Check for updates

Molecular evolution and phylogeographic analysis of wheat dwarf virus

Shiqing Wei[1], Linwen Liu[1], Guoliang Chen[1], Hui Yang[1], Liang Huang[2], Guoshu Gong[1], PeiGao Luo[1] and Min Zhang[1]*

[1]College of Agronomy, Sichuan Agricultural University, Chengdu, China, [2]State Key Laboratory for the Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China

Wheat dwarf virus (WDV) has caused considerable economic loss in the global production of grain crops. Knowledge of the evolutionary biology and population history of the pathogen remain poorly understood. We performed molecular evolution and worldwide phylodynamic analyses of the virus based on the genes in the protein-coding region of the entire viral genome. Our results showed that host-driven and geography-driven adaptation are major factors that affects the evolution of WDV. Bayesian phylogenetic analysis estimates that the average WDV substitution rate was 4.240×10^{-4} substitutions/site/year (95% credibility interval, $2.828 \times 10^{-4} - 5.723 \times 10^{-4}$), and the evolutionary rates of genes encoding proteins with virion-sense transcripts and genes encoding proteins with complementary-sense transcripts were different. The positively selected sites were detected in only two genes encoding proteins with complementary-sense, and WDV-barley are subject to stronger purifying selection than WDV-wheat. The time since the most recent common WDV ancestor was 1746 (95% credibility interval, 1517–1893) CE. Further analyses identified that the WDV-barley population and WDV-wheat population experienced dramatic expansion-decline episodes, and the expansion time of the WDV-barley population was earlier than that of the WDV-wheat population. Our phylogeographic analysis showed that the WDV population originating in Iran was subsequently introduced to Europe, and then spread from Eastern Europe to China.

KEYWORDS

wheat dwarf virus, viral population history, geography-driven adaptation, phylogeography, Bayesian tip-association significance, nature selection

Introduction

Wheat dwarf virus (WDV) was first described in the 1960s (Vacke, 1961) and to date it has been considered an important pathogen of economically important cereals, including wheat, barley and oats and widely scattered among European, African and Asian countries (Bisztray et al., 1989; Lindsten and Vacke, 1991; Bendahmane et al., 1995; Kvarnheden et al., 2002; Lemmetty and Huusela-Veistola, 2005; Köklü et al., 2007; Xie et al., 2007; Wang et al., 2008; Liu et al., 2012; Abt et al., 2020). WDV is transmitted via horizontal transmission among plants in a persistent, non-propagative manner by *Psammotettix alienus* (Lindblad and Areno, 2002). This virus invades the midgut of *P. alienus* through receptor mediated endocytosis and spreads through the digestive tract to haemolymph (Wang et al., 2014). However, *P. alienus* does not transmit virions to the eggs and nymphs of the first instar (Vacke, 1964). After

infection with WDV, the plants showed streaking, dwarfing of leaves, chlorosis and severe stunting of infected hosts, which resulted in serious crop yield reduction (Vacke, 1972). WDV causes an average yield loss of up to 35% in cereal crops globally, and in some cases up to 90% in localized epidemics (Vacke, 1964; Lindsten and Lindsten, 1999; Lindblad and Waern, 2002; Sirlova et al., 2005).

WDV is a member of the genus *Mastrevirus* in the family *Geminiviridae* (Schubert et al., 2007), which has a circular single-stranded DNA (ssDNA) genome of approximately 2.75 kb (Macdowell et al., 1985). This genome has four open reading frames (ORFs) and encodes four proteins, among which the virion-sense strand encodes movement protein (MP) and coat protein (CP), and the complementary-sense strand encodes two replication-associated proteins (RepA and Rep). When the virion of WDV circulates in the insect vector, the coat protein not only protects the viral genome, but also binds to ARF1 of *P. alienus* to help the virus spread from the gut to the haemolymph (Wang et al., 2019; Bennett and Agbandje-McKenna, 2020). In addition, coat protein facilitates nuclear and cell-to-cell transport of the genome via interaction with the MP and nuclear shuttle protein (NSP) (Macdowell et al., 1985). RepA binds the plant retinoblastoma-related protein pRBR to facilitate the expression of E2F transcription factors and activate genes encoding host DNA replication proteins and accessory factors (Hanley-Bowdoin et al., 2013; Sahu et al., 2013), and both RepA and Rep have RNA silencing suppressor activities (Liu et al., 2014). The WDV genome also contains a long intergenic region (LIR) and a short intergenic region (SIR), which are non-coding sequences that are important for the regulation of viral replication and gene expression (Hofer et al., 1992).

Phylogenetic studies of WDV genome sequences obtained from different host species confirmed that WDV can be divided into two main groups: WDV-wheat and WDV-barley strains (Lindsten and Vacke, 1991; Commandeur and Huth, 1999; Muhire et al., 2013; Schubert et al., 2014). Many experiments showed that WDV-wheat isolates can infect barley plants (Lindsten and Vacke, 1991; Commandeur and Huth, 1999; Kundu et al., 2009; Ramsell et al., 2009; Schubert et al., 2014), but there are many conflicting reports on whether WDV-barley isolates can infect wheat plants. Some field investigations and leafhopper-mediated transmission experiments have shown that the virion from barley plants could not infect wheat plants (Lindsten and Vacke, 1991; Tobias et al., 2011). However, WDV sequences belonging to those two strains were obtained from barley and wheat plants infected with WDV (Wang et al., 2008), and agricultural infecting clones of WDV-barley successfully infected wheat plants (Ramsell et al., 2009).

Although the mutation rate of DNA viruses is not as high as that of RNA viruses, DNA viruses have rapid replication dynamics and large virus populations. Genetic drift, natural selection and gene flow make the evolutionary characteristics and population genetic structure of DNA viruses are still more complicated (Zhan and McDonald, 2004). Stochastic changes in allele frequency in geographical populations lead to the random fixation of neutral alleles, thereby leading to non-adaptive differentiation (Gao et al., 2017). However, an increasing number of reports have shown that viruses have undergone obvious geographical adaptive differentiation via long evolutionary processes (Gao et al., 2017; Mao et al., 2019; Kawakubo et al., 2021). Natural selection is primarily expected to play a central role in the spatial population genetic structure of plant pathogens in agricultural ecosystems (Thrall et al., 2011). Recombination is also an important

evolutionary mechanism that produces diversity in most viruses (Pérez-Losada et al., 2015). Recombination is one of the important mechanisms driving WDV population diversity, and extensive intraspecific recombination generally occurs at the boundary of and within the genes encoding Rep and CP. Moreover, SIRs are also hot spots for recombination (Wu et al., 2008, 2015; Parizipouret et al., 2017).

Although there have been many phylogenetic and recombination analyses of WDV (Lindsten and Vacke, 1991; Commandeur and Huth, 1999; Wu et al., 2008; Muhire et al., 2013; Schubert et al., 2014; Wu et al., 2015; Parizipouret et al., 2017), there are few reports on population history and migration patterns on a given time scale. WDV is transmitted by vector insects and its distribution stretches across Eurasia. Whether there are geographical constraints on the evolution of WDV is unknown. In this study, we collected and analysed all of the sequence data from WDV isolates in GenBank to study the temporal dynamic, phylogeographic, and demographic history of WDV.

Materials and methods

Dataset

By October 2022, we had searched and obtained all complete genome sequences of each WDV gene from the GenBank database of the National Center for Biotechnology Information. Multiple alignment of nucleotide sequences was conducted with MAFFTv7 software (Katoh and Standley, 2013). To identify potential recombinant sequences of WDV, we calculated for signals of recombination using the RDP 4.95 suite (Martin et al., 2015) with seven software programs, namely, RDP, GENECONV, BOOTSCAN, Maximum Chi-Square (MAXCHI), CHIMAERA, 3SEQ and Sister Scanning (SISCAN). Only the events detected by at least four algorithms were accepted, with an associated p-value of 10^{-4}. For the accuracy of the results, recombinant isolates were removed and we eventually obtained 315 CP isolates, 299 MP isolates, 294 Rep isolates and 297 RepA isolates for phylogenetic analysis (Supplementary Table S1). Furthermore, we constructed a dataset containing 278 isolates of the whole–genome coding region of WDV.

Tests for temporal signals

We inferred the evolutionary timescale and substitution rate using a molecular clock calibrated by the sampling times of the sequences in which the temporal signal was evaluated by randomizing the sampling dates over clusters of tips and not over individual tips (Duchêne et al., 2015). The mean substitution rate estimated from the real sampling dates did not overlap with the 95% credibility intervals of rate estimates from 10 replicate datasets with cluster-permuted sampling dates. In addition, we evaluated the time structure of all datasets using newly developed BETS (Duchene et al., 2020). In this method, a heterogeneous model (M_{het}) and an isochronous model (M_{iso}) are introduced into the same dataset, and the marginal likelihood values estimated by generalized stepping stone sampling (Baele et al., 2016) are used to compare the fit of two competing data. The best fit model of the dataset was selected by Bayesian factor (BF;

Kass and Raftery, 1995), when (log) BF $\log[P(Y|M_{\text{het}})] - \log[P(Y|M_{\text{iso}})]$ is greater than 5, suggesting the presence of a sufficient temporal signal in the data.

Temporal dynamics of wheat dwarf virus

The substitution models (Supplementary Table S2) of the sequences were selected based on the Bayesian information criterion (BIC) using PartitionFinder (Lanfear et al., 2016) implemented in PhyloSuite (Zhang et al., 2020). Misspecification of the tree prior can lead to incorrect substitution rate (Müller et al., 2018). Therefore, marginal likelihood estimates based on path sampling (Baele et al., 2012) were used in BEAST 1.10 (Suchard et al., 2018) to find the best fit substitution tree prior (including the constant size, exponential growth, and Bayesian skyline coalescent) and best fit substitution clock model (both strict and relaxed clocks). All datasets were using uncorrelated lognormal relaxed clock and Bayesian skyline coalescent tree prior in the subsequent analysis (Supplementary Table S3). The Markov chain Monte Carlo (MCMC) analyses were run for 500 million generations, with sampling every 25,000 steps. Sufficient sampling was assessed by estimating the effective sample sizes values of all parameters and by inspecting traces implemented in Tracer 1.7 (Rambaut et al., 2018), and first 10% of the samples were discarded as burn-in.

Discrete phylogeographic analyses

An asymmetric substitution model with Bayesian random search variable selection options was implemented in BEAST based on the obtained optimal substitution rates for the whole–genome coding region (Supplementary Table S4), and Bayes factor (BF) were used to estimate significant diffusion rates between any two positioning states in 7 geographic regions, and the rates of viral migration across the discrete locations were calculated from the resulting log file in SpreaD3 (Bielejec et al., 2016). Significant migration pathways were determined based on the criteria of a BF > 3 and a mean indicator and mean indicator >0.5. The categories of rate support were as follows: decisively supported diffusion, BF > 1,000; very strongly supported diffusion, 1,000>BF ≥150; strongly supported diffusion, 150 >BF ≥ 10; and supported diffusion, 10 >BF ≥ 3 (Su et al., 2015).

Phylogeny-geography association and population structure analyses

A method accounting for phylogenetic uncertainty in investigating phylogeny-trait correlations with 1,000 random permutations of tip locations was implemented with BaTS 2.0 software to estimate the null distribution for each statistic (Parker et al., 2008); this was used to calculate three summary statistics (association index, *AI*; parsimony score, *PS*; and monophyletic clade, *MC*) from the posterior sample of trees produced with BEAST1.10 (Suchard et al., 2018) to identify the potential effects between the phylogeny and the geographic structure of WDV which the isolates only from wheat plant. The *MC* index assesses the association of geographic origin to phylogeny, and the *AI* and *PS* assess the association between trait and tree topology. High

MC scores and low *AI* index and *PS* values identified a strong phylogenetic trait association and low spatial admixture. In addition, we investigated the geographic regions of WDV using the discriminant analysis of principal components (DAPC) method independent of Hardy–Weinberg equilibrium and panmixia (Jombart et al., 2010).

Analysis of selection pressures

To investigate the difference in selection pressure between WDV-wheat and WDV-barley, we used the branch and site-specific models (Yang and Nielsen, 1998; Yang and Nielsen, 2002) of the CODEML algorithm (Yang, 2007) implemented in EasyCodeML (Gao et al., 2019). There are two models were used in the branch models, a ratio model assumed the ω-value for the entire tree, and a two-ratio model allowed the ω-value to vary between the two strains. For the site-specific models, three different sets of nested models (M3 vs. M0, M2a vs. M1a, and M8 vs. M7) were compared, and likelihood-ratio tests (LRTs) were applied to select the one that best fit the data. When the LRT was significant ($p < 0.01$), we used the Bayes empirical Bayes (BEB) method (Yang et al., 2005) to identify amino acid residues that four genes, CP, MP, Rep and RepA, are likely to have evolved under positive selection (posterior probability >0.95).

Results

The datasets containing all WDV-barley and WDV-wheat isolates and the datasets containing only WDV-wheat isolates passed the DRT, but the datasets containing only WDV-barley isolates did not pass the DRT (Supplementary Figure S1). Further BETS analysis results showed that the log-marginal likelihood value of M_{het} for all datasets was higher than the log-marginal likelihood value of M_{iso} (Supplementary Figure S1), indicating that that our datasets had sufficient time signals for reliable Bayesian tip analysis.

Evolutionary rates and timescales

Bayesian skyline coalescent tree priors and uncorrelated lognormal relaxed clocks provided the best fit to all of the datasets (Supplementary Table S3). We calculated the MRCA and evolutionary rate of the WDV CP, MP, Rep and RepA genes and WDV genome-wide coding region using five datasets. In addition, the rate of evolution and the MRCA of WDV-wheat and WDV-barley strains were also examined for each gene and for the genome-wide coding region. Our results showed that the evolutionary rate varied greatly among genes. The evolutionary rate of CP was 5.985×10^{-4} substitutions/site/year (95% credibility interval: 3.970×10^{-4}–8.243×10^{-4}), and the evolutionary rate of MP (1.309×10^{-3} substitutions/site/year, 95% credibility interval: 7.117×10^{-4}–1.993×10^{-3}) was more than twice that of CP (Supplementary Table S4). Although both the Rep gene and RepA gene (8.149×10^{-4} substitutions/site/year, 95% credibility interval: 5.937×10^{-4}–1.065×10^{-3}) encoded replication-related proteins, the evolutionary rate of RepA was significantly faster than that of Rep (Supplementary Table S4). The TMRCA obtained for different coding regions varied, and the TMRCA obtained from the whole–genome

coding region of WDV was 1746 CE (95% credibility interval: 1517 CE-1893 CE, Supplementary Table S4). The TMRCA obtained for CP, MP, Rep and RepA was 1837 CE (95% credibility interval: 1662–1939), 1885 CE (95% credibility interval: 1746–1980), 1897 CE (95% credibility interval: 1781–1969) and 1921 CE (95% credibility interval: 1831–1974), respectively (Supplementary Table S4). In addition, to assess the differences in evolutionary rates between WDV-wheat and WDV-barley, the existing 5 datasets were divided again according to strains to obtain a total of 10 datasets. These datasets were calibrated according to the evolutionary rates obtained from the datasets that passed the DRT, and the evolutionary rate of WDV-wheat was faster than that of WDV-barley, as measured by CP, Rep, RepA and the whole–genome coding region of WDV (Supplementary Table S4). Surprisingly, the evolutionary rate of MP in WDV-barley was significantly faster than that in WDV-wheat. However, the results obtained from the datasets of either coding region showed that the barley strain of WDV had an earlier TMRCA than the wheat strain.

The maximum clade credibility (MCC) trees inferred from the whole–genome coding region and the CP datasets shared very similar topologies, and all isolates could be separated into two major clades WDV-wheat and WDV-barley (Figures 1A,B). In both wheat and barley strains, isolates from Iran were clustered on one branch. Our Bayesian analysis placed the root of the trees for Iranian isolates with

a higher posterior probability than for isolates from other regions, which was supported by two datasets: one dataset containing the WDV coat protein coding region and another dataset containing the WDV whole–genome coding region (Figure 1).

Global migration pattern of wheat dwarf virus

Ten migration pathways within the spatial diffusion of WDV were supported by our Bayesian phylogeographic analysis based on the whole–genome coding region datasets, eight of which were supported by WDV-wheat dataset (Figure 2; Supplementary Table S5). There are four migration pathways from the middle reaches of the Yellow River (Shanxi and Shaanxi provinces) to Eastern Europe (Hungary, Poland, and Ukraine), Eastern China (Hebei and Shandong provinces), Northwest China (Gansu, Ningxia, Qinghai and Xinjiang provinces) and Southwest China (Sichuan, Yunnan and Guizhou provinces) supported by WDV-wheat dataset. There was also a migration pathway from China and from Eastern China to Northwest China supported by WDV-wheat dataset. There were two migration pathways from Iran spread to Western Europe (France, Germany, Spain, and UK) and Eastern Europe, and one migration pathway from Eastern Europe to the middle reaches of the Yellow River supported by WDV-wheat dataset (Figure 2; Supplementary Table S5). Moreover,
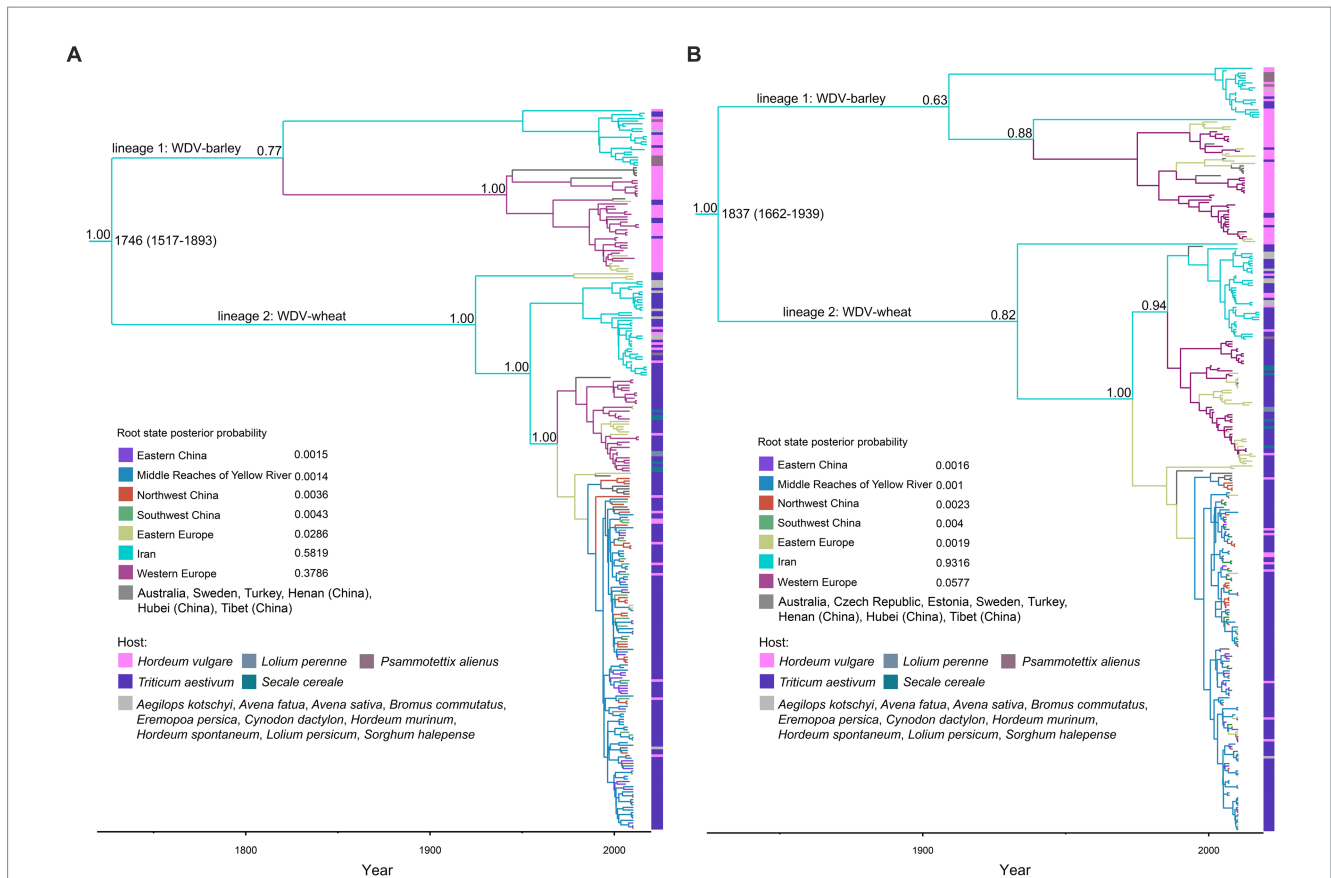
FIGURE 1
Time-scaled maximum clade credibility tree inferred from the whole–genome coding region (A) and coat protein (B) sequences of wheat dwarf virus. The tree topologies have been chosen to maximize the product of node posterior probabilities. Branch lengths are scaled according to time, as shown by the horizontal axis. Branch colors represent inferred location states. The inset panel is the root state posterior probabilities of the geographic regions.

FIGURE 2
The spatial dynamics analysis of wheat dwarf virus inferred from whole–genome coding region and coat protein. **(A)** Supported global spatial diffusion pathways and **(B)** histogram of the total number of location-state transitions inferred from whole genome coding region. ECh, Eastern China; EEu, Eastern Europe; MRY, Middle Reaches of Yellow River; NCh, Northwest China; SCh, Southwest China; WEu, Western Europe.

there were two migration pathways from Iran and Western Europe, which spread to Western Europe and Eastern Europe, respectively, supported by WDV-barley dataset (Figure 2; Supplementary Table S5).

Since the dataset of CP gene contained more isolates from Hungary, we further analysed the global migration pathways of WDV using CP gene. Not surprisingly, the results that we obtained were similar to those obtained from the whole–genome coding region of WDV, except for the migration pathway from Eastern China to Northwest China (Supplementary Table S5).

The inferred spatial dynamics of WDV suggest that Western Europe and the middle reaches of the Yellow River have acted as important sources for epidemics emerging in other regions. This was also supported by the state change counts, with migration from the middle reaches of the Yellow River and Western Europe being much greater than that from any other geographic region included in our analysis (Figure 2). Eastern Europe, Eastern China, Northwest China and Southwest China are the main immigration regions (Figure 2).

Spatial dynamics of wheat dwarf virus over time

We further summarize the WDV migration load and direction across time (Figure 3). The first wave of WDV emigrations from Iran to Eastern Europe had started before the 1840s and stopped in the early 1990s, the second wave of WDV emigrations also from Iran and spread to Western Europe had started after the 1860s and the migration stopped approximately in the 1960s (Figure 3A). However, the migrations from Western Europe to Eastern Europe, from Eastern Europe to the middle reaches of the Yellow River were not started until after the 1960s (Figure 3B). The migrations of the middle reaches of the Yellow River to Northwest China had started in the late 1970s, and the migrations of the middle reaches of the Yellow River to Eastern Europe, East China, and Southwest China began in the 1990s and began to decline in the late 2010s

(Figure 3C). In addition, there was a brief migration of Eastern China to Northwest China in the 2010s (Figure 3D).

Demographic history of wheat dwarf virus

A coalescence-based BSP generated by the whole–genome coding region data revealed the explicit demographic history of WDV, showing that WDV populations experienced a dramatic expansion in the early 1990s and then began to decline sharply for approximately 15 years (Figure 4). Our results show that the populations of WDV strains in barley and wheat remained stable until 1990, but it is clear that the population size of WDV-wheat is smaller than that of WDV-barley, and that WDV-wheat emerged later than WDV-barley. The population size of WDV-barley experienced expansion for approximately 7 years from the early 1990s to the 2000s, followed by population decline in approximately 2005. However, the population size of WDV-wheat began to expand in approximately 3 years in approximately 2003, began to decline sharply in less than 3 years in approximately 2007, and subsequently became consistent with the population size of WDV-barley (Figure 4).

Geographic structure patterns of wheat dwarf virus

The strain structure of WDV was apparent in the maximum clade credibility (MCC) trees and phylogenetic trees (Figure 1), but the geographical structure was not evident. Therefore, we investigated the global trait association of phylogeography structure with the evolution of WDV ($p_{AI} < 0.001$ and $p_{PS} < 0.001$, Table. 1). According to the *MC* statistic for single genes and whole–genome coding region of WDV, there were significant population segmentation was observed for the 4 geographic regions (Table 1). The DAPC scatter plots indicated that the population from Iran was relatively distinct from the other
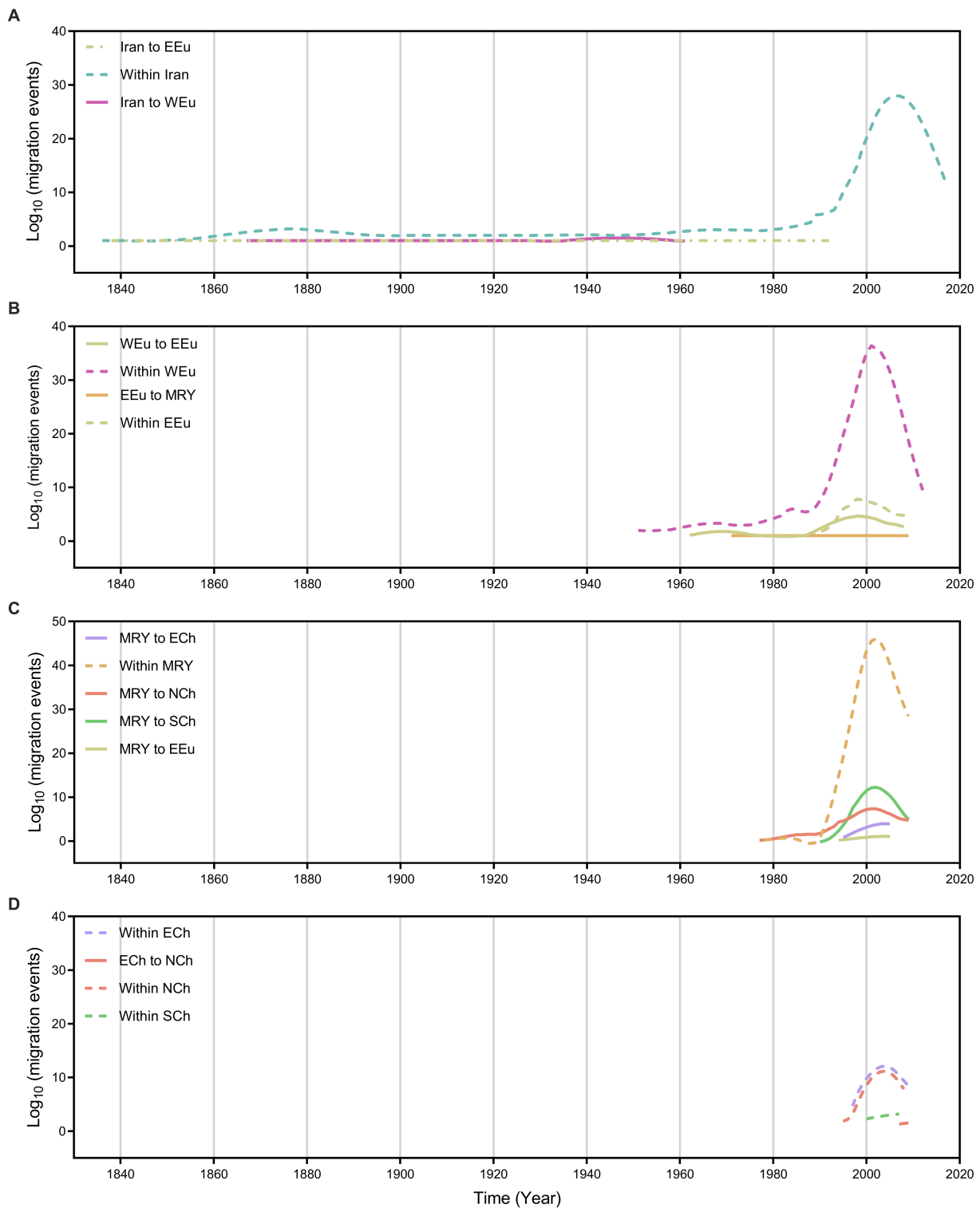
FIGURE 3
Inferred migration events (on $\log_{10}$ scale) of wheat dwarf virus over time (year) with each region. **(A)** Migration events from and within Iran, **(B)** Migration events from and within EEu and WEu, **(C)** Migration events from and within MRY, **(D)** Migration events from ECh and within ECh, NCh and SCh. ECh, Eastern China; EEu, Eastern Europe; MRY, Middle Reaches of Yellow River; NCh, Northwest China; SCh, Southwest China; WEu, Western Europe.

populations along the first discriminant function axis (Supplementary Figure S2). The isolates from Europe are clearly divided into groups: Eastern and Western European, while all isolates from China are significantly clustered together and separated from the

groups of Eastern and Western Europe on the second discriminant axis (Supplementary Figure S2). In addition, we conducted the global trait association of phylogeography structure and PCA analysis of the four WDV coding genes respectively, our results showed that the
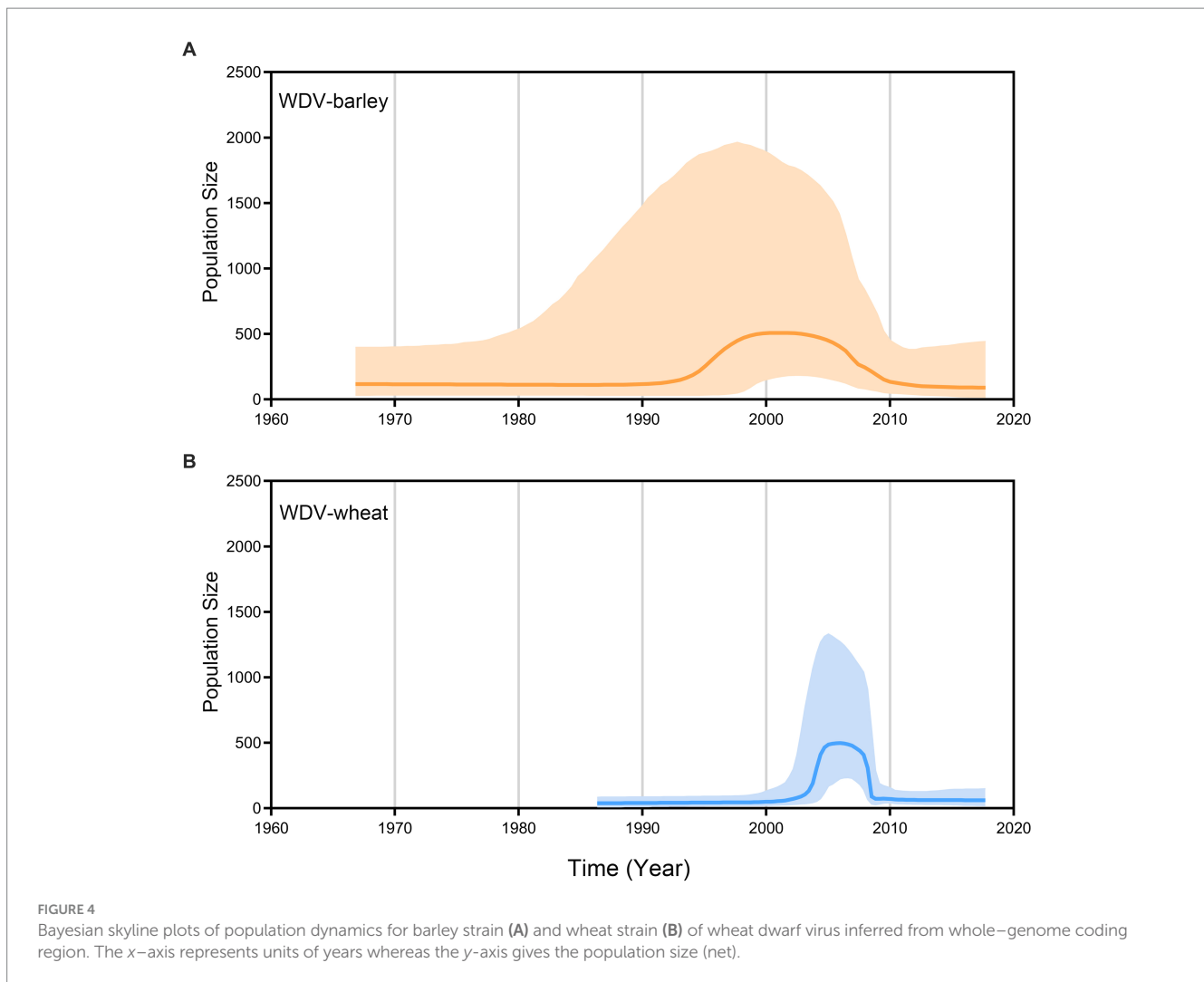
FIGURE 4
Bayesian skyline plots of population dynamics for barley strain **(A)** and wheat strain **(B)** of wheat dwarf virus inferred from whole−genome coding region. The x−axis represents units of years whereas the y-axis gives the population size (net).

TABLE 1 Analysis of the phylogeographic structure of wheat dwarf virus.

Coding region analysed	Statistic	Isolates	Observed mean (95% HPD CIs)	Null mean (95% HPD CIs)	p-value
Whole−genome coding region	AI		3.940 (3.219–4.658)	11.664 (10.583–12.729)	<0.001
	PS		31.132 (29.000–33.000)	61.153 (58.783–63.276)	<0.001
	MC (China)	114	4.349 (4.000–5.000)	1.290 (1.000–2.000)	<0.01
	MC (Eastern Europe)	14	5.698 (5.000–9.000)	1.686 (1.052–2.114)	<0.01
	MC (Iran)	27	17.060 (14.000–19.000)	5.870 (4.412–9.005)	<0.01
	MC (Western Europe)	24	3.094 (3.000–4.000)	1.853 (1.160–2.463)	<0.02

differentiation of the four WDV coding genes had strong geographical correlation (Supplementary Table S6; Supplementary Figures S3, S4).

Selection pressure analysis of wheat dwarf vieus

To analyse possible selection pressure in the WDV coding region, the ratio of nonsynonymous to synonymous sites (dN/dS) was calculated. Our resul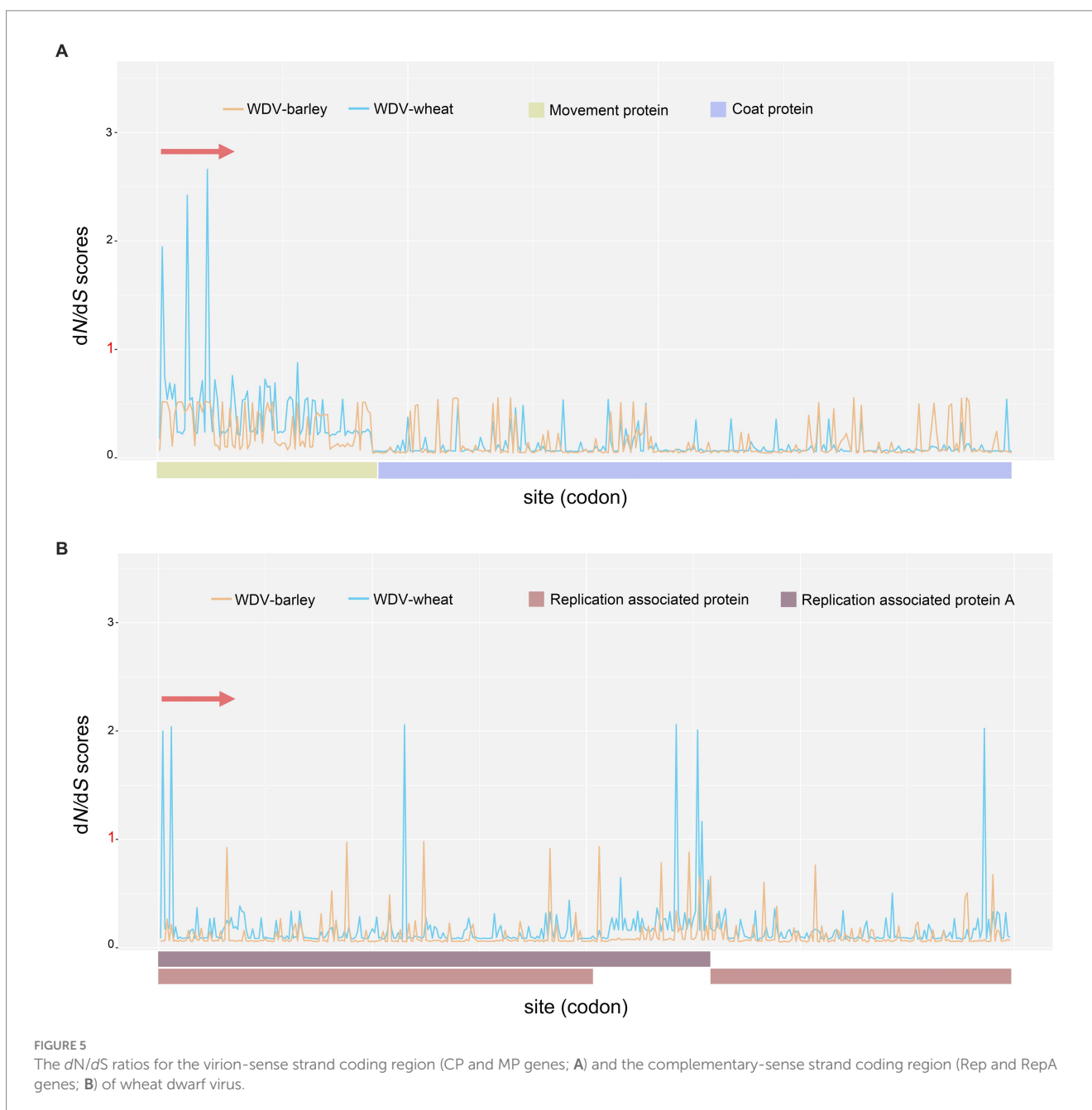ts indicated that negative selection was acting on the WDV whole−genome coding regions, and purifying selection was the main evolutionary force (Supplementary Table S7). Among the four protein-coding genes, the CP gene was subjected to the strongest selection pressure (Supplementary Table S7), and for the four protein-coding genes of WDV, WDV-barley was subjected to stronger purifying selection than WDV-wheat (Figure 5; Supplementary Table S7). Moreover, five positively selected sites were detected in two genes encoding replication proteins in WDV-wheat (Supplementary Table S8). Three positively selected sites were located in the overlapping coding regions of the Rep and RepA genes, among

which two positively selected sites were located at the N-terminus of the replicating proteins, while two and one positively selected sites were located at the C-terminus of RepA and Rep, respectively (Supplementary Table S8). Furthermore, no positively selected sites were detected in the CP and MP genes.

Discussion

We conducted a large-scale systematic dynamics analysis of the global population of WDV by screening all complete gene sequences retrieved from GenBank. Our MCC trees show that WDV has two distinct evolutionary lineages, WDV-wheat and WDV-barley (Figure 1), which is consistent with the findings of previous reports (Lindsten and Vacke, 1991; Commandeur and Huth, 1999; Muhire

et al., 2013; Schubert et al., 2014). Wheat is a hexaploid crop, while barley is a diploid crop. The differences between their genomes may lead to significant differences between barley and wheat in terms of resistance to WDV infection and proliferation in host cells. In response to different resistance mechanisms, WDV has host-specific host adaptations to better infect different hosts, resulting in the origin of different subspecies. Some reports showed that virions from WDV-barley and WDV-wheat strains do not infect barley and wheat plants (Abt et al., 2020), respectively, but other experiments showed that the virions of WDV-wheat could successfully infect barley plants (29), and the virions of WDV-barley could infect wheat plants (Lindsten and Vacke, 1991; Commandeur and Huth, 1999; Schubert et al., 2007; Kundu et al., 2009; Ramsell et al., 2009). According to the MCC trees that we reconstructed, some isolates of WDV-barley were derived from wheat, and some isolates of WDV-wheat has been

FIGURE 5
The dN/dS ratios for the virion-sense strand coding region (CP and MP genes; **A**) and the complementary-sense strand coding region (Rep and RepA genes; **B**) of wheat dwarf virus.

derived from barley (Figure 1). Based on these results, we can confirm that there do not appear to be host differences between the wheat and barley strains of WDV, although the virus appears in two distinct evolutionary groups due to host-driven adaptation.

Furthermore, our results also showed that in WDV-barley, all isolates from Iran (except for one isolate in the CP dataset) were clustered together, and the rest were from Europe (Figures 1A,B). Similarly, in WDV-wheat, all isolates from Iran (except for one isolate in the CP dataset) still clustered on the same branch (Figure 1). There is increasing evidence that the genetic variability of many viruses is related to the geographical origin of virus isolates (Gao et al., 2017; Mao et al., 2019; Rao et al., 2020). Our results also show that the evolution of WDV was closely related to geographical location (Table 1; Supplementary Figures S1–S3). Geographic, environmental, or ecological isolation leads to genetic differentiation (Wang et al., 2013; MacDonald et al., 2020), and the emergence of environmental isolation and ecological isolation (IBE) comes from the spatially divergent adaptation of populations to different environmental or ecological conditions (Wang et al., 2013; Sánchez-Ramírez et al., 2018; MacDonald et al., 2020). Populations in environmental/ecological isolation necessarily imply differential selection in relation to abiotic or biological factors, which promoting genomic differentiation and limiting gene flow within species, thereby affecting the generation and maintenance of spatial population structure (Zhan and McDonald, 2013; MacDonald et al., 2022). The rapid accumulation of adaptive genetic variation in plant pathogen populations results from directional selection imposed by physical environments, pathogen biology and the ways of human intervention (Thrall et al., 2011). There are many similarities in the control of field plant diseases caused by viruses in different countries and regions. However, different farming methods and chemical pressures may lead to different adaptations of the virus. In addition, an interesting result is that the isolates from China clustered significantly together (Supplementary Figures S2–S4), even though the isolates are from different regions with different elevations and climatic conditions. This implies that physical factors such as the altitude and climatic experienced by the pathogens may not be causes of WDV population diversity.

WDV is transversely transmitted among host plants by *P. alienus*, and the geographical environment is one of the barriers to insect virus transmission. Although 25 million hectares of wheat and barley are grown in China every year (Ministry of Agriculture and Rural Affairs of China), no barley strains have been reported in the country. The most common form of geographic isolation is isolation by distance (IBD), which describes a pattern in which the genetic distance between individuals or populations is positively correlated with their geographic isolation (Wright, 1943). The extreme isolation by distance between China and Europe prevented *P. alienus* from transmitting WDV-barley into China. It is also important to note that although only wheat strains of WDV were distributed in China, some isolates were still collected from barley plants in the field. This further indicates that there is no difference in host plants between the wheat and barley strains of WDV. In addition, we did not observe adaptive evolution of WDV in other infected hosts, such as *A. sativa* (hexaploid) and *Secale cereale* (diploid). These results suggest that after the host-driven adaptation evolution of WDV, host-related evolution occurred again in response to the adaptation of its infected hosts.

Our results indicate that selection pressure plays an important role in the evolution of WDV, and the genes in WDV are subject to different degrees of purifying selection. The CP gene is subject to the strongest purifying selection, while MP gene is subjected to the weakest purification selection among the four genes of WDV. Positively selected sites were detected in only two genes encoding replication-related proteins (Supplementary Table S8). As the structural proteins of virions, coat proteins play a very important role in the life history of viruses. Meanwhile, for the viruses of *Geminiviridae*, the nuclear localization sequences (NLSs) of coat proteins are key to the entry of viral genetic material into the nucleus, and coat protein facilitates nuclear and cell-to-cell transport of the genome via interaction with the MP and NSP (Bennett and Agbandje-McKenna, 2020). Under high mutation rates, more important genes are bound to be subject to stronger purifying selection to prevent the effects of harmful mutations. Furthermore, an interesting result is that WDV-barley are subject to stronger purifying selection than WDV-wheat (Supplementary Table S7), and although only five positively selected sites were detected, they were all detected in a wheat strain of WDV (Figure 5; Supplementary Table S8). This may be related to the wider geographical distribution (China) of wheat strains, and the variation in host genetics, fungicide applications and agricultural practices among regions as means of natural selection led to the differential evolution between WDV-wheat and WDV-barley. Although the purifying selection of the Rep and RepA genes was less than that of the MP gene, all five positive selection sites were detected in the Rep and RepA genes. Both RepA and Rep have RNA silencing suppressor activities (Liu et al., 2014), and the mutability of specific sites may better resist the host plant's defense mechanisms.

Because of their more stable genome structure, DNA viruses evolve at a slower rate than RNA viruses, which generally maintain a rate of approximately 10^{-4} substitutions/site/year (Pagan and Holmes, 2010; Fraile et al., 2011; Holmes, 2011). Unexpectedly, the evolutionary rate of WDV remained at 10^{-4} substitutions/site/year, and the results were similar to those of many RNA viruses. More surprisingly, the evolutionary rate of the MP gene was 1.309×10^{-3} substitutions/site/year (95% credibility interval: 7.117×10^{-4}–1.993×10^{-3}), which is even faster than that of some RNA viruses (Pagan and Holmes, 2010; Gao et al., 2020). Our results suggest that the evolutionary rate of WDV varies among genes, which is related to the function of the proteins encoded by the genes. Moreover, the TMRCAs of WDV obtained from the datasets with different coding regions were not the same, and the TMRCA of WDV obtained from the whole–genome coding region dataset was 1746 CE (95% credibility interval: 1517 CE–1893 CE). Furthermore, the TMRCA of each WDV gene obtained from the barley strain datasets was earlier than that from the wheat strain datasets, indicating that WDV may have first appeared in barley plants. In addition, the evolutionary rate of WDV-wheat was faster than that of WDV-barley (except for the mp gene, Supplementary Table S4). The difference in geographical distribution between wheat and barley strains of WDV may have contributed to this difference. This evolutionary difference between wheat and barley strains of WDV may be related to wheat strains having more complex plant–virus–environment relationships. Two host-driven adaptations of WDV resulted in distinct differentiation between WDV-wheat and WDV-barley. Moreover, only wheat strains were detected in China, which not only resulted in a larger population size of WDV-wheat than WDV-barley, but also made wheat strains face more complex environmental factors than barley strains.

Wheat and barley are the most important crops originating in the Fertile Crescent in Western Asia (Heun et al., 1997; Dubcovsky and

Dvorak, 2007; Allaby et al., 2008; Brown et al., 2009; Pankin et al., 2018), and parts of Iran belong to the Fertile Crescent in Western Asia. As expected, our results identify that it is a high probability that WDV originated in Iran, and that the root posterior probabilities for Iran are much higher than that for China, Eastern Europe and Western Europe, which is supported by the whole–genome coding region and CP datasets (Figure 1; Supplementary Table S5). In addition, we obtained the same result by analysing the CP gene, which contains more isolates from Hungary (Figure 1B; Supplementary Table S5). Furthermore, our results, including the population history and TMRCA of WDV strains in barley and wheat, suggest that WDV may have originated in barley (Figure 3; Supplementary Table S4). The outbreak of WDV in the Fertile Crescent in Western Asia resulted in two distinct evolutionary lineages due to host adaptation. Our analysis of WDV migration pathways suggests that wheat strain of WDV spread from the Fertile Crescent in Western Asia to Eastern Europe before the 1840s, while WDV (both WDV-barley and -wheat) from the Fertile Crescent in Western Asia to Western Europe after 30 years. Surprisingly, Eastern European strain of WDV-barley only spread from Western Europe in the early 1960s. The wheat strain of WDV then spread to China in the early 1970s and from the middle reaches of the Yellow River to other major wheat–producing areas in China (Figure 2; Supplementary Table S5). Our other study (Wei et al., 2023) showed that the middle reaches of the Yellow River (Shaanxi and Shanxi provinces) are also a major emigration region of BYDV (barley yellow dwarf virus).

We performed a molecular epidemiology of WDV based on the whole–genome coding region sequences and unveiled the evolutionary history of the virus. During the evolutionary history of WDV, there were two episodes of adaptive evolution in its infected host plants. The first episode, which occurred in the Fertile Crescent in Western Asia, led to the emergence of two different evolutionary lineages of WDV, namely, wheat and barley strains, and the second episode allowed the two different strains to successfully infect hexaploid and diploid host plants. The TMRCA and BSP results indicate that WDV first appeared in barley plants. The population of WDV-barley began to expand in the early 1990s, and the population of the wheat strain began to expand in the early 2000s, while the population of the barley strain began to decline. After 2010, the populations of WDV-barley and WDV-wheat decreased to basically the same size as the population before expansion. Through pedigree and geographic analysis, we found that WDV probably originated in the Fertile Crescent in Western Asia and spread to other regions, and that Western Europe and the middle reaches of the Yellow River were the main export regions for WDV. While these results have provided new understanding into the evolutionary history of WDV, we may need variant information collected after 2017 for further assess the evolutionary history of WDV.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

SW: Conceptualization, Data curation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. LL: Visualization, Writing – review & editing. GC: Visualization, Writing – review & editing. HY: Conceptualization, Writing – review & editing. LH: Software, Writing – review & editing. GG: Conceptualization, Writing – review & editing. PL: Conceptualization, Methodology, Writing – review & editing. MZ: Conceptualization, Funding acquisition, Project administration, Writing – original draft, Writing – review & editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2024.1314526/full#supplementary-material

References

Abt, I., Souquet, M., Angot, G., Mabon, R., Dallot, S., Thébaud, G., et al. (2020). Functional transcomplementation between wheat dwarf virus strains in wheat and barley. *Viruses* 12:34. doi: 10.3390/v12010034

Allaby, R. G., Fuller, D. Q., and Brown, T. A. (2008). The genetic expectations of a protracted model for the origins of domesticated crops. *Proc. Natl. Acad. Sci. U S A* 105, 13982–13986. doi: 10.1073/pnas.0803780105

Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M. A., and Alekseyenko, A. V. (2012). Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Mol. Biol. Evol.* 29, 2157–2167. doi: 10.1093/molbev/mss084

Baele, G., Lemey, P., and Suchard, M. A. (2016). Genealogical working distributions for Bayesian model testing with phylogenetic uncertainty. *Syst. Biol.* 65, 250–264. doi: 10.1093/sysbio/syv083

Bendahmane, M., Schalk, H. J., and Gronenborn, B. (1995). Identification and characterization of wheat dwarf virus from France using a rapid method for geminivirus DNA preparation. *Phytopathology* 85, 1449–1455. doi: 10.1094/Phyto-85-1449

Bennett, A., and Agbandje-McKenna, M. (2020). Geminivirus structure and assembly. *Adv. Virus Res.* 2020, 1–32. doi: 10.1016/bs.aivir.2020.09.005

Bielejec, F., Baele, G., Vrancken, B., Suchard, M. A., Rambaut, A., and Lemey, P. (2016). SpreaD3: interactive visualization of spatiotemporal history and trait evolutionary processes. *Mol. Biol. Evol.* 33, 2167–2169. doi: 10.1093/molbev/msw082

Bisztray, G. Y., Gáborjányi, R., and Vacke, J. (1989). Isolation and characterisation of wheat dwarf virus found for the first time in Hungary. *J. Plant Dis. Prot.* 96, 449–445.

Brown, T. A., Jones, M. K., Powell, W., and Allaby, R. G. (2009). The complex origins of domesticated crops in the Fertile Crescent. *Trends Ecol. Evol.* 24, 103–109. doi: 10.1016/j.tree.2008.09.008

Commandeur, U., and Huth, W. (1999). Differentiation of strains of wheat dwarf virus in infected wheat and barley plants by means of polymerase chain reaction. *J. Plant Dis. Prot.* 106, 550–552.

Dubcovsky, J., and Dvorak, J. (2007). Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* 316, 1862–1866. doi: 10.1126/science.1143986

Duchêne, S., Duchêne, D., Holmes, E. C., and Ho, S. Y. W. (2015). The performance of the date-randomization test in phylogenetic analyses of time-structured virus data. *Mol. Biol. Evol.* 32, 1895–1906. doi: 10.1093/molbev/msv056

Duchene, S., Lemey, P., Stadler, T., Ho, S. Y. W., Duchene, D. A., Dhanasekaran, V., et al. (2020). Bayesian evaluation of temporal signal in measurably evolving populations. *Mol. Biol. Evol.* 37, 3363–3379. doi: 10.1093/molbev/msaa163

Fraile, A., Pagan, I., Anastasio, G., Saez, E., and Garcia-Arenal, F. (2011). Rapid genetic diversification and high fitness penalties associated with pathogenicity evolution in a plant virus. *Mol. Biol. Evol.* 28, 1425–1437. doi: 10.1093/molbev/msq327

Gao, F. L., Chen, C., Arab, D. A., Du, Z., He, Y., and Ho, S. Y. W. (2019). Easycodeml: a visual tool for analysis of selection using codeML. *Ecol. Evol.* 9, 3891–3898. doi: 10.1002/ece3.5015

Gao, F. L., Kawakubo, S., Ho, S. Y. W., and Ohshima, K. (2020). The evolutionary history and global spatio-temporal dynamics of potato virus Y. *Virus Evol.* 6:veaa05. doi: 10.1093/ve/veaa056

Gao, F. L., Zou, W., Xie, L. X., and Zhan, J. S. (2017). Adaptive evolution and demographic history contribute to the divergent population genetic structure of potato virus Y between China and Japan. *Evol. Appl.* 10, 379–390. doi: 10.1111/eva.12459

Hanley-Bowdoin, L., Bejarano, E. R., Robertson, D., and Mansoor, S. (2013). Geminiviruses: masters at redirecting and reprogramming plant processes. *Nat. Rev. Microbiol.* 11, 777–788. doi: 10.1038/nrmicro3117

Heun, M., Schäfer-Pregl, R., Klawan, D., Castagna, R., Accerbi, M., Borghi, B., et al. (1997). Site of einkorn wheat domestication identified by DNA fingerprinting. *Science* 278, 1312–1314. doi: 10.1126/science.278.5341.1312

Hofer, J. M. I., Dekker, E. L., Reynolds, H. V., Woolston, C. J., Cox, B. S., and Mullineaux, P. M. (1992). Coordinate regulation of replication and virion sense gene expression in wheat dwarf virus. *Plant Cell* 4, 213–223. doi: 10.1105/tpc.4.2.213

Holmes, E. C. (2011). What does virus evolution tell us about virus origins? *J. Virol.* 85:9655. doi: 10.1128/JVI.05689-11

Jombart, T., Devillard, S., and Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 11:94. doi: 10.1186/1471-2156-11-94

Kass, R. E., and Raftery, A. E. (1995). Bayes factors. *J. Am. Stat. Assoc.* 90, 773–795. doi: 10.1080/01621459.1995.10476572

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kawakubo, S., Gao, F. L., Li, S. F., Tan, Z. Y., Huang, Y. K., and Adkar-Purushothama, C. R. (2021). Genomic analysis of the brassica pathogen turnip mosaic potyvirus reveals its spread along the former trade routes of the silk road. *Proc. Natl. Acad. Sci. USA* 118:e2021221118. doi: 10.1073/pnas.2021221118

Köklü, G., Ramsell, J. N. E., and Kvarnheden, A. (2007). The complete genome sequence for a Turkish isolate of wheat dwarf virus (WDV) from barley confirms the presence of two distinct WDV strains. *Virus Genes* 34, 359–366. doi: 10.1007/s11262-006-0029-0

Kundu, J., Gadiou, S., and Cervena, G. (2009). Discrimination and genetic diversity of wheat dwarf virus in the Czech Republic. *Virus Genes* 38, 468–474. doi: 10.1007/s11262-009-0352-3

Kvarnheden, A., Lindblad, M., Lindsten, K., and Valkonen, J. P. T. (2002). Genetic diversity of wheat dwarf virus. *Arch. Virol.* 147, 205–216. doi: 10.1007/s705-002-8313-x

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., and Calcott, B. (2016). PartitionFinder 2: new methods for selecting partitioned models of evolution formolecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* 34, 772–773. doi: 10.1093/molbev/msw260

Lemmetty, A., and Huusela-Veistola, E. (2005). First report of WDV in winter wheat in Finland. *Plant Dis.* 89:912. doi: 10.1094/PD-89-0912B

Lindblad, M., and Areno, P. (2002). Temporal and spatial population dynamics of *Psammotettix alienus*, a vector of wheat dwarf virus. *Int. J. Pest Manag.* 48, 233–238. doi: 10.1080/09670870110118731

Lindblad, M., and Waern, P. (2002). Correlation of wheat dwarf incidence to winter wheat cultivation practices. *Agric. Ecosyst. Environ.* 92, 115–122. doi: 10.1016/S0167-8809(01)00302-4

Lindsten, K., and Lindsten, B. (1999). Wheat dwarf–an old disease with new outbreaks in Sweden. *J. Plant Dis. Prot.* 106, 325–332.

Lindsten, K., and Vacke, J. (1991). A possible barley adapted strain of wheat dwarf virus (WDV). *Acta Phytopathol. Entomol. Hung.* 126, 175–118.

Liu, Y., Jin, W., Wang, L., and Wang, X. (2014). Replication-associated proteins encoded by wheat dwarf virus act as RNA silencing suppressors. *Virus Res.* 190, 34–39. doi: 10.1016/j.virusres.2014.06.014

Liu, Y., Wang, B., Vida, G., Cséplo-Károlyi, M., Wu, B. L., Wu, Y. H., et al. (2012). Genomic analysis of the natural population of wheat dwarf virus in wheat from China and Hungary. *J. Integr. Agric.* 11, 2020–2027. doi: 10.1016/S2095-3119(12)60459-6

MacDonald, Z. G., Dupuis, J. R., Davis, C. S., Acorn, J. H., Nielsen, S. E., and Sperling, F. A. (2020). Gene flow and climate-associated genetic variation in a vagile habitat specialist. *Mol. Ecol.* 29, 3889–3906. doi: 10.1111/mec.15604

MacDonald, Z. G., Snape, K. L., Roe, A. D., and Sperling, F. A. H. (2022). Host association, environment, and geography underlie genomic differentiation in a major forest pest. *Evol. Appl.* 15, 1749–1765. doi: 10.1111/eva.13466

Macdowell, S. W., Macdonald, H., Hamilton, W. D. O., Coutts, R. H. A., and Buck, K. W. (1985). The nucleotide sequence of cloned wheat dwarf virus DNA. *EMBO J.* 4, 2173–2180. doi: 10.1002/j.1460-2075.1985.tb03912.x

Mao, Y. Z., Sun, X. H., Shen, J. G., Gao, F. L., Qiu, G. W., Wang, T., et al. (2019). Molecular evolutionary analysis of potato virus Y infecting potato based on the VPg gene. *Front. Microbiol.* 10:1708. doi: 10.3389/fmicb.2019.01708

Martin, D. P., Murrell, B., Golden, M., Khoosal, A., and Muhire, B. (2015). RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol.* 1:vev003. doi: 10.1093/ve/vev003

Muhire, B., Martin, D. P., Brown, J. K., Navas-Castillo, J., Moriones, E., Zerbini, F. M., et al. (2013). Genome-wide pairwise-identity-based proposal for the classification of viruses in the genus *Mastrevirus* (family Geminiviridae). *Arch. Virol.* 158, 1411–1424. doi: 10.1007/s00705-012-1601-7

Müller, N. F., Rasmussen, D., and Stadler, T. (2018). MASCOT: parameter and state inference under the marginal structured coalescent approximation. *Bioinformatics* 34, 3843–3848. doi: 10.1093/bioinformatics/bty406

Pagan, I., and Holmes, E. C. (2010). Long-term evolution of the Luteoviridae: time scale and mode of virus speciation. *J. Virol.* 84, 6177–6187. doi: 10.1128/JVI.02160-09

Pankin, A., Altmüller, J., Becker, C., and von Korff, M. (2018). Targeted resequencing reveals genomic signatures of barley domestication. *New Phytol.* 218, 1247–1259. doi: 10.1111/nph.15077

Parizipouret, M. H. G., Schubert, J., Behjatnia, S. A. K., Afsharifar, A., Habekuß, A., and Wu, B. L. (2017). Phylogenetic analysis of wheat dwarf virus isolates from Iran. *Virus Genes* 53, 266–274. doi: 10.1007/s11262-016-1412-0

Parker, J., Rambaut, A., and Pybus, O. G. (2008). Correlating viral phenotypes with phylogeny: accounting for phylogenetic uncertainty. *Infect. Genet. Evol.* 8, 239–246. doi: 10.1016/j.meegid.2007.08.001

Pérez-Losada, M., Arenas, M., Galán, J. G., Palero, F., and González-Candelas, F. (2015). Recombination in viruses: mechanisms, methods of study, and evolutionary consequences. *Infect. Genet. Evol.* 30, 296–307. doi: 10.1016/j.meegid.2014.12.022

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. (2018). Posterior summarisation in Bayesian phylogenetics using tracer 1.7. *Syst. Biol.* 67, 901–904. doi: 10.1093/sysbio/syy032

Ramsell, J. N. E., Boulton, M. I., Martin, D. P., Valkonen, J. P. T., and Kvarheden, A. (2009). Studies on the host range of the barley strain of wheat dwarf virus using an agroinfectious viral clone. *Plant Pathol.* 58, 1161–1169. doi: 10.1111/j.1365-3059.2009.02146.x

Rao, S. F., Chen, X. W., Qiu, S. Y., Peng, J. J., Zheng, H. Y., and Lu, Y. W. (2020). Identification of two new isolates of Chilli veinal mottle virus from different regions in China: molecular diversity, Phylogenetic and recombination analysis. *Front. Microbiol.* 11:616171. doi: 10.3389/fmicb.2020.616171

Sahu, P. P., Sharma, N., Puranik, S., Muthamilarasan, M., and Prasad, M. (2013). Involvement of host regulatory pathways during geminivirus infection: a novel platform for generating durable resistance. *Funct. Integr. Genomics* 14, 47–58. doi: 10.1007/s10142-013-0346-z

Sánchez-Ramírez, S., Rico, Y., Berry, K. H., Edwards, T., Karl, A. E., and Henen, B. T. (2018). Landscape limits gene flow and drives population structure in Agassiz's desert tortoise (*Gopherus agassizii*). *Sci. Rep.* 8:11231. doi: 10.1038/s41598-018-29395-6

Schubert, J., Habekuss, A., Kazmaier, K., and Jeske, H. (2007). Surveying cereal-infecting geminiviruses in Germany—diagnostics and direct sequencing using rolling circle amplification. *Virus Res.* 127, 61–70. doi: 10.1016/j.virusres.2007.03.018

Schubert, J., Habekuss, A., Wu, B. L., Thieme, T., and Wang, X. F. (2014). Analysis of complete genomes of isolates of the wheat dwarf virus from new geographical locations and descriptions of their defective forms. *Virus Genes* 48, 133–139. doi: 10.1007/s11262-013-0989-9

Sirlova, L., Vacke, J., and Chaloupkova, M. (2005). Reaction of selected winter wheat varieties to autumnal infection with wheat dwarf virus. *Plant Prot. Sci.* 41, 1–7. doi: 10.17221/2732-PPS

Su, Y. C. F., Bahl, J., Joseph, U., Butt, K. M., Peck, H. A., Koay, E. S. C., et al. (2015). Phylodynamics of H1N1/2009 influenza reveals the transition from host adaptation to immune-driven selection. *Nat. Commun.* 6:7952. doi: 10.1038/ncomms8952

Suchard, M., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., and Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* 4:vey016. doi: 10.1093/ve/vey016

Thrall, P. H., Oakeshott, J. G., Fitt, G., Southerton, S., Burdon, J. J., Sheppard, A., et al. (2011). Evolution in agriculture: the application of evolutionary approaches to the management of biotic interactions in agro-ecosystems. *Evol. Appl.* 4, 200–215. doi: 10.1111/j.1752-4571.2010.00179.x

Tobias, I., Shevchenko, O., Kiss, B., Bysov, A., Snihur, H., Polischuk, V., et al. (2011). Comparison of the nucleotide sequences of wheat dwarf virus (WDV) isolates from Hungary and Ukraine. *Pol. J. Microbiol.* 60, 125–131. doi: 10.33073/pjm-2011-017

Vacke, J. (1961). Wheat dwarf virus disease. *Biol. Plant.* 3, 228–233. doi: 10.1007/BF02933566

Vacke, J. (1964). Some new findings on wheat dwarf virus. In: C. Blattný, editor. Proceedings of the 5th conference of the Czechoslovak plant virologists. Prague, Czechoslovakia: Publishing house of the Czechoslovak Academy of Sciences, Prague, Czechoslovakia; pp. 331–334.

Vacke, J. (1972). Host plants range and symptoms of wheat dwarf virus. *Věd Pr Výz Ústavú Rostl Výroby Praha-Ruzyně* 17, 151–162.

Wang, I. J., Glor, R. E., and Losos, J. B. (2013). Quantifying the roles of ecology and geography in spatial genetic divergence. *Ecol. Lett.* 16, 175–182. doi: 10.1111/ele.12025

Wang, H., Liu, Y., Zhang, L., Kundu, J. K., Liu, W., and Wang, X. (2019). ADP ribosylation factor 1 facilitates spread of wheat dwarf virus in its insect vector. *Cell. Microbiol.*:e13047. doi: 10.1111/cmi.13047

Wang, Y. J., Mao, Q. Z., Liu, W. W., Mar, T. T., Wei, T. Y., Liu, Y., et al. (2014). Localization and distribution of wheat dwarf virus in its vector leafhopper, *Psammotettix alienus*. *Phytopathology* 104, 897–904. doi: 10.1094/PHYTO-09-13-0251-R

Wang, X. F., Wu, B., and Wang, J. F. (2008). First report of wheat dwarf virus infecting barley in Yunnan, China. *J. Plant Pathol.* 90:400.

Wei, S. H., Chen, G. L., Yang, H., Huang, L., Gong, G. S., Luo, P. G., et al. (2023). Global molecular evolution and phylogeographic analysis of barley yellow dwarf viruses based on the cp and mp genes. *Virol. J.* 20:130. doi: 10.1186/s12985-023-02084-1

Wright, S. (1943). Isolation by distance. *Genetics* 28, 114–138. doi: 10.1093/genetics/28.2.114

Wu, B. L., Melcher, U., Guo, X. Y., Wang, X. F., Fan, L. J., and Zhou, G. H. (2008). Assessment of codivergence of *Mastreviruses* with their plant hosts. *BMC Evol. Biol.* 8:335. doi: 10.1186/1471-2148-8-335

Wu, B. L., Shang, X. N., Schubert, J., Habekuß, A., Elena, S. F., and Wang, X. F. (2015). Global-scale computational analysis of genomic sequences reveals the recombination pattern and coevolution dynamics of cereal-infecting *geminiviruses*. *Sci. Rep.* 5:8153. doi: 10.1038/srep08153

Xie, J., Wang, X., Liu, Y., Peng, Y., and Zhou, G. (2007). First report of the occurrence of wheat dwarf virus in wheat in China. *Plant Dis.* 91:111. doi: 10.1094/PD-91-0111B

Yang, Z. H. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088

Yang, Z. H., and Nielsen, R. (1998). Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *J. Mol. Evol.* 46, 409–418. doi: 10.1007/PL00006320

Yang, Z. H., and Nielsen, R. (2002). Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol. Biol. Evol.* 19, 908–917. doi: 10.1093/oxfordjournals.molbev.a004148

Yang, Z. H., Wong, W. S., and Nielsen, R. (2005). Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol. Biol. Evol.* 22, 1107–1118. doi: 10.1093/molbev/msi097

Zhan, J. S., and McDonald, B. A. (2004). The interaction among evolutionary forces in the pathogenic fungus Mycosphaerella graminicola. *Fungal Genet. Biol.* 41, 590–599. doi: 10.1016/j.fgb.2004.01.006

Zhan, J. S., and McDonald, B. A. (2013). Experimental measures of pathogen competition and relative fitness. *Annu. Rev. Phytopathol.* 51, 131–153. doi: 10.1146/annurev-phyto-082712-102302

Zhang, D., Gao, F. L., Jakovlić, I., Zou, H., Zhang, J., Li, W. X., et al. (2020). PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* 20, 348–355. doi: 10.1111/1755-0998.13096