Virology 498 (2016) 257–264

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Ongoing geographical spread of *Tomato yellow leaf curl virus* ☆

Batsirai Mabvakure [a], Darren P. Martin [b], Simona Kraberger [c,d], Leendert Cloete [a], Sharon van Brunschot [e,f], Andrew D.W. Geering [f], John E. Thomas [f], Kaveh Bananej [g], Jean-Michel Lett [h], Pierre Lefeuvre [h], Arvind Varsani [b,c,i,*], Gordon W. Harkins [a,**]

[a] South African National Bioinformatics Institute, University of the Western Cape, Bellville 7535, South Africa
[b] Department of Clinical Laboratory Sciences, University of Cape Town, Cape Town, South Africa
[c] School of Biological Sciences and Biomolecular Interaction Centre, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand
[d] Department of Microbiology, Immunology, and Pathology, Colorado State University, Fort Collins, CO, USA
[e] School of Biological Sciences, The University of Queensland, Brisbane, Queensland 4072, Australia
[f] The University of Queensland, Queensland Alliance for Agriculture and Food Innovation, Ecosciences Precinct, GPO Box 267, Brisbane, QLD 4001, Australia
[g] Department of Plant Virus Research, Iranian Research Institute of Plant Protection (IRIPP), Tehran 19395-1454, Iran
[h] CIRAD, UMR PVBMT, Pôle de Protection des Plantes, 7 chemin de l'IRAT, Saint-Pierre, Ile de la Réunion 97410, France
[i] The Center for Functional Microbiomics, The Biodesign Institute and School of Life sciences, Arizona State University, Tempe, AZ 85287, USA

ARTICLE INFO

Article history:
Received 11 July 2016
Returned to author for revisions
29 August 2016
Accepted 30 August 2016
Available online 15 September 2016

Keywords:
TYLCV
Phylogeography
Geminivirus
Tomato

ABSTRACT

Tomato yellow leaf curl virus (TYLCV) seriously impacts tomato production throughout tropical and sub-tropical regions of the world. It has a broad geographical distribution and continues to spread to new regions in the Indian and Pacific Oceans including Australia, New Caledonia and Mauritius. We undertook a temporally-scaled, phylogeographic analysis of all publicly available, full genome sequences of TYLCV, together with 70 new genome sequences from Australia, Iran and Mauritius. This revealed that whereas epidemics in Australia and China likely originated through multiple independent viral introductions from the East-Asian region around Japan and Korea, the New Caledonian epidemic was seeded by a variant from the Western Mediterranean region and the Mauritian epidemic by a variant from the neighbouring island of Reunion. Finally, we show that inter-continental scale movements of TYLCV to East Asia have, at least temporarily, ceased, whereas long-distance movements to the Americas and Australia are probably still ongoing.

1. Introduction

Tomato yellow leaf curl virus (TYLCV) is a monopartite begomovirus in the family *Geminiviridae* and is one of many closely related viruses that cause tomato yellow leaf curl disease (TYLCD) (Abhary et al., 2007; Navot et al., 1991). TYLCD was initially recognised in the Jordan Valley, Israel, in the 1930s, but it was not until the early 1960s that TYLCV was identified (Cohen and Nitzany, 1960, 1966). Subsequently, the virus has spread unabated into the Mediterranean basin and into most tropical and sub-tropical regions of the world and is recognised as one of the world's most devastating pathogens of tomato (Abhary et al., 2007; Delatte et al., 2007; Delatte et al., 2005; Diaz-Pendon et al., 2010; Duffy and Holmes, 2007; Kenyon et al., 2014; Lefeuvre et al., 2010; Moriones and Navas-Castillo, 2000; Péréfarres et al., 2012; Picó et al., 1996; Polston and Anderson, 1997; Stonor et al., 2003; Van Brunschot et al., 2010).

Although there are seven recognised strains of TYLCV (Brown et al., 2015), only two, the mild (Mld) and Israel (IL) strains, have ever been found outside of Iran. The global dissemination of TYLCV-Mld and TYLCV-IL from the Middle East or the Eastern Mediterranean (Duffy and Holmes, 2007; Lefeuvre et al., 2010) is attributed to the movement of infected planting material (Seal et al., 2006), together with spread of the Middle East-Asia Minor (MEAM1 formally referred to as the B biotype) and the Mediterranean (MED formally referred to as the Q biotype) cryptic species of its whitefly vector, *Bemisia tabaci* (Czosnek et al., 2002; Diaz-Pendon et al., 2010; Horowitz et al., 2007; Seal et al., 2006). Recent reports suggest that TYLCV is possibly unique amongst begomoviruses in that it is capable of both replicating within *B. tabaci* (Pakkianathan et al. (2015), as well as being seed

☆GenBank Accession #s: KX347094–KX347145, KX347155–KX347172.

* Corresponding author at: The Center for Functional Microbiomics, The Biodesign Institute and School of Life sciences, Arizona State University, Tempe, AZ 85287, USA.
** Corresponding author.
E-mail addresses: arvind.varsani@asu.edu (A. Varsani), gordon@sanbi.ac.za (G.W. Harkins).

http://dx.doi.org/10.1016/j.virol.2016.08.033
0042-6822/

transmitted in tomato (Kil et al., 2016). These characteristics may have contributed to it achieving a geographical range that is far broader than those of almost all other begomovirus species.

As with other begomoviruses, TYLCV is able to rapidly adapt to new environments as a consequence of its high rates of mutation and recombination (Delatte et al., 2005; Duffy and Holmes, 2007, 2008; Lefeuvre et al., 2010; Monci et al., 2002). For example, TYLCV-IL is a recombinant of TYLCV-Mld and *Tomato leaf curl Karnataka virus* (another tomato-infecting begomovirus), while other begomoviruses from the Mediterranean basin are recombinants of TYLCV-IL and TYLCV-Mld (Navas-Castillo et al., 2000).

The global spread of TYLCV began in the 1980s, after the emergence of the Mld and IL strains (Duffy and Holmes, 2008; Lefeuvre et al., 2010). The region centred on Iran harbours the highest diversity of TYLCV, although there has been little obvious movement of viruses out of this region since before the early 1980s (Lefeuvre et al., 2010). A previous phylogeographic study by Lefeuvre et al. (2010) included 91 coat protein and 82 full genome sequences of TYLCV, which had been generated over 22 years. However, this study was limited in geographical scope, as the virus isolates were primarily from the Mediterranean basin, the Middle East and the Americas, with Southeast Asia, the Pacific and Indian Ocean island nations/territories and Australia being greatly underrepresented. Furthermore, the analytical tools to account for the potentially confounding influences of recombination were not then available.

Here we analyse a much larger TYLCV sequence dataset comprising 414 full-genome sequences (70 of which are published here for the first time) sampled over 26 years from 33 countries to infer the historic global movement dynamics of TYLCV. Using fully probabilistic Bayesian modelling methods and accounting for recombination, we specifically focus on the contributions of south-western Pacific (Australia and New Caledonia) and south-eastern Indian Ocean (Mauritius and Reunion) states to the spread of TYLCV. Also, because of the intensified sampling for TYLCV over the past five years in various other parts of the world, we are also able to provide much more clarity on TYLCV movements into and across Asia, the Americas and the Caribbean.

2. Methods and materials

2.1. Sampling, TYLCV genome recovery and sequencing

Total DNA was extracted from tomato samples displaying leaf curl symptoms from Australia (n=52), Iran (n=12) and Mauritius (n=6). Circular DNA was enriched by rolling circle amplification (RCA) using Templiphi (GE Healthcare, USA). Unit length TYLCV genomes were recovered from the RCA concatemers using *Xmn*I, *Nco*I, *Bam*HI or *Sal*I restriction enzymes, and cloned into pJET 1.2 plasmid vector (ThermoFisher, USA) for *Xmn*I digested genomes and into pBluescript SK (Stratagen, USA) for *Nco*I, *Bam*HI or *Sal*I digested genomes. The recombinant plasmids were Sanger sequenced by primer walking at Macrogen Inc. (South Korea). Complete genome sequences were assembled using DNA Baser V4 (Heracle Biosoft S.R.L., Romania).

2.2. Construction of a recombination free dataset

A dataset of 435 full TYLCV genomes was assembled, which contained sequences of isolates sampled from 33 countries between 1988 and 2014 (Supplementary Table 1) including 356 full genome sequences retrieved from GenBank.

A preliminary multiple sequence alignment was generated using the slow, iterative refinement method (FFT-NS-I) implemented in MAFFT version 7 (Katoh and Standley, 2014). This alignment was then manually edited using IMPALE (available from http://web.cbio.uct.ac.za/~arjun/).

The resulting alignment was used for recombination analyses using the seven detection methods implemented in RDP version 4.36 (Martin et al., 2015) with default settings and a Bonferroni corrected *p*-value cut-off of 0.05. Events detected with three or more methods coupled with significant phylogenetic support were considered credible evidence of recombination. The breakpoint positions and recombinant sequence(s) inferred for every detected potential recombination event were manually checked and adjusted where necessary using the extensive phylogenetic and recombination signal analysis features available in RDP4.56 (Martin et al., 2015).

The final TYLCV recombination-free dataset (RF-dataset) comprised 414 TYLCV sequences, all generated following recombination analysis and the removal of (i) all tracts of sequence from the alignment that were detected to have been acquired through recombination (replaced in the alignment with gap characters), and (ii) 21 sequences from the TYLCV dataset that were inferred to have acquired $>30\%$ (or >810 nucleotides) of their genomes via recombination with non-TYLCV parental viruses.

2.3. Geographical clustering

Geographical clustering was done as described by Lefeuvre et al. (2010), using the centroid hierarchical clustering method (Rokach and Maimon, 2005) implemented in R (R Core Team, 2013) to determine the most appropriate regional grouping scheme for the phylogeographic analyses.

2.4. Identification of best-fit evolutionary models

The best-fit nucleotide substitution model was inferred using jModelTest (Posada, 2008) implemented in MEGA6 (Tamura et al., 2013) and the best-fit molecular clock, and demographic models were inferred using Path Sampling and Stepping stone methods with 100 path steps and a chain length of one million (Baele et al., 2012; Baele et al., 2013) using BEAST v1.8.1. (Drummond and Rambaut, 2007) and the BEAGLE high-performance library v2.1.2 (Ayres et al., 2012).

We used linear regression techniques available in TempEst (Rambaut et al., 2016) to visually examine the degree of divergence accumulation that had occurred over the sampling time interval as a proxy for temporal signal. This method explores the root-to-tip distances of the branches in the maximum likelihood tree as a function of sampling time. In this analysis, TempEst outputs the correlation coefficient and the coefficient of determination, for which higher values indicate strong temporal signal in the data, and improved fit of the data to the strict clock nucleotide substitution model, respectively.

2.5. Phylogeographic analyses

A discrete reversible diffusion model with the Bayesian stochastic search variable selection (BSSVS) procedure (Lemey et al., 2009), implemented in BEAST v1.8.1. (Drummond and Rambaut, 2007), was used to conduct Bayes factor (BF) tests that identified the statistically supported epidemiological links between the geographical regions considered (Lemey et al., 2009). Statistically supported links between locations were identified as those with an associated BF test statistic >5: where BF scores >100 were taken as representing decisive support for one or more movements between locations, BF scores >10 were taken as indicating strong support for movement(s), and BF scores <5 were taken as indicating negligible support (Kass and Raftery, 1995).

To determine whether the inference of the most probable root location (interpreted as the geographical origin of TYLCV) was biased towards locations/countries with the largest sample sizes, we compared the root probability results obtained with and without applying a tip swap location randomization procedure in BEAST v1.8.1 (Drummond and Rambaut, 2007).

The length of the Markov chains that were explored during these analyses had between 1×10^8 and 3×10^8 steps for the ten replicate runs of carried out with each model. When similar results were obtained with independent replicate runs of the chain for a particular model, the log and tree files were combined using LogCombiner (a computer program available in the BEAST v1.8.1 package; (Drummond and Rambaut, 2007). For all models, the runs were continued until effective sample size (ESS) values for all individual model parameters exceeded 200.

TreeAnnotator, which is also available as a component of the BEAST v1.8.1 package, was used to produce an annotated maximum clade credibility tree from the posterior distribution of trees produced during the MCMC analyses. This tree was visualized in FigTree v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/). The program SPREAD version 1.06 (Bielejec et al., 2011) was used to calculate BFs (Kass and Raftery, 1995) for potential TYLCV movements inferred using the discrete reversible diffusion model. SPREAD was also used to generate a key markup language (kml) file for the visualization of TYLCV movements using Google-Earth (Supplementary Data 1).

3. Results and discussion

3.1. Classification of geographical data into regions

Prior to analysing the movement dynamics of TYLCV, it was necessary to classify the 414 sequences in the RF-dataset based on their geographical origins. A hierarchical clustering method based on the geographical distances separating all these sequences indicated that they fell into twelve reasonably distinct geographical clusters that we named Africa (n=2), North & Central America (n=15), Australia (n=60), China (n=149), East Asia (n=64), Eastern Mediterranean (n=13), Western Mediterranean (n=17), Mauritius (n=6), Middle East (n=71), New Caledonia (n=6), Reunion Island (n=2) and Caribbean (n=9) (Fig. 1; Supplementary Table 1).

3.2. Estimation of best-fit evolutionary models and TYLCV nucleotide substitution rate estimates

The nucleotide substitution model that best fit the RF-dataset

	Country	Number
Africa	Egypt	2
Australia	Australia	60
China	China	149
East Asia	Korea	45
	Japan	19
Eastern Mediterranean	Israel	2
	Jordan	8
	Lebanon	2
	Turkey	1
Mauritius	Mauritius	6
Middle East	Iran	53
	Iraq	1
	Kuwait	2
	Oman	15
	Saudi Arabia	2
New Caledonia	New Caledonia	6
North & central America	Guatemala	1
	Mexico	6
	USA	8
Reunion Island	Reunion Island	2
Caribbean	Costa Rica	3
	Cuba	1
	Dominican Republic	1
	Grenada	2
	Puerto Rico	1
	Venezuela	1
Western Mediterranean	Italy	1
	Morocco	10
	Netherlands	1
	Portugal	1
	Spain	3
	Tunisia	1

	Root location probability	
Location	Normal	Randomised
Africa	0.0010	0.0010
Australia	0	0.0502
China	0	0.3883
East Asia	0	0.2968
Eastern Mediterranean	0.5500	0.2125
Mauritius	0.0020	0.0021
Middle East	0.4230	0.0171
New Caledonia	0.0010	0.0067
North & Central America	0	0.0093
Reunion	0.0010	0.0057
Caribbean	0.0010	0.0041
Western Mediterranean	0.0210	0.0062

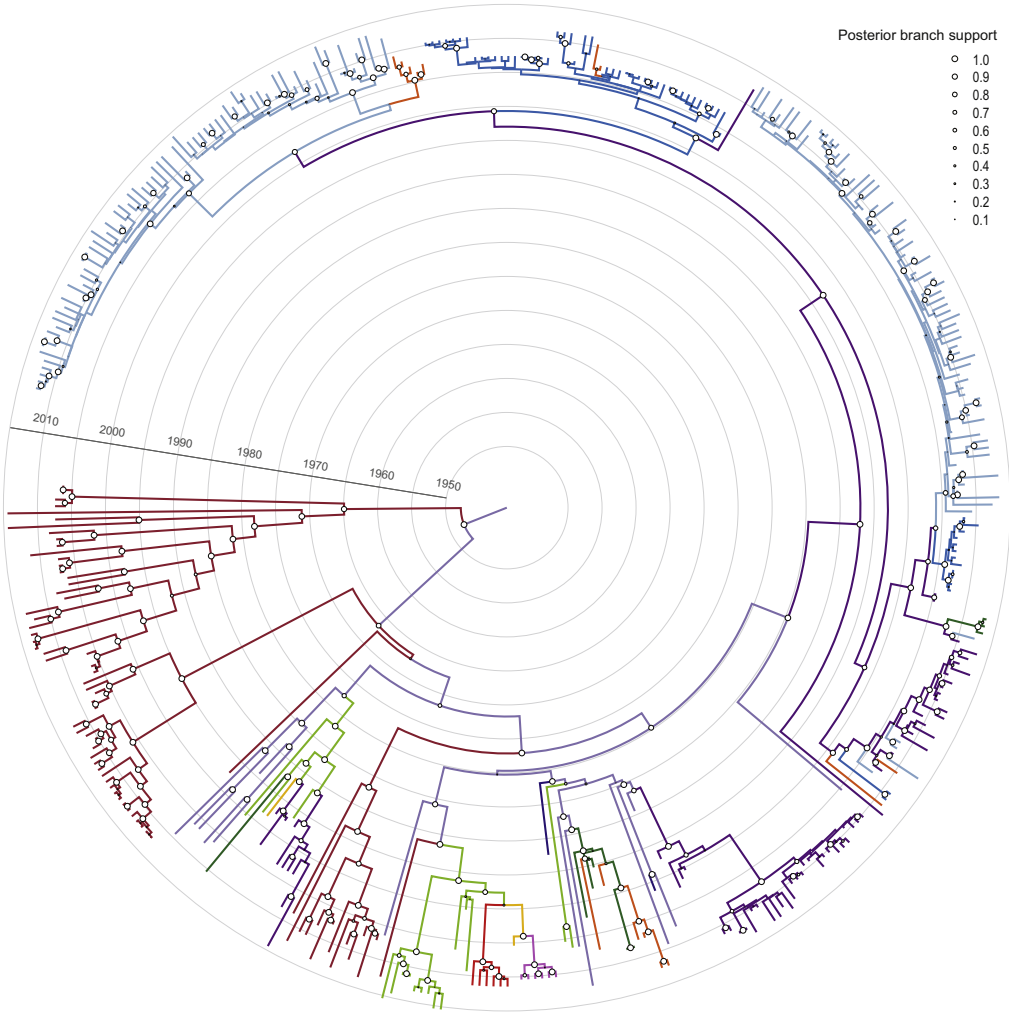

Fig. 1. Temporally scaled maximum clade credibility (MCC) tree constructed using 414 recombination-free TYLCV sequences. Branches are coloured according to where out of the twelve considered geographic regions the ancestral sequences represented by these branches most likely occurred. The posterior probability support of branches is represented by circles at the tip nodes of branches, with circle sizes being proportional to degrees of branch support. Also indicated are both the numbers of TYLCV samples from each region, and the probabilities of each region being the origin of the most recent common ancestor of all the sampled TYLCVs.

was the generalized time-reversible model with four gamma rate categories and a proportion of invariant sites (GTR+G4+I). The best fit molecular clock and demographic model combination, identified using marginal likelihood estimation (MLE) by the path sampling and stepping stone methods (Baele et al., 2012), was the uncorrelated lognormal relaxed molecular clock model with the Bayesian Gaussian Markov random field (GMRF) skygrid coalescent tree prior (Gill et al., 2013; Minin et al., 2008).

The correlation between root-to-tip divergence and sampling time for the RF-dataset, inferred using TempEst, yielded an r value of 0.2727, with a residual r^2 of 7.4374×10^{-2}, which indicated that although a strict molecular clock model was unlikely to fit our data very well, the data likely contained a detectable signal of sequence divergence throughout the sampling interval.

The mean TYLCV nucleotide substitution rate for the RF-dataset was determined to be 8.8929×10^{-4} (95% HPD 7.7457×10^{-4} to 9.9679×10^{-4}) subs/site/year, which is faster than previously reported rates for complete TYLCV genomes with recombination included (Duffy and Holmes, 2008; Lefeuvre et al., 2010; Yang et al., 2014), but is similar to those reported for a largely-recombination free TYLCV coat protein dataset (Lefeuvre et al., 2010).

The maximum clade credibility (MCC) tree for the RF-dataset (Fig. 1; Supplementary Fig. 1) indicated that the most recent common ancestor (MRCA) of the TYLCV isolates examined here occurred in the either the Eastern Mediterranean (p =0.55) or Middle East (p=0.42) around 1946 (95% HPD=1914–1971) (Fig. 1). It is unlikely that this inference is attributable to uneven sampling density among the locations, since the most probable locations of the MRCA inferred from MCC trees with randomized sampling locations were China, followed by East-Asia (Fig. 1). The modal location state estimates, indicated by the branch colours in the MCC tree, also reveal a reasonably strong spatial structure for this virus across its geographic range.

It must be stressed, however, that our conclusions regarding a Eastern Mediterranean or Middle-Eastern origin of the TYLCV MRCA have two associated caveats. First, discrete phylogeographic analyses such as we have performed are incapable of inferring any origin location outside of the regions from which the analysed sequences were sampled. Second, besides biases caused by uneven spatial sampling density, these analyses could have also been biased by uneven temporal sampling density across different locations. For example, it is possible that the fact that the oldest TYLCV sequences that we analysed were sampled from the Eastern Mediterranean may have unduly influenced the identification of this region as the most-probable location of the MRCA. However, it is noteworthy that there seemed to be no obvious over-all association between the temporal depth of sampling at particular locations and the inferred probability of those locations being the MRCA. Specifically, while the second most probable MRCA location, the Middle-East (p=0.423) had only a single sequence sampled prior to 2000, the two regions with the most samples collected prior to 2000, East Asia (n=5) and the Western Mediterranean (n=4), had associated posterior probabilities of being the MRCA location of 0.00 and 0.02, respectively.

3.3. Geographical dissemination of TYLCV

A total of 18 statistically supported epidemiological links were inferred across the twelve geographical regions that were

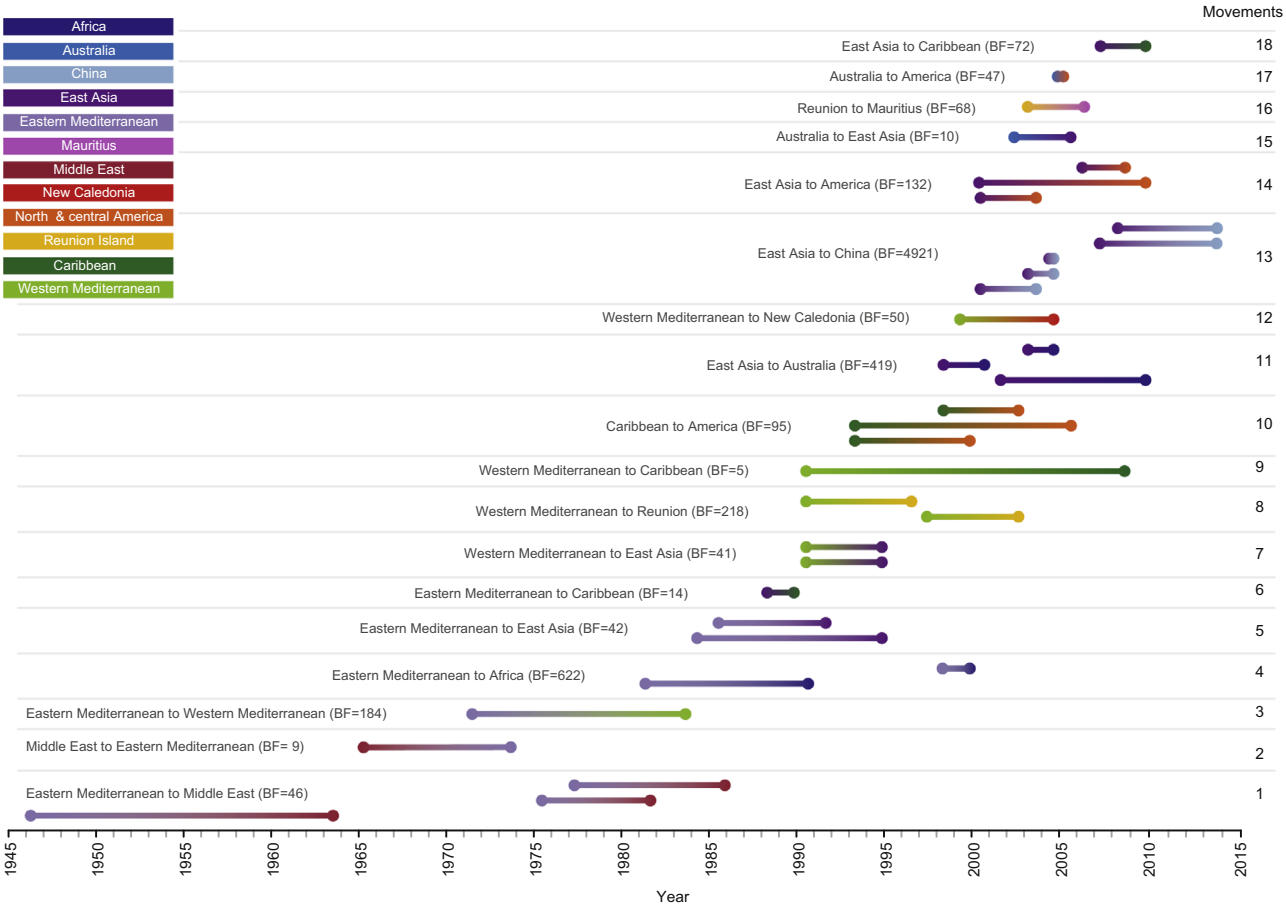

Fig. 2. Graph showing the timing of historic TYLCV movement events between the twelve analysed regions. Colours gradients across lines indicate potential movements of ancestral TYLCV variants from the region represented by the colour at the left of the line to the region represented by the colour on the right of the line.

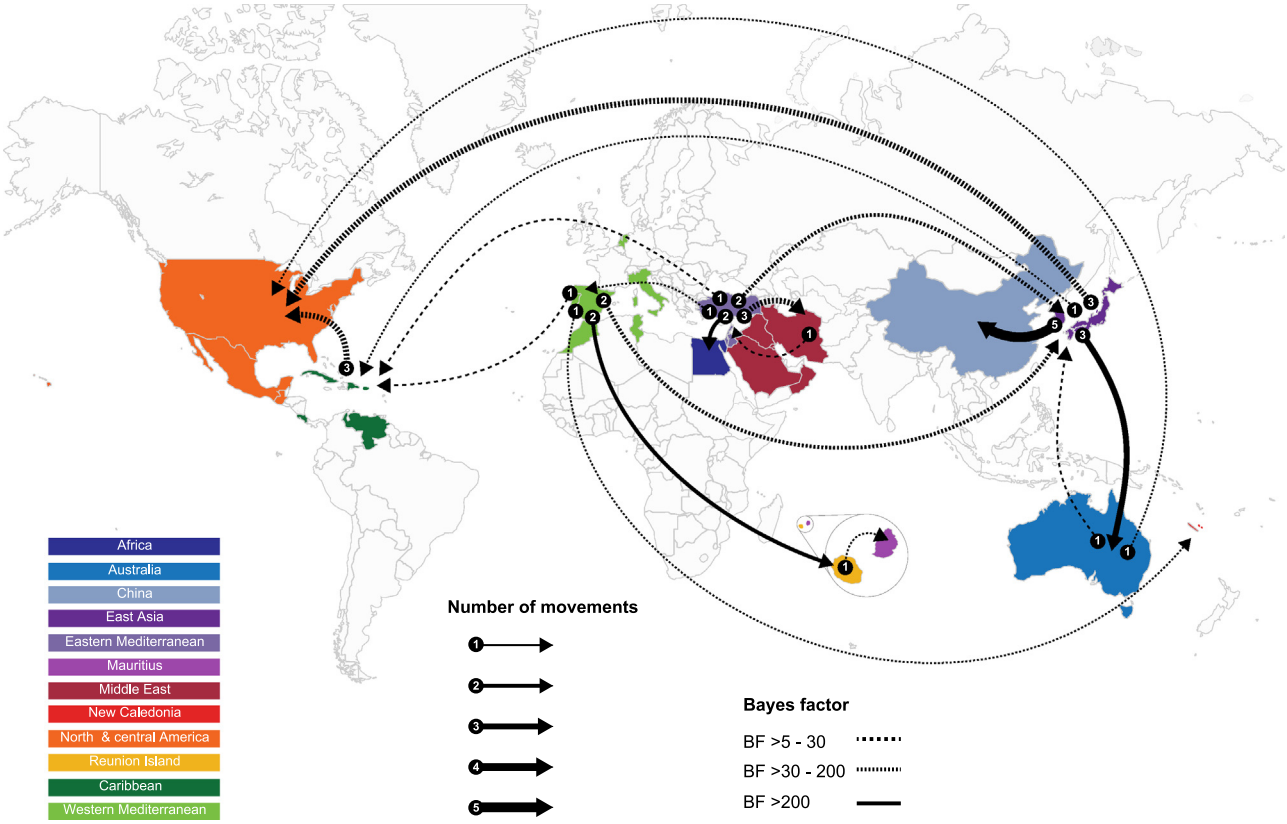

Fig. 3. A summary of the 18 statistically supported epidemiological links between the twelve geographic regions considered here. Dashed and solid lines represent the degree of Bayes factor support and the thickness of the lines represent the numbers of independent movements that were inferred to have occurred between locations.

considered (summarised in Figs. 2 and 3). The earliest detectable TYLCV dispersal event involved an unknown TYLCV strain moving from the Eastern Mediterranean to the Middle East between 1946 and 1964 (Figs. 1–3). This was followed by at least three more movements in the opposite direction between 1965 and 1974, 1971 and 1986 and 1988–2007. The estimated date of the existence of the MCRA of these isolates coincided with the first suspected cases of TYLCV in the Jordan Valley during the late 1950s to early 1960s (Cohen and Nitzany, 1960, 1966). This outbreak in the Eastern Mediterranean (Almusa, 1982; Makkouk, 1978; Makkouk et al., 1979) is thought to have represented the first opportunity for TYLCV to spread to the rest of the world (Lefeuvre et al., 2010).

Further movements from the Eastern Mediterranean included the dispersal of a TYLCV-Mld variant to the Western Mediterranean between 1971 and 1984 and two subsequent movements of the TYLCV-IL strain in the same direction between 1981 and 2001 and 1990 and 1995 (Figs. 1–3). While the first of these inferred movements is very close to the time of the first reports of TYLCV in the Western-Mediterranean around 1983, the timings of the other movements correspond to outbreaks in Italy and Spain between 1988 and 1992 (Credi et al., 1989; Gallitelli et al., 1991; Kheyr-Pour et al., 1991; Louro et al., 1996; Navas-Castillo et al., 1997; Noris et al., 1994).

Other inferred dispersal events of TYLCV-IL also correspond with actual observations, for example the arrival of the virus in Egypt during the early 1980s (Nakhla et al., 1993), in Japan in 1996, in China in 2006 and in South Korea in 2008 (Kenyon et al., 2014; Kim et al., 2011; Lee et al., 2010; Ueda et al., 2012).

Our analysis indicates that the Caribbean has possibly been a major portal for the introduction of TYLCV into the Western Hemisphere. TYLCV likely first entered the Caribbean from the Eastern Mediterranean between 1988 and 1991, a timeframe coinciding with serious TYLCV outbreaks in the Dominican Republic in the early 1990s (Czosnek and Laterrot, 1997; Polston et al., 1999; Salati et al., 2002). It was previously suggested that this introduction was via the Dominican Republic or Cuba, possibly in a shipment of tomato seedlings from Israel (Polston et al., 1999). In 1993, TYLCD decimated tomato production in the Dominican Republic (Polston et al., 1999) and from this focal point, the virus moved into Jamaica and Cuba (Mcglashan et al., 1994; Polston et al., 1999; Roye et al., 1999; Zubiaur et al., 2004).

From there, TYLCV quickly spread to the USA, where it was identified in Virginia, Tennessee and South Carolina in the mid-1990s (Polston et al., 1995), then in Florida (Polston and Anderson, 1997), Georgia (Momol et al., 1999), Mississippi (Ingram and Henn, 2001) and in Central America (Mexico) (Ascencio-Ibáñez et al., 1999; Banuelos-Hernandez et al., 2012; Brown and Idris, 2006). Other independent introductions of TYLCV into the Caribbean likely occurred when TYLCV-IL moved from East-Asia between 2006 and 2011 (Figs. 1–3), and TYLCV-Mld arrived there from the Western Mediterranean between 1990 and 2009 (Bird et al., 2001; Mcglashan et al., 1994; Roye et al., 1999; Zambrano et al., 2007).

Besides the introduction of TYLCV-IL into the Caribbean, other long distance movements of TYLCV from the Western Mediterranean were also detected in our analyses. One of these was to East Asia (between 1990 and 1995), preceding the first reports of TYLCV in Japan in 1996 (Kato et al., 1998; Kenyon et al., 2014), one to New Caledonia (between 1998 and 2006) (Péréfarres et al., 2012) and two to Reunion Island: a TYLCV-Mld movement between 1987 and 1997, coinciding with the first report of TYLCV there in 1997; (Peterschmitt et al., 1999) and a TYLCV-IL movement between 1998 and 2005 (Delatte et al., 2007). Although we were unable to obtain TYLCV samples from France (TYLCV is presently thought to have been eradicated there), it is plausible

that France was the Western Mediterranean country from which TYLCV entered Reunion Island and New Caledonia, since Reunion is a French overseas department and New Caledonia, a French territory (with both islands maintaining highly connected trade and transport links with metropolitan France). The importance of such links is highlighted by the fact that we also detected the movement of TYLCV-Mld from Reunion to the neighbouring island of Mauritius between 2003 and 2007 (Lobin et al., 2010).

TYLCV-IL was first reported in Australia in the peri-urban areas of Brisbane in 2006 and subsequently in the production areas of Bundaberg and Gatton (Van Brunschot et al., 2010). Consistent with this infection history, our analyses indicated that there were at least three separate introductions of TYLCV-IL from East Asia into Australia, two of which, occurred before this date: one between 1998 and 2001 from Japan to Brisbane, a second between 2003 and 2004 from Japan to Bundaberg, and a third between 2001 and 2009, which forms a monophyletic clade on the tree that is distinct from the two other monophyletic clades comprising the remaining Australian sequences from Brisbane and Bundaberg (Fig. 1).

We also inferred one movement of a Brisbane group TYLCV-IL virus from Australia to East Asia between 2000 and 2011 (most likely to Japan) and another to the USA (to either Hawaii or California) (Melzer et al., 2010; Rojas and Kon, 2007) between 2004 and 2005. These movements coincided with outbreaks of TYLCV in Arizona (Idris et al., 2007) and Texas (Isakeit et al., 2007), which were previously thought to have an independent, possibly East-Asian, origin to the TYLCV-IL variants found in the Caribbean (Duffy and Holmes, 2008; Lefeuvre et al., 2010).

We also detected two movements of TYLCV-IL from East Asia to the North American region (but also possibly directly to Hawaii), between 2006 and 2009 and between 2000 and 2010, which is consistent with the first reports of TYLCV-IL in Hawaii in 2009 (Melzer et al., 2010). An additional movement from East Asia, which further underlines the importance of this region as a major hub of global TYLCV dissemination, was to the Caribbean between 2006 and 2011.

Probably due to the proximity of Japan, South Korea and China, we inferred seven independent short range movements of TYLCV-IL from Japan/Korea into China between 2000 and 2012. These movements coincide with the first reported cases of TYLCV in China (around Shanghai) in 2006 (Yongping et al., 2008). Despite the rapid spread of TYLCV within China (Kenyon et al., 2014), there were no statistically supported movements of TYLCV from China to any of the other eleven regions studied. However, because the sequences from China make up ~36% of the total sample, and of these, 80% were collected between 2010 and 2014, it is possible that the absence of statistically supported movements out of China is in fact an artefact of the sampling scheme. Specifically, the relatively large proportion of samples originating from China increases the probability of detecting rare and infrequent relatively recent movements into this region.

4. Concluding remarks

While the phylogeographic analyses that we have performed broadly confirm the findings of similar analyses with smaller datasets, they substantially clarify the movement dynamics of TYLCV in the Western Hemisphere and Far East regions. We conclude that introductions of TYLCV into Australia and China have likely been from the East-Asia region, whereas the introduction to New Caledonia was likely from the Western Mediterranean. The fact that this association between New Caledonia and the Western Mediterranean is mirrored by that between the Western Mediterranean and Reunion Island, another region that is politically tied to France, suggests that metropolitan France might be the actual origin of TYLCV on these islands.

Regarding other inter-continental scale movements, we conclude that there have been at least nine independent introductions of TYLCV to the American/Caribbean region from the East-Asian, Australian, and Western Mediterranean regions, and at least five introductions of TYLCV to the East-Asian region from the Eastern Mediterranean and Western Mediterranean regions. Although there is no evidence of any movements to East Asia from the Mediterranean basin since 1995, at least five movements of TYLCV into the American/Caribbean region have occurred since the year 2000, suggesting that the flow of TYLCV variants into the Americas from elsewhere in the world is likely ongoing. The recent discovery that at least some TYLCV-IL variants are likely seed-transmissible (Kil et al., 2016) should be seriously investigated as a potential contributor to these movements. In this regard it would be of great interest to compare the seed transmission potential of all the main TYLCV lineages so as to determine whether a potentially causal association exists between the seed-transmissibility of particular lineages and their geographical ranges.

Crucially, the large numbers of inferred TYLCV movements over the past two decades, with multiple independent movements into and out of many of the regions analysed, strongly suggests that not enough is presently being done to control the ongoing spread of this major crop pathogen. Despite its already near cosmopolitan distribution, it is important that new containment strategies are implemented that account for the seed-transmissibility of TYLCV. Besides containing the current geographical range of TYLCV, such strategies will be crucial for impeding movements across this range of arising pathogenic and/or resistance-breaking variants of the virus.

Acknowledgements

We thank Michel Peterschmitt for sharing his TYLCV sequences from tomato samples collected in Morocco. BM, DPM, AV and GWH are supported by the National Research Foundation of South Africa (Grant No. TTK1207122745). SvB was supported by the Co-operative Research Centre for National Plant Biosecurity (CRC60070) and currently by the Cotton Research and Development Corporation, Australia (Grant No UQ1305). JML and PL are supported by the Région Réunion and the European Union (FEDER; Action # 1.03).

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.virol.2016.08.033. These data include Supplementary Figure 1, Supplementary Table 1 and Google earth animated kml showing the dispersal dynamics in real time (years).

References

Abhary, M., Patil, B., Fauquet, C.M., 2007. Molecular biodiversity, taxonomy, and nomenclature of tomato yellow leaf curl-like viruses. In: Czosnek, H. (Ed.), Tomato Yellow Leaf Curl Virus Disease, pp. 85–118.

Almusa, A., 1982. Incidence, economic importance, and control of tomato yellow leaf curl in Jordan. Plant Dis. 66, 561–563.

Ascencio-Ibáñez, J.T., Diaz-Plaza, R., Méndez-Lozano, J., Monsalve-Fonnegra, Z.I., Argüello-Astorga, G.R., Rivera-Bustamante, R., 1999. First report of tomato yellow leaf curl virus in Yucatan, Mexico. Plant Dis. 83, 1178.

Ayres, D.L., Darling, A., Zwickl, D.J., Beerli, P., Holder, M.T., Lewis, P.O., Huelsenbeck, J.P., Ronquist, F., Swofford, D.L., Cummings, M.P., Rambaut, A., Suchard, M.A., 2012. BEAGLE: an application programming interface and high-performance

computing library for statistical phylogenetics. Syst. Biol. 61, 170–173.
Baele, G., Li, W.L., Drummond, A.J., Suchard, M.A., Lemey, P., 2013. Accurate model selection of relaxed molecular clocks in bayesian phylogenetics. Mol. Biol. Evol. 30, 239–243.
Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M.A., Alekseyenko, A.V., 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. Mol. Biol. Evol. 29, 2157–2167.
Banuelos-Hernandez, B., Mauricio-Castillo, J.A., Cardenas-Conejo, Y., Guevara-Gonzalez, R.G., Arguello-Astorga, G.R., 2012. A new strain of tomato severe leaf curl virus and a unique variant of tomato yellow leaf curl virus from Mexico. Arch. Virol. 157, 1835–1841.
Bielejec, F., Rambaut, A., Suchard, M.A., Lemey, P., 2011. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. Bioinformatics 27, 2910–2912.
Bird, J., Idris, A.M., Rogan, D., Brown, J.K., 2001. Introduction of the Exotic Tomato yellow leaf curl virus-Israel in Tomato to Puerto Rico. Plant Dis. 85, 1028.
Brown, J.K., Idris, A.M., 2006. Introduction of the exotic monopartite tomato yellow leaf curl virus into west coast Mexico. Plant Dis. 90, 1360.
Brown, J.K., Zerbini, F.M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J. C., Fiallo-Olive, E., Briddon, R.W., Hernandez-Zepeda, C., Idris, A., Malathi, V.G., Martin, D.P., Rivera-Bustamante, R., Ueda, S., Varsani, A., 2015. Revision of Begomovirus taxonomy based on pairwise sequence comparisons. Arch. Virol. 160, 1593–1619.
Cohen, S., Nitzany, F.E., 1966. Transmission and host range of tomato yellow leaf curl virus. Phytopathology 56, 1127–1131.
Cohen, S., Nitzany, F.E., 1960. Curly Top Virus of Tomatoes: Its Identification and Mode of Transmission, Israeli Plant Protection and Inspection Services.
Credi, R., Betti, L., Canova, A., 1989. Association of a geminivirus with a severe disease of tomato in Sicily. Phytopathol. Mediterr. 25, 223–226.
Czosnek, H., Laterrot, H., 1997. A worldwide survey of tomato yellow leaf curl viruses. Arch. Virol. 142, 1391–1406.
Czosnek, H., Ghanim, M., Ghanim, M., 2002. The circulative pathway of begomoviruses in the whitefly vector Bemisia tabaci - insights from studies with Tomato yellow leaf curl virus. Ann. Appl. Biol. 140, 215–231.
Delatte, H., Holota, H., Naze, F., Peterschmitt, M., Reynaud, B., Lett, J.M., 2005. The presence of both recombinant and nonrecombinant strains of Tomato yellow leaf curl virus on tomato in Réunion Island. Plant Pathol. 54, 262.
Delatte, H., Holota, H., Moury, B., Reynaud, B., Lett, J.M., Peterschmitt, M., 2007. Evidence for a founder effect after introduction of Tomato yellow leaf curl virus-mild in an insular environment. J. Mol. Evol. 65, 112–118.
Diaz-Pendon, J.A., Canizares Mc Fau - Moriones, E., Moriones E Fau - Bejarano, E.R., Bejarano Er Fau - Czosnek, H., Czosnek H Fau - Navas-Castillo, J., Navas-Castillo, J., 2010. Tomato yellow leaf curl viruses: menage a trois between the virus complex, the plant and the whitefly vector. Mol. Plant Pathol. 11, 441–450.
Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC. Evolut. Biol. 7, 214.
Duffy, S., Holmes, E.C., 2007. Multiple introductions of the Old World begomovirus Tomato yellow leaf curl virus into the New World. Appl. Environ. Microbiol. 73, 7114–7117.
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.
Gallitelli, D., Luisoni, E., Martinelli, G.P., Caciagli, P., Milne, R.G., Accotto, G.P., Antignus, Y., 1991. Tomato yellow leaf curl disease in Sardinia. Inf. Fitopatol. 41, 42–46.
Gill, M.S., Lemey, P., Faria, N.R., Rambaut, A., Shapiro, B., Suchard, M.A., 2013. Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. Mol. Biol. Evol. 30, 713–724.
Horowitz, R., Denholm, I., Morin, S., 2007. Resistance to Insecticides in the TYLCV vector, Bemisia tabaci. In: Czosnek, H. (Ed.), Tomato Yellow Leaf Curl Virus Disease: Management, Molecular Biology, Breeding for Resistance. Springer, Netherlands, Dordrecht, pp. 305–325.
Idris, A.M., Guerrero, J.C., Brown, J.K., 2007. Two distinct isolates of tomato yellow leaf curl virus threaten tomato production in Arizona and Sonora, Mexico. Plant Dis. 91, 910.
Ingram, D.M., Henn, A., 2001. First report of tomato yellow leaf curl virus in Mississippi. Plant Dis., 85.
Isakeit, T., Idris, A.M., Sunter, G., Black, M.C., Brown, J.K., 2007. Tomato yellow leaf curl virus in tomato in texas, originating from transplant facilities. Plant Dis. 91, 466.
Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 90, 773–795.
Kato, K., Onuki, M., Fuji, S., Hanada, K., 1998. The first occurrence of Tomato yellow leaf curl virus in tomato (Lycopersicon esculentum Mill.) in Japan. Annu. Phyopathol. Soc. Jpn. 64, 552–559.
Katoh, K., Standley, D.M., 2014. MAFFT: iterative refinement and additional methods. Methods Mol. Biol. 1079, 131–146.
Kenyon, L., Tsai, W.S., Shih, S.L., Lee, L.M., 2014. Emergence and diversity of begomoviruses infecting solanaceous crops in East and Southeast Asia. Virus Res 186, 104–113.
Kheyr-Pour, A., Bendahmane, M., Matzeit, V., Accotto, G.P., Crespi, S., Gronenborn, B., 1991. Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monopartite geminivirus. Nucleic Acids Res. 19, 6763–6769.
Kil, E.J., Kim, S., Lee, Y.J., Byun, H.S., Park, J., Seo, H., Kim, C.S., Shim, J.K., Lee, J.H., Kim, J.K., Lee, K.Y., Choi, H.S., Lee, S., 2016. Tomato yellow leaf curl virus (TYLCV-IL): a seed-transmissible geminivirus in tomatoes. Sci. Rep. 6, 19013.
Kim, S.H., Oh, S., Oh, T.K., Park, J.S., Kim, S.C., Kim, S.H., Kim, Y.S., Hong, J.K., Sim, S.Y., Park, K.S., Lee, H.G., Kim, K.J., Choi, C.W., 2011. Genetic diversity of tomato-infecting Tomato yellow leaf curl virus (TYLCV) isolates in Korea. Virus Genes 42, 117–127.
Lee, H., Song, W., Kwak, H.R., Kim, J.D., Park, J., Auh, C.K., Kim, D.H., Lee, K.Y., Lee, S., Choi, H.S., 2010. Phylogenetic analysis and inflow route of Tomato yellow leaf curl virus (TYLCV) and Bemisia tabaci in Korea. Mol. Cells 30, 467–476.
Lefeuvre, P., Martin, D.P., Harkins, G., Lemey, P., Gray, A.J., Meredith, S., Lakay, F., Monjane, A., Lett, J.M., Varsani, A., Heydarnejad, J., 2010. The spread of tomato yellow leaf curl virus from the Middle East to the world. PLoS Pathog. 6, e1001164.
Lemey, P., Rambaut, A., Drummond, A.J., Suchard, M.A., 2009. Bayesian phylogeography finds its roots. PLoS Comput. Biol. 5, e1000520.
Lobin, K., Druffel, K.L., Pappu, H.R., Benimadhu, S.P., 2010. First report of tomato yellow leaf curl virus in tomato in Mauritius. Plant Dis. 94, 1261.
Louro, D., Noris, E., Veratti, F., Accotto, G.P., 1996. First report of tomato yellow leaf curl virus in Portugal. Plant Dis. 80, 1079.
Makkouk, K.M., 1978. Study on tomato viruses in Jordan valley with special emphasis on tomato yellow leaf curl. Plant Dis. Rep. 62, 259–262.
Makkouk, K.M., Shehab, S., Majdalani, S.E., 1979. Tomato yellow leaf curl – incidence, yield losses and transmission in Lebanon. Phytopathol. Z. 96, 263–267.
Martin, D.P., Murrell, B., Golden, M., Khoosal, A., Muhire, B., 2015. RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol. 1, vev003.
Mcglashan, D., Polston, J.E., Bois, D., 1994. Tomato yellow leaf curl geminivirus in Jamaica. Plant Dis. 78, 1219.
Melzer, M.J., Ogata, D.Y., Fukuda, S.K., Shimabuku, R., Borth, W.B., Sether, D.M., Hu, J. S., 2010. First Report of Tomato yellow leaf curl virus in Hawaii. Plant Dis. 94, 641.
Minin, V.N., Bloomquist, E.W., Suchard, M.A., 2008. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. Mol. Biol. Evol. 25, 1459–1471.
Momol, M.T., Simone, G.W., Dankers, W., Sprenkel, R.K., Olson, S.M., Momol, E.A., Polston, J.E., Hiebert, E., 1999. First report of tomato yellow leaf curl virus in tomato in South Georgia. Plant Dis. 83, 487.
Monci, F., Sanchez-Campos, S., Navas-Castillo, J., Moriones, E., 2002. A natural recombinant between the geminiviruses Tomato yellow leaf curl Sardinia virus and Tomato yellow leaf curl virus exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. Virology 303, 317–326.
Moriones, E., Navas-Castillo, J., 2000. Tomato yellow leaf curl virus, an emerging virus complex causing epidemics worldwide. Virus Res 71, 123–134.
Nakhla, M.K., Mazyad, H.M., Maxwell, D.P., 1993. Molecular characterization of four tomato yellow leaf curl virus isolates from Egypt and development of detection methods. Phytopathol. Mediterr. 32, 163–173.
Navas-Castillo, J., Sanchez-Campos, S., Diaz, J.A., Saez-Alonso, E., Moriones, E., 1997. First report of Tomato yellow leaf curl virus is in Spain: Coexistence of two different geminiviruses in the same epidemic outbreak. Plant Dis. 81, 1461.
Navas-Castillo, J., Sanchez-Campos, S., Noris, E., Louro, D., Accotto, G.P., Moriones, E., 2000. Natural recombination between Tomato yellow leaf curl virus-Is and Tomato leaf curl virus. J. Gen. Virol. 81, 2797–2801.
Navot, N., Pichersky, E., Zeidan, M., Zamir, D., Czosnek, H., 1991. Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a single genomic component. Virology 185, 151–161.
Noris, E., Hidalgo, E., Accotto, G.P., Moriones, E., 1994. High Similarity among the Tomato yellow leaf curl virus isolates from the west mediterranean basin – the nucleotide-sequence of an infectious clone from Spain. Arch. Virol. 135, 165–170.
Pakkianathan, B.C., Kontsedalov, S., Lebedev, G., Mahadav, A., Zeidan, M., Czosnek, H., Ghanim, M., 2015. Replication of Tomato yellow leaf curl virus in its whitefly vector, Bemisia tabaci. J. Virol. 89, 9791–9803.
Péréfarres, F., De Bruyn, A., Kraberger, S., Hoareau, M., Barjon, F., Lefeuvre, P., Pellegrin, F., Caplong, P., Varsani, A., Lett, J.M., 2012. Occurrence of the Israel strain of Tomato yellow leaf curl virus in New Caledonia and Loyalty Islands. New Dis. Rep. 25, 6.
Peterschmitt, M., Granier, M., Mekdoud, R., Dalmon, A., Gambin, O., Vayssieres, J.F., Reynaud, B., 1999. First report of tomato yellow leaf curl virus in Réunion Island. Plant Dis. 83, 303.
Picó, B., Díez, M.J., F., N, 1996. Viral diseases causing the greatest economic losses to the Tomato crop. II. The Tomato yellow leaf curl virus—a review. Sci. Hortic. 67, 151–196.
Polston, J.E., Anderson, P.K., 1997. The emergence of whitefly-transmitted geminiviruses in tomato in the western hemisphere. Plant Dis. 81, 1358–1369.
Polston, J.E., Keinath, A.P., Chellemi, D.O., 1995. Occurrence of Tomato Mottle Geminivirus in South-Carolina, Tennessee, and Virginia. Plant Dis. 79, 539.
Polston, J.E., McGovern, R.J., Brown, L.G., 1999. Introduction of tomato yellow leaf curl virus in Florida and implications for the spread of this and other geminiviruses of tomato. Plant Dis. 83, 984–988.
Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.
R Core Team, 2013. R: A Language and Environment for Statistical Computing.
Rambaut, A., Lam, T.T., Max Carvalho, L., Pybus, O.G., 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol. 2.
Rojas, M.R., Kon, T., 2007. First report of tomato yellow leaf curl virus associated with tomato yellow leaf curl disease in California. Plant Dis. 91, 1056.
Rokach, L., Maimon, O., 2005. Clustering methods. In: Maimon, O., Rokach, L. (Eds.), Data Mining and Knowledge Discovery Handbook. Springer, US, Boston, MA, pp. 321–352.
Roye, M.E., Wernecke, M.E., McLaughlin, W.A., Nakhla, M.K., Maxwell, D.P., 1999.

Tomato dwarf leaf curl virus, a new bipartite geminivirus associated with tomatoes and peppers in Jamaica and mixed infection with tomato yellow leaf curl virus. Plant Pathol. 48, 370–378.
Salati, R., Nahkla, M.K., Rojas, M.R., Guzman, P., Jaquez, J., Maxwell, D.P., Gilbertson, R.L., 2002. Tomato yellow leaf curl virus in the Dominican Republic: characterization of an infectious clone, virus monitoring in whiteflies, and identification of reservoir hosts. Phytopathology 92, 487–496.
Seal, S.E., vandenBosch, F., Jeger, M.J., 2006. Factors influencing begomovirus evolution and their increasing global significance: Implications for sustainable control. Crit. Rev. Plant Sci. 25, 23–46.
Stonor, J., Hart, P., Gunther, M., DeBarro, P., Rezaian, M.A., 2003. Tomato leaf curl geminivirus in Australia: occurrence, detection, sequence diversity and host range. Plant Pathol. 52, 379–388.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.
Ueda, S., Onuki, M., Yamashita, M., Yamato, Y., 2012. Pathogenicity and insect transmission of a begomovirus complex between tomato yellow leaf curl virus and Ageratum yellow vein betasatellite. Virus Genes 44, 338–344.
Van Brunschot, S.L., Persley, D.M., Geering, A.D.W., Campbell, P.R., Thomas, J.E., 2010. Tomato yellow leaf curl virus in Australia: distribution, detection and discovery of naturally occurring defective DNA molecules. Australas. Plant Pathol. 39, 412–423.
Yang, X.L., Zhou, M.N., Qian, Y.J., Xie, Y., Zhou, X.P., 2014. Molecular variability and evolution of a natural population of tomato yellow leaf curl virus in Shanghai. China J. Zhejiang Univ. Sci. B 15, 133–142.
Yongping, Z., Weimin, Z., Huimei, C., Yang, Q., Kun, S., Yanhui, W., Longying, Z., Li, Y., Zhang, H., 2008. Molecular identification and the complete nucleotide sequence of TYLCV isolate from Shanghai of China. Virus Genes 36, 547–551.
Zambrano, K., Carballo, O., Geraud, F., Chirinos, D., Fernandez, C., Marys, E., 2007. First report of Tomato yellow leaf curl virus in Venezuela. Plant Dis. 91, 768.
Zubiaur, Y.M., Fonseca, D., Quinones, M., Palenzuela, I., 2004. Presence of Tomato yellow leaf curl virus infecting squash (Curcubita pepo) in Cuba. Plant Dis. 88, 572.