Journal of General Virology (2012), **93**, 419–429 DOI 10.1099/vir.0.036574-0

The evolutionary rate of citrus tristeza virus ranks among the rates of the slowest RNA viruses

Gonçalo Silva, Natália Marques and Gustavo Nolasco

Plant Virology Laboratory, Center for Biodiversity, Functional and Integrative Genomics (BioFig), Universidade do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal

Correspondence
Gustavo Nolasco
gnolasco@ualg.pt

Received 19 July 2011
Accepted 5 November 2011

Citrus tristeza virus (CTV) has been studied intensively at the molecular level. However, knowledge regarding the dynamics of its evolution is practically non-existent. In the past, diverse authors have referred to CTV as a highly variable virus, implying rapid evolution. Others have, in recent times, referred to CTV as an exceptionally slowly evolving virus. In this work, we used the capsid protein (CP) gene to estimate the rate of evolution. This was obtained from a large set of heterochronous CP gene sequences using a Bayesian coalescent approach. The best-fitting evolutionary and population models pointed to an evolutionary rate of 1.58×10^{-4} nt per site year^{-1} (95 % highest posterior density, 1.73×10^{-5}–3.16×10^{-4} nt per site year^{-1}). For an unbiased comparison with other plant and animal viruses, the evolutionary rate of synonymous substitutions was considered. In a series of 88 synonymous evolutionary rates, ranging from 5.2×10^{-6} to 6.2×10^{-2} nt per site year^{-1}, CTV ranks in the 10th percentile, embedded among the slowest animal RNA viruses. At the time of citrus dissemination to Europe and the New World, the major clades that led to the current phylogenetic groups were already defined, which may explain the absence nowadays of geographical speciation.

INTRODUCTION

Members of the species *Citrus tristeza virus* (CTV) (genus *Closterovirus*, family *Closteroviridae*) cause one of the most economically damaging virus diseases in citrus worldwide. Since the first major epidemic of 'tristeza' that occurred in Argentina and Brazil about 80 years ago, huge economic losses have occurred in the most important citrus industries (Moreno *et al.*, 2008). The virus exists as multiple strains. Some, commonly referred to as mild strains, do not produce noticeable symptoms under field conditions. Others may have a devastating effect on trees grafted on sour orange; these may die after a short period of time, a situation known as the 'quick decline' or tristeza syndrome. Other strains produce the stem-pitting syndrome, characterized by the appearance of elongated pits in the branches of sweet oranges, grapefruit and mandarins, irrespective of the rootstock. Trees affected by this syndrome do not die, but become devalued economically due to low yield. A relationship between the clustering pattern of the capsid protein (CP) gene and symptoms induced has been established (Nolasco *et al.*, 2009). CTV has a single-stranded, positive-sense RNA molecule of about 19.3 kb encapsidated in flexuous filamentous particles about 2000 nm long (Bar-Joseph *et al.*, 1989) by two capsid proteins, a major 25 kDa (CP) covering about 95 % of the particle length and a minor 27 kDa capsid protein that covers only one extremity (Febres *et al.*, 1996). This large genome encodes 12 ORFs, potentially encoding at least 19 protein products (Karasev, 2000). The genome organization and expression mechanisms – reviewed by Dolja *et al.* (2006) – share similarities with those of the family *Coronaviridae*, a group of animal viruses.

It has been assumed that RNA viruses have a high evolutionary rate, in part due to the lack of proofreading activity of the RNA-dependent RNA polymerase (Drake & Holland, 1999). Within this concept frame, the early estimates of plant virus evolutionary rates appear contrasting, pointing to a higher genetic stability (García-Arenal *et al.*, 2003). For instance, it has been reported that tobamovirus isolates from herbarium specimens collected over a period of 100 years showed a remarkable genetic stability, pointing to an evolutionary rate of 1×10^{-8} nucleotide substitutions per site year^{-1} (Fraile *et al.*, 1997). A long-term evolutionary study based on the comparison of sequences of turnip yellow mosaic virus, which co-diverged along with its hosts over 12 000 years ago, points to a rate of about 10^{-6} nt per site year^{-1} for the CP gene (Blok *et al.*, 1987). More recently, short-term rates have been estimated through statistical inference methods by applying Bayesian coalescent concepts and software introduced by Drummond *et al.* (2003) to genomic sequences collected at different moments. Most nucleotide-substitution rates in animal RNA viruses are in the range of 10^{-2} to 10^{-5} nt per site year^{-1} (Hanada *et al.*, 2004; Jenkins *et al.*, 2002). Short-term evolutionary rates of plant viruses have been estimated for a few cases based on serial sampling, heterochronous sampling or node dating (reviewed by Gibbs *et al.*, 2010) and are in the order of

Table 1. Heterochronous dataset used in the Bayesian coalescent framework

Haplotype	Collection date	Geographical origin	GenBank accession no.
Q8_1	2003	Albania	EU660916
Q8_3	2003	Albania	EU660918
Q8_5	2003	Albania	EU660917
59M_33	2003	Albania	EU579396
F5	2004	Angola	DQ660346
O2	2004	Angola	DQ660348
O6	2004	Angola	DQ660349
O7	2004	Angola	DQ660350
P2	2004	Angola	DQ660351
P9	2004	Angola	DQ660352
Q4	2004	Angola	DQ660353
Q8	2004	Angola	DQ660354
Q14	2004	Angola	DQ660355
BaraoB_4	2003	Brazil	EU579363
BaraoB_6	2003	Brazil	EU579362
I141_55	2003	Brazil	EU579356
I141_59	2003	Brazil	EU579358
PeraIAC_31	2003	Brazil	EU579354
B274*	1991	Colombia	
399_38	2005	Croatia	
440_4	2005	Croatia	EU579415
440_6	2005	Croatia	EU579416
440_10	2005	Croatia	EU579417
443_4	2003	Croatia	AY791844
446_7	2003	Croatia	EU579424
L192GR	2010	Greece	HQ270182
M8GR	2009	Greece	HQ270183
CTV1	2000	Greece	HQ270184
B165*	1991	India	
CTV_B2	2007	India	EU869297
CTV_B3	2007	India	EU869298
CTV_B4	2007	India	EU869299
CTV_B5	2007	India	EU869300
K18	2010	India	HM853684
2_93	1995	Madeira Island	AF184116
2_98	1995	Madeira Island	AF184117
8_1	1998	Madeira Island	
8_2	1998	Madeira Island	
8_3	1998	Madeira Island	
8_6	2005	Madeira Island	
13A	1995	Madeira Island	
13C	1995	Madeira Island	AF184113
13_117	1995	Madeira Island	
15_118	1995	Madeira Island	AY660009
15a	1995	Madeira Island	
7_1	1998	Madeira Island	
7_2	1998	Madeira Island	
8a	1997	Madeira Island	
7_3	1998	Madeira Island	
10c	2000	Madeira Island	
10e	2000	Madeira Island	
15c	1995	Madeira Island	
67i	1998	Madeira Island	
67J	1998	Madeira Island	
67N	1998	Madeira Island	
AMJ31	2010	Malaysia	HM131219

Table 1. cont.

Haplotype	Collection date	Geographical origin	GenBank accession no.
398A	2000	Morocco	
398_2	2000	Morocco	EU579434
398_3	2000	Morocco	EU579433
398_4	2000	Morocco	EU579431
398_6	2000	Morocco	EU579432
P1_3	2003	Morocco	
MZ3_8	2003	Morocco	
MZ3_5	2003	Morocco	
108	2008	Pakistan	HQ329188
109	2008	Pakistan	HQ329189
135	2008	Pakistan	HQ329191
142	2008	Pakistan	HQ329192
143	2008	Pakistan	HQ329193
146	2008	Pakistan	HQ329194
159	2008	Pakistan	HQ329195
160	2008	Pakistan	HQ329196
179	2008	Pakistan	HQ329197
183	2008	Pakistan	HQ329198
28_121	1995	Portugal	
28C	1995	Portugal	AF184118
34c	1995	Portugal	
40d	1998	Portugal	
399_5	2000	Portugal	
QN_1	2005	Portugal	
3220_42	2005	Portugal	
225A	2000	Portugal	
Tr11_a	1995	Reunion Island	AY660010
124_18	2006	Reunion Island	
6_5	2004	S. Tomé e Príncipe	DQ660340
6_7	2004	S. Tomé e Príncipe	DQ660341
15_8	2004	S. Tomé e Príncipe	DQ660342
15_10	2004	S. Tomé e Príncipe	DQ660343
15_11	2004	S. Tomé e Príncipe	DQ660344
15_14	2004	S. Tomé e Príncipe	DQ660345
B7*	1991	South Africa	
19_121	1995	Spain	AF184114
19a	1995	Spain	
19b	1995	Spain	
20_2	1998	Spain	EU579429
25_120	1995	Spain	AF184115
P2_5	2003	Spain	
P2_8	2003	Spain	
R1_2	2003	Spain	
SYR_A2	2008	Syria	FN667582
T36	1990	USA	M76485
T3*	1991	USA	
T30	1991	USA	AF260651
CanTho1	2007	Vietnam	FN811556
DongThap3	2007	Vietnam	FN811557
B249*	1991	Venezuela	
065_17	2003	Venezuela	EU579366

*Sequences kindly provided by Dr C. L. Niblett, Venganza, Inc., Raleigh, NC, USA. The collection year may range from 1988 to 1991.

magnitude of 10^{-4} nt per site year^{-1}, thus falling within the range of animal RNA virus evolutionary rates. A new estimate of the tobamovirus evolutionary rate using Bayesian methods also falls within this range (Pagán *et al.*, 2010).

Although having been studied intensively at the molecular level, knowledge regarding CTV evolution is practically non-existent. In the past, diverse authors have referred to CTV as a highly variable virus. Later, Albiach-Martí *et al.* (2000) and Lbida *et al.* (2004) indicated that isolates with a common origin that were separated geographically for several decades depicted very similar sequences. In a recent work, the nucleotide changes were determined directly in the progeny of a CTV T36 isolate infectious clone, resulting in an evolutionary rate in the range 4×10^{5}–8×10^{-5} nt per site year^{-1} (Weng *et al.*, 2010).

For a more in-depth analysis of CTV evolutionary dynamics, we set up a long-lasting experiment in which isolates with characterized CP genes were maintained in a greenhouse and reanalysed later, over a period of 10 years, (i) for recombination events (to be presented in a forthcoming paper) and (ii) to obtain an estimate of the CP gene evolutionary rate. However, when analysed using a Bayesian coalescent framework, the presence of a temporal signal could not be demonstrated for these isolates. To overcome this, we used a set of 107 CP gene sequences from isolates collected worldwide between 1990 and 2010. With this dataset, it was possible to demonstrate the existence of a temporal signal and to estimate the evolutionary rate. The results obtained show that CTV is, to date, the most slowly evolving plant RNA virus.

RESULTS

Inference of evolutionary rates

The evolutionary rate and the time to the most recent common ancestor were estimated using Bayesian coalescent methods from the set of 107 sequences, which gathers the CP gene sequences obtained in this study and sequences retrieved from GenBank (Table 1), obtained at different moments over 20 years (between 1990 and 2010).

The best-fitting nucleotide-substitution model set was, according to MEGA5 software, the Kimura two-parameter (K80) model with five classes of gamma rate heterogeneity. Several other models derived from the more general Hasegawa–Kishino–Yano (HK85) model ranked down from the HK85 model with gamma rate heterogeneity in seventh place. Both K80 and HK85 models were implemented in 50 million generation runs in diverse clock, codon partition and demographic model arrangements and ordered according to their Bayes factor (BF) (Table 2). An additional 50 million generation runs were done for the first five arrangements and combined with the previous ones, increasing the effective sample size for all of the parameters to values >200. The best-fitting combination was implemented under two partitions (first plus second codon positions, and third codon position), a Bayesian skyline, uncorrelated lognormal clock, and K80 substitution model with five gamma rate heterogeneity classes. The mean rate of evolution for this model is 1.58×10^{-4} nt per site year^{-1} [95 % highest posterior density (HPD), 1.73×10^{-5}–3.16×10^{-4} nt per site year^{-1}] and the mean root height was 838 years (95 % HPD, 122–1450 years). Considering the BF table presented by Kass & Raftery (1995), there is very strong evidence ($2 \ln BF > 10$) against the models implemented with a single partition or under a strict clock, and strong evidence ($10 > 2 \ln BF > 6$) against the constant-size demographic model. There is positive evidence ($6 > 2 \ln BF > 2$) favouring five instead of four gamma rate classes. However, the best-fitting combination is not significantly different from the next sub-optimal combination, which was implemented with the HK85 instead of the K80 model. The mean rates of both estimates differ by <2 % and the upper tail of their 95 % HPD interval extends for less than one order of magnitude; the lower tail of the best-fitting combination extends for a little less than one order of magnitude. Although a 3 year incertitude could not be excluded for five of the haplotypes (Table 1), this did not alter the estimates, as confirmed by repeating the best-fitting Markov chain Monte Carlo (MCMC) runs.

To obtain an estimate free of purifying-selection effects that act upon the CP gene, the evolutionary rate of synonymous substitutions was estimated by the substitution rate of the partition constituting the third codon. The ratio of each partition's evolutionary rate to the joint evolutionary rate is given by the μ parameter in the BEAST software package, in this case 1.859 with a 95 % HPD ranging from 1.700 to 2.028. This implicates a synonymous evolutionary rate of 2.94×10^{-4} nt per site year^{-1} (95 % HPD, 2.69×10^{-4}–3.20×10^{-4} nt per site year^{-1}).

Assessment of temporal signal

The validity of the above results holds only if it can be proved that the evolutionary rates found result from temporal signal contained in the dataset and are not a result of random nucleotide changes due to the methodological approach. To assess this issue, the evolutionary rate of the real dataset was compared with the rates obtained from datasets where each sequence sampling date was shuffled randomly with the dates from the other sequences, as described by Ramsden *et al.* (2009) and Duffy & Holmes (2009). Five randomized datasets were constructed. To verify that any randomized dataset would not cause artefacts due to a particular combination of the reshuffled tip dates, the mean difference between the real and reshuffled tip dates was computed. The values obtained were very similar, respectively 5.5, 5.1, 6.0, 5.7 and 5.9 years for randomized datasets 1–5, indicating a similar degree of mixing. These evolutionary rates were estimated using the same conditions as for the best-fitting model. Loss of

Table 2. Summary of the results obtained with the diverse models, ranked according to BF

Abbreviations: UnclLN, relaxed molecular clock with lognormal distribution rates uncorrelated across branches; Strict, strict molecular clock; SK, Bayesian skyline; CS, constant size; K80, Kimura two-parameter; HK85, Hasegawa–Kishino–Yano; gamma classes, no. of gamma rate heterogeneity classes; HPD 95 %, values limiting the HPD interval at 95 %; ln BF, natural logarithm of the BF.

Clock model	Demographic model	Substitution model	Gamma classes	ln BF	Mean rate (nt per site year^{-1}) (HPD 95 %)	Root height (years) (HPD 95 %)
Two partitions (first + second codon positions and third codon position)						
UnclLN	SK	K80	5	0	1.58×10^{-4} (1.73×10^{-5}–3.16×10^{-4})	838 (122–1450)
UnclLN	SK	HK85	5	0.250	1.61×10^{-4} (1.24×10^{-5}–3.19×10^{-4})	550 (110–1274)
UnclLN	SK	HK85	4	1.465	1.62×10^{-4} (2.86×10^{-6}–3.17×10^{-4})	634 (120–1413)
UnclLN	SK	K80	4	1.899	1.56×10^{-4} (5.00×10^{-6}–3.17×10^{-4})	669 (115–1675)
UnclLN	CS	K80	5	3.255	3.01×10^{-4} (1.62×10^{-4}–4.41×10^{-4})	250 (124–411)
UnclLN	CS	K80	4	3.720	2.98×10^{-4} (1.57×10^{-4}–4.49×10^{-4})	253 (118–409)
UnclLN	CS	HK85	5	3.920	2.95×10^{-4} (1.69×10^{-4}–4.40×10^{-4})	255 (125–411)
UnclLN	CS	HK85	4	4.183	3.00×10^{-4} (1.58×10^{-4}–4.36×10^{-4})	251 (126–413)
Strict	SK	K80	5	15.345	9.32×10^{-5} (2.70×10^{-6}–1.92×10^{-4})	1213 (190–3684)
Strict	SK	HK85	5	15.520	9.65×10^{-5} (1.63×10^{-6}–1.93×10^{-4})	1190 (178–3454)
Strict	SK	HK85	4	16.220	1.04×10^{-4} (1.82×10^{-6}–2.05×10^{-4})	1006 (179–2413)
Strict	CS	K80	5	17.773	2.42×10^{-4} (1.32×10^{-4}–3.48×10^{-4})	275 (160–425)
Strict	CS	HK85	5	18.920	2.38×10^{-4} (1.23×10^{-4}–3.50×10^{-4})	281 (162–450)
Strict	CS	HK85	4	19.120	2.34×10^{-4} (1.22×10^{-4}–3.41×10^{-4})	284 (163–438)
One partition (first + second + third codon positions)						
UnclLN	SK	HK85	5	69.420	1.42×10^{-4} (1.75×10^{-5}–2.80×10^{-4})	689 (150–1590)
UnclLN	SK	HK85	4	70.320	1.56×10^{-4} (1.57×10^{-5}–3.08×10^{-4})	612 (119–1598)
UnclLN	CS	HK85	5	72.920	3.03×10^{-4} (1.62×10^{-4}–4.54×10^{-4})	253 (123–416)
UnclLN	CS	HK85	4	73.920	3.09×10^{-4} (1.60×10^{-4}–4.58×10^{-4})	248 (115–407)
Strict	SK	HK85	4	83.020	1.01×10^{-4} (1.01×10^{-6}–2.03×10^{-4})	1363 (156–3602)
Strict	SK	HK85	5	83.720	1.08×10^{-4} (1.04×10^{-6}–2.17×10^{-4})	1018 (143–2631)
Strict	CS	HK85	4	86.820	2.44×10^{-4} (1.32×10^{-4}–3.54×10^{-4})	274 (163–424)
Strict	CS	HK85	5	87.120	2.43×10^{-4} (1.31×10^{-4}–3.60×10^{-4})	279 (156–434)

temporal structure in the randomized datasets is confirmed by an increased extension of the lower tail, with values of the 95 % LPD ranging from 2.3×10^{-7} downward to 8.0×10^{-9} nt per site year^{-1} (Fig. 1). Evidence for temporal structure in the real dataset should appear as a significant difference between the mean rates of the real and randomized datasets. Although there is a partial superposition of the lower tails of the 95 % HPD intervals, the mean rate of the real data (1.58×10^{-4} nt per site year^{-1}) is outside the upper limit of the 95 % HPD of the randomized datasets, respectively 1.38, 1.35, 1.46, 1.36 and 0.84×10^{-4} nt per site year^{-1}. This result is considered significant for the existence of temporal structure according to Duffy & Holmes (2009) and Pagán *et al.* (2010).

Phylogenetic reconstruction

A maximum clade credibility tree (each node having the highest probability of occurrence among all of the trees) was constructed for the best-fitting model (Fig. 2). The major characteristics of the topology obtained are the same as presented previously (Nolasco *et al.*, 2009), where seven phylogenetic groups are well-defined, with posterior probabilities of 1. Two major branches separate at the root, leading to groups 1, 2 and M and groups 3a, 3b, 4 and 5. Most of the diversification occurring within the terminal groups probably started in the last 100–300 years, except for groups 1 and M, whose individualization appears more recent.

DISCUSSION

Temporal structure present in the dataset

Diverse causes can contribute to the loss of temporal signal in the dataset. As stated in the Introduction, the preliminary results (not shown) obtained directly from greenhouse-maintained isolates did not show enough temporal structure. This may have occurred due to the limited number of isolates, short time span (10 years) and/or artificial conditions of selection of virus variants. Other causes for disappearance of the temporal signal may include the use of inappropriate priors (Duffy & Holmes, 2009) or unnoticed recombination events (Lefeuvre *et al.*, 2011). Additionally, masking of the temporal signal may also result from random misincorporated nucleotides during reverse transcription, PCR or sequencing. These factors may be very difficult to control if sequences have been obtained in different conditions or laboratories and

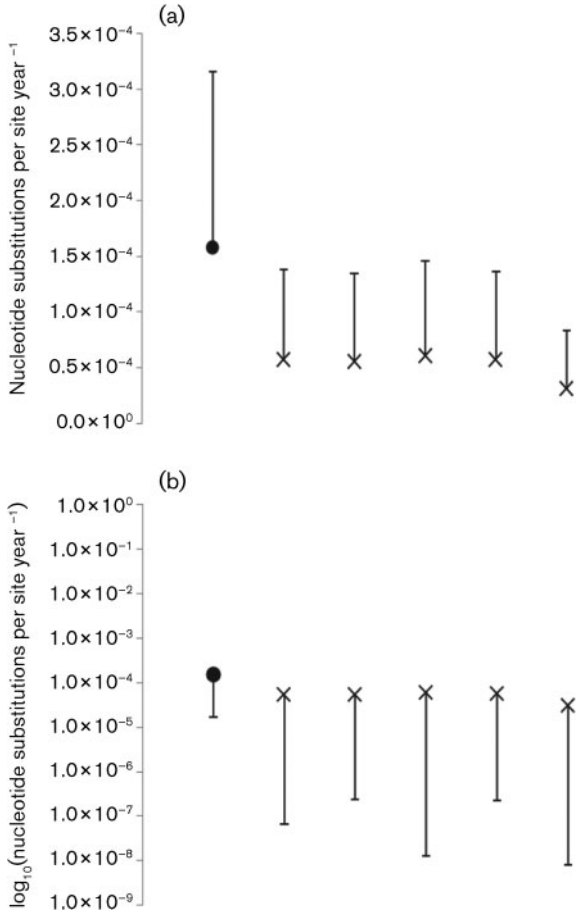

Fig. 1. Mean substitution rate and 95 % HPD intervals estimated from the real (●) and randomized (×) datasets using the same MCMC running conditions. (a) Upper tails of the 95 % HPD intervals plotted on a linear scale. (b) Lower tails of the 95 % HPD intervals plotted on a logarithmic scale.

gathered from GenBank, as for a large part of the dataset analysed in this work. This may be crucial if the sequences were collected over a short time span and the evolutionary rate is low, as discussed by Firth *et al.* (2010). Instead of trying to evaluate the effect of each factor individually, we globally assessed the existence of temporal structure in the dataset by comparing the results obtained with those from randomized datasets in which the temporal signal is inexistent, as suggested by Ramsden *et al.* (2009) and Duffy & Holmes (2009). As stated in the Results, significant evidence of temporal structure was found, according to the criteria advanced by Duffy & Holmes (2009) and Pagán *et al.* (2010).

Estimates of CTV evolutionary rate

Previously published studies have shown that haplotypes composing CTV isolates known to be geographically separated for several decades had minimal nucleotide changes. Those conclusions had been obtained from casual comparison among group M haplotypes (Albiach-Martí *et al.*, 2000) or group 5 haplotypes (Lbida *et al.*, 2004). In the present study, a preliminary attempt to estimate directly the evolutionary rate in greenhouse-maintained isolates failed due to the lack of temporal structure (results not shown). Besides the reasons discussed above, the lack of temporal signal may reside in the limitations of this kind of experiment: small number of varying sites, absence of bottlenecks associated with aphid transmission (Nolasco *et al.*, 2008) and/or short period for the purifying selection to have time to eliminate lethal mutations. Weng *et al.* (2010) took another direct approach by sequencing the progeny of an infectious clone that was sampled 1–6 years later. However, the values found, 4×10^{-5}–8×10^{-5} nt per site year^{-1}, appear to be too small for such a short period of sampling and these authors investigated the existence of temporal structure in their dataset.

To overcome the limitations of the direct estimation, a Bayesian coalescent approach based on heterochronous data was taken. For this, dated sequences retrieved from GenBank or from previous work were used, covering an interval of about 20 years.

The coalescent methods used assume a panmitic population free of recombination and selection. Besides absence of recombination in the dataset, which was verified previously, several studies (e.g. Nolasco *et al.*, 2009) support the notion that CTV, at least in the period sampled in this study, behaves as a panmitic population at the worldwide level. Existence of selection was investigated through FEL software implemented in the DataMonkey server (http://www.datamonkey.org). At the 0.05 significance level, out of 210 codons analysed, two codons were detected to be under positive selection and 45 were under negative selection. Although this constitutes a violation of the assumption of neutral selection, it is usually considered that the Bayesian MCMC analysis is sufficiently robust to withstand violations of neutrality; examples of successful analyses conducted in genes subjected to selection pressures are common (Duffy & Holmes, 2008; Simmons *et al.*, 2008; Wu *et al.*, 2011).

As seen in Table 2, strong evidence favouring two partitions, a relaxed molecular clock and a Bayesian skyline demographic model was obtained against one partition: the strict clock and constant-size population. The effect of the substitution model, K80 or HK85, was small. The mean evolutionary rate estimated under the best-fitting model (K80) was 1.58×10^{-4} nt per site year^{-1}. The upper and lower 95 % HPD tails extended for less than one order of magnitude in each direction, which can be considered a narrow distribution (Firth *et al.*, 2010). In comparison with the previous direct estimate obtained from the progeny of the infectious clone, the value obtained in the present work cannot be considered significantly different, as the lower tail of the 95 % HPD completely overlaps the interval of that estimate. It is, however, smaller than values of short-term rates obtained for other plant viruses (reviewed by Gibbs *et al.*, 2010).

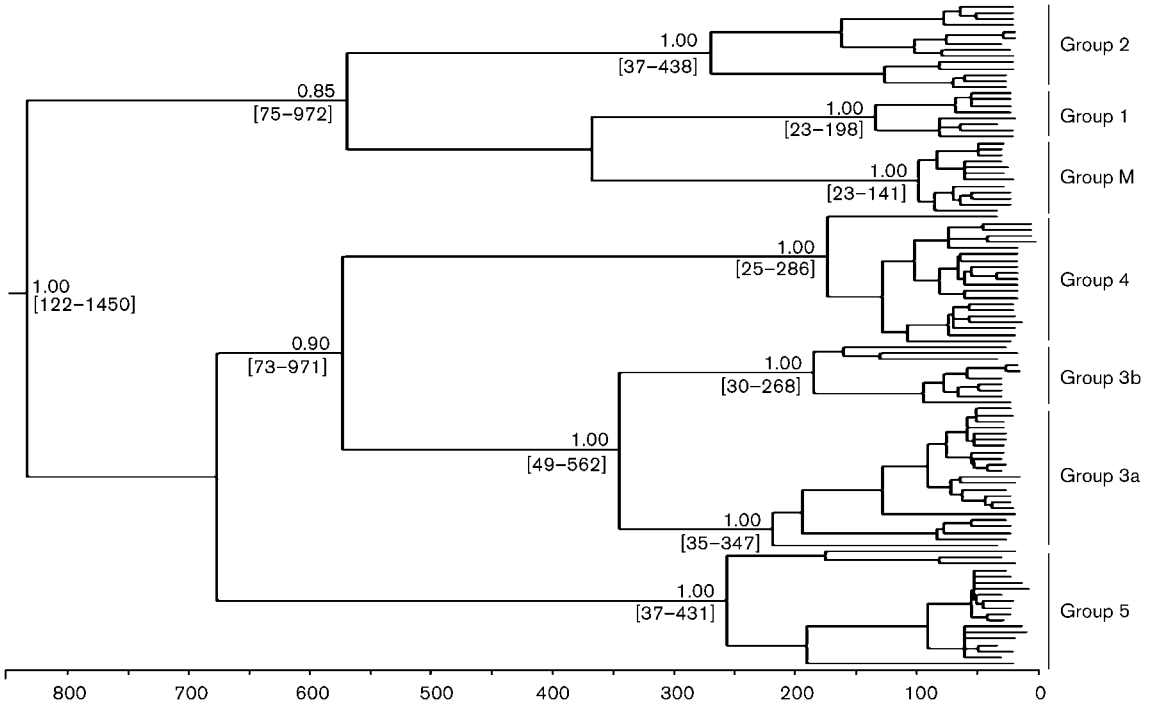

Fig. 2. Maximum clade credibility tree inferred from CP gene sequences of 107 CTV isolates collected between 1990 and 2010 under the evolutionary model with the highest BF. The mean node height (years) before the present is represented according to the horizontal scale at the bottom. For the major nodes, the 95 % HPD height interval (in brackets) and the posterior probability, if >0.85, are indicated. The seven phylogenetic groups are also indicated. Posterior probabilities and 95 % HPD intervals of nodes inside groups are not indicated, independent of their value.

Comparison with evolutionary rates of other RNA viruses

It is known that the capsid proteins of vector-borne plant viruses are subjected to a strong purifying selection (Chare & Holmes, 2004), a fact that has also been reported for CTV (Rubio *et al.*, 2001). For an unbiased comparison with the evolutionary rates of other RNA viruses, independent of constraints that might act at the protein level, Jenkins *et al.* (2002) and Hanada *et al.* (2004) considered the evolutionary rates of synonymous substitutions. The synonymous evolutionary rate for CTV was, in the present work, estimated to be 2.94×10^{-4} nt per site year^{-1}. Considering the data for animal viruses (Hanada *et al.*, 2004; Jenkins *et al.*, 2002) and the sole plant virus for which this rate was estimated under similar assumptions (rice yellow mottle virus; Fargette *et al.*, 2008) and eliminating those cases where this rate was not significantly different from zero (Jenkins *et al.*, 2002) or where new estimates changed the previous ones significantly (Ramsden *et al.*, 2008), a series of 88 synonymous evolutionary rates was gathered, ranging from 5.2×10^{-6} to 6.2×10^{-2} nt per site year^{-1}. In this series, CTV ranks in the 10th percentile (ranging from 9th to 11th percentile if the 95 % HPD interval is considered) counted from the most slowly evolving viruses, embedded among the slow animal RNA viruses, whereas the sole plant virus for which there is comparable data, rice yellow mottle virus (Fargette *et al.*, 2008), ranks in the 32nd percentile. Thus, although having a low evolutionary rate, CTV is not an exception when compared with RNA viruses, in contrast to what was suggested by Weng *et al.* (2010).

Possible mechanisms underlying the low evolutionary rate

A weak negative correlation has been found between RNA virus genome size and evolutionary rate (Jenkins *et al.*, 2002) or mutation rate (Sanjuán *et al.*, 2010). As such, the low evolutionary rate found in this work would not be surprising, considering that CTV has a 19.3 kb genome, which is the largest single RNA genome molecule besides those of members of the families *Coronaviridae* and *Roniviridae*, which range from 26 to 32 kb. Interestingly, the coronaviruses encode a 3′–5′ exoribonuclease (Minskaia *et al.*, 2006) that provides a proofreading activity in genome replication (Eckerle *et al.*, 2010). According to Gorbalenya *et al.* (2006), this would compensate for the increased probability of incorporating errors due to the large genome size. Homologues of this exoribonuclease are also present in members of the family *Roniviridae* (Gorbalenya *et al.*, 2006). Although the closteroviruses share diverse particular aspects of genome organization and replication with the coronaviruses (Dolja

et al., 2006), none of the conserved motifs characteristic of this family of exonucleases, referred to by Eckerle *et al.* (2010), could be predicted from the translation of the CTV genome (results not shown). This suggests that the errors expected to occur during replication in view of the large size of the genome of CTV shall be compensated in some other way. Referring to animal viruses, Drake & Holland (1999) considered that differences in the replication frequency should be the main source of variation of the synonymous evolutionary rate. French & Stenger (2003), while reviewing some arguments that had been advanced by Luria (1951), proposed that, for plant viruses, for which a proofreading activity has not been found, the best replicating strategy would be a slow, linear replication (the minus strand being used as the template for most progeny strands), also known as the stamp-machine model. In CTV, the ratio between genomic plus and minus strands has been estimated to lie between 10 : 1 and 50 : 1 (Navas-Castillo *et al.*, 1997; Satyanarayana *et al.*, 2002) 3–4 days after infection. These values are more likely to occur if the replication is, at that stage, linear instead of exponential [in that case, the ratio should approach 1 : 1 – see for instance Baltimore *et al.* (1966)]. Interestingly, the ratio of plus to minus strands is controlled by p23 (Satyanarayana *et al.*, 2002), an early protein in CTV infection (Navas-Castillo *et al.*, 1997), which is also implicated in the intracellular suppression of post-transcriptional gene silencing (Lu *et al.*, 2004). By reducing the relative amount of the minus strand, p23 impedes the virus from replicating in an exponential fashion, thus contributing to minimizing the replication errors.

Other factors may also operate at the host level, contributing to maintaining a slow rate. The virus restriction to access phloem tissues results in a lighter viral load that contributes to a lower effective population size. Additionally, the impaired ability to move from cell to cell limits the virus's colonizing ability to systemic movement (Folimonova *et al.*, 2008), which is known in diverse virus systems to originate strong genetic bottlenecks (Li & Roossinck, 2004; Sacristán *et al.*, 2003). Other kinds of mechanism may also be operating. Work developed on Borna disease virus (Formella *et al.*, 2000) showed that superinfection of cells by closely related variants leads to the elimination of the incoming variant, contributing to maintaining the virus's stability. Folimonova *et al.* (2010) have shown recently that superinfection exclusion was also working in citrus infection by CTV and proposed that this mechanism could shape CTV evolution by eliminating closely related variants that arise from replication errors.

Taken together, these data suggest that the low evolutionary rate of CTV is maintained through mechanisms that operate at diverse levels, which are not different from what has been found in other RNA viruses.

CTV evolution and the citriculture history

CTV is restricted to citrus, as well as to some other species from the family Rutaceae, and it is generally accepted that the virus co-evolved with the host species in South-East Asia (Bar-Joseph *et al.*, 1989). It has also been accepted that, in the past, citrus was propagated solely through seed; CTV that is not seed-transmissible would have relied singly on aphids for transmission before the onset of modern citriculture (Moreno *et al.*, 2008). It is, however, hard to understand how long-distance dissemination (across South-East Asia) could have occurred in ancient times based only on aphids, which are essentially local spreading agents. On the other hand, a report from 1936 on citriculture in South China (Condit *et al.*, 1936) indicates that seed propagation was reserved only for some kinds of citrus; air layering was used extensively to obtain rootstocks, and inarching has been used for citrus propagation for many centuries. It is thus conceivable that, before trade relationships were established with Europe, CTV might have been disseminated across South-East Asia and along the shores and islands of the Indian Ocean through vegetative-propagation materials. Establishing maritime trade routes between Europe and eastern Africa and Asia quickly led to the entrance of sweet oranges and other new kinds of citrus in Europe, about 400–500 years ago. This was quickly followed by citrus dissemination in the Americas. Lack of adequate citrus-cultivation practices, such as the ones used in China at that time (Webber *et al.*, 1967), or long-lasting overseas travels may have promoted the use of seeds as the prevailing propagation material, resulting in a delay of CTV entrance in the Mediterranean and probably in the Americas and Australia until the mid-19th century. During this period, a major separation had already occurred (95 % HPD, 122–1450years) between the two branches that led to groups 3a, 3b, 4 and 5 and to groups 1, 2 and M. Probably an even higher degree of speciation had already taken place before, as there is evidence that group M variants already existed as a group 100 years ago (Albiach-Martí *et al.*, 2000) and group 5 existed 80 years ago (Lbida *et al.*, 2004). Indirect evidence for higher speciation comes also from the generalized introduction of sour orange (*Citrus aurantium*) as rootstock in the second half of the 19th century as a consequence of phytophthora root rot (Webber *et al.*, 1967). At least isolates from groups 3a and 1 originate a quick decline followed by death of trees grafted on this rootstock and it is hard to conceive how these groups could have undergone their speciation after the introduction of sour orange. On the other hand, before the introduction of sour orange, the symptomatology induced by all CTV groups would appear essentially as mild in the frame of a non-sophisticated citriculture, thus not acting as a negative-selection factor by farmers. The use of grafting for citrus propagation started in the first half of the 19th century (Webber *et al.*, 1967) and became the definitive method for citrus propagation by the second half of the 19th century. Long-distance transport of vegetative materials for citrus propagation accompanied this tendency, resulting in the worldwide dissemination of CTV. Accepting that world-wide dissemination has occurred after major clade separation, coupled to the low evolutionary rate, may explain why the CTV population is panmitic nowadays.

It is interesting that branch separation at the root level occurs between a clade that comprises strains that cause, to diverse degrees, stem pitting in sweet orange, grapefruit and mandarins, i.e. groups 4, 5, 3a and 3b, grossly corresponding, in the grouping scheme defined by Hilf *et al.* (2005), to the VT and T3 groups; and a clade comprising strains that do not cause symptoms in these hosts (if not grafted on sour orange), i.e. groups 1, 2 and M, corresponding to the T30 and T36 groups, in the terminology of Hilf *et al.* (2005). However, the reasons for this are not clear.

Overall, the present study demonstrates that CTV lineages maintain a high genetic stability over time. These results enable the development of long-term control measures.

METHODS

Virus isolates. CTV isolates from different geographical origins and with different biological characteristics were gathered from field surveys or brought to the laboratory by participants in ring tests in previous studies. The infected material was frozen prior to molecular analysis.

Amplification and characterization of the CP gene. RNA from infected tissue was extracted using an RNeasy Plant Mini kit (Qiagen). cDNA synthesis was done using random primers. Briefly, 5 µl total RNA was mixed with 1 µl random primers [0.5 µg μl^{-1}, random $p(dN)_6$; Roche], denatured for 5 min at 95 °C and transferred quickly to ice. Reverse transcription was done for 1 h at 39 °C using SuperScript III reverse transcriptase (Invitrogen).

The entire CP gene was amplified from the cDNA with primers CTV1 and CTV10, described by Sequeira & Nolasco (2002). The PCR products were TA-cloned into the pGEM-T Easy vector (Promega) or into the pCRII vector (Invitrogen) according to the manufacturers' instructions. Several of the white colonies were picked and checked by PCR with primers CTV1 and CTV10 for the presence of the CP gene. PCR products that produced a single band of the right size were characterized further by single-strand conformation polymorphism (SSCP). Haplotypes that produced different SSCP patterns were selected and sequenced by Macrogen (Korea). Haplotypes isolated before 2000 may have been obtained by slightly different methods. In that case, sequences were obtained by classical radioactive labelling followed by electrophoresis and autoradiography.

Sequence analysis. Chromatograms were assembled with the help of a CTV CP reference sequence using the TraceEditPro (Ridom GmbH) software. For the greenhouse-maintained isolates, as well as for the CP gene sequences retrieved from GenBank, the first and final 21 nt (primer regions) were removed and the sequences were aligned. The search for recombination evidence amongst sequences was done using the RDP3 software package (Martin *et al.*, 2010). Recombinant sequences were excluded from the dataset.

Phylogenetic reconstruction using Bayesian MCMC. For this study were used: (i) the CP gene sequences obtained, together with (ii) CP gene sequences retrieved from GenBank and (iii) other CP gene sequences obtained from previous work in the laboratory. Only sequences for which the collection year was known were considered. Overall, 107 CP gene sequences of diverse worldwide haplotypes (Table 1) belonging to different phylogenetic groups and collected between 1990 and 2010 were gathered. The search for the best-fitting nucleotide-substitution model was done with the help of the Bayesian information criterion implemented in the MEGA5 software (Tamura *et al.*, 2011). Phylogenetic reconstruction was done using a Bayesian MCMC approach implemented in the BEAST package (Drummond & Rambaut, 2007). A uniform distribution ranging from 10^{-9} to 10^{-2} was used as a reduced-information prior for the mean evolution rate; it was considered that this range was large enough to encompass the evolutionary rates of most viruses, and that it could accommodate the lower tail expansion in the eventuality of the absence of temporal signal. Constant-size and Bayesian skyline (Drummond *et al.*, 2005) demographical models were combined with a strict or a relaxed molecular clock. In the latter case, the evolutionary rates among branches were uncorrelated and the rate in each branch was drawn independently from a lognormal distribution (Drummond *et al.*, 2006). The sequence dataset was analysed either as a single partition or as two partitions defined by the nucleotide positions in each codon (first plus second, and third nucleotide positions). The evolutionary rate of the third nucleotide position was used as the synonymous-substitution rate (Fargette *et al.*, 2008). The ratio of each partition's evolutionary rate to the joint evolutionary rate was given by the μ parameter in the BEAST software. The fitting of the diverse models considered to the dataset was compared through the BF (Suchard *et al.*, 2001), categorized according to Kass & Raftery (1995). The MCMC for each analysis was run for at least 50 million generations with sampling every 5000 generations, until the effective sampling size (ESS), after 10 % burn-in, reached a value >100; in some cases (see Results), the length of the chain was extended to 100 million generations by combining two 50 million generation runs in order to get a final ESS >200. To assess whether the dataset used contained temporal structure, five datasets were constructed, maintaining the sequences but randomizing the tip dates. These were analysed in BEAST using the same conditions as for the best-fitting model for the real data (Duffy & Holmes, 2009; Ramsden *et al.*, 2009). Evidence for time-structured data was considered significant if the mean evolutionary rate of the real dataset was outside the 95 % HPD interval of the randomized dataset's evolutionary rates (Duffy & Holmes, 2009).

ACKNOWLEDGEMENTS

This work was supported by the Foundation for Science and Technology (FCT), Portugal (projects PEst-OE/BIA/UI4046/2011 and PTDC/AGR-GPL/99512/2008). G. S. is the recipient of FCT grants SFRH/BD/62248/2009 and SFRH/BD/37310/2007.

REFERENCES

Albiach-Martí, M. R., Mawassi, M., Gowda, S., Satyanarayana, T., Hilf, M. E., Shanker, S., Almira, E. C., Vives, M. C., López, C. & other authors (2000). Sequences of *Citrus tristeza virus* separated in time and space are essentially identical. *J Virol* **74**, 6856–6865.

Baltimore, D., Girard, M. & Darnell, J. E. (1966). Aspects of the synthesis of poliovirus RNA and the formation of virus particles. *Virology* **29**, 179–189.

Bar-Joseph, M., Marcus, R. & Lee, R. F. (1989). The continuous challenge of citrus tristeza virus control. *Annu Rev Phytopathol* **27**, 291–316.

Blok, J., Mackenzie, A., Guy, P. & Gibbs, A. (1987). Nucleotide sequence comparisons of turnip yellow mosaic virus isolates from Australia and Europe. *Arch Virol* **97**, 283–295.

Chare, E. R. & Holmes, E. C. (2004). Selection pressures in the capsid genes of plant RNA viruses reflect mode of transmission. *J Gen Virol* **85**, 3149–3157.

Condit, I. J., Benemerito, A. N. & Chen, W. H. (1936). Citrus fruits and their culture in Kwangtung Province, South China. AgNIC (Agriculture Network Information Center). Riverside, CA: University of California. http://websites.lib.ucr.edu/agnic/condit.pdf

Dolja, V. V., Kreuze, J. F. & Valkonen, J. P. (2006). Comparative and functional genomics of closteroviruses. *Virus Res* **117**, 38–51.

Drake, J. W. & Holland, J. J. (1999). Mutation rates among RNA viruses. *Proc Natl Acad Sci U S A* **96**, 13910–13913.

Drummond, A. J. & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**, 214.

Drummond, A. J., Pybus, O. G., Rambaut, A., Forsberg, R. & Rodrigo, A. G. (2003). Measurably evolving populations. *Trends Ecol Evol* **18**, 481–488.

Drummond, A. J., Rambaut, A., Shapiro, B. & Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol Biol Evol* **22**, 1185–1192.

Drummond, A. J., Ho, S. Y., Phillips, M. J. & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biol* **4**, e88.

Duffy, S. & Holmes, E. C. (2008). Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus tomato yellow leaf curl virus. *J Virol* **82**, 957–965.

Duffy, S. & Holmes, E. C. (2009). Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. *J Gen Virol* **90**, 1539–1547.

Eckerle, L. D., Becker, M. M., Halpin, R. A., Li, K., Venter, E., Lu, X., Scherbakova, S., Graham, R. L., Baric, R. S. & other authors (2010). Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. *PLoS Pathog* **6**, e1000896.

Fargette, D., Pinel, A., Rakotomalala, M., Sangu, E., Traoré, O., Sérémé, D., Sorho, F., Issaka, S., Hébrard, E. & other authors (2008). Rice yellow mottle virus, an RNA plant virus, evolves as rapidly as most RNA animal viruses. *J Virol* **82**, 3584–3589.

Febres, V. J., Ashoulin, L., Mawassi, M., Frank, A., Bar-Joseph, M., Manjunath, K. L., Lee, R. F. & Niblett, C. L. (1996). The P27 protein is present at one end of citrus tristeza virus particles. *Phytopathology* **86**, 1331–1335.

Firth, C., Kitchen, A., Shapiro, B., Suchard, M. A., Holmes, E. C. & Rambaut, A. (2010). Using time-structured data to estimate evolutionary rates of double-stranded DNA viruses. *Mol Biol Evol* **27**, 2038–2051.

Folimonova, S. Y., Folimonov, A. S., Tatineni, S. & Dawson, W. O. (2008). Citrus tristeza virus: survival at the edge of the movement continuum. *J Virol* **82**, 6546–6556.

Folimonova, S. Y., Robertson, C. J., Shilts, T., Folimonov, A. S., Hilf, M. E., Garnsey, S. M. & Dawson, W. O. (2010). Infection with strains of *Citrus tristeza virus* does not exclude superinfection by other strains of the virus. *J Virol* **84**, 1314–1325.

Formella, S., Jehle, C., Sauder, C., Staeheli, P. & Schwemmle, M. (2000). Sequence variability of Borna disease virus: resistance to superinfection may contribute to high genome stability in persistently infected cells. *J Virol* **74**, 7878–7883.

Fraile, A., Escriu, F., Aranda, M. A., Malpica, J. M., Gibbs, A. J. & García-Arenal, F. (1997). A century of tobamovirus evolution in an Australian population of *Nicotiana glauca*. *J Virol* **71**, 8316–8320.

French, R. & Stenger, D. C. (2003). Evolution of wheat streak mosaic virus: dynamics of population growth within plants may explain limited variation. *Annu Rev Phytopathol* **41**, 199–214.

García-Arenal, F., Fraile, A. & Malpica, J. M. (2003). Variation and evolution of plant virus populations. *Int Microbiol* **6**, 225–232.

Gibbs, A. J., Fargette, D., García-Arenal, F. & Gibbs, M. J. (2010). Time – the emerging dimension of plant virus studies. *J Gen Virol* **91**, 13–22.

Gorbalenya, A. E., Enjuanes, L., Ziebuhr, J. & Snijder, E. J. (2006). *Nidovirales*: evolving the largest RNA virus genome. *Virus Res* **117**, 17–37.

Hanada, K., Suzuki, Y. & Gojobori, T. (2004). A large variation in the rates of synonymous substitution for RNA viruses and its relationship to a diversity of viral infection and transmission modes. *Mol Biol Evol* **21**, 1074–1080.

Hilf, M. E., Mavrodieva, V. A. & Garnsey, S. M. (2005). Genetic marker analysis of a global collection of isolates of *Citrus tristeza virus*: characterization and distribution of CTV genotypes and association with symptoms. *Phytopathology* **95**, 909–917.

Jenkins, G. M., Rambaut, A., Pybus, O. G. & Holmes, E. C. (2002). Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. *J Mol Evol* **54**, 156–165.

Karasev, A. V. (2000). Genetic diversity and evolution of closteroviruses. *Annu Rev Phytopathol* **38**, 293–324.

Kass, R. E. & Raftery, A. E. (1995). Bayes factors. *J American Statistic Assoc* **90**, 773–795.

Lbida, B., Fonseca, F., Santos, C., Zemzami, M., Bennani, A. & Nolasco, G. (2004). Genomic variability of *Citrus tristeza virus* (CTV) isolates introduced into Morocco. *Phytopathol Mediterr* **43**, 205–210.

Lefeuvre, P., Harkins, G. W., Lett, J. M., Briddon, R. W., Chase, M. W., Moury, B. & Martin, D. P. (2011). Evolutionary time-scale of the begomoviruses: evidence from integrated sequences in the *Nicotiana* genome. *PLoS One* **6**, e19193.

Li, H. & Roossinck, M. J. (2004). Genetic bottlenecks reduce population variation in an experimental RNA virus population. *J Virol* **78**, 10582–10587.

Lu, R., Folimonov, A., Shintaku, M., Li, W. X., Falk, B. W., Dawson, W. O. & Ding, S. W. (2004). Three distinct suppressors of RNA silencing encoded by a 20-kb viral RNA genome. *Proc Natl Acad Sci U S A* **101**, 15742–15747.

Luria, S. E. (1951). The frequency distribution of spontaneous bacteriophage mutants as evidence for the exponential rate of phage reproduction. *Cold Spring Harb Symp Quant Biol* **16**, 463–470.

Martin, D. P., Lemey, P., Lott, M., Moulton, V., Posada, D. & Lefeuvre, P. (2010). RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* **26**, 2462–2463.

Minskaia, E., Hertzig, T., Gorbalenya, A. E., Campanacci, V., Cambillau, C., Canard, B. & Ziebuhr, J. (2006). Discovery of an RNA virus 3′→5′ exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proc Natl Acad Sci U S A* **103**, 5108–5113.

Moreno, P., Ambrós, S., Albiach-Martí, M. R., Guerri, J. & Peña, L. (2008). *Citrus tristeza virus*: a pathogen that changed the course of the citrus industry. *Mol Plant Pathol* **9**, 251–268.

Navas-Castillo, J., Albiach-Martí, M. R., Gowda, S., Hilf, M. E., Garnsey, S. M. & Dawson, W. O. (1997). Kinetics of accumulation of citrus tristeza virus RNAs. *Virology* **228**, 92–97.

Nolasco, G., Fonseca, F. & Silva, G. (2008). Occurrence of genetic bottlenecks during citrus tristeza virus acquisition by *Toxoptera citricida* under field conditions. *Arch Virol* **153**, 259–271.

Nolasco, G., Santos, C., Silva, G. & Fonseca, F. (2009). Development of an asymmetric PCR-ELISA typing method for citrus tristeza virus based on the coat protein gene. *J Virol Methods* **155**, 97–108.

Pagán, I., Firth, C. & Holmes, E. C. (2010). Phylogenetic analysis reveals rapid evolutionary dynamics in the plant RNA virus genus tobamovirus. *J Mol Evol* **71**, 298–307.

Ramsden, C., Melo, F. L., Figueiredo, L. M., Holmes, E. C., Zanotto, P. M. & VGDN Consortium (2008). High rates of molecular evolution in hantaviruses. *Mol Biol Evol* **25**, 1488–1492.

Ramsden, C., Holmes, E. C. & Charleston, M. A. (2009). Hantavirus evolution in relation to its rodent and insectivore hosts: no evidence for codivergence. *Mol Biol Evol* **26**, 143–153.

Rubio, L., Ayllón, M. A., Kong, P., Fernández, A., Polek, M., Guerri, J., Moreno, P. & Falk, B. W. (2001). Genetic variation of *Citrus tristeza virus* isolates from California and Spain: evidence for mixed infections and recombination. *J Virol* **75**, 8054–8062.

Sacristán, S., Malpica, J. M., Fraile, A. & García-Arenal, F. (2003). Estimation of population bottlenecks during systemic movement of tobacco mosaic virus in tobacco plants. *J Virol* **77**, 9906–9911.

Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M. & Belshaw, R. (2010). Viral mutation rates. *J Virol* **84**, 9733–9748.

Satyanarayana, T., Gowda, S., Ayllón, M. A., Albiach-Martí, M. R., Rabindran, S. & Dawson, W. O. (2002). The p23 protein of citrus tristeza virus controls asymmetrical RNA accumulation. *J Virol* **76**, 473–483.

Sequeira, Z. & Nolasco, G. (2002). Bacterial expressed coat protein: development of a single antiserum for routine detection of *Citrus tristeza virus*. *Phytopathol Mediterr* **41**, 55–62.

Simmons, H. E., Holmes, E. C. & Stephenson, A. G. (2008). Rapid evolutionary dynamics of zucchini yellow mosaic virus. *J Gen Virol* **89**, 1081–1085.

Suchard, M. A., Weiss, R. E. & Sinsheimer, J. S. (2001). Bayesian selection of continuous-time Markov chain evolutionary models. *Mol Biol Evol* **18**, 1001–1013.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731–2739.

Webber, H. J., Reuther, W. & Lawton, H. W. (1967). History and development of the citrus industry. In *The Citrus Industry*, vol. I (History, World Distribution, Botany, and Varieties). Edited by W. Reuther, H. J. Webber & L. D. Batchelor. Riverside, CA: Division of Agricultural Sciences, University of California. http://websites.lib.ucr.edu/agnic/webber/Vol1/Chapter1.htm

Weng, Z., Liu, X., Gowda, S., Barthelson, R. A., Galbraith, D. W., Dawson, W. O. & Xiong, Z. (2010). Extreme genome stability of *Citrus tristeza virus*. In *Abstracts of the XVIII Conference of the IOCV, 7–12 November 2010, Campinas, SP, Brazil*, pp. 1–129. http://revistalaranja.centrodecitricultura.br/edicoes/down.php?idedicao=27&arquivo=v31_3_%2025.pdf

Wu, B., Blanchard-Letort, A., Liu, Y., Zhou, G., Wang, X. & Elena, S. F. (2011). Dynamics of molecular evolution and phylogeography of *Barley yellow dwarf virus*-PAV. *PLoS One* **6**, e16896.