a proportion of selected codons (M2a and M8) and a model for d_N/d_S heterogeneity among aa residues (M3) were applied for tests of selective pressure on N protein residues. The log likelihood values (L_n in Table 1) indicated that positive selection models (M2a, M3, M8) fitted the tested region better than neutral models (M0, M1a, M7). The nested comparisons between neutral and positive models, including M0 versus M3 (M0/M3), M1a/M2a and M7/M8, confirm the better fitness of positive models, suggesting that positive selection occurs at certain sites during the evolution of the TW N protein ($p < 0.001$ in these three comparisons, chi-square test) (Table 1). Both aa 64 and aa 123 of the TW N protein were consistently highlighted by positive selection models (M2a, M3, M8) as putative selection sites with high posterior probability values (p^a and p^b values of both sites >0.95 in M3 and M8 models, respectively) (Table 1).

The detail profile of these two positively selected sites of all tested IBVs in this study was summarized in Table 2. Using the available database of fully sequenced N proteins (72 IBV strains), we determined that Thr residues are present at the aa 64 and 123 positions in 80.8% (21/26) and

Table 2

A list of positively selected sites of N protein of IBVs.

TW group				CH group				US group			
Strain	Genbank	aa 64	aa 123	Strain	Genbank	aa 64	aa 123	Strain	Genbank	aa 64	aa 123
TW1171/92	ABG36802	T	T	BJ	AAP92682	P	V	ArkDPI	AAX39764	T	V
TW2296/95	AAT39490	T	T	CK/CH/LHB/961	ABC02826	S	V	Beaudette	AAA70242	P	T
TW2575/98	ABG36794	Q	V	CK/CH/LTJ/951	ABC02831	N	V	Cal99	AAS00087	I	V
TW3071/03	ACJ50190	T	T	LDT3	AAU14255	N	V	Connecticut	AAX39765	T	V
TW3263/04	ACJ50182	T	T	LX4	AAQ21592	N	V	CU-T2	AAC57080	T	T
TW3468/07	ACJ50166	T	T	S14	AAT70779	N	V	Gray	ACB24054	T	V
TW97-4	AAR15907	T	T	SD0611	ABY71197	N	V	H120	ACJ50206	L	T
TP/64	ACF77105	T	T	SH	ABF19054	N	V	Mass 41	AAW33793	P	T
TW2992/02	AAT39492	P	V	A2	EU526388	N	I	Ark99	AAA91855	T	V
TW2993/02	AAT39485	T	T	CK/CH/LAH/031	ABF48587	S	V	ArkDPI11	ACB59379	T	V
TW3025/02	AAT39486	T	T	CK/CH/LAH/991	ABC02823	S	V	CU82994	AAX39766	I	V
TW3051/02	AAT39487	T	T	CK/CH/LDL/011	ABC75294	T	T	H52	ACJ12841	P	T
TW3339/05	ACT31368	T	T	CK/CH/LDL/04II	ABC75302	P	V	K01-1699	AAX39767	I	V
TW3368/05	AEO79892	T	T	CK/CH/LDL/971	ABN04183	T	T	Md27	ACM45236	P	V
TW3369/05	AEO79893	T	T	CK/CH/LDL/981	ABC02824	T	T	S04-8842	AAX39768	T	V
TW3370/05	AEO79894	T	T	CK/CH/LHLJ/04V	ABC02599	N	V				
TW3371/05	ACT731372	T	T	CK/CH/LSC/951	ABC02829	T	V				
TW3372/05	AEO79898	T	V	GE	ABD34802	T	V				
TW3373/05	ACT731374	T	V	GX1-98	AAP32083	T	V				
TW3374/05	ACJ50174	P	V	GX2-98	AAP32046	T	V				
TW3376/05	AEO79895	T	V	HD	ABF19056	N	V				
TW3382/06	ACT31320	L	T	IBVQ	ABF19053	L	V				
TW3384/06	AEO79896	T	V	J	AAY85823	T	V				
TW3385/06	AEO79897	T	V	LC2	ABD34801	N	V				
T03/01	AAT39484	T	V	QXIBV	AAF06352	N	V				
T07/02	AAT39491	T	T	SAIB14	AAO46050	P	T				
				SAIBK	DQ288927	T	V				
				SAIBwj	AAM82282	P	T				
				SC021202	ACH72802	S	T				
				WF	ABF19055	N	V				
				ZJ971	AAK27161	L	T				

Strains in gray are recruited for BI and CODEML analyses.

65.4% (17/26) of the TW IBV strains, respectively (Table 2). However, Thr residues are present at the aa 64 and 123 positions in only 46.7% (7/15) and 33.3% (5/15) of the US IBV strains, respectively. In the CH IBVs, the prevalence of Thr residues at these two positions is further reduced to 29.0% (9/31) and 22.6% (7/31), respectively. This information supports the CODEML prediction that both positions have undergone positive selection in TW IBVs.

To further demonstrate the result of positively selected residues in Table 1 (24 IBV strains), 58 IBV strains in Table 2 were recruited and few strains with partial sequences were excluded for CODEML analysis. As shown in Table S2, both aa sites at 64 and 123 are predicted to be positively selected using M8 model (Probability > 0.95), strengthening the finding of the occurrence of positive selection at aa 64 and aa 123 of N protein among TW IBVs.

Notably, the selected sites are located within the putative crossover region of the N protein in the TW IBVs (Fig. 1C), linking the RNA recombination with the positive selection events. This observation also reflects the notion that particular residues of TW IBV recombinants have evolved under positive selection pressure in vaccinated flocks. In addition, the putative positively selected residues are located in the NTD, suggesting that viral progeny with strong RNA-binding affinities may have been selected during the adaptive evolution of IBV in Taiwan.

3.3. Stereography of the positively selected residues in the N protein

The NTD of the IBV N protein participates in the binding to viral RNA and the formation of the RNP complex. The crystal structure of the NTD, based on the Beaudette strain, shows a U-shaped conformation composed of a five-stranded antiparallel β sheet with positively charged amino acids clustered throughout the groove (Fan et al., 2005). Flexible loops and turns are around the inner core of the β sheet of the NTD. The positively selected sites, aa 64 and 123, are located in the external α turn and loop, respectively (arrow heads, Fig. 2A). Notably, neither site is close to the RNA-binding groove (arrow, Fig. 2A). The locations of these two residues indicate that they do not directly participate in viral RNA binding. Software predictions by RNABindR (Terribilini et al., 2007), including Ensemble, PSSMSeq and PSSMStruct algorithms, also support this observation (data not shown).

Regarding the aa 64 position, Pro is present in the Beaudette model strain, and Thr is present in the TW TP/64 strain. Interestingly, replacing Pro with Thr at residue 64 transforms the α -turn conformation into a looped structure after protein modeling (Fig. 2B), indicating the structural flexibility of this NTD residue. This result also suggests that the epitope character of this region of the TW

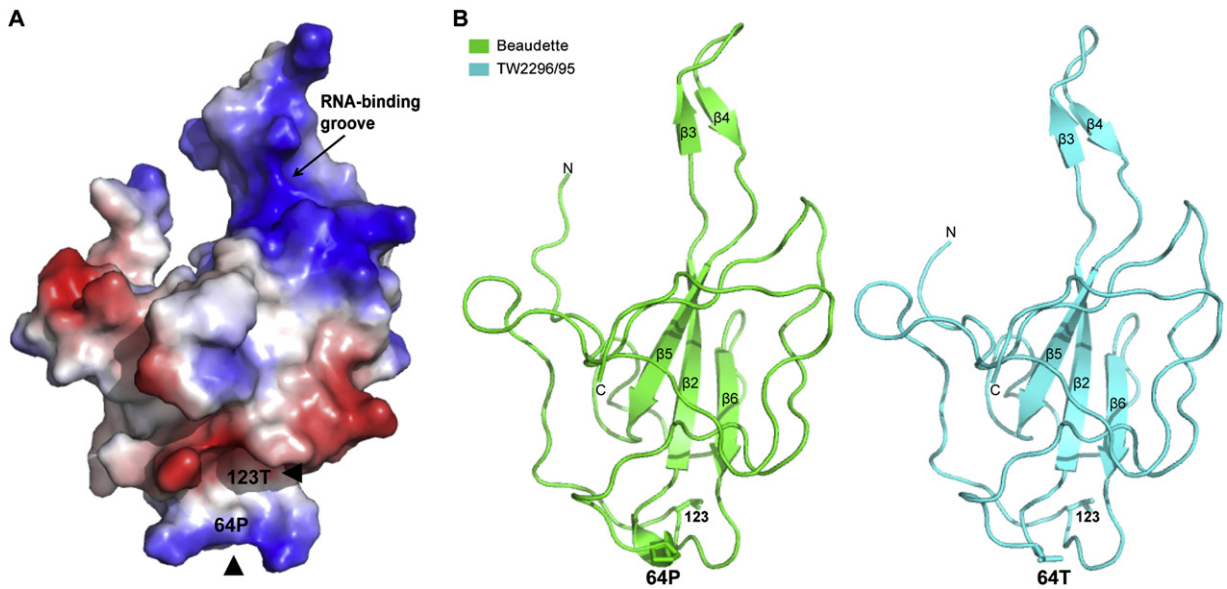


Fig. 2. Stereography of the positively selected residues of the IBV N protein. (A) Distribution of electrostatic potential on the surface of the NTD was calculated using the APBS module in PyMOL and based on the Beaudette sequence (PDB file: 2BTL). The RNA-binding groove (arrow) and the selected sites (arrowheads) are indicated. (B) Protein modeling of TW2296/95 IBV (isolated in 1995), simulated by the SWISS-MODEL program, is shown in the right side of panel B (in sky blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

IBV N protein may be vulnerable to alteration under immunological pressure. For the other selected site, aa 123, where a Thr is present in the TW TP/64 strain and a Val is present in the Beaudette strain, no obvious conformational change was noted after homology modeling. Nevertheless, a single mutational change from a hydrophilic amino acid (Thr) to a hydrophobic one (Val) could modulate the surface charge of the protein and its efficiency at forming high-order oligomers (Fan et al., 2005).

3.4. Functional test of the positively selected sites

To investigate the importance of the selected amino acid residues, the TW TP/64 strain, which was the first IBV strain to be identified in Taiwan in 1964, was chosen to represent wild type TW IBV and subjected to an analysis of RNA-binding activity. The TP/64 strain shares high genetic similarity to most TW IBV isolates (Fig. 1A and B and Table 2). Three N protein mutants, including T64P (ACA to CCA, Thr to Pro), T64N (ACA to AAT, Thr to Asp), and T123V (ACT to GTT, Thr to Val), were generated by site-directed mutagenesis and confirmed by sequencing (Fig. 3A). The production of full-length N protein of TP/64 (arrow head in Fig. 3B) was successfully induced by adding isopropyl β -D-1-thiogalactopyranoside (IPTG) in *E. coli* (Fig. 3B). The N protein of TP/64 (wild type, WT) and the mutant N proteins (T64P, T64N, T123V) were purified by Ni-column and further examined by the staining with Coomassie blue in SDS-PAGE (Fig. 3C). The detected size of full-length N protein was around 57 kDa (Fig. 3B and C), higher than the theoretical prediction (45 kDa). The upper shift of band location of expressed N protein might be caused by the intrinsic charged aa components of the N protein but not post-translational modification in *E. coli* due to the same detected molecular

weight (MW) of N protein produced by an in vitro transcription and translation system (Fig. S4). In addition, this upper shift of the N protein in SDS-PAGE were reported in a previous IBV N protein study (Yu et al., 2010) and other coronavirus N proteins (Hurst et al., 2009, 2010).

The RNA-binding capacity of the N protein was determined by surface plasmon resonance (SPR) analysis. Previous studies showed that the coronavirus N protein has a high affinity for the TRS sequence (Grossoehme et al., 2009). A repeated TRS sequence has been used as a probe in the SPR experiments to measure the interaction of the N protein with viral RNA (Huang et al., 2009; Nelson et al., 2000). The applied viral RNA probe consisted of repeated IBV TRS sequences, 5'-(CUUAA-CAA)₄-3' and was biotin-labeled. One hundred resonance units (RU) were immobilized onto a streptavidin-coated biosensor chip for detecting the binding capacity of purified IBV N proteins. Compared to the wild type N protein, all three mutants showed significantly reduced TRS-binding capacity at 0.1 and 0.5 μ M protein concentrations (Fig. 4A–C,) ($p < 0.01$, one-way ANOVA, Tukey's post hoc analysis). When the protein concentrations were elevated from 0.05 μ M to 0.5 μ M, proportional increases in their RNA-binding capacities were detected (Fig. 4C). The T64P variant, which mimicked the 64 aa position of US Beaudette strain, showed about half the RNA-binding capacity of the wild type TW N protein (Fig. 4C) ($p < 0.01$, one-way ANOVA, Tukey's post hoc analysis). Likewise, the T64N and T123V variants, which represented the most common configurations of aa 64 and aa 123 in the CH N protein (52.2% and 91.3%, respectively) (Table 2), showed only 30–40% of the RNA-binding capacity of the TW TP/64 strain (Fig. 4C) ($p < 0.01$, one-way ANOVA, Tukey's post hoc analysis). The ANOVA analyses also

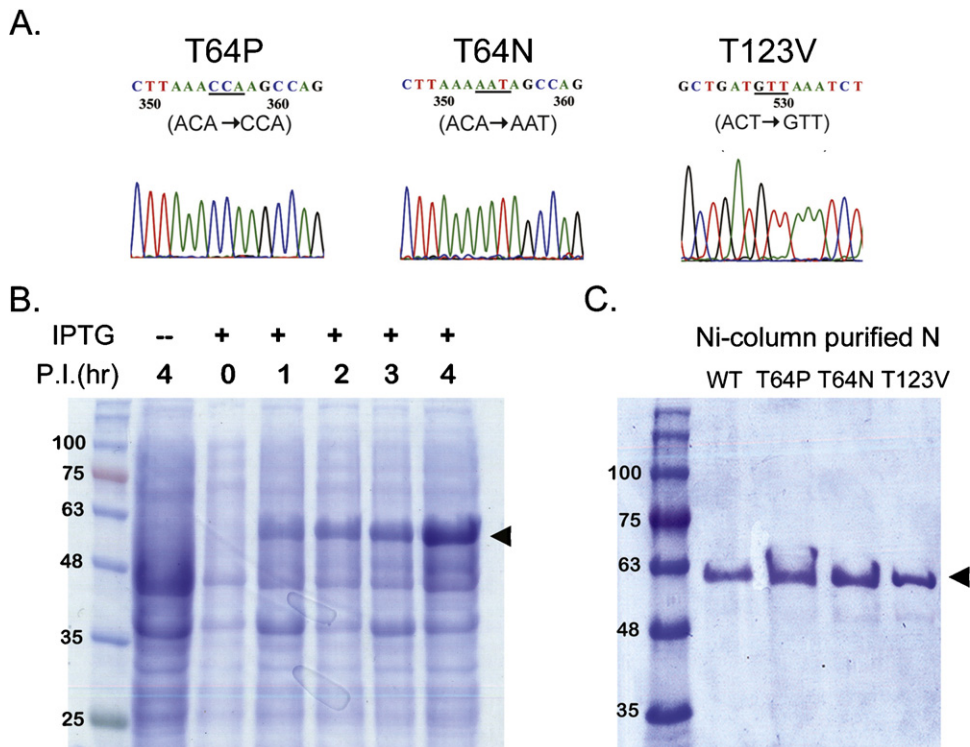


Fig. 3. Expression of wild type and mutant N proteins. (A) Nucleotide (NT) sequencing of the mutant N proteins, including the T64P (ACA → CCA), T64N (ACA → AAT), and T123V (ACT → GTT) variants of TW IBV TP/64. The altered codons are underlined. (B and C) The full-length N gene in the pET-28 plasmid was transferred into *E. coli* BL21 (DE3). The N protein expression was induced by adding 1 mM IPTG for 4 h at 25 °C when the OD₆₀₀ value reached 0.4–0.6 (arrow head) (B). The cells were collected and suspended in the lysis buffer (described in Section 2). The crude cell lysates were applied to resin affinity columns and N protein fractions were eluted with a buffer containing 8 M urea. Concentrations of the purified N proteins were determined by the Bradford assay and by 10% SDS-PAGE stained with Coomassie brilliant blue (C). WT, wild type N protein from TW TP/64 strain.

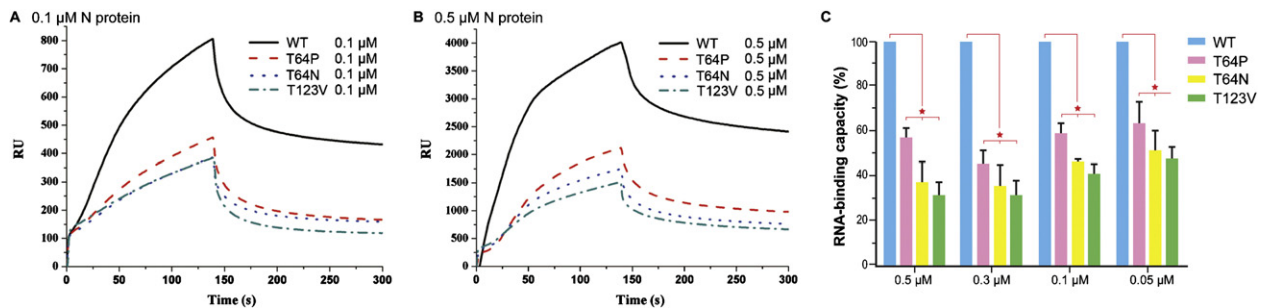


Fig. 4. RNA-binding capacities of TW N proteins. (A and B) IBV N protein interactions with synthetic TRS repeats (5'-(CUUAACAA)₄-3') were measured by SPR assays. The sensorgrams illustrate the N proteins' binding capacities for TRS repeats on immobilized chips when applied at 0.1 μM (A) or 0.5 μM (B). WT, wild type TW N proteins (black). The N mutants included the T64P (red), T64N (blue), and T123V (green) variants. (C) The relative binding activity ratios summarized from the maximal RNA-binding capacity are presented for each tested protein. *, $p < 0.01$, one-way ANOVA, Tukey's post hoc test. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

revealed that the RNA-binding activity between the pairs of mutants (T64P, T64N and T123V) has no significant difference. Taken together, these results indicate that both aa 64 and 123 are critical for 5'-(CUUAACAA)₄-3' binding, and the modulation of these two residues may affect the binding affinity of the N protein for viral genomic or subgenomic RNA.

4. Discussion

While the IBV N protein has generally been conserved and negatively selected during viral evolution, we have identified two positively selected sites at aa 64 and 123 in the N protein of IBVs. These two residues are located in the putative recombinant region of the NTD domain and are

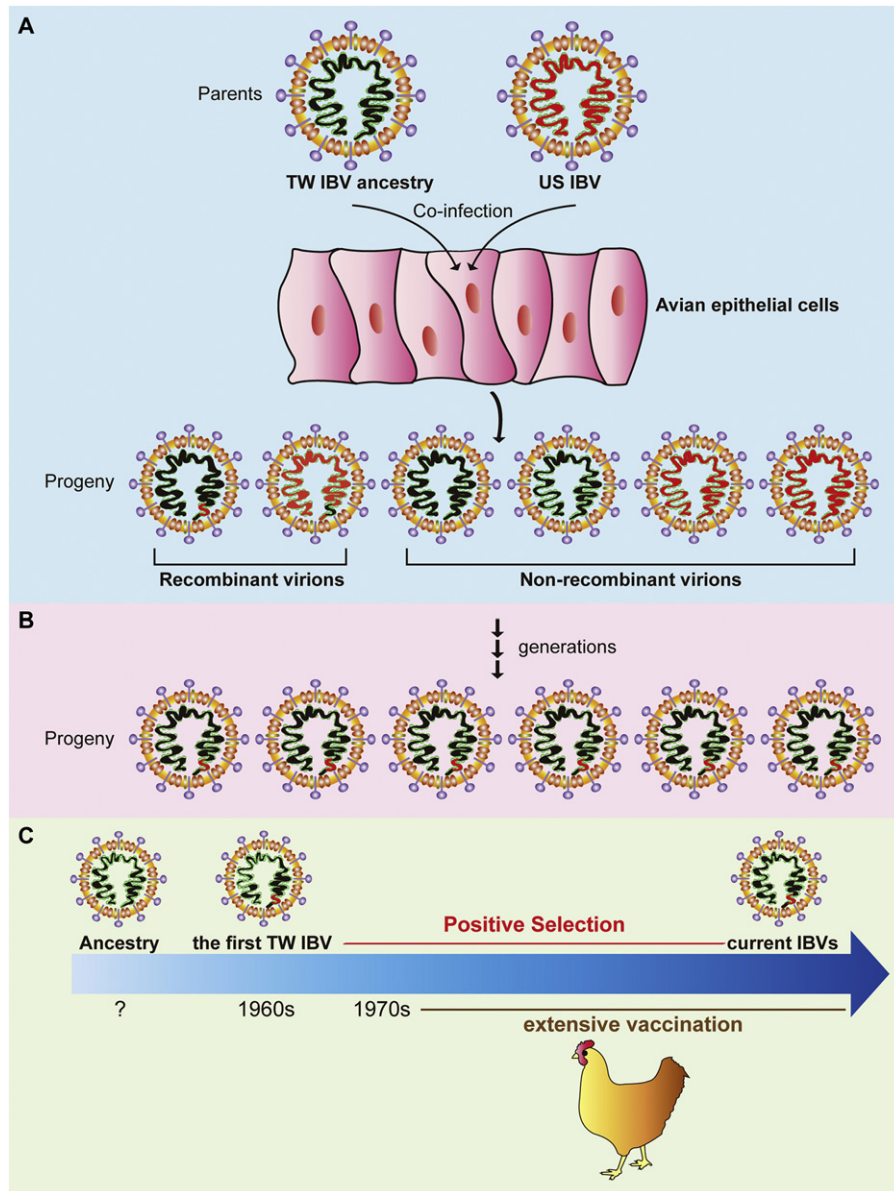


Fig. 5. Diagram illustrating the proposed evolution of IBV in Taiwan. (A) The original genetic recombination occurred in a coinfecting epithelial cell between a TW IBV ancestor (CH IBV-like) and a US IBV strain. (B) After several generations of viral amplification, the recombinant IBV has been fixed and become a dominant population in Taiwan. (C) The first TW IBV was isolated in 1964 (TP/64 strain), while the exact time of the emerging of recombinant IBV founder is unclear. We propose that in vaccinated fowl, the positive selection pressure on the N protein potentiates the viral fitness of recombinant TW IBVs.

critical for binding to TRS repeats. To our knowledge, this is the first report on viral evolution linking an RNA recombination event with positive selection. This study also provides the first functional assay to illustrate the importance of positively selected sites in the N protein for coronaviral RNA binding.

Sporadic genetic recombination in the sequences of interests will change branch lengths and the topology of phylogenetic tree (Yang et al., 2000). These two parameters are assumed to be constant across the tested sequences for the analyses of CODEML, especially when the positive selection is evaluated by branch-site method based on a

likelihood ratio test (Scheffler et al., 2006). To avoid false result of the positive selection in this study, the possible recombinant IBV strains, such as TW3374/05, TW2992/02, TW3381/06 and TW3382/06, are excluded (Fig. 1A and B) and the selected strains are chosen from the population in a same clade. In addition, the candidate of positive selected sites (Table 1) is evaluated by a codon-substitution model but not a branch-site likelihood model. Finally, the fidelity of the mathematic prediction by CODEML is further validated by the experimental function assay, demonstrating that the positively selected sites are critical for the viral RNA-binding activity.

In addition to forming part of the RNP, the coronaviral N protein participates in the formation of the replication–transcriptional complex. Specifically, the N protein's NTD binds to the TRS of the leading sequence and regulates TRS–cTRS (complementary TRS) helical unwinding (Grossoehme et al., 2009), suggesting its critical role in genomic duplication and subgenomic expression. In this study, we showed that substituting either of the positively selected residues with the amino acid present in most CH IBVs dramatically reduced the binding capacity of the N protein for synthetic TRS repeats (Fig. 4). Although neither selected site is located in the RNA-binding groove of the NTD, the modified residues may alter the secondary structure or surface charge distribution of the N protein and consequently affect RNA–NTD interactions. Here, phylogenetic evaluation of viral proteins not only helps us to reconstruct the evolutionary paths of viral species but also provides new insights into functional residues in viral proteins.

During viral propagation, variants with enhanced cellular tropism, viral transmission or replicative advantage show enhanced fitness in infected hosts (Domingo and Holland, 1997). Amino acid mutations in human immunodeficiency virus (HIV) *gag* (Banke et al., 2009) and *pol* genes (Huang et al., 2002), which are mainly involved in genomic duplication, were positively selected in patients and conferred fitness under the pressure of anti-HIV drug treatment. In addition, positively selected residues in capsid proteins have been reported in rabbit hemorrhagic disease virus (Esteves et al., 2008), hepatitis C virus (Kurbanov et al., 2010) and foot-and mouth disease virus (Haydon et al., 2001). However, the importance of these selected sites has not yet been functionally evaluated. In this study, SPR binding assays of point-mutant IBV N proteins demonstrated that the positively selected residues in the TW IBV N protein may improve binding efficiency for viral TRS repeats. Sophisticated approaches using viral replicons or recombinant infectious clone should further illuminate the detailed roles of these selected sites in coronaviral propagation.

Given that vaccine-based immunization imposes strong selective pressure on viral evolution, we did not rule out the possibility that the avian immune system has reshaped the recombinant N protein of TW IBV through positive selection. Both S and N proteins are known to be major antigenic determinants of IBV (Cavanagh, 2003). Administration of N proteins via intraperitoneal injection can elicit protective adaptive immunity against an IBV challenge (Cavanagh, 2003). We speculate that the TW IBV, sharing epitopes in the NTD of the N protein with the US-serotype IBV because of an RNA recombination event, may have become more vulnerable to attack by the adaptive immune system in fowl vaccinated against the Connecticut or other US-like IBV strains. To counteract this adverse effect, it is possible that residues in the α -turn (aa 63–67) and surrounding peptides located in an externally exposed loop (or a turn) have been positively selected to attenuate antigenic recognition by host B lymphocytes. This speculation is supported by the fact that almost all the test TW IBVs were isolated from vaccinated flocks and were under immune selection pressure. In addition, a recent study on MHV infection, which emphasized that

epitope-escape coronaviral strains can be quickly selected by genetic deletion or mutation under strong immunological pressure (Chua et al., 2004). In addition, it has been suggested that the α -turn region (aa 63–67) may form an antigenic epitope in the IBV N protein (Ignjatovic and Sapats, 2005). Future works will aim to determine whether the mutations at aa 64 and 123 of the N protein in adapted strains result in altered epitope determinants and consequently provide competitive benefits for quasispecies in hosts by allowing the selected TW strains to escape immune surveillance.

Taken together, our data support the conclusion that positive selection has occurred in specific residues of the recombinant IBV N protein. This selection may promote the viral fitness in infected hosts, at least in part, by modulating the RNA-binding capacity of the N protein's NTD. Diagram illustrating the proposed evolution of IBV in Taiwan is provided in Fig. 5. Further investigation is required to illustrate the details that how the recombinant configuration of the N gene in the original founder was maintained and consequently fixed in the current TW IBV population.

Acknowledgments

We are grateful to Dr. Drena Dobbs and Michael Terribilini at Iowa State University for supplying RNABindR RNA-binding predictions for the NTD of the TW IBV N protein. This work was supported by National Science Council in Taiwan (NSC 99-2321-B-005-012-MY3). This work was also supported in part by the Ministry of Education, Taiwan, R.O.C. under the ATU plan.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.vetmic.2012.10.020>.

References

- Arnold, K., Bordoli, L., Kopp, J., Schwede, T., 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22, 195–201.
- Banke, S., Lillemark, M.R., Gerstoft, J., Obel, N., Jorgensen, L.B., 2009. Positive selection pressure introduces secondary mutations at Gag cleavage sites in human immunodeficiency virus type 1 harboring major protease resistance mutations. *J. Virol.* 83, 8916–8924.
- Cavanagh, D., 2003. Severe acute respiratory syndrome vaccine development: experiences of vaccination against avian infectious bronchitis coronavirus. *Avian Pathol.* 32, 567–582.
- Cavanagh, D., 2007. Coronavirus avian infectious bronchitis virus. *Vet. Res.* 38, 281–297.
- Chua, M.M., MacNamara, K.C., San Mateo, L., Shen, H., Weiss, S.R., 2004. Effects of an epitope-specific CD8+ T-cell response on murine coronavirus central nervous system disease: protection from virus replication and antigen spread and selection of epitope escape mutants. *J. Virol.* 78, 1150–1159.
- Cowley, T.J., Long, S.Y., Weiss, S.R., 2010. The murine coronavirus nucleocapsid gene is a determinant of virulence. *J. Virol.* 84, 1752–1763.
- Dolz, R., Pujols, J., Ordonez, G., Porta, R., Majo, N., 2008. Molecular epidemiology and evolution of avian infectious bronchitis virus in Spain over a fourteen-year period. *Virology* 374, 50–59.
- Domingo, E., Holland, J.J., 1997. RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* 51, 151–178.

- Esteves, P.J., Abrantes, J., Carneiro, M., Muller, A., Thompson, G., van der Loo, W., 2008. Detection of positive selection in the major capsid protein VP60 of the rabbit haemorrhagic disease virus (RHDV). *Virus Res.* 137, 253–256.
- Fan, H., Ooi, A., Tan, Y.W., Wang, S., Fang, S., Liu, D.X., Lescar, J., 2005. The nucleocapsid protein of coronavirus infectious bronchitis virus: crystal structure of its N-terminal domain and multimerization properties. *Structure* 13, 1859–1868.
- Grossoehme, N.E., Li, L., Keane, S.C., Liu, P., Dann 3rd, C.E., Leibowitz, J.L., Giedroc, D.P., 2009. Coronavirus N protein N-terminal domain (NTD) specifically binds the transcriptional regulatory sequence (TRS) and melts TRS-cTRS RNA duplexes. *J. Mol. Biol.* 394, 544–557.
- Haydon, D.T., Bastos, A.D., Knowles, N.J., Samuel, A.R., 2001. Evidence for positive selection in foot-and-mouth disease virus capsid genes from field isolates. *Genetics* 157, 7–15.
- Holmes, E.C., Rambaut, A., 2004. Viral evolution and the emergence of SARS coronavirus. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 359, 1059–1065.
- Huang, C.Y., Hsu, Y.L., Chiang, W.L., Hou, M.H., 2009. Elucidation of the stability and functional regions of the human coronavirus OC43 nucleocapsid protein. *Protein Sci.* 18, 2209–2218.
- Huang, K.J., Alter, G.M., Wooley, D.P., 2002. The reverse transcriptase sequence of human immunodeficiency virus type 1 is under positive evolutionary selection within the central nervous system. *J. Neurovirol.* 8, 281–294.
- Hurst, K.R., Koetzner, C.A., Masters, P.S., 2009. Identification of in vivo-interacting domains of the murine coronavirus nucleocapsid protein. *J. Virol.* 83, 7221–7234.
- Hurst, K.R., Ye, R., Goebel, S.J., Jayaraman, P., Masters, P.S., 2010. An interaction between the nucleocapsid protein and a component of the replicase–transcriptase complex is crucial for the infectivity of coronavirus genomic RNA. *J. Virol.* 84, 10276–10288.
- Ignjatovic, J., Sapats, S., 2005. Identification of previously unknown antigenic epitopes on the S and N proteins of avian infectious bronchitis virus. *Arch. Virol.* 150, 1813–1831.
- Jayaram, H., Fan, H., Bowman, B.R., Ooi, A., Jayaram, J., Collisson, E.W., Lescar, J., Prasad, B.V., 2006. X-ray structures of the N- and C-terminal domains of a coronavirus nucleocapsid protein: implications for nucleocapsid formation. *J. Virol.* 80, 6612–6620.
- Keck, J.G., Stohlman, S.A., Soe, L.H., Makino, S., Lai, M.M., 1987. Multiple recombination sites at the 5'-end of murine coronavirus RNA. *Virology* 156, 331–341.
- Kopecny-Bromberg, S.A., Martinez-Sobrido, L., Frieman, M., Baric, R.A., Palese, P., 2007. Severe acute respiratory syndrome coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins function as interferon antagonists. *J. Virol.* 81, 548–557.
- Kuo, S.M., Wang, C.H., Hou, M.H., Huang, Y.P., Kao, H.W., Su, H.L., 2010. Evolution of infectious bronchitis virus in Taiwan: characterisation of RNA recombination in the nucleocapsid gene. *Vet. Microbiol.* 144, 293–302.
- Kurbanov, F., Tanaka, Y., Matsuura, K., Sugauchi, F., Elkady, A., Khan, A., Hasegawa, I., Ohno, T., Tokuda, H., Mizokami, M., 2010. Positive selection of core 70Q variant genotype 1b hepatitis C virus strains induced by pegylated interferon and ribavirin. *J. Infect. Dis.* 201, 1663–1671.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291.
- Makino, S., Keck, J.G., Stohlman, S.A., Lai, M.M., 1986. High-frequency RNA recombination of murine coronaviruses. *J. Virol.* 57, 729–737.
- Martin, D., Rybicki, E., 2000. RDP: detection of recombination amongst aligned sequences. *Bioinformatics* 16, 562–563.
- Mourez, T., Vabret, A., Han, Y., Dina, J., Legrand, L., Corbet, S., Freymuth, F., 2007. Baculovirus expression of HCoV-OC43 nucleocapsid protein and development of a Western blot assay for detection of human antibodies against HCoV-OC43. *J. Virol. Methods* 139, 175–180.
- Nelson, G.W., Stohlman, S.A., Tahara, S.M., 2000. High affinity interaction between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA. *J. Gen. Virol.* 81, 181–188.
- Scheffler, K., Martin, D.P., Seoighe, C., 2006. Robust inference of positive selection from recombining coding sequences. *Bioinformatics* 22, 2493–2499.
- Spencer, K.A., Hiscox, J.A., 2006. Characterisation of the RNA binding properties of the coronavirus infectious bronchitis virus nucleocapsid protein amino-terminal region. *FEBS Lett.* 580, 5993–5998.
- Surjit, M., Liu, B., Chow, V.T., Lal, S.K., 2006. The nucleocapsid protein of severe acute respiratory syndrome-coronavirus inhibits the activity of cyclin–cyclin-dependent kinase complex and blocks S phase progression in mammalian cells. *J. Biol. Chem.* 281, 10669–10681.
- Terribilini, M., Sander, J.D., Lee, J.H., Zaback, P., Jernigan, R.L., Honavar, V., Dobbs, D., 2007. RNABindR: a server for analyzing and predicting RNA-binding sites in proteins. *Nucleic Acids Res.* 35, W578–W584.
- Yang, Z., 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591.
- Yang, Z., Nielsen, R., Goldman, N., Pedersen, A.M., 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155, 431–449.
- Yu, D., Han, Z., Xu, J., Shao, Y., Li, H., Kong, X., Liu, S., 2010. A novel B-cell epitope of avian infectious bronchitis virus N protein. *Viral Immunol.* 23, 189–199.