Genetic and evolutionary characterization of RABVs from China using the phosphoprotein gene

Lihua Wang†, Hui Wu†, Xiaoyan Tao, Hao Li, Simon Rayner, Guodong Liang and Qing Tang*

Abstract

Background: While the function of the phosphoprotein (P) gene of the rabies virus (RABV) has been well studied in laboratory adapted RABVs, the genetic diversity and evolution characteristics of the P gene of street RABVs remain unclear. The objective of the present study was to investigate the mutation and evolution of P genes in Chinese street RABVs.

Results: The P gene of 77 RABVs from brain samples of dogs and wild animals collected in eight Chinese provinces through 2003 to 2008 were sequenced. The open reading frame (ORF) of the P genes was 894 nucleotides (nt) in length, with 85-99% (80-89%) amino acid (nucleotide) identity compared with the laboratory RABVs and vaccine strains. Phylogenetic analysis based on the P gene revealed that Chinese RABVs strains could be divided into two distinct clades, and several RABV variants were found to co-circulating in the same province. Two conserved (CD1, 2) and two variable (VD1, 2) domains were identified by comparing the deduced primary sequences of the encoded P proteins. Two sequence motifs, one believed to confer binding to the cytoplasmic dynein light chain LC8 and a lysine-rich sequence were conserved throughout the Chinese RABVs. In contrast, the isolates exhibited lower conservation of one phosphate acceptor and one internal translation initiation site identified in the P protein of the rabies challenge virus standard (CVS) strain. Bayesian coalescent analysis showed that the P gene in Chinese RABVs have a substitution rate (3.305x10^-4 substitutions per site per year) and evolution history (592 years ago) similar to values for the glycoprotein (G) and nucleoprotein (N) reported previously.

Conclusion: Several substitutions were found in the P gene of Chinese RABVs strains compared to the laboratory adapted and vaccine strains, whether these variations could affect the biological characteristics of Chinese RABVs need to be further investigated. The substitution rate and evolution history of P gene is similar to G and N gene, combine the topology of phylogenetic tree based on the P gene is similar to the G and N gene trees, indicate that the P, G and N genes are equally valid for examining the phylogenetics of RABVs.

Keywords: Rabies virus, Phosphoprotein gene, Genetic diversity, Molecular evolution

Introduction

Rabies is a lethal neurological disease caused by infection with members of the genus lyssavirus. Eleven distinct lyssavirus species are currently recognized worldwide [1]. In China, only the classical rabies virus (RABV) is known to circulate in dogs, which serve as the principal reservoir and transmitter of rabies to humans and domestic animals [2,3]. RABV has a non-segmented negative sense RNA genome comprised of five genes in the order 3′-N-P-M-G-L-5′ [4]. The relatively divergent P gene [5-7] encodes a multifunctional phosphoprotein (P protein) [8] and has been extensively investigated using laboratory adapted RABV strains. Five serine residues of the challenge virus standard (CVS) strain have been identified as phosphate acceptor sites [9]. Also, P is a critical component of the viral polymerase responsible for transcription and replication through its binding to the N and L proteins [10-12]. Two independent N binding sites, one located within amino acids (aa) 66–176 at the N-terminal half of the protein and the other located to amino acids 268–297 within 50 residues of the C-terminus, have been found in the P protein [10,11]. Via N-P complexes, the nonspecific
aggregation of N can be prevented and can keep N in a suitable form for specific encapsidation [13]. The short lysine-rich motif FSKKYKF (aa 214–220) is an important component of the C-terminal N protein binding domain of P [14]. P is associated with the genome expression process by acting as an intermediary for the attachment of the L polymerase core to the N-RNA template [15]. In addition, the first 19 N-terminal residues of P confer L protein binding [10]. P also specifically interacts with many host cell components. It has been reported that the sequence (K/R)XTQT represents a conserved cytoplasmic dynein light chain (LC8) binding motif, an element of the microtubule-associated motors involved in minus-end directed axonal transport, through which it may play some role in viral retrograde transport [16-18]. P interferes with the host’s innate immune system through inhibition of the activities of interferon regulatory factor 3 (IRF3) [19] and signal transducer and activator of transcription 1 (STAT1) [20,21], thereby abrogating the cellular type I interferon pathway. P also binds to the promyelocytic leukemia (PML) protein, which has many possible functions in nuclear trafficking, viral defense mechanisms and apoptosis [22], suggesting that P acts as an antagonist towards antiviral PML function [23].

Since all functional studies on the RABV P protein have been performed using a limited number of laboratory strains, the relevance of the results to field isolates is unclear. In this study we sequenced the P gene of Chinese RABV street strains collected in most rabies endemic areas of China and investigated the genetic diversity, sequence characteristics and estimated the overall substitution rate of the P gene. In addition, the phylogeny and evolution history of Chinese RABVs based on P gene were examined.

**Results**

**Length and identity of P gene in Chinese RABV street strains**

77 RABV positive brain specimens were detected by direct fluorescent antibody (DFA) and subjected to RT-PCR for determination of the P gene of RABV street strains. These specimens were from field captured dogs and ferret badgers in eight provinces which had high (Guangxi, Guizhou and Hunan provinces), middle (Jiangsu and Shandong provinces) and low (Anhui, Shanghai, and Zhejiang provinces) incidences of rabies (Figure 1). The open reading frame (ORF) of the P gene, corresponding to nt 1514–2407 of the PV strain (M13215), was determined for all 77 RABV isolates. The ORF of the P gene of all Chinese RABVs were 894 nt in length and sequences were submitted to GenBank (HM582519–HM582595). The species of origin, the year of isolation, and geographical location of these sequences are summarized in Table 1.

![Figure 1](http://www.virologyj.com/content/10/1/14)

**Figure 1 Locations of specimen collection in this study.** AH, GX, GZ, HuN, JS, SD, SH, ZJ, indicate Anhui, Guangxi, Guizhou, Hunan, Jiangsu, Shandong, Shanghai, Zhejiang provinces of China, where the specimen were collected in this study during 2003 to 2008. I and II indicate the presence of isolates corresponding to Clade I and II as classified by the phylogenetic tree.
| Genus/isolates | Host | Origin | Year | GenBank acc. no. | Genus/isolates | Host | Origin | Year | GenBank acc. no. |
|---------------|------|--------|------|-----------------|---------------|------|--------|------|-----------------|
| AH8           | Dog  | AH     | 2005 | HM582562        | SH9           | Dog  | SH     | 2004 | HM582564        |
| AH12          | Dog  | AH     | 2005 | HM582567        | SH15          | Dog  | SH     | 2004 | HM582555        |
| GX0801        | Dog  | GX     | 2008 | HM582588        | SH16          | Dog  | SH     | 2004 | HM582584        |
| GX0802        | Dog  | GX     | 2008 | HM582589        | SH17          | Dog  | SH     | 2004 | HM582554        |
| GX0803        | Dog  | GX     | 2008 | HM582590        | SH19          | Dog  | SH     | 2004 | HM582550        |
| GX0804        | Dog  | GX     | 2008 | HM582591        | SH20          | Dog  | SH     | 2004 | HM582532        |
| GX0805        | Dog  | GX     | 2008 | HM582592        | SH24          | Dog  | SH     | 2003 | HM582536        |
| GX0806        | Dog  | GX     | 2008 | HM582593        | SH25          | Dog  | SH     | 2003 | HM582553        |
| GX0807        | Dog  | GX     | 2008 | HM582594        | SH27          | Dog  | SH     | 2003 | HM582529        |
| GZ12          | Dog  | GZ     | 2005 | HM582545        | SH26          | Dog  | SH     | 2003 | HM582583        |
| GZ13          | Dog  | GZ     | 2005 | HM582537        | D01           | Dog  | ZJ     | 2008 | HM582573        |
| GZ14          | Dog  | GZ     | 2005 | HM582551        | D02           | Dog  | ZJ     | 2008 | HM582570        |
| GZ15          | Dog  | GZ     | 2005 | HM582552        | D04           | Dog  | ZJ     | 2008 | FJ032321        |
| GZ16          | Dog  | GZ     | 2005 | HM582534        | D08           | Dog  | ZJ     | 2008 | FJ032322        |
| GZ17          | Dog  | GZ     | 2005 | HM582530        | F02           | CFB  | JX     | 2008 | GJ647092        |
| GZ21          | Dog  | GZ     | 2005 | HM582572        | 8743THA       | Human | Thailand | 1983 | EU293121        |
| HN27          | Dog  | HuN    | 2005 | HM582582        | INRV          | Human | India   | 2005 | Ay956319        |
| HN29          | Dog  | HuN    | 2005 | HM582520        | NNV-RAB-H     | Human | India   | 2006 | EF47215         |
| HN30          | Dog  | HuN    | 2005 | HM582542        | CVS           | Challenge virus standard |   | X55727 |
| JS29          | Dog  | JS     | 2006 | HM582563        | aG            | Vaccine strain | China   | DQ646875    |
| JS34          | Dog  | JS     | 2006 | HM582565        | CTN           | Vaccine strain | China   | FJ959397    |
nucleotide and amino acid sequence identities of the P gene were 80.2-100% and 85.2-100% respectively. When compared with the vaccine strains, the P gene of the 77 Chinese RABVs had 85.0-99.2% (80.0-89.5%) amino acid (nucleotide) identity, respectively.

Variation of functionally significant sequence motifs and residues

Based on the identity analysis, an amino acid alignment of the 77 Chinese RABVs isolates and representative sequences of laboratory and vaccine strains was generated and investigated for mutations (Figure 2). In total, seventy two amino acid substitutions throughout the P protein were observed in the Chinese RABVs isolates relative to the PV vaccine strain (M13215). Based on the location of the mutations, the protein had both highly conserved and highly variable regions that have been previously shown to be associated with viral function. Specifically, there were two conserved domains at residues 1–50 (CD1) and 184–279 (CD2) and two variable domains at residues 51–80 (VD1) and 126–178 (VD2) (Figure 2). The first 19 aa residues at the N-terminal, shown to be associated with L binding [10], are completely conserved. The short lysine-rich segment FSKKYKF (209-216aa) thought to be an important component of the C-terminal N protein binding domain [14], is also highly conserved in all Chinese isolates. Within region VD2, the cytoplasmic dynein LC8 binding motif (K/R) XTQT [18] is conserved with Chinese RABVs, and all the strains contain the motif KSTQT (located between 144 and 148 aa). Interestingly, the STAT-1 binding sites, located in the last 30 aa residues of the C-terminal [20] showed limited conservation in Chinese isolates. The internal translation initiation sites 20, 53, 69, and 83 in the P protein of the rabies challenge virus standard (CVS) strain [24] are at the same position in the Chinese RABVs isolates. Three of them (Met20, Met53, and Met83) are completely conserved in Chinese RABVs. For the remainder, the mutation Met69 to Val69 occurred in isolate GZ8 and mutation Met69 to Ala69 occurred in isolates HN29, GX0802, GZ7, GX16. Four (Ser64, Ser65, Ser210, Ser271) of five serine residues reported to function as phosphate acceptors in the P protein of the rabies challenge virus standard (CVS) strain [9] were absolutely conserved. For the fifth residues mutation Ser63 to Phe63 or Ser63 to Leu63 was observed in all the Chinese isolates with the exception of isolate SH19.

Phylogenetic analysis of RABVs in China

A phylogenetic analysis of 113 (77 collected in this study, with an additional 36 samples downloaded from GenBank) RABV P gene sequences was performed. The Neighbor-joining tree is shown in Figure 3 with bootstrap values shown for the main groupings. The sequences of Chinese isolates were divided into two major clades, named clade I and II (Figure 3). Most of the 77 isolates collected in this study are placed in Clade I (bootstrap value = 98). These isolates are mainly from Anhui, Guangxi, Guizhou, Hunan, Jiangsu, Jiangxi, Inner Mongolia, Shandong, Shanghai, Zhejiang provinces of China, respectively; CFB: Chinese Ferret Badgers.
Figure 2 (See legend on next page.)
Substitution rates and evolution history analysis of P gene

By using a Bayesian Markov chain Monte Carlo method, the evolutionary history, including evolutionary rates of populations (nucleotide substitutions per site per year) and TMRCA (the most recent common ancestor) were analyzed based on 58 P gene sequences (Only sequences with an homology less than 98% and with full background information in terms of location and isolation time were used in the calculation). The estimated mean rate of nucleotide substitution for the P gene of Chinese RABVs was 3.305 x 10^-4 substitutions per site per year (95% HPD values, 1.127 - 6.209 x 10^-4 substitutions per site per year).

Bayesian coalescent analysis estimated the most recent common ancestor (TMRCA) to have originated 592 years ago (95% HPD, 142 – 2621 years) (Figure 4).

Discussion and conclusion

The N gene (the most conserved and abundant mRNA in infected cells) and G gene (plays a crucial role in viral neurotropism and pathogenicity) have been widely targeted for genetic, molecular epidemiology and evolutionary analysis of RABVs [4, 25–28]. In contrast, for the P gene, only a few laboratory [29, 30] and wild-type RABV strains [31], an ABLV isolate [32] and Mokola virus [33] have been genetically characterized. In this study we

Figure 2 Alignment of the P amino acid sequences of street strains collected in this study, vaccine strains and standard challenge virus strain CVS11. Dots indicate amino acids that are in agreement with the reference sequence (PV vaccine strain (M13215)) on the first line. Box a and e conserved domains 1 and 2; box b and c variable domains 1 and 2; box d: Dynein light chain (LC8)-binding motif; solid underline shows L protein binding region(1–19 aa) and the lysine-rich motif (209–216 aa), respectively; dashed underline shows N protein binding site; triangles indicate the positions of methionine residues and confirmed translation initiation in the CVS strain; arrows indicate the positions of serine residues identified as phosphoacceptors in the P protein of the CVS strain.

Figure 3 Neighbor-joining phylogenetic tree (P-distance) for the P gene of RABVs collected in this study, vaccine strains and representative strains of lyssavirus. Numbers indicate the bootstrap value from 1000 replicates. Clade I and clade II are indicated.
attempted to characterize the genetic and evolutionary properties of the \textit{P} gene of Chinese street RABVs. 77 \textit{P} genes from brain samples of dogs and wild animals in eight provinces through 2003 to 2008 were sequenced and subjected to molecular and phylogenetic analysis.

Several substitutions were found in the Chinese RABVs strains compared to the laboratory adapted and vaccine strains. The nucleotide ($\geq 80.2\%$) and amino acid sequence identities ($\geq 85.2\%$) of the \textit{P} gene were lower than the corresponding values for the \textit{N} ($\geq 87.6\%$ and 95.4\%) and \textit{G} gene ($\geq 87\%$ and 93.8\%) [26,28]. Consistent with the wild type RABVs strains isolated in North America [31], two conserved (CD1, 2) and two variable (VD1, 2) domains were identified in Chinese RABVs. The observed substitutions are mainly located in the middle of \textit{P}, while the \textit{N} and \textit{C} terminal are relatively well conserved. As reported previously, the need to retain overall negative charge rather than primary sequence would explain the VD1 region’s high level of diversity [6]. The poorly conserved VD2 might indicate a function as a spacer/hinge segment analogous to the hinge region of the \textit{P} gene in Vesicular stomatitis virus (VSV) located between two functionally important domains [34]. Two sequence motifs, one believed to confer binding to the cytoplasmic dynein light chain LC8, and a lysine-rich sequence probably contributing to \textit{N} protein binding [14], were conserved throughout Chinese RABVs samples, while the STAT-1 binding sites [20], internal translation initiation sites and phosphate acceptor sites showed different degrees of variation. Whether these variations could affect the biological characteristics of Chinese RABVs need to be further investigated.

There have been several previous estimates of RABVs substitution rates for the \textit{G} gene ($1.2-6.5 \times 10^{-4}$ substitutions per site per year) and the \textit{N} gene ($1.1-5.6 \times 10^{-4}$ substitutions per site per year) based on dog, fox and mongoose RABVs samples collected worldwide [25,27,35-38]. In this study, Bayesian coalescent analysis showed that mean substitution rate of the \textit{P} gene for the Chinese RABVs isolates is $3.305 \times 10^{-4}$ substitutions per site per year, which indicates that the genome RNA of RABVs circulating worldwide is stable. The TMRCA of cosmopolitan canine RABV variants has previously been estimated to be between 284 and 504 years ago [39]. The mean divergence time estimated based on the \textit{G} gene is 583 years ago for RABVs circulating globally [25,35], and 596 years ago for RABVs for current Chinese RABVs [27]. Using a similar
analysis, we estimated the average TMRCA of RABVs circulating in China based on the P gene to be 592 years ago, which was in accordance with previous reports for RABVs.

Previous phylogenetic studies based on the G and N genes [26,28,39,40] showed that RABVs in China can be classified into distinct clades or groups. The phylogenetic analysis in this report based on the P gene revealed that Chinese RABVs could be divided into two distinct clades, and that isolates from more than one clade RABV variants are currently co-circulating in the same Chinese provinces. Also, RABVs in Clades I are grouped with RABVs from Thailand, and RABVs in clade II are grouped with RABVs from India. The topology of the phylogenetic tree based on the P gene is similar to the G and N gene trees [26,28,39,40]. This indicates that the P, G and N genes are equally valid for examining the phylogenetics of RABVs and is consistent with observations that the N, P, M, G and L genes of RABVs interact and evolve in a co-operative manner to effect virus infection and evolution [41,42].

Methods

Viral specimens sampling

Brain specimens were collected as part of a national surveillance program from dogs used as meat in restaurants and from suspected rabid Ferret badgers from eight provinces (Anhui, Guangxi, Guizhou, Hunan, Jiangsu, Shandong, Shanghai and Zhejiang) in China from 2003 to 2008 (Figure 1).

Detection and sequencing of RABV

All specimens were examined by using a direct immune fluorescence assay (DFA) [26] with a fluorescent-labeled monoclonal antibody against the RABV N protein (Rabies DFA Reagent; Chemicon Europe Ltd., Chandlers Ford, UK). For all identified RABV specimens, RNA was extracted from tissue of rabies-infected brains (0.1 g) with TRizol Reagent (Invitrogen, Carlsbad, CA, USA) and used as template for cDNA synthesis with Ready-To-Go YouPrime First-Strand Beads (Amersham Pharmacia Biotechnology, Chalfont St. Giles, UK) and a rabies P gene specific primer: Pfor 5’-GAACCATCCCAAYATG AG -3’ (corresponding to bases 1500–1519 of the positive sense genome sequence of the PV strain). The ORF sequence of the P gene, encoding regions corresponding to bases 1514 to 2407 of the total genetic sequence of the PV strain, was amplified with primers Pfor and Prev 5’- GATACGGTTCG CAGAAARTCTAC -3’ (corresponding to bases 2496 to 2517 of the positive sense genome sequence of the PV strain). PCR products were purified by using the QIAquick PCR Purification Kit (QIAGEN Ltd., Crawley, UK) and sequenced with an ABI PRISM 3100 DNA sequencer (Applied Biosystems, Foster City, CA, USA).

Sequence alignment and phylogenetic analysis

P gene sequences of lyssaviruses deposited in GenBank were downloaded and combined with the newly sequenced samples to form the dataset used in this study. Alignment of nucleotide sequences and deduced amino acid sequence were performed by using the ClustalX program, version 2.1 [43]. Genetic identities were determined using the BioEdit program [44] and MegaAlign software version 5 (DNAStar, Inc., Madison, WI, USA). Phylogenetic and evolutionary analyses were conducted using Mega 3.1 [45]. Neighbor-joining (NJ) phylogenetic trees were constructed using evolutionary distance correction statistics [46,47]. Bootstrap analysis was performed using 1000 replications and values greater than 70% were regarded as strong evidence for particular phylogenetic groupings.

Bayesian Markov chain Monte Carlo (MCMC) evolutionary analysis

Evolutionary history, including evolutionary rates of populations (nucleotide substitutions per site per year) and TMRCA (the most recent common ancestor) were inferred by using the Bayesian Markov chain Monte Carlo (MCMC) method available in the BEAST software package (http://beast.bio.ed.ac.uk/Main_Page) [48]. Briefly, an input file for BEAST was generated by using the BEAUTi program with sequences dated according to the year of isolation. Sequences with homology greater than 98% were removed from the analysis using TCOFFEE. The best-fit model of nucleotide substitution for Bayesian analysis was selected with Modeltest 3.7 [49]. The general time reversible (GTR) substitution model, incorporating a proportion of invariable sites (I) and a gamma distribution of rate variation among sites (C4) was used for the BEAST analysis. Both strict and relaxed (uncorrelated exponential and lognormal) molecular clocks [50] were considered to explore the extent of variation in the rate of nucleotide substitution. The BEAST output was assessed using the TRACER program. The maximum clade credibility (MCC) tree was generated using Figtree (available from http://beast.bio.ed.ac.uk).

Competing interests

The authors declare that they have no competing interests.

Authors’ contributions

LHW did genetic mutation, phylogenetic and evolution analysis and drafted the manuscript; HW carried out nucleic acid detection and sequencing; YXT and HL participated in the collection of samples; SR participated the genetic mutation, phylogenetic and evolution analysis; GDL participated in the design of experiments; QT conceived of the study, and participated in its design and coordination. All authors read and approved the final manuscript.

Authors’ information

Dr. Lihua Wang, Ph.D., is an associate professor at the State Key Laboratory for Infectious Disease Prevention and Control, the Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention. His current research focuses on molecular epidemiology of Rabies virus, development reverse genetic system of rabies virus and basic research related to rabies.
The phosphoprotein of rabies virus is involved in binding to the viral nucleocapsid. The nucleocapsid contains the genomic RNA and the phosphoprotein plays a key role in nuclear import and packaging of the viral genome. The phosphoprotein is essential for viral replication and virion formation. It interacts with other viral proteins, including the matrix protein and the nucleoprotein, to facilitate the assembly of the virion. The phosphoprotein is also involved in the regulation of the viral life cycle, including the inhibition of interferon production and the induction of cell death. The study of the phosphoprotein provides insights into the pathogenesis and epidemiology of rabies, which is a zoonotic disease caused by lyssaviruses.

References:

1. Internationale Kommission zur Taxonomie der Virusarten: ICTV Taxonomy of Viruses. 2009. Version 4. ICTV Bulletins 8:365–401. http://talk.ictvonline.org/files/icvdocs/ICTV Taxonomy of Viruses.pdf

2. Jacob Y, Ceccaldi PE, Tordo N. The nucleocapsid of lyssaviruses: structure and function. Virus Res 2003, 91:51–113.

3. Gupta AK, Blondel D, Choudhary S. The role of the phosphoprotein in the entry and replication of lyssaviruses. Virus Res 2005, 109:137–153.

4. Lo KW, Naisbitt S, Fan JS, Sheng M, Zhang M. The 8-kDa dynactin light chain binds to its targets via a conserved (R/K)RXTGQ motif. J Biol Chem 2001, 276:10459–10466.

5. Bzdok K, Finke S, Conzelmann KK. Identification of the rabies virus alpha/beta interferon antagonist: phosphoprotein R interferes with phosphorylation of interferon regulatory factor 3. J Virol 2005, 79:7673–7681.

6. Vidy A, Chelbi-Alix MK, Blondel D. Rabies virus P protein interacts with STAT1 and inhibits interferon signal transduction pathways. J Virol 2005, 79:14411–14420.

7. Vidy A, Bougrini J, Chelbi-Alix MK, Blondel D. The nucleocapsid of rabies virus interacts with the PML repressor to prevent cell death. J Virol 2006, 80:6196–6206.

Acknowledgements

We thank the staff of the provincial CDCs (Anhui, Guangxi, Guizhou, Hunan, Jiangsu, Shanxi, Shanghai and Zhejiang) for helping with field investigations and sample collection. This work was supported by the National Department Public Benefit Research Foundation (200803014), Major Program of National Natural Science Foundation of China (30630409), Key Technologies Research and Development Program of China (2009ZX10004-705) and Grant from NIIID (National Institute of Infectious Diseases, Japan).

Author details

1. State Key Laboratory for Infectious Disease Prevention and Control, Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention, 155 Changbai St., Changping Dist, Beijing 102206, China.

2. State Key Laboratory for Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Hubei 430071, China.

Received: 28 June 2012 Accepted: 7 December 2012

Published: 7 January 2013

Page 9 of 10
42. Shimizu K, Ito N, Mita T, Yamada K, Hosokawa-Muto J, Sugiyama M, Minamoto N. Involvement of nucleoprotein, phosphoprotein, and matrix protein genes of rabies virus in virulence for adult mice. *Virus Res* 2007, 123:154–160.

43. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 1997, 25:4876–4882.

44. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp* 1999, 41:95–98.

45. Kumar S, Tamura K, Nei M. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* 2004, 5:96–101.

46. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 1980, 15:111–120.

47. Tajima F, Nei M. Estimation of evolutionary distance between nucleotide sequences. *Mol Biol Evol* 1984, 1:269–281.

48. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 2007, 7:214.

49. Posada D, Crandall KA. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 1998, 14(9):817–818.

50. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating with confidence. *PLoS Biol* 2006, 4:e88.

doi:10.1186/1743-422X-10-14

Cite this article as: Wang et al.: Genetic and evolutionary characterization of RABVs from China using the phosphoprotein gene. *Virology Journal* 2013, 10:14.