BMC Bioinformatics

CrossMark

Phylodynamic reconstruction of the spatiotemporal transmission and demographic history of coxsackievirus B2

Hui-Wen Huang[1,2†], Yao-Shen Chen[3,4,5†], Jeff Yi-Fu Chen[6], Po-Liang Lu[7,8], Yung-Cheng Lin[9], Bao-Chen Chen[4], Li-Chiu Chou[8], Chu-Feng Wang[8], Hui-Ju Su[8], Yi-Chien Huang[10], Yong-Ying Shi[10], Hsiu-Lin Chen[11,12], Bintou Sanno-Duanda[10,13], Tsi-Shu Huang[4], Kuei-Hsiang Lin[7], Yu-Chang Tyan[14,15] and Pei-Yu Chu[8,10*]

Abstract

Background: Studies regarding coxsackievirus (CV) tend to focus on epidemic outbreaks, an imbalanced topology is considered to be an indication of acute infection with partial cross-immunity. In enteroviruses, a clear understanding of the characteristics of tree topology, transmission, and its demographic dynamics in viral succession and circulation are essential for identifying prevalence trends in endemic pathogens such as coxsackievirus B2 (CV-B2). This study applied a novel Bayesian evolutionary approach to elucidate the phylodynamic characteristics of CV-B2. A dataset containing 51 VP1 sequences and a dataset containing 34 partial 3D^{pol} sequencing were analyzed, where each dataset included Taiwan sequences isolated during 1988–2013.

Results: Four and five genotypes were determined based on the 846-nucleotide VP1 and 441-nucleotide 3D^{pol} (6641–7087) regions, respectively, with spatiotemporally structured topologies in both trees. Some strains with tree discordance indicated the occurrence of recombination in the region between the VP1 and 3D^{pol} genes. The similarities of VP1 and 3D^{pol} gene were 80.0 %–96.8 % and 74.7 %–91.9 %, respectively. Analyses of population dynamics using VP1 dataset indicated that the endemic CV-B2 has a small effective population size. The balance indices, high similarity, and low evolutionary rate in the VP1 region indicated mild herd immunity selection in the major capsid region.

Conclusions: Phylodynamic analysis can reveal demographic trends and herd immunity in endemic pathogens.

Keywords: Coxsackievirus B2, Phylodynamic, Spatiotemporal transmission, Demographic history, Molecular epidemiology

Background

Coxsackievirus (CV) B2 belongs to the species *Enterovirus* B (EV-B) in the genus *Enterovirus*, family *Picornaviridae*. Long-term surveillance of circulating EV serotypes has identified three main outbreak patterns: epidemic, endemic and sporadic [1]. However, most studies of EV have focused on 'epidemic' pathogens rather than on 'sporadic' or 'endemic' pathogens such as CV-B2; thus, no phylogenetic analysis of CV-B2 has been reported. In terms of epidemiology, a virus is considered 'endemic' if its population remains constant within a population in a geographic area; it is considered 'sporadic' if it appears often but not regularly. The circulation model of individual serotypes differs by season, geographic location, herd immunity and viral genotypes [1, 2]. Thus, analysis history of spatiotemporal transmission and viral population dynamics can reveal trends in their patterns of circulation and succession.

In Taiwan, episodes of unusually high CV-B2 activity were detected in 1999, 2002, 2003 and 2006, and

* Correspondence: peiyuchu@kmu.edu.tw
†Equal contributors
8Department of Laboratory Medicine, Kaohsiung Medical University Hospital, Kaohsiung, ROC, Taiwan
10Department of Medical Laboratory Science and Biotechnology, College of Health Sciences, Kaohsiung Medical University, Kaohsiung, ROC, Taiwan
Full list of author information is available at the end of the article

Huang et al. BMC Bioinformatics (2015) 16:302

Page 2 of 9

infections are still reported almost annually to the Taiwan Centers for Disease Control (Additional file 1). CV-B2 has an endemic circulation pattern worldwide where it comprises 1.5 %–6.0 % of the annual reported EV rates (Additional file 2) [1, 3–7]. The peak activity of CV-B2 (>10 %) is rare, but it is commonly identified as one of the top 15 most frequently detected EV serotypes globally [3, 4]. CV-B2 shares many common properties and clinical features with other CV-Bs but they have different outbreak tempos, and overwhelming infection may develop in infants (<1 year) [8]. CV-B infection is also associated with myocarditis, polymyositis and chronic autoimmune diseases such as insulin-dependent diabetes mellitus [9]. Moreover, outbreaks of emerging and re-emerging infectious diseases are strongly associated with climate change and human activity [10]. Increased CV-B1 activity levels and severe infections of young infants have been reported in the USA (2007–2008) and South Korea (2008–2009) [11, 12]. Therefore, identifying the phylodynamics of CV-B2 is essential for bridging relationships between successive prevalent serotypes and identifying EV circulation trends.

The Bayesian Evolutionary Analysis Sampling Trees (BEAST) program combined various models to infer complex relationships among viral phylodynamics [13]. In addition to genetic variation, viral spatial and temporal components are fundamental issues that require clarification as the dispersal pattern expands. A clear understanding of the tree topology, transmission history, and demographic dynamics of an endemic pathogen such as CV-B2 during viral succession and circulation is needed to identify its prevalence trends. In EVs, the VP1 gene sequence encodes the major serotypic epitope, thereby making it the target choice for molecular typing [14, 15]. Recombination, which is reportedly a common phenomenon in the EV family, can be identified based on phylogenetic incongruencies in the VP1 and $3D^{pol}$ gene regions [16–18] because $3D^{pol}$ is located in the last codon region of the EV genome. To explore the phylogenetic structure of endemic pathogens such as CV-B2, both VP1 and $3D^{pol}$ regions were ananlyzed by the BEAST program to elucidate its phylodynamic properties and reconstruct its spatiotemporal transmission history.

Methods

Ethics statement

Twenty virus strains were randomly selected from CV-B2-positive viral stock. These virus strains were collected during 1988–2013 by the Taiwan Centers for Disease Control, Taiwan (CDC-TW) and two medical centers in Southern Taiwan, Kaohsiung Veterans General Hospital and Kaohsiung Medical University Hospital. This study was approved by the ethics committees of both

hospitals. All samples were de-identified and analyzed anonymously. Informed consent was waived because we only conducted experiments on viral isolates obtained from clinically necessary laboratory procedure and no harm to the patients was anticipated because none of the patients' medical history was studied.

Specimens, viral RNA extraction, reverse transcription-polymerase chain reaction (RT-PCR) and sequencing

Among the twenty CV-B2 strains, fifteen were isolated from throat swabs and five were isolated from rectal swabs or stool. Forty percent of the infected patients were infant younger than 1 year (mean age: 2 years; range: 1 month – 11 years). The male-female ratio was 1.86:1. Viral RNA was extracted using a QIAmp viral RNA purification kit according to the manufacturer instructions (Qiagen, Chatsworth, CA, USA). The RT-PCR and sequencing were performed as described previously [19, 20]. Table 1 lists the degenerate primer pairs used for amplifying and sequencing the VP1 and partial $3D^{pol}$ genes. The purified PCR products were sequenced with an ABI Prism Ready Reaction Dideoxy Terminator cycle sequencing kit (Model 3730, Version3.4, Applied Biosystems, Foster City, CA, USA). The obtained sequence data were submitted to GenBank under accession numbers AB862097, AB862101, AB862107, AB862115–AB862116, LC055763–LC055766, LC055768–LC055778, and LC057297–LC057316.

Model selection and sequence variation detection

Multiple sequence alignments were performed with T-coffee [21, 22]. In addition to the 20 CV-B2 strains obtained in this study, the sequences available in GenBank were also included. Sequences were excluded if they had a nonsense or frame-shift mutation pattern in a single strain. The final VP1 dataset comprised 51 sequences and the partial $3D^{pol}$ dataset comprised 34 sequences (Additional file 3).

Table 1 Primer sets used for amplification and sequencing of Coxsackievirus B2

Gene	Primer[a]	Sequence	Position[b]	Reference
VP3	2400F	GCTTTGTGTCTGCMTGYAATGA	2372-2393	CDC-TW[c]
VP1	222R	CICCIGGIGGIAYRWACAT	2909-2891	[20]
VP1	292F	MIGCIGYIGARACNGG	2550-2565	[20]
2A	011R	GCICCIGAYTGITGICCRAA	3327-3308	[20]
$3D^{pol}$	PY-03F	GTYACMTATGTGAARGATG	6383-6398	This study
$3D^{pol}$	PY-04R	CTTCATTGGCATTACTGGATG	7104-7085	This study

[a]F: Forward primer, R: reverse primer
[b]Numbering system used for the Coxsackievirus B2 strain (Accession No. AF085363)
[c]Primer designed by the Centers for Disease Control, Taiwan

Huang *et al. BMC Bioinformatics* (2015) 16:302

Page 3 of 9

The jModelTest v 2.1.4 program [23] was used to select the best-fit model for both datasets according to Akaike's information criterion. A 4-category Generalised Time Reversible [24] model with a Gamma distribution (GTR + G) was used for VP1 and $3D^{pol}$ (G = 0.1570 and G = 0.1380, respectively). For the BEAST estimation, eight model compositions were compared: two substitution models (GTR + G and Shapiro-Rambaut-Drummond-2006 (SRD06) [25]), two population models [constant (CON) population size and Bayesian skyline plot (BSP)], and two relaxed molecular clock models [uncorrelated log-normal distribution and uncorrelated exponential distribution (uced)] [26]. The Bayes Markov chain Monte Carlo (BMCMC) analysis results were sampled regularly until convergence was reached.

Recombination was estimated using the Recombination Detection Program (RDP v3.44) with the default settings [27]. Pairwise comparisons of nucleotide and amino acid (aa) sequences were calculated based on *p*-distance matrices implemented in the Molecular Evolutionary Genetics Analysis version 6 (MEGA6) program. The ratio of non-synonymous/synonymous substitutions rate (d*N*/d*S*) was used as an indicator of the selection force acting on the coding sequences and was detected by using Single Likelihood Ancestor Counting (SLAC) method via the Datamonkey website [28].

Phylodynamic analyses

Phylogenetic analyses were performed as described previously [19]. The BMCMC tree analyses were performed using BEAST v.1.8.2 [13]. Nodal reliability of BMCMC trees were estimated according to posterior probability (PP). The breakpoint for significant support was set to a PP value > 0.9. The BEAST program was used to co-estimate the nt substitution rate and population growth model in the Tracer v.1.6 program [29]. Parameters with effective sample size (ESS) values > 200 were considered reliable. Briefly, each run of VP1 and $3D^{pol}$ had a chain length of 6 million generations, and sampling was performed once every 6000 generations. When the major parameters with ESS value > 200 were considered as candidates for the best model composition, the best model for both the VP1 and $3D^{pol}$ regions was SRD06 + UCED + CON (Additional file 4). All estimation parameters were shown as mean and 95% highest posterior density (HPD). Demographic changes over time (Neτ) were also explored by BSP method [30]. The FigTree v.1.4.2 program was used to construct and visualize the maximum clade credibility tree. The SPREAD program [31] was used to estimate the route of virus transmission and to calculate the Bayes factors (BFs) associated with major virus dispersal routes. A major virus dispersal route was defined as a route in which at least two locations had BF values > 3. Furthermore, the TreeStat program in the
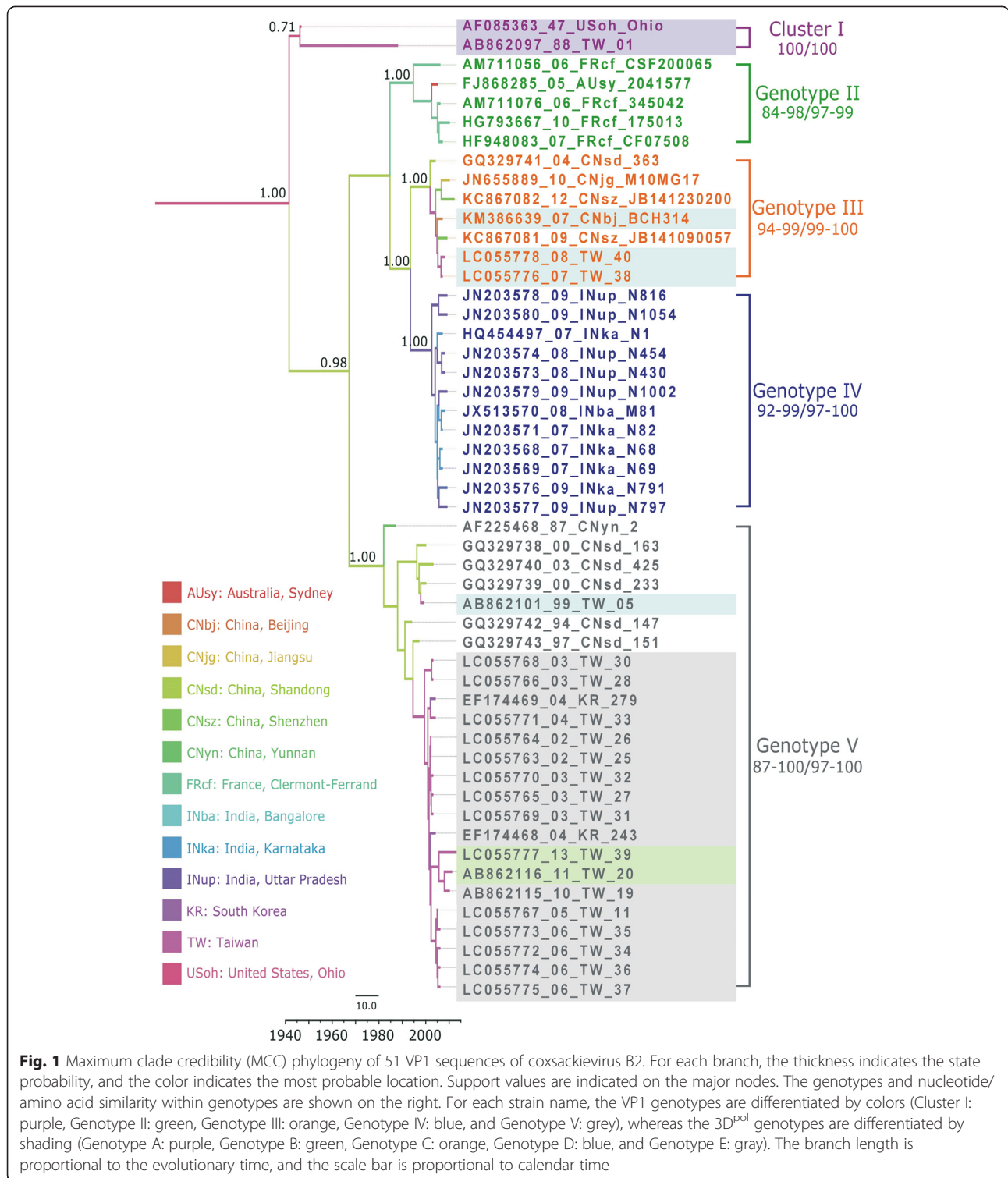
BEAST package was used to summarize statistical tests of neutrality (Fu & Li's *D*) and to calculate three tree-balance statistical measures: cherry count (*Cn*), Colless's tree imbalance (*Ic*), and B_1 [32].

Results and discussion

No recombination events were detected in the VP1 or $3D^{pol}$ region. High support values (defined as PP > 0.9) were obtained for four genotypes in the VP1 tree (Genotypes II–V; GII–GV) and for five genotypes in the $3D^{pol}$ tree (Genotypes A–E; GA–GE) (Figs. 1 and 2). Both the VP1 and $3D^{pol}$ trees contained spatiotemporally structured clusters, but the VP1 and $3D^{pol}$ trees were clearly incongruent. Based on a genetic discrimination rate of at least 15%, five clusters were obtained in the VP1 region (Fig. 1). In the ascent cluster (cluster 1), the prototype Ohio strain (isolated in the US in 1947) and Taiwan strain 01 (isolated in 1988) were clustered together in VP1 and in $3D^{pol}$ (GA). However, the support value was not statistically significant for VP1. In both the VP1 and partial $3D^{pol}$ regions, Taiwan strain 01 had the same sequence as the prototype Ohio strain. Strains GII (1994–2013), GIV (1999–2003), and GV (2007–2009) were isolated in Asia, where GIV was geographically distributed in India. By contrast, GIII was isolated in France (2006–2010) and Australia (2005). The analysis of the historical transmission routes for VP1 by using the Spread program showed that only the Taiwan-South Korea route had a BF of 6.67. The demographic history determined by BSP showed that the CV-B2 viral population had maintained a stable level with only a slight decrease since 1947, where the median Neτ was 3.57–3.42 (Fig. 3a). Compared with the data reported in year 2000, the Neτ for CV-B2 was much lower than the median Neτ values of 4×10^3 reported for EV68 [33] and 30 for CV-B5 [34].

Phylogenetic comparisons of the VP1 topology with those of other EVs revealed similar features: (1) a serotype-based monophyletic origin; (2) different CV-B2 genotypes evolved and diverged sequentially, where each genotype had a different geographic distribution and circulation half-life; (3) the prevalent lineages were chronologically replaced by emerging different lineages that co-existed at the same time and in the same location and (4) tree topologies with long internal branches (i.e., the branch from the node near the root) and intricate short terminal branches [35]. In particular, the long internal branches and short terminal branches indicated a stable pathogen population size [36], which supported BSP estimation of the CV-B2 viral population in this study.

Five genotypes are depicted in the $3D^{pol}$ tree (Fig. 1b). The ancestral America strains (1947–1996) were clustered in GA–C, especially in GB. By contrast, most of

Huang *et al. BMC Bioinformatics* (2015) 16:302

Page 4 of 9

Fig. 1 Maximum clade credibility (MCC) phylogeny of 51 VP1 sequences of coxsackievirus B2. For each branch, the thickness indicates the state probability, and the color indicates the most probable location. Support values are indicated on the major nodes. The genotypes and nucleotide/ amino acid similarity within genotypes are shown on the right. For each strain name, the VP1 genotypes are differentiated by colors (Cluster I: purple, Genotype II: green, Genotype III: orange, Genotype IV: blue, and Genotype V: grey), whereas the 3D^{pol} genotypes are differentiated by shading (Genotype A: purple, Genotype B: green, Genotype C: orange, Genotype D: blue, and Genotype E: gray). The branch length is proportional to the evolutionary time, and the scale bar is proportional to calendar time

the Taiwan strains were clustered in GD–GE with the other Asian strains. Interestingly, GB in the 3D^{pol} tree included two Taiwan strains isolated during 2011–2013 (clustered in the VP1 region of GV). This indicated that those strains with GV sequence in VP1 region has endemic in Taiwan since 2002–2013, but the 3D^{pol} gene might have switched from GE to GB, because the 3D^{pol} gene in the latter two isolates (2011 and 2013) clustered

Huang *et al. BMC Bioinformatics* (2015) 16:302

Page 5 of 9
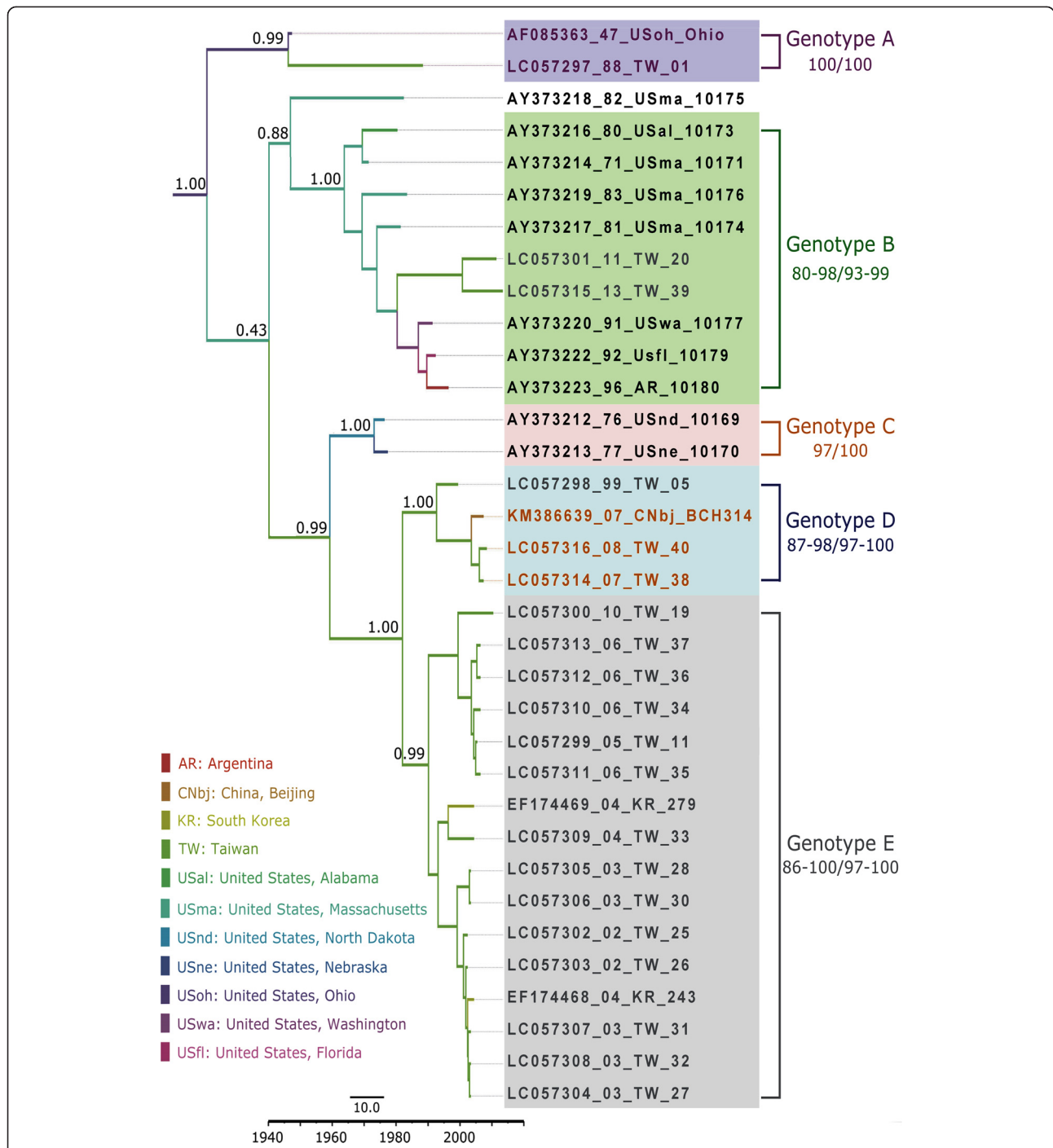
Fig. 2 Maximum clade credibility (MCC) phylogeny of 34 3D^{pol} sequences coxsackievirus B2. For each branch, the thickness indicates the state probability and the color indicates the most probable location. Support values are indicated on the major nodes. The genotypes and nucleotide/amino acid similarity within genotypes are shown on the right. For each strain name, the VP1 genotypes are differentiated by colors (Cluster I: purple, Genotype II: green, Genotype III: orange, Genotype IV: blue, and Genotype V: grey), whereas the 3D^{pol} genotypes are differentiated by shading (Genotype A: purple, Genotype B: green, Genotype C: orange, Genotype D: blue, and Genotype E: gray). The branch length is proportional to the evolutionary time, and the scale bar is proportional to calendar time

together with US strains in GB. Therefore, changes in the trends in CV-B2 in Taiwan should be monitored continuously in Taiwan.

In contrast to the low (< 4) median Neτ value estimated by VP1 region of CV-B2 in this study, the demographic history analysis of the 3D^{pol} region by BSP

Huang *et al. BMC Bioinformatics* (2015) 16:302

Page 6 of 9
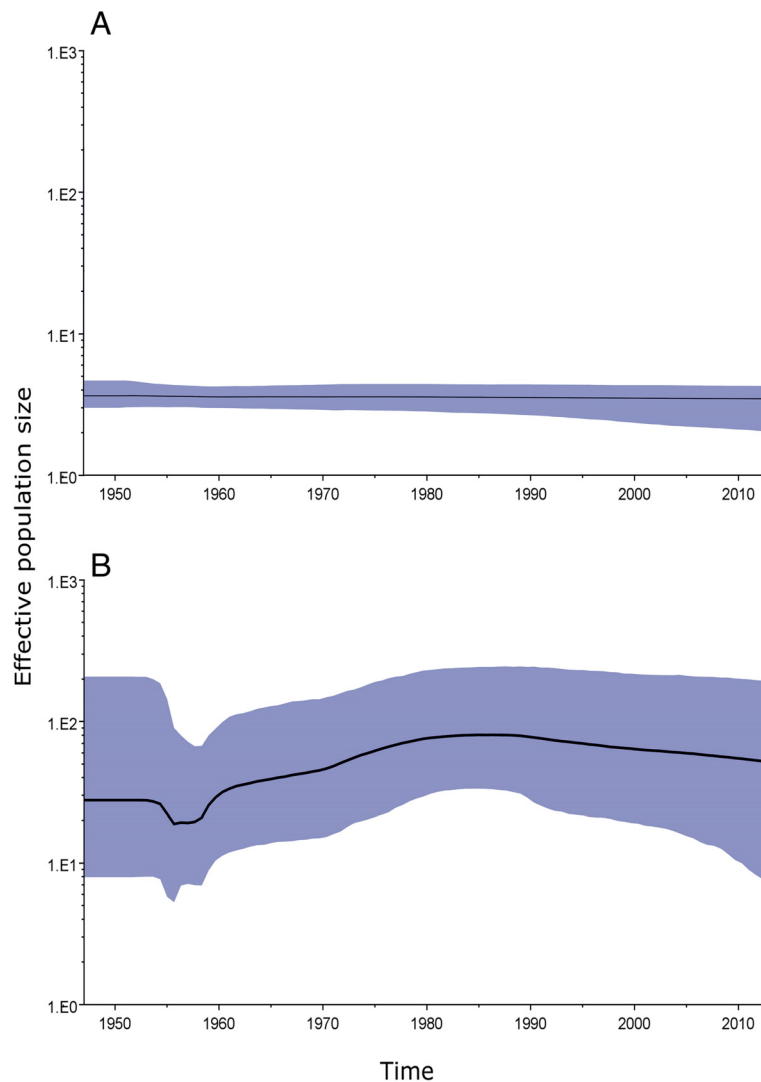
Fig. 3 Bayesian skyline plot of (**a**) 51 VP1 sequences and (**b**) 34 3D^{Pol} sequences. The x-axis is the time scale (years) and the y-axis is the logarithmic Neτ scale (Ne is the effective population size and τ is the generation time). The thick solid line indicates the median estimates and the shaded area indicates the 95% highest posterior density

showed that CV-B2 emerged during 1947, where it had a larger Neτ value of 27.8 before it began to bottleneck in 1955. After peaking in 1985 the Neτ values decreased gradually from 80.4 to 51.7 during 1985–2013 (Fig. 3b). Only capsid-based analyses reveal that serotypes of EVs are monophyletic; however, when the analysis is based on regions outside the capsid region, EVs are monophyletic species rather than serotypes [37]. Thus, the 3D^{pol}-based estimation of Neτ was relevant to EV-B species (CV-B2 is one of the EV-B serotypes). One proposed explanation for this phenomenon is independent evolution of VP1 and 3D^{pol} [38], which is consistent with the discordant topologies of VP1 and 3D^{pol} determined in the present study. Our results indicate that recombination

may occur within discordant strains, and that mutation in the capsid region is a result of immune escape.

However, this study detected high (>95 %) aa sequence similarity. Based on the *p*-distances calculated in this study, the estimated maximum differences in the 846-nt (2456–3301) and 282-aa sequences of the VP1 gene were 80.0 % and 96.8 %, respectively. The estimated maximum differences in the 441-nt (6641–7087) and 147-aa sequences of the partial 3D^{pol} gene were 74.7 % and 91.9 %, respectively. Further, the estimated evolutionary rates of VP1 and 3D^{pol} were 5.42×10^{-3} (95 % HPD: 3.38×10^{-3}–6.98×10^{-3}) and 6.02×10^{-3} (3.72×10^{-3}–8.31×10^{-3}) substitutions/site/year (s/s/y), respectively. The VP1 mutation rate of CV-B2 is within the mutation rate range reported

Huang *et al. BMC Bioinformatics* (2015) 16:302

Page 7 of 9

in a previous study of EVs (range from 3.40×10^{-3} to 1.19×10^{-2} s/s/y) [39]. In the present study, the estimated evolution rate of $3D^{pol}$ (6.02×10^{-3} s/s/y) was not only higher than that of VP1 but also higher than the estimated evolution rate reported for $3D^{pol}$ in previous studies of EVs (from 5.53×10^{-3} to 1.17×10^{-2} ns/s/y) [39]. The dN/dS values for the VP1 and $3D^{pol}$ region were 0.0334 and 0.0265, respectively, but neither region contained a positive selection site. In summary, the low and slightly decreasing Neτ value, low evolution rate, and negative selection indicate that the variation in this VP1 region was fixed by adaptive fitness.

Phylodynamic analysis depicts interaction among epidemics, evolution and selection and reveals the demographic and spatiotemporal transmission signatures [26]. A phylogenetic tree with an imbalanced trunk and short terminal branches suggests acute infection with partial cross-immunity, for example, the phylogenetic tree topology for seasonal influenza virus is ladder-like. In contrast, the topology of a chronic infection such as hepatitis C virus is star-like with long terminal branches [26]. The tree balance is impacted by selection; therefore, a ladder-like backbone indicates the trajectory of fitness. Table 2 summarizes the results of the statistical analyses of tree shape and balance obtained in the present study. Fu and Li's D value is a classic summary statistic for testing neutrality, which indicates the normalize difference between the external branch lengths and total tree length. As the symmetry of a phylogeny increases, the values for B_1 and Cn are expected to increase, whereas the value of Ic is expected to decrease. The distributions of the mean Ic indices among

human and zoonotic RNA viruses reportedly peak between 0.15 and 0.2 [32]. The values for the VP1 tree were within this range in the present study, but $3D^{pol}$ had higher tree balance indices compared with VP1.

A major challenge in modern phylodynamics is qualitatively and quantitatively describing a tree topology to elucidate the epidemiological, evolutionary, and demographic characteristics of a pathogen. To investigate the evolutionary history of an endemic or sporadic pathogen, a BMCMC sampling framework was used to infer time scaled phylogenies of CV-B2 during 1947–2013. Sampling bias was unavoidable in this study because the incomplete GenBank data precluded sampling of all strains isolated in earlier time periods (particularly before the 1980s). An even greater problem was that the sequences tended to accumulate in the same isolation years and locations because CV-B2 outbreaks are rare. On average, CVB2 has been implicated in 1.5 %–6.0 % of the reported surveillance results for known EV serotypes, and thus the low frequency of CV-B2 outbreaks may be related to the stable low Neτ values estimates obtained in this study. Furthermore, low herd immunity selection may have an important role in endemic EVs because the VP1 has high aa sequence similarity and a low rate of evolution. Compare to $3D^{pol}$ region, the VP1 region has a more balanced topology.

Conclusions

An endemic or sporadic EV, CV-B2, was analyzed in a phylodynamic analysis, and four genotypes in the VP1 region and five genotypes in the $3D^{pol}$ region were determined. Recombination events were not detected in the VP1 sequences or $3D^{pol}$ sequences, but VP1 and $3D^{pol}$ trees were clearly incongruent, thereby indicating the occurrence of recombination events involving the region between VP1 and $3D^{pol}$. Although the random branching topology observed in this study is typically interpreted as the rapid evolution of an RNA virus, the imbalanced tree topology and the high similarity of the VP1 sequences indicate that endemic CV-B2 is characterized by a low viral population and low herd immune selection.

Availability of supporting data
See additional files for supporting data.

Additional files

> **Additional file 1: Annual reported rate of coxsackievirus B2 (CV-B2) in Taiwan.** (TIFF 686 kb)
>
> **Additional file 2: Global surveillance of coxsackievirus B2.** (PDF 6 kb)
>
> **Additional file 3: List of sampled coxsackievirus B2 (CV-B2) strains.** (PDF 14 kb)
>
> **Additional file 4: Models compared by an extension of air carrier investment model (ACIM).** (PDF 10 kb)

Table 2 Summary statistics used in tests of demographic neutrality and tree balance

Gene region	Statistical measures	Mean	Median	95 % HPD[a]
VP1	D^b	-2.9149	-2.933	-3.5863, -2.1837
	B_1^c	24.9094	24.9591	22.9639, 26.9044
	Cn^d	15.7291	16	14, 18
	Ic^d	0.1969	0.1902	0.1682, 0.2384
3D	D	-2.1294	-2.1457	-2.818, -1.4114
	B_1	16.7549	16.6385	15.5468, 18.0199
	Cn	11.0685	11	10, 12
	Ic	0.238	0.2386	0.1875, 0.2822

[a]HPD Highest posterior density
[b]*D* Fu and Li's D: Normalized difference between the external branch lengths and total tree length. This is a classic summary statistic used for testing neutrality
[c]*B₁* Maximum number of nodes between an internal node and the tips of the tree. A larger value indicates a more balanced tree
[d]*Cn* Cherry count, the number of internal nodes with exactly two terminal branches. A larger number indicates a more balanced tree
[e]*Ic* Colless's tree imbalance index with a range of [0, 1]. A larger number indicates a more imbalanced tree

Huang *et al. BMC Bioinformatics* (2015) 16:302

Page 8 of 9

Author details
[1]Department of Anesthesiology, Kaohsiung Chang Gung Memorial Hospital and Chang Gung University College of Medicine, Kaohsiung, ROC, Taiwan. [2]Department of Biological Sciences, National Sun Yat-Sen University, Kaohsiung, ROC, Taiwan. [3]Division of Infectious Diseases, Kaohsiung Veterans General Hospital, Kaohsiung, ROC, Taiwan. [4]Division of Microbiology, Department of Pathology and Laboratory Medicine, Kaohsiung Veterans General Hospital, Kaohsiung, ROC, Taiwan. [5]Department of Internal Medicine, National Yang-Ming Medical University, Taipei, ROC, Taiwan. [6]Department of Biotechnology, Kaohsiung Medical University, Kaohsiung, ROC, Taiwan. [7]School of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, ROC, Taiwan. [8]Department of Laboratory Medicine, Kaohsiung Medical University Hospital, Kaohsiung, ROC, Taiwan. [9]Institute of Bioscience and Biotechnology, National Taiwan Ocean University, Keelung, ROC, Taiwan. [10]Department of Medical Laboratory Science and Biotechnology, College of Health Sciences, Kaohsiung Medical University, Kaohsiung, ROC, Taiwan. [11]Department of Pediatrics, Kaohsiung Medical University Hospital, Kaohsiung, ROC, Taiwan. [12]Department of Respiratory Therapy, College of Medicine, Kaohsiung Medical University, Kaohsiung, ROC, Taiwan. [13]Department of laboratory medicine, Edward Francis Small Teaching Hospital, Banjul, Gambia. [14]Department of Medical Imaging and Radiological Sciences, Kaohsiung Medical University, Kaohsiung, Taiwan. [15]Center for Infectious Disease and Cancer Research, Kaohsiung Medical University, Kaohsiung, Taiwan.

References
1. Khetsuriani N, Ashley L-F, Oberst MS, Pallansch MA. Enterovirus surveillance–United States, 1970-2005. MMWR Surveill Summ. 2006;15(55):1–20.
2. Di Tommaso P, Orobitg M, Guirado F, Cores F, Espinosa T, Notredame C. Cloud-Coffee: implementation of a parallel consistency-based multiple alignment algorithm in the T-Coffee package and its benchmarking on the Amazon Elastic-Cloud. Bioinformatics. 2010;26(15):1903–4.
3. Hyeon JY, Hwang S, Kim H, Song J, Ahn J, Kang B, et al. Accuracy of diagnostic methods and surveillance sensitivity for human enterovirus, South Korea, 1999-2011. Emerg Infect Dis. 2013;19(8):1268–75.
4. Tryfonos C, Richter J, Koptides D, Yiangou M, Christodoulou CG. Molecular typing and epidemiology of enteroviruses in Cyprus, 2003-2007. J Med Microbiol. 2011;60(Pt 10):1433–40.
5. Antona D, Lévêque N, Chomel J, Dubrou S, Lévy-Bruhl D, Lina B. Surveillance of enteroviruses in France, 2000–2004. Eur J Clin Microbiol Infect Dis. 2007;26:403–12.
6. Roth B, Enders M, Arents A, Pfitzner A, Terletskaia-Ladwig E. Epidemiologic aspects and laboratory features of enterovirus infections in Western Germany, 2000-2005. J Med Virol. 2007;79(7):956–62.
7. Bahri O, Rezig D, Nejma-Oueslati BB, Yahia AB, Sassi JB, Hogga N, et al. Enteroviruses in Tunisia: virological surveillance over 12 years (1992-2003). J Med Microbiol. 2005;54(Pt 1):63–9.
8. Sojka M, Wsolova L, Petrovicova A. Coxsackieviral infections involved in aseptic meningitis: a study inSlovakia from 2005 to 2009. Euro Surveill. 2011;16(30):pii:19927.
9. Knipe D, Howley P, Griffin D, Lamb R, Martin M, Roizman B, et al. Fields Virology. 2001. p. 723–76.
10. Morand S, Owers KA, Waret-Szkuta A, McIntyre KM, Baylis M. Climate variability and outbreaks of infectious diseases in Europe. Sci Rep. 2013;3:1774.
11. Wikswo ME, Khetsuriani N, Fowlkes AL, Zheng X, Penaranda S, Verma N, et al. Increased activity of Coxsackievirus B1 strains associated with severe disease among young infants in the United States, 2007-2008. Clin Infect Dis. 2009;49(5):e44–51.
12. Kim H, Kang B, Hwang S, Hong J, Chung J, Kim S, et al. Molecular characteristics of human coxsackievirus B1 infection in Korea, 2008-2009. J Med Virol. 2013;85(1):110–5.
13. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007;7:214.
14. Oberste MS, Maher K, Kilpatrick DR, Flemister MR, Brown BA, Pallansch MA. Typing of human enteroviruses by partial sequencing of VP1. J Clin Microbiol. 1999;37(5):1288–93.
15. Oberste MS, Maher K, Flemister MR, Marchetti G, Kilpatrick DR, Pallansch MA. Comparison of classic and molecular approaches for the identification of untypeable enteroviruses. J Clin Microbiol. 2000;38(3):1170–4.
16. Lindberg AM, Andersson P, Savolainen C, Mulders MN, Hovi T. Evolution of the genome of Human enterovirus B: incongruence between phylogenies of the VP1 and 3CD regions indicates frequent recombination within the species. J Gen Virol. 2003;84(Pt 5):1223–35.
17. Lukashev AN, Lashkevich VA, Ivanova OE, Koroleva GA, Hinkkanen AE, Ilonen J. Recombination in circulating Human enterovirus B: independent evolution of structural and non-structural genome regions. J Gen Virol. 2005;86(Pt 12):3281–90.
18. Zhang T, Du J, Xue Y, Su H, Yang F, Jin Q. Epidemics and Frequent Recombination within Species in Outbreaks of Human Enterovirus B-Associated Hand, Foot and Mouth Disease in Shandong China in 2010 and 2011. PLoS One. 2013;8(6):e67157.
19. Chu PY, Lu PL, Tsai YL, Hsi E, Yao CY, Chen YH, et al. Spatiotemporal phylogenetic analysis and molecular characterization of coxsackievirus A4. Infect Genet Evol. 2011;11(6):1426–35.
20. Oberste MS, Nix WA, Maher K, Pallansch MA. Improved molecular identification of enteroviruses by RT-PCR and amplicon sequencing. J Clin Virol. 2003;26(3):375–7.
21. Di Tommaso P, Moretti S, Xenarios I, Orobitg M, Montanyola A, Chang JM, et al. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. Nucleic Acids Res. 2011;39(Web Server issue):W13–17.
22. Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol. 2000;302(1):205–17.
23. Butan C, Filman DJ, Hogle JM. Cryo-electron microscopy reconstruction shows poliovirus 135S particles poised for membrane interaction and RNA release. J Virol. 2014;88(3):1758–70.
24. Tavaré S: Some Probabilistic and Statistical Problems in the Analysis of DNA Sequences. In: Lectures on Mathematics in the Life Sciences. vol. 17: American Mathematical Society, New York; 1986: 57-86.
25. Shapiro B, Rambaut A, Drummond AJ. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Mol Biol Evol. 2006;23(1):7–9.

Huang *et al. BMC Bioinformatics* (2015) 16:302

Page 9 of 9

26. Grenfell BT, Pybus OG, Gog JR, Wood JL, Daly JM, Mumford JA, et al. Unifying the epidemiological and evolutionary dynamics of pathogens. Science. 2004;303(5656):327–32.

27. Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics. 2010;26(19):2462–3.

28. Pond SL, Frost SD. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics. 2005;21(10):2531–3.

29. Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. Mol Biol Evol. 2012;29(9):2157–67.

30. Morimoto K, Satoh-Yamaguchi K, Hamaguchi A, Inoue Y, Takeuchi M, Okada M, et al. Interaction of cancer cells with platelets mediated by Necl-5/poliovirus receptor enhances cancer cell metastasis to the lungs. Oncogene. 2008;27(3):264–73.

31. Bielejec F, Rambaut A, Suchard MA, Lemey P. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. Bioinformatics. 2011;27(20):2910–2.

32. Drummond AJ, Suchard MA. Fully Bayesian tests of neutrality using genealogical summary statistics. BMC Genet. 2008;9:68.

33. Meijer A, van der Sanden S, Snijders BE, Jaramillo-Gutierrez G, Bont L, van der Ent CK, et al. Emergence and epidemic occurrence of enterovirus 68 respiratory infections in The Netherlands in 2010. Virology. 2012;423(1):49–57.

34. Henquell C, Mirand A, Richter J, Schuffenecker I, Bottiger B, Diedrich S, et al. Phylogenetic patterns of human coxsackievirus B5 arise from population dynamics between two genogroups and reveal evolutionary factors of molecular adaptation and transmission. J Virol. 2013;87(22):12249–59.

35. Chu PY, Ke GM, Chen YS, Lu PL, Chen HL, Lee MS, et al. Molecular epidemiology of Coxsackievirus B3. Infect Genet Evol. 2010;10(6):777–84.

36. Volz EM, Koelle K, Bedford T. Viral phylodynamics. PLoS Comput Biol. 2013;9(3):e1002947.

37. Oberste MS, Maher K, Pallansch MA. Evidence for frequent recombination within species human enterovirus B based on complete genomic sequences of all thirty-seven serotypes. J Virol. 2004;78(2):855–67.

38. Oberste MS. Comparative genomics of the coxsackie B viruses and related enteroviruses. Curr Top Microbiol Immunol. 2008;323:33–47.

39. Hicks AL, Duffy S. Genus-specific substitution rate variability among picornaviruses. J Virol. 2011;85(15):7942–7.