



HAL
open science

Passage in egg culture is a major cause of apparent positive selection in influenza B haemagglutinin.

Derek Gatherer

► **To cite this version:**

Derek Gatherer. Passage in egg culture is a major cause of apparent positive selection in influenza B haemagglutinin.. *Journal of Medical Virology*, 2009, 82 (1), pp.123. 10.1002/jmv.21648 . hal-00541147

HAL Id: hal-00541147

<https://hal.science/hal-00541147>

Submitted on 30 Nov 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TABLE I

Site 197-199	<i>Egg-cultured</i>	<i>Egg-free</i>
With N-glycosylation	14	50
Without N-glycosylation	49	9

Site 233-235	<i>Egg-cultured</i>	<i>Egg-free</i>
With N-glycosylation	29	22
Without N-glycosylation	34	37

Table I: Presence or absence of N-glycosylation site at positions 197-199 and 233-235 in egg-cultured and egg-free strains. Chi-square test significant only for site 197-199 ($p < 0.0001$).

TABLE II

Site 197-199	<i>Yamagata</i>	<i>Victoria</i>
With N-glycosylation	46	18
Without N-glycosylation	33	25

Site 233-235	<i>Yamagata</i>	<i>Victoria</i>
With N-glycosylation	13	38
Without N-glycosylation	66	5

Table II: Presence or absence of N-glycosylation sites at positions 197-199 and 233-235 in Yamagata and Victoria strains. Chi-square test significant only for site 233-235 ($p < 0.0001$).

TABLE III

EGG-CULTURED				
M8	Site	Pr($\omega > 1$)	ω	SE
	75	0.994*	9.843	1.092
	116	0.774	7.793	3.827
	129	0.916	9.115	2.638
	197	1.000*	9.902	0.820
	199	1.000*	9.902	0.820
M2	75	0.952*	9.797	2.043
	197	1.000*	10.246	0.522
	199	1.000*	10.246	0.522

EGG-FREE				
M8	Site	Pr($\omega > 1$)	ω	SE
	48	0.801	1.523	0.613
	56	0.908	1.657	0.527
	73	0.857	1.595	0.583
	75	0.977*	1.729	0.462
	116	0.821	1.549	0.602
	129	0.884	1.629	0.559
	146	0.817	1.544	0.604
	199	0.994*	1.743	0.445
M2	56	0.727	1.662	0.651
	75	0.879	1.835	0.720
	129	0.725	1.676	0.692
	199	0.944	1.880	0.699

1
2
3 Table III: Positive selection in influenza B HA in egg-cultured and egg-free
4 strains. $\Pr(\omega > 1)$: probability that the site in question is subject to positive
5 selection; SE: standard error of omega. Sites 197 and 199 are part of the N-
6 glycosylation site commonly lost in egg-cultured strains. All sites with a
7 $\Pr(\omega > 1)$ of greater than 0.7 are listed. Asterisks indicate $\Pr(\omega > 1)$ of
8 > 0.95
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

TABLE IV

M2	$p \geq 0.5$	$p \geq 0.75$	$p \geq 0.95$	$p \geq 0.99$	<i>Max. omega</i>
<i>Egg-cultured</i>	4	3	3	2	10.246
<i>Egg-free</i>	16	2	0	0	1.880

M8	$p \geq 0.5$	$p \geq 0.75$	$p \geq 0.95$	$p \geq 0.99$	<i>Max. omega</i>
<i>Egg-cultured</i>	5	5	3	3	9.902
<i>Egg-free</i>	22	8	2	1	1.743

Table IV: Summary of number of candidate positively selected sites and the strength of selection, in egg-cultured and egg-free strains. Model M2 allows for 3 selective categories, namely constraint, neutrality and positive selection, and M8 for a beta distribution of 10 constrained or neutral categories plus an eleventh category of positive selection. $p \geq x$: number of sites positively selected at probability of greater than x . *Max. omega*: largest sitewise omega value found in the data.

1
2
3 1 **Passage in egg culture is a major cause of apparent positive selection in**
4
5 2 **influenza B haemagglutinin.**
6
7 3
8
9 4

10 *Derek Gatherer*

11 *MRC Virology Unit*

12 *University of Glasgow*

13 *Church Street*

14 *Glasgow*

15 *G11 5JR, UK*
16
17
18
19
20
21
22
23

24 11 Correspondence to:

25
26 12 Derek Gatherer, MRC Virology Unit, Institute of Virology, Church Street, Glasgow G11
27 13 5JR, UK.

28 14 Tel: +44 141 330 6268

29 15 Fax: +44 141 337 2236

30 16 email: d.gatherer@mrcvu.gla.ac.uk
31
32
33
34
35
36
37
38

39 18 Running title: Positive selection in Influenza B
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 19
4
5 20
6 **Abstract**
7 21
8
9
10 22
11 23
12 24
13 25
14 26
15 27
16 28
17 29
18 30
19 31
20 32
21 33
22 34
23 35
24 36
25 37
26 38
27 39
28 40
29 41
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Several studies have identified residues apparently under positive selection in influenza B virus haemagglutinin. Host immune evasion is the main mechanism proposed to exert this selection pressure. However, these reports have not considered the culture history of the strains used to calculate positive selection. This paper shows that passage of influenza B virus through egg culture is a strong contributory factor to the strength and statistical significance of positive selection on haemagglutinin. Non-synonymous mutations resulting in the loss of the N-glycosylation site at positions 197-199 of haemagglutinin have been positively selected to a far greater degree in egg-cultured strains than in other strains. Once egg-cultured strains are removed from the analysis, positive selection is found to be far weaker, less statistically significant and more diffusely localised along the protein. Caution should therefore be exercised both in claims for the existence of positive selection in influenza B haemagglutinin, and in attribution of host immune evasion as its cause. The major cause of molecular adaptation in influenza B haemagglutinin proteins may well be laboratory eggs rather than natural hosts.

Keywords: N-glycosylation, immune evasion, non-synonymous substitution, codeml, omega, adaptation

42 Introduction

43
44 Influenza A (Family *Orthomyxoviridae*, Genus *Influenzavirus A*, Species *Influenza A*
45 *virus*) and the related influenza B (Family *Orthomyxoviridae*, Genus *Influenzavirus B*,
46 Species *Influenza B virus*) are major public health concerns. A lack of lifelong immunity
47 in human hosts results in seasonal epidemics, with annual vaccination the best current
48 strategy for limiting mortality and the economic impact of these viruses.

49
50 The earliest viral isolates confirmed as influenza B by haemagglutinin (HA) gene
51 sequencing date from 1940 (strains B/Lee/0/1940 to B/Lee/3/1940). Other isolates
52 document the continued presence of the virus during the subsequent two decades
53 (strains B/Maryland/0/59 and B/Russia/0/69). Since the late 1980s, systematic seasonal
54 sequencing of new strains has been performed, revealing that all influenza B HA gene
55 sequences of the last 20 years belong to one or other of two lineages, designated
56 Yamagata and Victoria after their original isolates (B/Yamagata/16/88 and
57 B/Victoria/2/87 respectively). HA protein sequences within each of the two lineages are
58 more than 97% identical, with sequence identity in inter-lineage comparisons at 88-90%
59 on average. The original Lee strain HA from 1940 is approximately 91% identical to
60 modern samples of either lineage. By comparison, influenza A HA sequences are only
61 22-23% identical to influenza B HA.

62
63 It has been proposed that influenza B, like influenza A, achieves host immune evasion by
64 rapid evolution of HA. Mutations causing amino acid substitutions at certain residues of
65 HA are hypothesised to afford a selective advantage to the virus, resulting in a
66 statistically significant excess of non-synonymous over synonymous substitutions at the
67 codons corresponding to these residues [Shen et al., 2009]. Positive selection is the
68 term given to such selection for non-synonymous substitutions. The ratio of the rates of
69 non-synonymous to synonymous substitution is termed omega. An omega of less than 1
70 indicates selective constraint and suppression of non-synonymous substitutions. In

1
2
3 71 neutrally evolving proteins omega will be close to 1, and positive selection produces
4
5 72 omegas which may be considerably greater than 1. PAML [Yang, 1997] provides a suite
6
7 73 of tools for the calculation of omega and determination of its statistical significance.
8
9 74 Omega may be calculated as an average over the whole length of a gene, or in a
10
11 75 sitewise manner. The latter has proved useful for identifying individual residues that are
12
13 76 under selective pressure.
14

15
16 77
17
18 78 Previous studies have detected codons with omega significantly greater than 1 in HA
19
20 79 genes from influenza B, and attribute this positive selection to immune evasion on the
21
22 80 grounds that the detected sites are within HA epitope regions [Nunes et al., 2004; Nunes
23
24 81 et al., 2008; Pechirra et al., 2005; Shen et al., 2009]. However, some of the candidate
25
26 82 positively selected residues are associated with an N-glycosylation site, and mutations
27
28 83 causing the loss of that site are often found in laboratory strains passaged in
29
30 84 embryonated hen eggs [Robertson et al., 1985; Saito et al., 2004], where the absence
31
32 85 of the site can result in increased binding to the plasma membranes of chorio-allantoic
33
34 86 cells [Gambaryan et al., 1999]. By contrast, loss of this N-glycosylation site can also
35
36 87 occur in strains cultured only in Madin Darby canine kidney (MDCK) cells [Ikonen et al.,
37
38 88 2005] and some egg-cultured strains can retain it [Rota et al., 1992]. A full survey of
39
40 89 the association of N-glycosylation site presence with egg culture has not been
41
42 90 undertaken so far.
43
44

45 91
46
47 92 The possibility therefore remains that host immune evasion is being confused with egg
48
49 93 culture adaptation as a cause of positive selection in influenza B HA genes. This study
50
51 94 analyses positive selection in an extensive set of influenza B HA genes, separating genes
52
53 95 from egg-cultured strains from those of cell-cultured strains. Studies published
54
55 96 previously of positive selection in influenza B HA genes have not performed this analysis.
56
57 97 A survey of the association of N-glycosylation site presence with egg culture is also
58
59 98 described. The relative contribution of egg culture to detected positive selection may
60
99 therefore be more fully understood.

1
2
3 100
4
5 101
6
7 102
8
9 103
10
11 104
12
13 105
14
15 106
16
17 107
18
19 108
20
21 109
22
23 110
24
25 111
26
27 112
28
29 113
30
31 114
32
33 115
34
35 116
36
37 117
38
39 118
40
41 119
42
43 120
44
45 121
46
47 122
48
49 123
50
51 124
52
53 125
54
55 126
56
57 127
58
59 128
60

Materials & Methods

Identification of culture history

259 full length non-redundant influenza B HA coding sequences, dating from 1940 to 2007, were retrieved in GenBank format from the NCBI Influenza Virus Resource (<http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html>). GenBank format includes the "/lab_host" tag, annotating the culture history of the strain. Of the 259 sequences, 133 had a "/lab_host" tag. Of these, 63 sequences were listed as having had contact with egg culture (Supplementary Data Table 1). For instance, entry CY018373 (B/Michigan/04/2006) has "/lab_host" tag: "Egg 2 passage(s)". 70 sequences had no mention of eggs, for example CY018301 (B/Houston/B720/2004) has "/lab_host" tag: "RhMK 1 passage(s)", indicating the use of a rhesus monkey kidney cell line. This initial set of 70 sequences was reduced to 59 by removal of those with unknown or ambiguous history (Supplementary Data Table 2). For example, entry CY018725 (B/Paraguay/636/2003) has "/lab_host" tag: "Unknown 1 MDCK 2 passage(s)", and was therefore omitted from the final set of 59 putative egg-free sequences, on the grounds that "Unknown" could indicate previous egg culture. The egg-cultured strains range in date from 1987 to 2006 and the egg-free strains from 1992 to 2006. The proportions of egg-cultured and egg-free strains are roughly equal for most years after 1997. The lineages of the egg-free strains are 23 Victoria and 36 Yamagata strains. For egg-cultured strains, there are 20 Victoria and 43 Yamagata strains.

N-glycosylation site analysis

Due to indels in HA, the N-glycosylation site discussed in the present paper may elsewhere be referred to as occurring at residues 197 to 199 [Ikonen et al., 2005; Kinnunen et al., 1992; Nunes et al., 2004; Nunes et al., 2008; Pechirra et al., 2005; Rota et al., 1992; Rota et al., 1990; Shaw et al., 2002], 196 to 198 [Robertson et al., 1985; Saito et al., 2004] or 194 to 196 [Shen et al., 2009]. In this study, the majority

1
2
3 129 system is adopted: 197-199. For each HA, the presence or absence of the N-
4
5 130 glycosylation consensus sequence (NxS/T, where x is any amino acid) was recorded at
6
7 131 197-199 and the nearby 233-235, which has no previous association with positive
8
9 132 selection and therefore serves as a control. A 2x2 chi-square test was used to
10
11 133 determine the statistical significance of the proportion of N-glycosylation sites in the
12
13 134 egg-cultured and egg-free sets, and also the proportion of N-glycosylation sites in the
14
15 135 Yamagata and Victoria lineages.
16
17
18 136

19 20 137 *Quantitation of positive selection*

21
22 138 Sequences were translated using transeq (EMBOSS) [Rice et al., 2000], then aligned
23
24 139 using MAFFT [Kato et al., 2005]. The protein alignment was then used as a basis for
25
26 140 alignment of the DNA sequences. PAML's CODEML mode [Yang, 1997] was used to
27
28 141 quantify positive selection in these alignments. Comparisons between models M1 (2
29
30 142 discrete categories: constraint plus neutrality only) and M2 (3 discrete categories:
31
32 143 constraint, neutrality and positive selection), between models M7 (10 beta-distributed
33
34 144 categories: constraint plus neutrality only) and M8 (11 categories: as M7 with 1 further
35
36 145 discrete category for positive selection), and between models M8 and M8a (as M8 but
37
38 146 with maximum omega fixed at 1), were used to assess the statistical significance of
39
40 147 detected positive selection.
41
42
43 148

44 45 149 **Results**

46 47 150 48 49 151 *N-glycosylation and egg culture*

50
51 152
52
53 153 The presence or absence of N-glycosylation sites at positions 197-199 and 233-235 was
54
55 154 examined in HA sequences annotated as egg-cultured and those annotated as
56
57 155 mammalian cell-cultured with no mention of "unknown" culture conditions (the "egg-
58
59 156 free" strains). The first null hypothesis is that culture conditions do not affect the
60
157 presence or absence of N-glycosylation sites at either position. Table I shows the

1
2
3 158 statistics for the occurrence of the relevant N-glycosylation sites. A 2x2 chi-square test
4
5 159 is significant only for site 197-199. Thus the null hypothesis, that culture conditions do
6
7 160 not affect N-glycosylation state, can be rejected for site 197-199 only. Egg culture is
8
9 161 therefore demonstrated to be a likely cause of loss of N-glycosylation site 197-199 but to
10
11 162 be unlikely to have any effect on site 233-235. The second null hypothesis is that a
12
13 163 strain's ancestry does not affect the presence or absence of N-glycosylation sites at
14
15 164 either position. Table II similarly shows the N-glycosylation site statistics divided
16
17 165 between Yamagata and Victoria strains. This time, a 2x2 chi-square test is significant
18
19 166 only for site 233-235. Although N-glycosylation is slightly less common in Victoria
20
21 167 strains at position 197-199, this is not statistically significant. The second null
22
23 168 hypothesis, that lineage does not affect N-glycosylation state, can be rejected for site
24
25 169 233-235 only. At that site, Victoria strains are predominantly N-glycosylated, whereas
26
27 170 Yamagata strains are predominantly non-N-glycosylated. Therefore, the presence of the
28
29 171 N-glycosylation site of HA at position 233-235 is mainly an aspect of the ancestry of the
30
31 172 HA sequence and is not affected by culture conditions. The converse applies for the site
32
33 173 at position 197-199 where the relative absence of the N-glycosylation site is a result of
34
35 174 selective pressure from culture conditions and there is no correlation with ancestry of the
36
37 175 strain. The possibility that the absence of the N-glycosylation site at positions 197-199
38
39 176 is merely a result of greater presence of one or other of the 2 lineages in egg culture is
40
41 177 also thereby excluded.
42
43
44
45 178

46 47 179 *Positively selected sites on HA*

48
49 180 Table III shows the PAML output for egg-cultured and egg-free strains. Each set was
50
51 181 analysed using Bayes Empirical Bayes (BEB) [Yang et al., 2005], under models M2 (3
52
53 182 discrete categories, $\omega < 1$, $\omega = 1$ and $\omega > 1$) and M8 (10 beta-distributed
54
55 183 rate categories from ω of zero to 1, plus 1 discrete category of $\omega > 1$). Table
56
57 184 IV summarizes the frequency and intensity of positively selected sites for the two models
58
59 185 in egg-cultured and egg-free strains.
60

186

1
2
3 187 For egg-cultured strains, position 75 and the N-glycosylation site 197-199 are strongly
4
5 188 positively selected ($\omega > 9$) at high probability ($p > 0.99$), using model M8. Positions
6
7 189 116 and 129 have slightly weaker selection ($\omega > 7.5$) at lower statistical
8
9 190 probabilities ($p > 0.7$ for 116 and $p > 0.9$ for 129). Model M2 detects only positions 75,
10
11 191 197 and 199. This confirms previous results [Nunes et al., 2004; Nunes et al., 2008;
12
13 192 Pechirra et al., 2005]. However, in egg-free strains detected positive selection is far
14
15 193 weaker. Using model M8, positions 75 and 199 have weak positive selection only
16
17 194 ($\omega < 1.8$). With model M2, statistical probabilities of positive selection are also
18
19 195 weakened. The apparent positive selection detected in previous studies evaporates to a
20
21 196 large degree when egg-cultured strains are excluded from the analysis.
22
23
24 197

25
26 198 One intriguing further observation (Table IV) is that egg-free strains have more
27
28 199 candidate positively selected sites at lower significance levels ($0.5 < p < 0.95$).
29
30 200 Although caution must be exercised in interpretation of this weaker signal, it may
31
32 201 indicate more subtle positive selection going on at a wider variety of HA sites in
33
34 202 mammalian cell culture or in the human host. In egg-culture, very strong positive
35
36 203 selection focuses on a rather small number of sites, of which N-glycosylation site 197-
37
38 204 199 is prominent.
39
40
41 205

42 206 **Discussion**

43
44 207
45
46
47 208 Detection of the actual N-glycosylation state of an HA protein requires laboratory
48
49 209 analysis that is not performed here. However, the assumption is made that the N-
50
51 210 glycosylation consensus sequence is a direct indicator of a true N-glycosylation site.
52

53 211
54
55 212 The N-glycosylation site at position 197-199 was chosen as it has previously been
56
57 213 described as suppressed in strains cultured in eggs, and also to be a site subject to
58
59 214 positive selection. A neighbouring N-glycosylation site at positions 233-235 was also
60

1
2
3 215 chosen as a control since it has not been implicated in either positive selection or
4
5 216 culture-dependent loss.
6

7 217

8
9 218 The sites detected in the present study as positively selected at $p > 0.95$: namely 75, 197
10
11 219 and 199, are exactly the same as those detected by Nunes et al. [2004; 2008] and
12
13 220 Pechirra et al. [2005], but differ to some extent from those of Shen et al. [2009]. These
14
15 221 latter authors point out that their study divides the data set into Yamagata and Victoria
16
17 222 lineages, and into time series windows. The residues identified as positively selected
18
19 223 differ in each data set as a result of the lineage-specific and time-variable selective
20
21 224 pressures.
22

23
24 225

25
26 226 The possibility cannot be excluded that the egg-free dataset contains some
27
28 227 contaminating egg-cultured sequences that have been incorrectly annotated. Therefore,
29
30 228 even the residual levels of positive selection seen in egg-free sequences (Table III) may
31
32 229 be misleadingly high. If a truly egg-free set were to be obtainable, it is predicted that
33
34 230 the signal of positive selection would be even weaker.
35

36 231

37
38 232 In summary, caution is necessary to avoid over-interpretation of positive selection
39
40 233 scores in influenza B HA sequences. Although positive selection has now been
41
42 234 demonstrated by several studies, it may be considerably weaker and less significant
43
44 235 once a correction is made for egg-culture. The correction made in the present study
45
46 236 may even be only a partial one, as it cannot be guaranteed that all strains without
47
48 237 annotation for egg culture or unknown culture conditions are truly egg-free in their
49
50 238 culture history. Recent papers on positive selection in influenza B have failed to address
51
52 239 this important issue. Really strong omega values (around 9 or 10) are only found in
53
54 240 egg-cultured strains. Inclusion of egg-cultured strains in an analysis may elevate overall
55
56 241 significance levels and omega values. Under such conditions it is easy to erroneously
57
58 242 attribute a high positive selection score to conditions in the host organism, particularly
59
60 243 immune evasion. Perhaps the biggest factor in apparent positive selection in influenza B

1
2
3 244 HA is adaptation to egg culture rather than to any host. It is suggested that future
4
5 245 analysis of positive selection should only be carried out on sequences free from culture
6
7 246 conditions of any kind, e.g. directly amplified by PCR from clinical material. Only then
8
9 247 will a true picture be seen of positive selection and its relation to host immune evasion in
10
11 248 influenza B.

12 249

13
14
15
16 250 **Acknowledgments**

17
18 251 The author thanks Prof. Duncan McGeoch (MRC Virology Unit, Glasgow) for sustained
19
20 252 encouragement and support over the last 6 years, and also Prof. Bill Carman and Dr Joy
21
22 253 Kean (West of Scotland Specialist Virology Centre, Glasgow) for pointing him in the
23
24 254 direction of influenza B.

25
26 255

27
28 256 Gambaryan AS, Robertson JS, Matrosovich MN. 1999. Effects of egg-adaptation on the
29 257 receptor-binding properties of human influenza A and B viruses. *Virology*
30 258 258(2):232-239.

31
32 259 Ikonen N, Pyhala R, Axelin T, Kleemola M, Korpela H. 2005. Reappearance of influenza
33 260 B/Victoria/2/87-lineage viruses: epidemic activity, genetic diversity and vaccination
34 261 efficacy in the Finnish Defence Forces. *Epidemiol Infect* 133(2):263-271.

35 262 Katoh K, Kuma K, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of
36 263 multiple sequence alignment. *Nucleic Acids Res* 33(2):511-518.

37 264 Kinnunen L, Ikonen N, Poyry T, Pyhala R. 1992. Evolution of influenza B/Victoria/2/87-like
38 265 viruses: occurrence of a genetically conserved virus under conditions of low epidemic
39 266 activity. *J Gen Virol* 73 (Pt 3):733-736.

40 267 Nunes B, Pechirra P, Canto e Castro L, Rebelo-de-Andrade H. 2004. Adaptive evolution on
41 268 HA1 subunit of influenza B virus. *International Congress Series* 1263:691-694.

42 269 Nunes B, Pechirra P, Coelho A, Ribeiro C, Arraiolos A, Rebelo-de-Andrade H. 2008.
43 270 Heterogeneous selective pressure acting on influenza B Victoria- and Yamagata-like
44 271 hemagglutinins. *J Mol Evol* 67(4):427-435.

45 272 Pechirra P, Nunes B, Coelho A, Ribeiro C, Goncalves P, Pedro S, Castro LC, Rebelo-de-
46 273 Andrade H. 2005. Molecular characterization of the HA gene of influenza type B
47 274 viruses. *J Med Virol* 77(4):541-549.

48 275 Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European Molecular Biology Open
49 276 Software Suite. *Trends Genet* 16(6):276-277.

50 277 Robertson JS, Naeve CW, Webster RG, Bootman JS, Newman R, Schild GC. 1985.
51 278 Alterations in the hemagglutinin associated with adaptation of influenza B virus to
52 279 growth in eggs. *Virology* 143(1):166-174.

53 280 Rota PA, Hemphill ML, Whistler T, Regnery HL, Kendal AP. 1992. Antigenic and genetic
54 281 characterization of the haemagglutinins of recent cocirculating strains of influenza B
55 282 virus. *J Gen Virol* 73 (Pt 10):2737-2742.

- 1
2
3 283 Rota PA, Wallis TR, Harmon MW, Rota JS, Kendal AP, Nerome K. 1990. Cocirculation of
4 284 two distinct evolutionary lineages of influenza type B virus since 1983. *Virology*
5 285 175(1):59-68.
6
7 286 Saito T, Nakaya Y, Suzuki T, Ito R, Saito H, Takao S, Sahara K, Odagiri T, Murata T, Usui
8 287 T, Suzuki Y, Tashiro M. 2004. Antigenic alteration of influenza B virus associated
9 288 with loss of a glycosylation site due to host-cell adaptation. *J Med Virol* 74(2):336-
10 289 343.
11 290 Shaw MW, Xu X, Li Y, Normand S, Ueki RT, Kunimoto GY, Hall H, Klimov A, Cox NJ,
12 291 Subbarao K. 2002. Reappearance and global spread of variants of influenza
13 292 B/Victoria/2/87 lineage viruses in the 2000-2001 and 2001-2002 seasons. *Virology*
14 293 303(1):1-8.
15 294 Shen J, Kirk BD, Ma J, Wang Q. 2009. Diversifying selective pressure on influenza B virus
16 295 hemagglutinin. *J Med Virol* 81(1):114-124.
17 296 Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood.
18 297 *Comput Appl Biosci* 13(5):555-556.
19 298 Yang Z, Wong WS, Nielsen R. 2005. Bayes empirical bayes inference of amino acid sites
20 299 under positive selection. *Mol Biol Evol* 22(4):1107-1118.
21 300
22 301
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60