VIRUS EVOLUTION

Virus Evolution, 2022, **8(2)**, 1–11

DOI: https://doi.org/10.1093/ve/veac060
Advance access publication date: 24 June 2022
Research Article

The recombinogenic history of turnip mosaic potyvirus reveals its introduction to Japan in the 19th century

Shusuke Kawakubo,[1,†,‡] Yasuhiro Tomitaka,[1,2] Kenta Tomimura,[1,3] Ryoko Koga,[1] Hiroki Matsuoka,[1] Seiji Uematsu,[4,5] Kazuo Yamashita,[6,7] Simon Y. W. Ho,[8,§] and Kazusato Ohshima[1,9,*,**]

[1]Laboratory of Plant Virology, Department of Biological Resource Science, Faculty of Agriculture, Saga University, 1-banchi, Honjo-machi, Saga 840-8502, Japan, [2]Institute for Plant Protection, National Agriculture and Food Research Organization, 2-1-18 Kannondai, Tsukuba, Ibaraki 305-8666, Japan, [3]Institute of Fruit Tree and Tea Science, National Agriculture and Food Research Organization, 485-6 Okitsu Nakacho, Shimizu, Shizuoka 424-0292, Japan, [4]Laboratory of Agro-Environmental Science, Warm Region Horticulture Institute, Chiba Prefectural Agriculture and Forestry Research Center, 1762 Yamamoto, Tateyama, Chiba 294-0014, Japan, [5]Laboratory of Molecular and Cellular Biology, Department of Bioregulation and Bio-interaction, Tokyo University of Agriculture and Technology, 3-5-8 Saiwai cho, Fuchu, Tokyo 183-8509, Japan, [6]Vegetable Research Institute, Aomori Prefectural Industrial Technology Research Center, 91 Yanagisawa, Inuotose, Rokunohe, Aomori 033-0071, Japan, [7]Fukuchi Garlic R&S, 4-92 Akane, Fukuda, Nanbu cho, Aomori 039-0815, Japan, [8]School of Life and Environmental Sciences, University of Sydney, Sydney, NSW 2006, Australia and [9]The United Graduate School of Agricultural Sciences, Kagoshima University, 1-21-24 Korimoto, Kagoshima 890-0065, Japan

[†]Present address: Graduate School of Agriculture, Hokkaido University, Kita 9 Nishi 9, Kita-ku, Sapporo, 060-8589, Japan.
[‡]https://orcid.org/0000-0002-5330-5532
[§]https://orcid.org/0000-0002-0361-2307
[**]https://orcid.org//0000-0002-8592-6042
*Corresponding author: E-mail: ohshimak@cc.saga-u.ac.jp

Abstract

Characterizing the detailed spatial and temporal dynamics of plant pathogens can provide valuable information for crop protection strategies. However, the epidemiological characteristics and evolutionary trajectories of pathogens can differ markedly from one country to another. The most widespread and important virus of brassica vegetables, turnip mosaic virus (TuMV), causes serious plant diseases in Japan. We collected 317 isolates of TuMV from *Raphanus* and *Brassica* plants throughout Japan over nearly five decades. Genomic sequences from these isolates were combined with published sequences. We identified a total of eighty-eight independent recombination events in Japanese TuMV genomes and found eighty-two recombination-type patterns in Japan. We assessed the evolution of TuMV through space and time using whole and partial genome sequences of both nonrecombinants and recombinants. Our results suggest that TuMV was introduced into Japan after the country emerged from its isolationist policy (1639–1854) in the Edo period and then dispersed to other parts of Japan in the 20th century. The results of our analyses reveal the complex structure of the TuMV population in Japan and emphasize the importance of identifying recombination events in the genome. Our study also provides an example of surveying the epidemiology of a virus that is highly recombinogenic.

Key words: turnip mosaic virus; potyvirus; evolution; recombination; space and time; Japan.

1. Introduction

Turnip mosaic virus (TuMV) is one of the most widespread viruses, having the widest host range (Walsh and Jenner 2002). It causes serious yield losses of brassica vegetables (e.g. broccoli, cabbage, radish, and turnip) and ornamental plants across the world (Ohshima, Yamaguchi, and Hirota et al. 2002; Walsh and Jenner 2002; Nellist et al. 2022) and is transmitted by aphids in a non-persistent manner. TuMV was first described in 1921 in the USA in *Brassica rapa* (Gardner and Kendrick 1921; Schultz 1921) and in the UK in *Brassica oleracea* (Smith 1935). The first descriptions of the pathogen in Japan were from *Raphanus* and *Brassica* plants (Yoshii 1951), and in the following decades, the Japanese isolates were biologically classified (Yoshii 1963; Sako 1981). There were also some reports of mosaic symptoms in *Brassica* plants (Takimoto 1930) including Chinese cabbage (*Brassica pekinensis*) and *Raphanus* plants (Ishiyama and Misawa 1942) before 1950, but the causal viruses were unknown at the time.

The major domesticated hosts of TuMV in Japan are *Raphanus* and *Brassica* plants. There are descriptions of Japanese radish (daikon, *Raphanus sativus*) in an early Japanese chronicle of myths, 'Kojiki' (712 CE) and 'Nihon Shoki' (720 CE), sometimes translated as the Chronicles of Japan (Aoba 2013). Japanese oilseed rape and Japanese mustard, including Japanese radish, were introduced in the Yayoi period (10th century BCE to 3rd century CE). Turnip (*B. rapa*) was also described in 'Nihon Shoki' and had probably been introduced from China. Heading leaf cabbage was introduced in the 1850s, toward the end of the Edo period (1603–1868 CE), and has been widely grown since about 1900. Non-heading

leaf Chinese cabbage (*B. pekinensis*) was introduced from China at the beginning of the Meiji period and became heading leaf Chinese cabbage around 1895 CE (Nishi 1981). The heading leaf Chinese cabbage was widely grown at the beginning of the Taisho period (1912–1926 CE) and has become a major crop vegetable since the 1950s (Shimizu 2008, 2017). Therefore, the various host plants of TuMV in Japan were introduced in different eras over the past two to three millennia.

TuMV is a member of the genus *Potyvirus*. This genus includes 190 virus species that infect a wide range of both cultivated crops and wild plants. *Potyvirus* is the largest genus of *Potyviridae*, the largest family of plant RNA viruses (Gibbs and Ohshima 2010; Gibbs et al. 2020; Inoue-Nagata, Jordan, and Kreuze et al. 2022). Potyviruses have flexuous filamentous particles 700–750 nm long, each of which contains a single copy of the genome. The genome is a single-stranded positive sense RNA molecule of about 9,000–12,000 nucleotides. It is translated into one large polyprotein that hydrolyzes itself into at least ten mature proteins. Moreover, an overlapping pretty interesting *Potyviridae* ORF (PIPO) exists in the +2 reading frame within the protein 3 (P3) coding region (Chung et al. 2008).

TuMV is characterized by extensive recombination events, a feature that is common among potyviruses (Ohshima, Yamaguchi, and Hirota et al. 2002; Chare and Holmes 2006; Gibbs and Ohshima 2010; Gibbs et al. 2020). Recombination hotspots in TuMV genomes were the first to have been reported for potyviruses (Ohshima, Tomitaka, and Wood et al. 2007). Furthermore, two major classes of recombinants have been found in the TuMV population: intralineage recombinants (i.e. those parents from the same major group or subgroup lineages) or interlineage recombinants (i.e. those parents from different major group or subgroup lineages).

The evolutionary history of such a highly recombinogenic virus can be difficult to reconstruct because standard phylogenetic approaches can be seriously misled by recombination (Schierup and Hein 2000; Martin, Lemey, and Posada 2011). Phylodynamic inference requires that the recombination-type pattern of each isolate be taken into account, while analysis might need to be limited to partial genomic regions. To minimize the impacts of substitution saturation and the uncertainty in ancestral location states, it is best to focus the virus-sampling efforts to a relatively small and well-defined geographic area (Schluter et al. 1997; Dellicour, Rose, and Faria et al. 2017).

Given the large impacts of recombination on genomes, it is unquestionable that accounting for this process is essential when reconstructing the evolutionary history of a virus. Since recombination can cause much larger genetic differences than point mutations, the process of recombination in animal viruses has been often studied for its clinical importance (Hostager, Ragonnet-Cronin, and Murrell et al. 2019; Zhang, Chen, and Shan et al. 2020). In the case of DNA plant viruses, inferring the date and location of recombination events has only been reported in tomato yellow leaf curl virus (Lefeuvre, Martin, and Harkins et al. 2010; Fiallo-Olivé et al. 2019). However, reconstruction of the dispersal history of each recombinant has not been reported in any plant viruses because of the more limited amount of genomic data.

Ages of virus recombination events have previously been inferred by multiple methods, including partitioning into genomic regions and evaluating their dates separately (Bertrand et al. 2012; Ding et al. 2017). Here, we infer the age of the most recent common ancestors of isolates sharing the same recombination events, as has previously been attempted for some viruses (Wittmann, Biek, and Hassanin et al. 2007; Visser, Bellstedt, and Pirie 2012; Yasaka, Ohba, and Schwinghamer et al. 2015; Yasaka, Fukagawa, and Ikematsu et al. 2017). Investigating the dispersal history using entire nonrecombinants and recombinants themselves should lead to better characterization of virus dispersals.

In this study, we examine the phylodynamic history of TuMV in Japan by analyzing a large data set collected from symptomatic brassica plants over a period of more than a half century. By using a phylogeographic approach, our goals are to understand the genetic diversity and spatio-temporal dispersal of TuMV in Japan, as an example of a plant virus on a whole-country scale. Our results provide molecular evolutionary information for a major pathogen of important agricultural crops, while also illustrating how epidemiological and evolutionary inference can be performed effectively for a highly recombinogenic virus.

2. Materials and methods

2.1 Collection of virus isolates

Japan is composed of the five main islands of Hokkaido, Honshu, Shikoku, Kyushu, and Okinawa, and the largest island Honshu is further divided into five districts, Tohoku, Kanto, Chubu, Kinki, and Chugoku (Supplementary Fig. S1). During the crop-producing seasons from 1960 to 2017, we collected plant samples showing mosaic or vein-clearing symptoms in *Raphanus* and *Brassica* plants (Supplementary Table S1). We intended to collect TuMV samples to include all parts of Japan, but it was relatively difficult to collect TuMV in the northeastern parts of Hokkaido Prefecture. However, we were able to collect TuMV isolates from all major islands and districts of Japan, which spans approximately 2,800 km north to south and 3,000 km east to west.

We tested for the presence of TuMV in all collected samples by direct double-antibody sandwich enzyme-linked immunosorbent assay (Clark and Adams 1977). All isolates were mechanically inoculated to *Chenopodium quinoa* using 0.01 M potassium phosphate buffer (PPB) (pH 7.0) and serially cloned through single lesions at least twice. We then propagated the biologically cloned isolates in *B. rapa* cv. Hakatasuwari (turnip) or *Nicotiana benthamiana* for two to three weeks before extracting viral RNA for sequencing. The old sample collections were stored in a deep freezer and not stored as live plant materials. For instance, one isolate was collected in 1960 by a colleague from the National Agriculture and Food Research Organization (Tsukuba, Tokyo) and stored in a deep freezer. Therefore, any mutations that might have occurred during viral propagation in susceptible hosts are likely to be small in number, so they would only have a minimal impact on our subsequent evolutionary analyses (Ohshima, Yamaguchi, and Hirota et al. 2002; Tan et al. 2005; Ohshima et al. 2010). Systemically infected leaves were homogenized in 0.01 M PPB, and the isolates were mechanically inoculated to young plants of *Raphanus* and *Brassica* plants. The host-infecting type of each isolate was examined as described previously (Ohshima, Yamaguchi, and Hirota et al. 2002). Inoculated plants were kept at 25°C for at least four weeks in a glasshouse.

2.2 Genome sequencing

Viral RNA was extracted from leaves of TuMV-infected turnip or *N. benthamiana* using Isogen II (Nippon Gene, Japan), then reverse transcribed by PrimeScript Moloney murine leukemia virus reverse transcriptase (Takara Bio, Japan) and amplified using high-fidelity Platinum *Pfx* DNA polymerase (Invitrogen/ThermoFisher Scientific, Japan). We separated the products by electrophoresis in agarose gels and purified them using the QIAquick Gel Extraction Kit (Qiagen K. K., Japan). We sometimes obtained

smaller unexpected bands possibly partial TuMV genomes or plant genomes, which would affect subsequent sequencing. Therefore, we removed these whenever they appeared.

The complete genome sequence of each isolate was determined using three or four independent Reverse transcription and polymerase chain reaction (RT-PCR) products with overlaps of 200–350 nucleotides. Each RT-PCR product was sequenced by primer walking using a BigDye Terminator version 3.1 Cycle Sequencing Ready Reaction kit (Life Technologies/ThermoFisher Scientific, Japan) with Applied Biosystems 3130 Genetic Analyzer, as described previously (Yasaka, Fukagawa, and Ikematsu et al. 2017). We assembled the sequence data using BioEdit version 5.0.9 (Hall 1999).

We sequenced the genomes of 317 isolates from Japan. Nearly all of these had lengths of 9,798–9,799 nucleotides, excluding 5′-end thirty-five nucleotide primer sequences used for amplification. All of the reported motifs in potyvirus polyprotein-encoded protein were found. Our data set included a total of 370 TuMV genome sequences from Japan, comprising the 317 determined in this study and fifty-three sequences available on GenBank. These were combined with representative sequences of global TuMV diversity that were published previously (Ohshima, Tomitaka, and Wood et al. 2007; Nguyen, Tomitaka, and Ho et al. 2013b; Yasaka, Ohba, and Schwinghamer et al. 2015; Yasaka, Fukagawa, and Ikematsu et al. 2017; Kawakubo, Gao, and Li et al. 2021).

The nucleotide sequences of the polyprotein-coding sequences were aligned via their amino acid translations using CLUSTAL X version 2 (Larkin, Blackshields, and Brown et al. 2007) with TransAlign (Weiller 1999). These alignments were then reassembled to form full genome sequences by adding the aligned 5′ and 3′ noncoding regions of RNA. The total dataset comprised 9,631 aligned nucleotides, excluding the thirty-five nucleotides that were used as primers for 5′-end RT-PCR amplification.

2.3 Recombination analysis

We inferred the phylogenetic relationships of the aligned full genomic sequences, including recombinant sequences, using the neighbor-net method in SPLITSTREE version 4.11.3 (Huson and Bryant 2006). Furthermore, all isolates collected in Japan were examined for evidence of recombination using RDP (Martin and Rybicki 2000), GENECONV (Sawyer 1999), BOOTSCAN (Salminen et al. 1995), MAXCHI (Smith 1992), CHIMAERA (Posada and Crandall 2001), and SISCAN (Gibbs, Armstrong, and Gibbs 2000), implemented in the RDP5 Beta 5.05 program (Martin, Varsani, and Roumagnac et al. 2020).

We first examined the full genomic sequences of Japanese isolates with those of published sequences for recombination events. We then examined the seven partial genomic sequences of nt 1–1,500 (5′ end to helper-component proteinase [HC-Pro] protein-coding region), nt 1–3,000 (5′ end to P3 protein-coding region), nt 1,501–4,500 (HC-Pro to cylindrical inclusion [CI] protein-coding regions), nt 3,001–6,000 (P3 to genome-linked viral protein [VPg] protein-coding regions), nt 4,501–7,500 (CI to nuclear inclusion b [NIb] protein-coding regions), nt 6,001–9,834 (VPg protein-coding region to 3′ end), and 7,501–9,834 (NIb protein-coding region to 3′ end). The partial and shorter sequences sometimes showed clearer evidence of recombination than those of full genomic sequences.

We performed the recombination analyses using default settings for the various programs and a Bonferroni-corrected *P* value cutoff of 0.01. After we examined all events with an associated *P* value of $<10^{-6}$ (i.e. the most likely recombination events), we discarded the likely interlineage recombinants from all of the data sets of full and partial genomic sequences. Focusing on the remaining intralineage recombinants allowed us to obtain clearer evidence of recombination. Finally, ambiguous recombination events were reconfirmed only using the sequence of the likely recombinant, two sequences of likely parental isolates, and two sequences of likely parental nonrecombinants obtained by RDP5 program.

We also examined the most likely recombination events and used a phylogenetic approach to identify the phylogenetic groups of the parent/donor sequences. We decided that evidence of a recombination event was 'clear' when it was detected (*P* value of $<10^{-6}$) by three or more different programs. To identify whether individual recombination events were from the same origin or not, we referenced the beginning or ending breakpoints of the 99 per cent confidence interval of likely recombinants and the selected major and minor parents by RDP5. Furthermore, we used SimPlot (Lole, Bollinger, and Paranjape et al. 1999) and original SISCAN version 2 (Gibbs, Armstrong, and Gibbs 2000) to confirm the recombination events of crossover sites using likely recombinant and the selected parental sequences by RDP5.

The estimated recombination-type patterns were then classified by the phylogenetic groups of the sequences in each genome: Asian-BR × basal-BR interlineage recombinant, Asian-BR × world-B3 interlineage recombinant, basal-BR intralineage recombinant, basal-BR × world-B3 interlineage recombinant, world-B2 × world-B3 interlineage recombinant, world-B3 intralineage recombinant, and Asian-BR × world-B3 × basal-BR interlineage recombinant. The recombination-type patterns were then plotted against the sampling year of each isolate by using ggplot2 (Wickham 2016). We then created recombination-free partial genomic sequences of partial HC-Pro (HC-Pro*), P3 (P3*), NIb (NIb*) and complete coat protein (CP) coding regions and checked that there was no phylogenetic incongruence for recombination events.

2.4 Phylogenetic analysis

We used maximum-likelihood to infer phylogenetic trees from full genomic sequences or polyprotein-coding sequences for three data sets: (1) 370 total Japanese isolates plus thirty previously reported nonrecombinant isolates, representing ten known TuMV phylogenetic groups and subgroups; (2) 370 total Japanese isolates only; and (3) 184 nonrecombinant sequences, comprising thirty-one Japanese isolates found in this study and 153 previously reported isolates collected globally (Kawakubo, Gao, and Li et al. 2021). The phylogeny was inferred using PhyML version 3.1 (Guindon et al. 2010) with the GTR + G + I substitution model, as chosen using the Akaike information criterion in jModeltest (Darriba et al. 2012). Node support was evaluated using bootstrapping with 1,000 replicates. To place the root of the TuMV phylogeny, we included outgroup genome sequences or polyprotein-coding sequences from narcissus late season yellows virus (NLSYV; accession numbers JQ326210, JX156421, and NC_023628), narcissus yellow stripe virus (NYSV; JQ395042, JQ911732, and NC_011541), Japanese yam mosaic virus (JYMV; AB016500 and KJ701427), wild onion symptomless virus (NC_030391), and scallion mosaic virus (ScaMV; NC_003399).

2.5 Phylogeographic analysis

We performed phylogeographic analyses based on four recombination-free genomic regions: HC-Pro*, P3*, NIb*, and CP (858–933 nucleotides) and polyprotein-coding sequences (9,486 nucleotides). For polyprotein-coding sequences, we assumed that the isolates sharing the same recombination-type pattern could

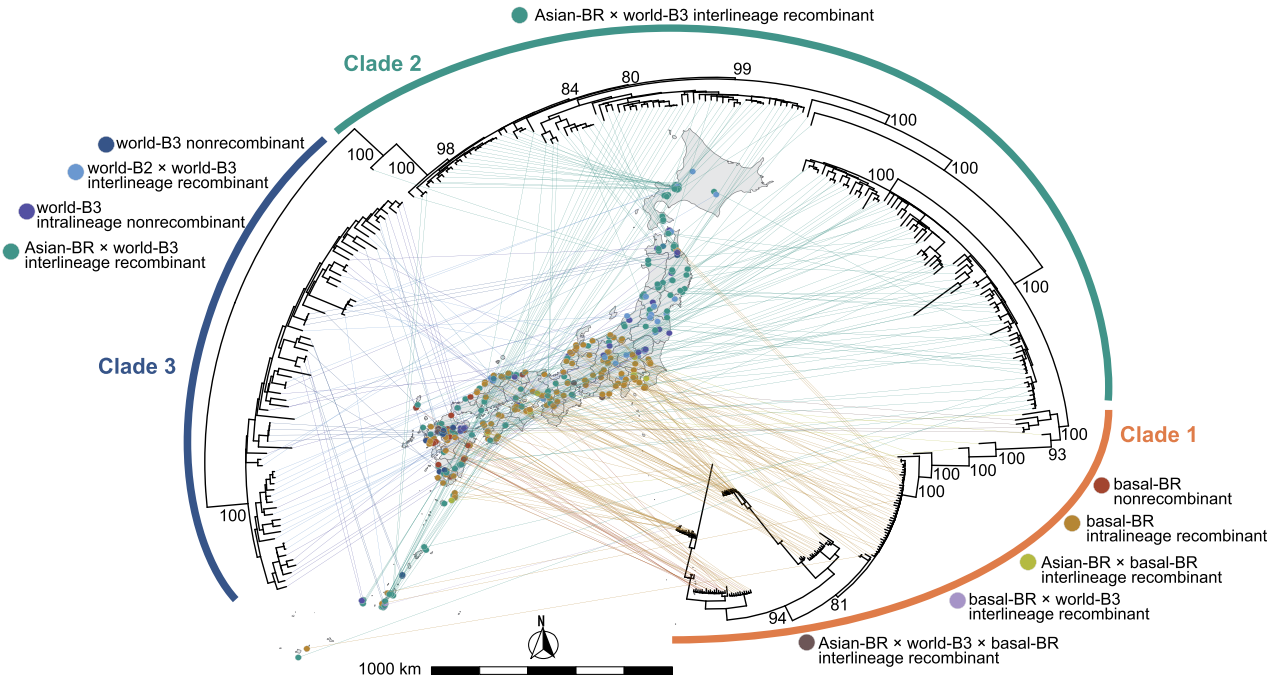

Figure 1. Sampling map of turnip mosaic virus (TuMV) Japanese isolates and maximum-likelihood tree inferred from polyprotein-coding sequences of 370 Japanese isolates. The aligned polyprotein-coding sequences are 9,486 nucleotides in length. Tree tips are connected to the sampling locations of the isolates. Outer lines show three major clades of Clade 1, 2, and 3. Colors of points on the map and connecting lines indicate recombination-type patterns, as shown in the key. Numbers at internal nodes indicate bootstrap percentages, based on 1,000 pseudoreplicates.

be treated as a monophyletic group. The phylogeographic dispersal history was inferred in both discrete and continuous space, using BEAST version 1.10.4 (Suchard et al. 2018). Posterior distributions of parameters were estimated using Markov chain Monte Carlo analyses, with samples drawn every 30,000 steps over 300 million steps. Each analysis was run at least twice independently to check for convergence. We checked for adequate mixing and sufficient sampling of each parameter (effective sample size ≥ 200) using Tracer version 1.7.1 (Rambaut et al. 2018). For our phylogeographic analysis in discrete space, each isolate was annotated with the sampling location based on recognized district divisions in Japan (Supplementary Fig. S1). Dispersal patterns were inferred using an asymmetric diffusion model and evaluated with Bayesian stochastic search variable selection (Lemey et al. 2009). We summarized the results in SpreaD3 (Bielejec et al. 2016) and accepted dispersals with Bayes factor ≥ 3 and posterior probability ≥ 0.5 as supported migration events. For our phylogeographic analyses in continuous space, we visualized the 95 per cent credible area of the inferred ancestral location using the *seraphim* package (Dellicour et al. 2016), based on a subsample of 100 post-burn-in trees.

For each recombination-type pattern, we analyzed dispersal speed using *seraphim*. We summarized the mean posterior estimates and 95 per cent credible intervals of mean branch dispersal history and visualized the epidemic wavefront from the root location.

3. Results

3.1 Virus isolates and host-infecting type

Our TuMV collection includes 317 isolates of TuMV throughout Japan between 1960 and 2017 (Supplementary Fig. S1), and we combined these with data from fifty-three Japanese isolates that we had collected for previous studies. The original hosts, sampling location, collecting year, and host-infecting type (pathogenicity) of the 370 isolates are summarized in Supplementary Table S1. It was easier to collect TuMV isolates from infected *Raphanus* plants (mainly Japanese radish: daikon), whose leaves generally show clear mosaic, rather than *Brassica* plants, which generally show mild symptoms in Japan. Hence, 295 of the 370 TuMV isolates (80 per cent) were collected from diseased *Raphanus* plants. Of the 317 Japanese isolates, 295 infected both *Brassica* and *Raphanus* plants, indicating that these isolates were *Brassica/Raphanus* [BR] host-infecting type. Ten of the isolates had been collected about half a century ago (1960–1974); these mostly infected *Brassica* plants but not *Raphanus* plants, indicating that these were *Brassica* [B] host-infecting type. On the other hand, three isolates that had been collected from *Brassica* plants only occasionally infected *Raphanus* plants, indicating that these isolates were B(R) host-infecting type.

3.2 Phylogenetic relationships inferred using nonrecombinant and recombinant sequences

A phylogenetic network was inferred using neighbor-net from the full genomic sequences including recombinant sequences (Supplementary Fig. S2). Since the network and our earlier analyses showed that recombination has been very common in TuMV, we analyzed the full genomic sequences using maximum-likelihood to infer the phylogeny of the 370 Japanese isolates (Supplementary Fig. S3). The tree shows that most Japanese isolates belong to the BR host-infecting type and are identified as recombinants (Supplementary Fig. S4).

The Japanese isolates of TuMV can be split into three major clades (Clades 1, 2, and 3), although these clades comprise both nonrecombinants and recombinants (Fig. 1). The basal-BR isolates and their relatives fell into Clade 1. The basal-BR group isolates infect both *Brassica* and *Raphanus* plants, and those are

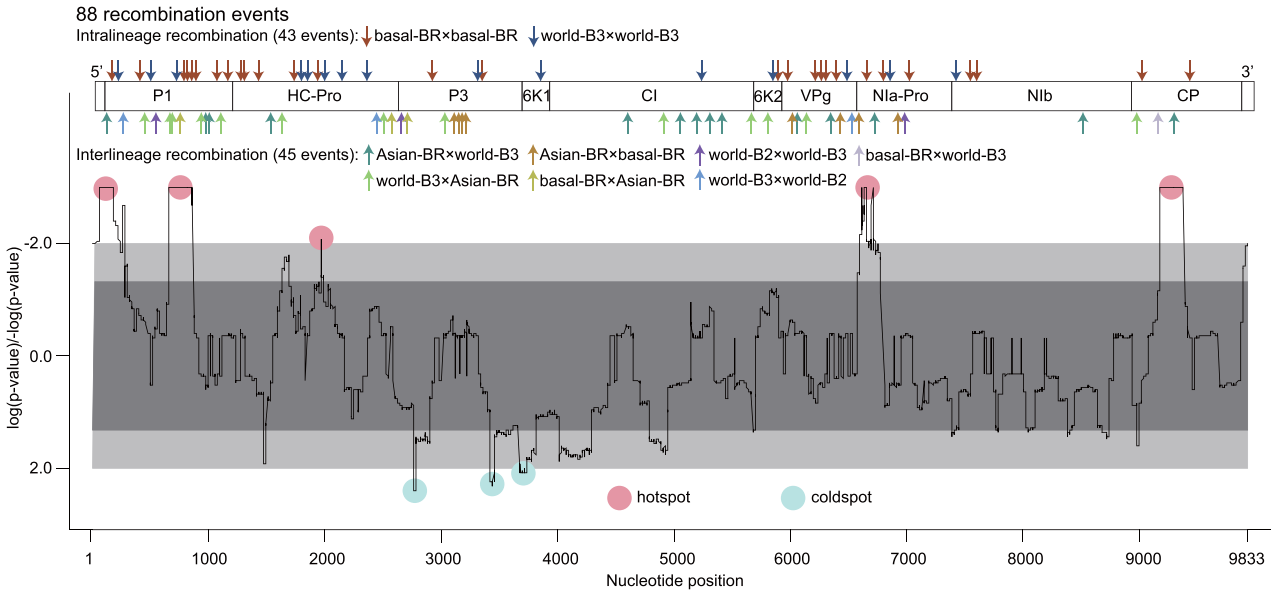

Figure 2. A total of eighty-eight recombination events found in Japanese turnip mosaic virus (TuMV) genome with recombination frequency distribution across the genome. *P* value for recombination frequency was plotted against TuMV genome corresponding to the nucleotide position of UK1 isolate (Jenner et al., 2000). Dark and light grey shadow indicates the range of 95 per cent and 99 per cent confidence intervals, respectively.

characterized as a highly virulent strain to *Raphanus*. The basal-BR group has been the epidemic phylogenetic group in China, Japan, and South Korea from 2000 (Tomimura et al. 2003). This group dispersed from Iran to Northwest China after 1975 and then to Central China (Kawakubo, Gao, and Li et al. 2021). The Asian-BR and world-B3 recombinants fell into Clade 2. The Asian-BR group isolates infect both *Brassica* and *Raphanus* plants, and this lineage might have originated in Turkey or Italy. Furthermore, Northeast China might have served as an important hub for the spread of TuMV to other parts of Asia, including Japan (Yasaka, Fukagawa, and Ikematsu et al. 2017; Kawakubo, Gao, and Li et al. 2021). The world-B3 group and its relatives fell into Clade 3. The spread of the world-B3 subgroup involved independent circulations on the Asian and European sides of Eurasia. Most European isolates do not infect *Raphanus* plants, whereas Asian isolates infect this plant species (Kawakubo, Gao, and Li et al. 2021). The tree reveals that world-B nonrecombinants and recombinants, and Asian-BR × world-B3 recombinants, are widely distributed in Japan. In contrast, basal-BR nonrecombinants and recombinants have a more limited distribution in the southwest of the country.

3.3 Recombination events

Our recombination analyses identified a large number of recombination events among the TuMV isolates (Supplementary Fig. S4). The large number of complete sequences in our data set allowed us to identify trends in recombination sites across the TuMV genome. We plotted the recombination frequency with corresponding hotspots and coldspots across the TuMV genome using 370 Japanese isolates (Fig. 2) and 914 worldwide isolates (Supplementary Fig. S5). The recombination hotspots were distributed in P1, HC-Pro, NIa-Pro, and CP protein-coding regions in Japanese TuMV genomes (Ohshima, Tomitaka, and Wood et al. 2007), whereas coldspots occurred in the P3 protein-coding region. The locations of hotspots and coldspots in the TuMV genomes were similar between those from Japan and worldwide.

A total of eighty-eight independent recombination events were found in the TuMV isolates from Japan (Fig. 2 and Supplementary Table S2A). Of these, eighty-four events were unique, whereas four events were shared with recombinants from outside Japan (Ohshima, Tomitaka, and Wood et al. 2007; Yasaka, Ohba, and Schwinghamer et al. 2015; Yasaka, Fukagawa, and Ikematsu et al. 2017; Kawakubo, Gao, and Li et al. 2021). The recombination events were widely distributed across the genome, and many intralineage recombination events were located in P1, HC-Pro, VPg, and NIb protein-coding regions. Interlineage recombination events were located throughout the genome except in the NIb protein-coding region. The intralineage recombination events for basal-BR and world-B3 did not share the same positions. A recombination event around nt 703–790 and nt 722–785 in the P1 protein-coding region was found in seventy-two and seventy-eight isolates and accounted for 47 per cent and 51 per cent of all world-B3 × Asian-BR interlineage recombinants. About 45 per cent and 29 per cent of Asian-BR × world-B3 interlineage recombinants showed evidence of a recombination event around nt 4,982–5,016 and nt 9,177–9,234 in the CI and CP coding regions, respectively. About 41 per cent of all Japanese isolates were Asian-BR × world-B3 interlineage recombinants.

3.4 Recombination-type patterns

We identified a total of thirty-four distinct type patterns, including seven intralineage and twenty-seven interlineage recombinants (Supplementary Table S2B). We found that 283 (76 per cent) isolates were multiple recombinants (more than one recombination event in the genome) (Supplementary Table S2C). Among the 370 Japanese isolates analyzed in our study, eighty-two recombination-type patterns were found in Japan. Only two of these were found outside Japan: one was also found in Taiwan (recombination-type pattern 71; RTP71), whereas the other was found in East Asia, Europe, and the USA (recombination-type pattern 81; RTP81) (Supplementary Fig. S4). Therefore, eighty out of eighty-two (98 per cent) recombination-type patterns are unique to Japan. We found that 141 were intralineage recombinants and 198 were interlineage recombinants (Supplementary Table S2B). The intralineage recombinants were

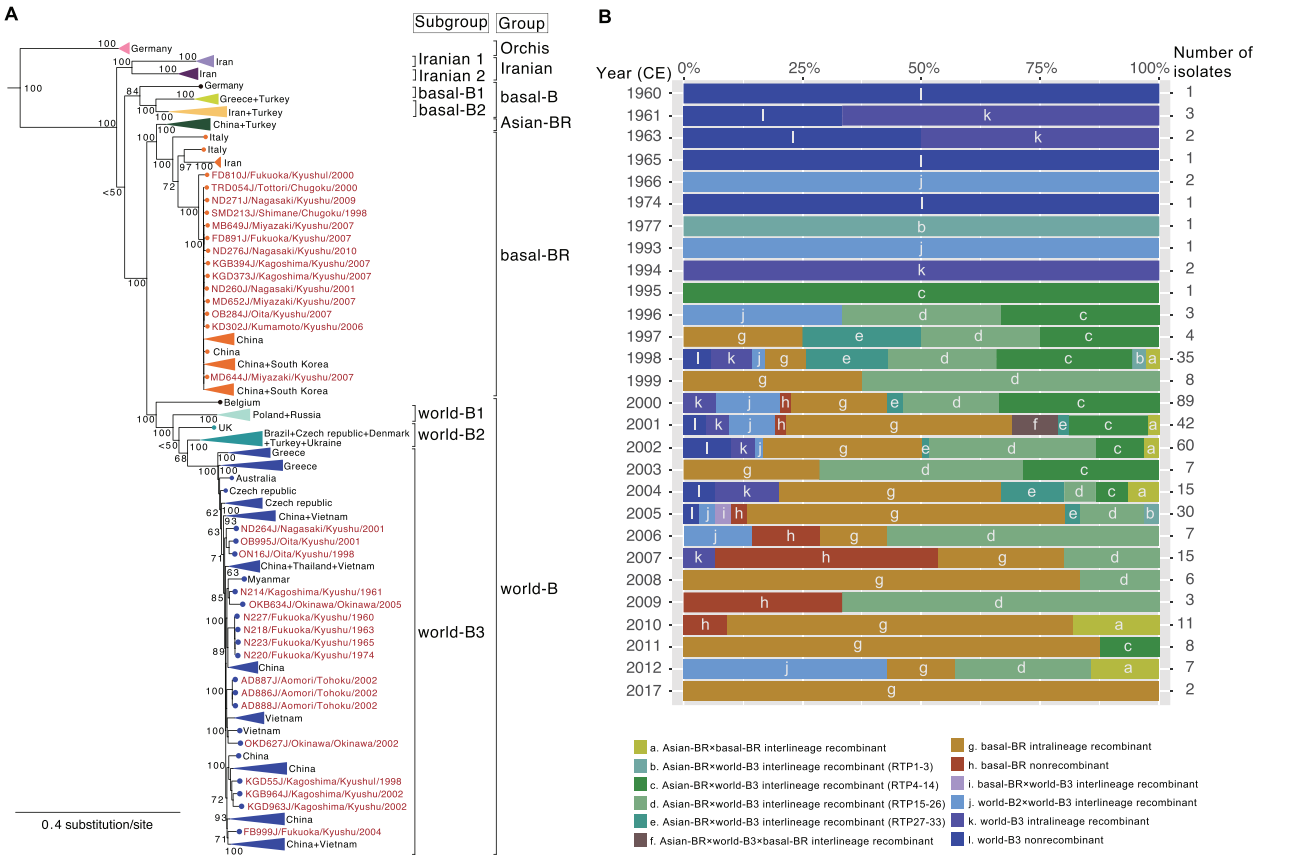

Figure 3. Phylogenetic tree of turnip mosaic virus (TuMV) inferred by maximum-likelihood from polyprotein-coding sequences of nonrecombinant isolates and sampling overview of Japanese isolates. (A) Tree of 184 nonrecombinant isolates including thirty-one Japanese nonrecombinant isolates newly found in this study. Of these, 153 isolates had been identified as nonrecombinant isolates in earlier studies (Yasaka, Ohba, and Schwinghamer et al. 2015; Yasaka, Fukagawa, and Ikematsu et al. 2017; Kawakubo, Gao, and Li et al. 2021). Numbers at internal nodes indicate bootstrap percentages, based on 1,000 pseudoreplicates. Triangles denote collapsed clades. Tip labels in red font indicate Japanese isolates, showing isolate name/prefecture of collection/district/year of collection. Tip labels in black font indicate isolates collected outside Japan, showing the country name of their collection. The polyprotein-coding sequences of the isolates of NLSYV (accession numbers JQ326210, JX156421, and NC_023628), NYSV (JQ395042, JQ911732, and NC_011541), JYMV (AB016500 and KJ701427), wild onion symptomless virus (NC_030391), and ScaMV (NC_003399) were used as outgroup taxa. The aligned polyprotein-coding sequences are 8,943 nucleotides in length. (B) Total numbers of TuMV nonrecombinant isolates and different recombination-type pattern (RTP) isolates in each sampling year. The classification of RTP is summarized in Supplementary Fig. S4. The Asian-BR × world-B3 interlineage recombinants are classified by the recombination type patterns as RTP1-3, RTP4-14, RTP15-26, and RTP27-33 based on the location of parental sequences of Asian-BR and world-B3.

derived from basal-BR or world-B3 parents, whereas interlineage recombinants were derived from Asian-BR, basal-BR, and/or world-B2 and world-B3 parents. The most dominant recombinants (14 per cent) in Japan were intralineage recombinants of basal-BR × basal-BR and basal-BR × basal-BR × basal-BR. In comparison, the interlineage recombinants of Asian-BR and world-B3 parents were found in 153 (41 per cent) of 370 isolates, with thirteen distinct type patterns, and those of world-B3 × Asian-BR × world-B3 (9 per cent) and Asian-BR × world-B3 × Asian-BR × basal-BR (12 per cent). Therefore, only thirty-one (8 per cent) isolates of basal-BR or world-B3 nonrecombinants were found in Japan, with 339 out of 370 (92 per cent) isolates being recombinants. The details of recombination-type patterns and the level of support for recombination events are shown in Supplementary Table S2A.

3.5 Phylogenetic relationships inferred using nonrecombinant sequences

Our previous maximum-likelihood phylogenetic analysis of nonrecombinant TuMV sequences revealed that the global diversity of the virus falls into six major phylogenetic groups. The TuMV population in Japan includes nonrecombinants from basal-BR group and world-B3 subgroup, but none in Orchis, Iranian, basal-B, and Asian-BR groups nor world-B1 or B2 subgroups (Fig. 3A). Japanese basal-BR isolates clustered with those from China and South Korea. Japanese world-B3 nonrecombinants split into seven minor sublineages, and Japanese isolates clustered with Chinese isolates, with Chinese and Vietnamese isolates, or on their own. Japanese nonrecombinant isolates of basal-BR group were distributed in Chugoku and Kyushu districts, whereas those of world-B3 group seemed to be mostly distributed in Kyushu island and some in Okinawa islands and the north Tohoku district. Thus, isolates from the world-B3 group are distributed in the north and south ends of Japan. The Japanese basal-BR and world-B3 nonrecombinants are largely found in Kyushu island.

3.6 Genetic populations in each sampling year

To characterize the annual transitions in the genetic composition of the TuMV population in Japan, we investigated and summarized total numbers of nonrecombinant and recombination-

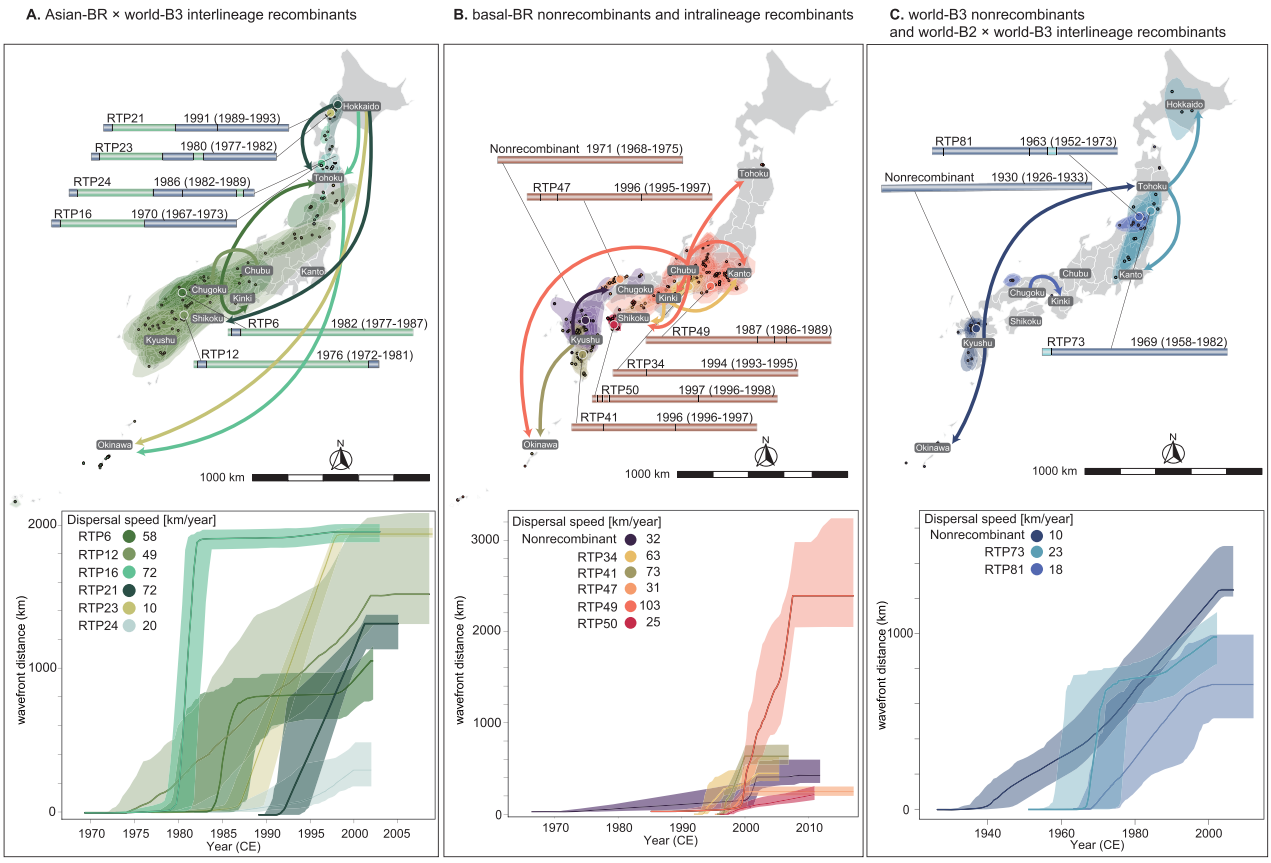

Figure 4. Dispersal history of nonrecombinants and recombinants of turnip mosaic virus. (A) Asian-BR × world-B3 interlineage recombinants; (B) basal-BR nonrecombinants and intralineage recombinants; and (C) world-B3 nonrecombinants and world-B2 × world-B3 interlineage recombinants. The polyprotein-coding sequences (aligned length 9,486 nucleotides) were used for this analysis. Upper panels show the year of the emergence and inferred dispersal routes. Inferred year of the emergence and 95 per cent credible intervals intervals are indicated above each recombination-type pattern (RTP). Large circles indicate the inferred root location for each recombination-type pattern, whereas small circles indicate the inferred locations of ancestral internal nodes and the sampling locations of the tips. Arrows indicate supported migration routes in discrete space. The 95 per cent credible interval areas of inferred internal nodes are projected based on 100 trees subsampled from the converged posterior samples and shown as different colored shadows by different recombination-type pattern. We only analyzed recombination-type patterns with more than six isolates, and the isolates of each recombination-type pattern are summarized in Supplementary Fig. S4. Lower panels show wavefront distance and mean dispersal speed averaged over their dispersal time. At each time point, the plot shows the spatial distance from the inferred root location of emergence. Dark lines and colored shadows indicate mean values and 95 per cent credible intervals, respectively. The dispersal speed of each recombination-type pattern in every ten years is summarized in Supplementary Table S7.

type pattern isolates in each sampling year (Fig. 3B). Although our Japanese TuMV collection started in 1960, world-B3 nonrecombinant and intralineage recombinants first appeared in the early 1960s, and the interlineage recombinants of world-B3 with world-B2 or Asian-BR recombinants then appeared. Therefore, the parental Asian-BR group first appeared in 1977 in Japan. We found basal-BR nonrecombinants and intralineage recombinants in 1997–2000. These formed an increasingly large proportion of the Japanese TuMV population to the present, while Asian-BR and world-B3 related recombinants have been decreasing.

3.7 Dispersal pathways

After the examination of temporal signal and model selection in each dataset (Supplementary Fig. S6, Tables S3 and S4), we inferred the dispersal pathways within Japan using a discrete phylogeographic model (Supplementary Fig. S7 and Table S5A). The analysis was based on recombination-free protein-coding regions of HC-Pro*, P3*, NIb*, and CP. For the Asian-BR phylogenetic group, we found support for seven dispersal pathways from Tohoku to Hokkaido, Chubu, Kinki, Chugoku, Shikoku, Kyushu, and Okinawa in HC-Pro* data set (Supplementary Fig. S7A). The dispersal pathways from Tohoku to the other parts of Japan were also supported by P3* and NIb* data sets (Supplementary Fig. S7B–C); hence, Tohoku district acted as a hub for the dispersal of Asian-BR.

For the basal-BR phylogenetic group, two dispersal pathways from Chubu to Kanto and Shikoku were supported by HC-Pro*, NIb*, and CP data sets. Dispersal pathways of the world-B3 phylogenetic group inferred from HC-Pro* and P3* data set seemed different from NIb* and CP data sets. The analysis of the NIb* and CP data set from the world-B3 phylogenetic group highlighted the role of Tohoku district as a hub, with six supported dispersal pathways to Hokkaido, Kanto, Chubu, Chugoku, Kyushu, and Okinawa by both data set (Supplementary Fig. S7C–D). However, the analysis of HC-Pro* and P3* data sets supported only a few dispersion from Tohoku district, while one dispersion from Kyushu to Kinki was supported by both HC-Pro* and P3* data sets. (Supplementary Fig. S7A–B).

We also examined the dispersal of TuMV using a continuous phylogeographic model based on sequence data from HC-Pro*, P3*, NIb*, and CP (Supplementary Fig. S8 and Table S6).

The analysis of HC-Pro* showed that Asian-BR isolates were introduced around the mid-1900s and then spread from multiple areas such as Tohoku, Kinki, and Kyushu districts until the 2000s. The basal-BR isolates spread rapidly in the 2000s, and their distribution is limited in the south and west parts of Japan. The world-B3 isolates spread from the early 1900s to the 2000s and their introduction also occurred in multiple areas of Japan, including Tohoku, Chugoku, and Kyushu. However, analyses using different protein-coding regions yielded results that had some discrepancies. For instance, the analysis using CP showed that the Asian-BR isolates have only spread to Honshu and Shikoku islands, while the analysis using HC-Pro*, P3*, and NIb* showed the dispersion to Hokkaido, Kyushu, and Okinawa islands.

3.8 Dating of initial recombinant dispersal

We inferred the initial dispersal time of thirteen recombination-type patterns that were found in more than six isolates in Japan (Fig. 4 and Supplementary Fig. S9). For Asian-BR × world-B3 interlineage recombinants, the initial dispersal times of recombination-type patterns were dated to around 1970–1990. Moreover, recombination events that were shared across different recombination-type patterns indicated that they evolved from similar ancestral recombinants. For example, the isolates of four recombination-type patterns (RTP16, 21, 23, and 24) shared the recombination event found around nt 722–785 in the P1 protein-coding region. Among these, recombination-type pattern 16 (RTP16) was the oldest, dated at 1970 CE (95 per cent credible interval: 1967–1973). Therefore, this recombination event occurred at least as early as the 1970s, and the other recombination-type patterns evolved from the same ancestral recombinants afterwards.

In contrast, the initial dispersal times of basal-BR intralineage recombinants were inferred as being more recent than those of Asian-BR × world-B3 interlineage recombinants, which were around 1990–2000. The world-B2 × world-B3 interlineage recombinants dispersed around 1950–1980. In addition, we inferred the age of the most recent common ancestors of nonrecombinants of basal-BR and world-B3 collected in Japan. The nonrecombinants of Japanese basal-BR dispersed in 1971 (95 per cent credible interval: 1968–1975) and those of world-B3 dispersed in 1930 (95 per cent credible interval: 1926–1933).

3.9 Spatial dynamics of recombinant and nonrecombinant isolates

Besides
recombination-free partial genomic sequences, we inferred the dispersal history of thirteen recombination-type patterns (Fig. 4). For Asian-BR × world-B3 interlineage recombinants, we examined six recombination-type pattern isolates. Most of them were distributed in a wide area of Japan, whereas some recombinants were only locally distributed. The distributions of basal-BR intralineage recombinants were narrower than those of Asian-BR × world-B3 interlineage recombinants. The isolates of basal-BR intralineage recombinants have not spread to the north of Kanto district even though there were three isolates found in Tohoku district. Although world-B2 × world-B3 interlineage recombinants were both distributed mainly in the north of Honshu island, world-B3 nonrecombinants occurred in Kyushu and Okinawa districts.

We also inferred dispersal pathways of recombinant isolates in discrete space (Fig. 4 and Supplementary Table S5B). For Asian-BR × world-B3 interlineage recombinants, we found support for eight dispersal pathways. These dispersals did not appear to be dependent on distance; for instance, recombination-type pattern 23 (RTP23) spread from Hokkaido to Okinawa islands and recombination-type pattern 6 (RTP6) spread from Chugoku to Tohoku districts. For basal-BR intralineage recombinants, we found support for eight dispersal pathways in addition to one dispersal of nonrecombinants from Kyushu Island. The recombination-type pattern 49 (RTP49) spread from Chubu district to the rest of Japan along five pathways, but dispersals of the other basal-BR intralineage recombination-type patterns were more related to geographic distance than those of Asian-BR × world-B3 interlineage recombinants. The world-B2 × world-B3 recombination-type pattern 73 (RTP73) spread from Tohoku to Hokkaido and Kanto district, while recombination-type pattern 81 (RTP81) only spread from Chugoku to Kinki. The isolates sharing the same recombination-type patterns tended to be geographically localized, in some cases being limited to a single coded location; these recombination-type patterns were not amenable to discrete phylogeographic analysis (Supplementary Table S5B).

We also evaluated the dispersal speed of each recombination-type pattern (Fig. 4 and Supplementary Table S7). The interlineage recombinants of Asian-BR × world-B3 spread around 1970–2000, and their mean dispersal speeds were 10–72 km/year. In contrast, basal-BR intralineage recombinants spread during 1995–2010, and their mean dispersal speed was 25–103 km/year. The world-B2 × world-B3 interlineage recombinants spread after 1950, with a lower dispersal speed of 18–23 km/year. We interpret these estimates with caution, given that they might have been affected by the large disparities in sample sizes.

4. Discussion

Our epidemiological genomic analysis of TuMV represents the most detailed study of a plant pathogen on a whole-country scale so far. In particular, the large number of complete sequences we have generated allowed us to identify trends in the pattern and frequency of occurrence of recombination sites and hotspot regions. The first report of evidence on TuMV recombination was published in 2002 (Ohshima, Yamaguchi, and Hirota et al. 2002). Since then, several reports (Tomimura et al. 2003; Tan et al. 2004; Tomimura, Špak, and Katis et al. 2004; Ohshima, Tomitaka, and Wood et al. 2007; Yasaka, Ohba, and Schwinghamer et al. 2015; Yasaka, Fukagawa, and Ikematsu et al. 2017; Kawakubo, Gao, and Li et al. 2021) on recombination in TuMV provided evidence for interlineage as well as intralineage recombination. This study has identified eighty-eight independent recombination events, and eighty-two recombination-type patterns with two basal-BR group and world-B3 subgroup were nonrecombinants, unique to the TuMV population in Japan; characterized the composition of nonrecombinant and recombination-type pattern isolates in each sampling year; and inferred the routes and timing of dispersals of this plant pathogen in Japan. Our results also indicate that those patterns were generated in a short period, and these findings agree with our previous studies in showing that recombination in TuMV is locally specific (Tan et al. 2004; Ohshima, Tomitaka, and Wood et al. 2007; Yasaka, Ohba, and Schwinghamer et al. 2015).

There have been other reports of phylogeographic analyses of island systems. In another potyvirus species, zucchini yellow mosaic virus, all of the isolates in Papua New Guinea were found to be recombinants but possibly arose from a single introduction (Maina et al. 2019). In rice yellow mottle virus, a species of sobemovirus, isolates of a unique recombination-type pattern

in Madagascar were most likely introduced from eastern Africa (Rakotomalala, Vrancken, and Pinel-Galzi et al. 2019).

Overall, TuMV seems to be excessively recombinogenic, with recombinant lineages occurring in high frequency compared with other plant viruses. In Asia, *Raphanus* vegetables are widely grown compared with other parts in the world, and *Brassica* and *Raphanus* vegetables are generally grown separately in farm fields. One distinctive feature of *Brassica* and *Raphanus* production in Japan is that these vegetables are often mixed planted in home gardens and small fields compared with other East Asian and South-East Asian countries. Those vegetables are mainly *Brassica* (Chinese cabbage, *B. pekinensis*; Japanese mustard spinach, *B. rapa* var. *perviridis*; leaf mustard, *B. juncea* var. *integrifolia*; cabbage, *B. oleracea* var. *capitata*) and *Raphanus* (*R. sativus*). This situation might be conducive to recombination in TuMV because of the increased likelihood that the viruses could encounter another isolate from various host plants.

Our oldest collections of TuMV isolates from 1960 to 1965 were all world-B3 nonrecombinants or world-B3 intralineage recombinants (Fig. 3B and Supplementary Table S1), indicating that the first introductions to Japan might have been from the world-B3 phyloge-
netic group. This seems to have occurred in the 1900s (95 per cent credible interval: 1924–1986) from the UK (Kawakubo, Gao, and Li et al. 2021), although further sampling of genomes from archival material will refine this date estimate. The first introduction of world-B3 group isolates corresponds well with our inference that this population is the oldest in Japan (Figs 3B and 4, and Supplementary Figs S7 and S8, and Table S5). Our results also reveal the rapid transition of the dominant population from multiple introductions of each phylogenetic group of TuMV.

The Asian-BR nonrecombinants were found mostly in China and Turkey (Kawakubo, Gao, and Li et al. 2021). However, no nonrecombinants and intralineage recombinants have yet been found in Japan, while many Asian-BR × world-B3 recombination-type pattern isolates were found over the forty years of sampling (Figs 3B and 4A), as well as in South and Southeast Asia (Nguyen, Tran, and Ohshima et al. 2013a; Kawakubo, Gao, and Li et al. 2021).

The reason why the basal-BR group nonrecombinants and intralineage recombinants were suddenly introduced to Japan is unknown. The group probably originated in Italy and spread to other parts from the 18th to 20th century (Kawakubo, Gao, and Li et al. 2021). The basal-BR group isolates were first found in *R. sativus* (Japanese radish) plants with severe symptoms before the 2000s in Japan (Tomitaka and Ohshima 2006; Tomitaka, Yamashita, and Ohshima 2007), then as a highly virulent strain in Chinese radish after the 2000s in China (Wang et al. 2009; Zhu, Sun, and Wang et al. 2016; Li, Zhu, and Yin et al. 2017) and South Korea (Gong et al. 2019), and more recently in the USA (Tomimura et al. 2003; Nguyen, Tomitaka, and Ho et al. 2013b) and Brazil (Rodrigues, Chaves, and Kitajima et al. 2021). Our present study (Figs 3B and 4B) corresponded well with the earlier studies of timescale of basal-BR introduction to Japan. In Japan, the spread of basal-BR group isolates has been limited to the southwest, unlike Asian-BR and world-B3 group isolates (Figs 1 and 4). This suggests that the basal-BR group is an emerging group not only in Japan but also across the globe and is now becoming an epidemic strain.

This study has provided evidence that the emergences of phylogenetic groups occurred in different periods in Japan. The introduction of TuMV into Japan is unlikely to have occurred, while the country was in its isolationist state (1639–1854 CE), when it was closed to foreign trades except for the Netherlands. Toward the end of the Edo Period (1603-1868 CE), the Japan–US Treaty of Peace and Amity and similar treaties reopened the ports to international trade. There were several introductions of TuMV to Japan, and these are likely to have been dependent on the introduction and cultivation of the host plants (Shimizu 2008, 2017). Although the long-distance spread of TuMV is related to human-mediated trade networks (Yasaka, Ohba, and Schwinghamer et al. 2015; Kawakubo, Gao, and Li et al. 2021), aphid transmission is likely to be a driver at smaller spatial scales (Adachi et al. 2018).

In conclusion, our study has confirmed that recombination has occurred frequently in the TuMV genome. Recombination is considered to be a major determinant of change in viral virulence and has been implicated in the emergence of new viral biological strains. Recombination is common in DNA and RNA viruses of both animals and plants (Worobey and Holmes 1999; Froissart, Roze, and Uzest et al. 2005; Ohshima, Tomitaka, and Wood et al. 2007; Martin, Lemey, and Posada 2011; Yasaka, Nguyen, and Ho et al. 2014; Yasaka, Ohba, and Schwinghamer et al. 2015; Rakotomalala, Vrancken, and Pinel-Galzi et al. 2019; Gong, Sui, and Li 2022). RNA viruses are characterized by high mutation and recombination rates, which allow rapid adaptation to new environments.

Our study has also demonstrated how virus recombination can be integrated into a phylogeographic framework, while showing how genomic diversity has been shaped by host and ecological constraints in a relatively closed spatial system. This information provides a platform for designing strategies to mitigate the evolution and spread of virus variants in agricultural crops in Japan. Future analyses would benefit from the development of more sophisticated methods that can explicitly incorporate the recombination process into phylogeographic analysis.

Data availability

Genomic data are available on GenBank with accession numbers LC639359–LC639675.

Supplementary data

Supplementary data are available at *VEVOLU* online.

Acknowledgements

We thank Tamaki Hamamoto, Yuka Yamaguchi, Ryo Hirota, Akemi Sato, Yoshiteru Minematsu, Toshio Shinohara, Kayo Maeda, Eri Okuzono, Kumiko Miki, Rika Kodama, Rie Koba, Hiromi Setoguchi, Seiko Tomari, Mutsumi Kaminaga, Eri Muraguchi, Yui Nakamizu, Masatoshi Yabu, Toshiyuki Furukawa, Hiromi Eguchi, Mika Imamura, Atsuki Seto, Mika Kajita, Yumi Nagano, Kiho Ohba, Kanako Motomura, Asuka Baba, Hanako Taira, Sumire Asada, and the students in the Laboratory of Plant Virology (Saga University) for their careful technical assistance to examine host reactions and determine the genomic sequences. Some of the computational analyses were performed on the National Institute of Genetics (NIG) supercomputer at the Research Organization of Information and Systems (ROIS) National Institute of Genetics. All the genomic sequences of turnip mosaic virus isolates were determined at the Laboratory of Plant Virology and Analytical Research Center for Experimental Sciences, Saga University.

Funding

This work was in part funded by the Japanese Society for the Promotion of Science KAKENHI Grant Numbers 24405026, 18KT0092,

and 21K05601. The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest: The authors declare no conflict of interest.

Author contributions

K.O. designed research; K.O. and S.K. wrote original draft; S.K., Y.T., K.T., H.M., R.K., S.U., K.Y., and K.O. performed research; S.K. and K.O. analyzed data; and S.K., S.Y.W.H., and K.O. edited the paper.

References

Adachi, S. et al. (2018) 'Effects of Infection by *Turnip Mosaic Virus* on the Population Growth of Generalist and Specialist Aphid Vectors on Turnip Plants', *PLoS One*, 13: e0200784.

Aoba, T. (2013) Daikon, in Nihon No Yasai Bunkashi Jiten, Rural Culture Association, p148, pp504, ISBN978-4-89694-160-9 (In Japanese).

Bertrand, Y. et al. (2012) 'First Dating of a Recombination Event in Mammalian Tick-borne Flaviviruses', *PLoS One*, 7: e31981.

Bielejec, F. et al. (2016) 'SpreaD3: Interactive Visualization of Spatiotemporal History and Trait Evolutionary Processes', *Molecular Biology and Evolution*, 33: 2167–9.

Chare, E. R., and Holmes, E. C. (2006) 'A Phylogenetic Survey of Recombination Frequency in Plant RNA Viruses', *Archives of Virology*, 151: 933–46.

Chung, B. Y.-W. et al. (2008) 'An Overlapping Essential Gene in the Potyviridae', *Proceedings of the National Academy of Sciences of the United States of America*, 105: 5897–902.

Clark, M. F., and Adams, A. N. (1977) 'Characteristics of the Microplate Method of Enzyme-Linked Immunosorbent Assay for the Detection of Plant Viruses', *Journal of General Virology*, 34: 475–83.

Darriba, D. et al. (2012) 'jModelTest 2: More Models, New Heuristics and Parallel Computing', *Nature Methods*, 9: 772.

Dellicour, S. et al. (2017) 'Using Viral Gene Sequences to Compare and Explain the Heterogeneous Spatial Dynamics of Virus Epidemics', *Molecular Biology and Evolution*, 34: 2563–71.

——— et al. (2016) 'SERAPHIM: Studying Environmental Rasters and Phylogenetically Informed Movements', *Bioinformatics*, 32: 3204–6.

Ding, N.-Z. et al. (2017) 'A Permanent Host Shift of Rabies Virus from *Chiroptera* to *Carnivora* Associated with Recombination', *Scientific Reports*, 7: 289.

Fiallo-Olivé, E. et al. (2019) 'Recurrent Speciation of a Tomato Yellow Leaf Curl Geminivirus in Portugal by Recombination', *Scientific Reports*, 9: 1332.

Froissart, R. et al. (2005) 'Recombination Every Day: Abundant Recombination in a Virus during a Single Multi-cellular Host Infection', *PLoS Biology*, 3: e89.

Gardner, M. W., and Kendrick, J. B. (1921) 'Turnip Mosaic', *Journal of Agricultural Research*, 22: 123–4.

Gibbs, A. J. et al. (2020) 'The Potyviruses: An Evolutionary Synthesis Is Emerging', *Viruses*, 12: 132.

Gibbs, A. J., and Ohshima, K. (2010) 'Potyviruses and the Digital Revolution', *Annual Review of Phytopathology*, 48: 205–23.

Gibbs, M. J., Armstrong, J. S., and Gibbs, A. J. (2000) 'Sister-Scanning: A Monte Carlo Procedure for Assessing Signals in Recombinant Sequences', *Bioinformatics*, 16: 573–82.

Gong, J. et al. (2019) 'Sequence Variations among Seventeen New Radish Isolates of *Turnip Mosaic Virus* Showing Differential Pathogenicity and Infectivity in *Nicotiana Benthamiana, Brassica Rapa*, and *Raphanus Sativus*', *Phytopathology*, 109: 904–12.

Gong, Y., Sui, L., and Li, Y. (2022) 'Recombination in Papillomavirus: Controversy and Possibility', *Virus Research*, 314: 198756.

Guindon, S. et al. (2010) 'New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0', *Systematic Biology*, 59: 307–21.

Hall, T. A. (1999) 'BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT', *Nucleic Acids Symposium Series*, 41: 95–8.

Hostager, R. et al. (2019) 'Hepatitis C Virus Genotype 1 and 2 Recombinant Genomes and the Phylogeographic History of the 2k/1b Lineage', *Virus Evolution*, 5: vez041.

Huson, D. H., and Bryant, D. (2006) 'Application of Phylogenetic Networks in Evolutionary Studies', *Molecular Biology and Evolution*, 23: 254–67.

Inoue-Nagata, A. K. et al. (2022) 'ICTV Virus Taxonomy Profile: *Potyviridae* 2022', *The Journal of General Virology*, 103: 001738.

Ishiyama, S., and Misawa, T. (1942) 'Stunt Disease of Japanese Radish', *Japanese Journal of Phytopathology*, 12: 116–30.

Jenner, C.E. et al. (2000) 'The cylindrical inclusion gene of turnip mosaic virus encodes a pathogenic determinant to the brassica resistance gene TuRB01', *Molecular Plant-Microbe Interactions*, 13: 1102–1108.

Kawakubo, S. et al. (2021) 'Genomic Analysis of the Brassica Pathogen Turnip Mosaic Potyvirus Reveals Its Spread along the Former Trade Routes of the Silk Road', *Proceedings of National Academy of Sciences of the United States of America*, 118: e2021221118.

Larkin, M. A. et al. (2007) 'Clustal W and Clustal X Version 2.0', *Bioinformatics*, 23: 2947–8.

Lefeuvre, P. et al. (2010) 'The Spread of Tomato Yellow Leaf Curl Virus from the Middle East to the World', *PLoS Pathogens*, 28: e1001164.

Lemey, P. et al. (2009) 'Bayesian Phylogeography Finds Its Roots', *PLoS Computational Biology*, 5: e1000520.

Li, X. et al. (2017) 'The Genetic Structure of *Turnip Mosaic Virus* Population Reveals the Rapid Expansion of a New Emergent Lineage in China', *Virology Journal*, 14: 165.

Lole, K. S. et al. (1999) 'Full-Length Human Immunodeficiency Virus Type 1 Genomes from Subtype C-Infected Seroconverters in India, with Evidence of Intersubtype Recombination', *Journal of Virology*, 73: 152–60.

Maina, S. et al. (2019) '*Zucchini Yellow Mosaic Virus* Genomic Sequences from Papua New Guinea: Lack of Genetic Connectivity with Northern Australian or East Timorese Genomes, and New Recombination Findings', *Plant Disease*, 103: 1326–36.

Martin, D., and Rybicki, E. (2000) 'RDP: Detection of Recombination Amongst Aligned Sequences', *Bioinformatics*, 16: 562–3.

Martin, D. P., Lemey, P., and Posada, D. (2011) 'Analysing Recombination in Nucleotide Sequences', *Molecular Ecology Resources*, 11: 943–55.

Martin, D. P. et al. (2020) 'RDP5: A Computer Program for Analyzing Recombination In, and Removing Signals of Recombination From, Nucleotide Sequence Datasets', *Virus Evolution*, 7: veaa087.

Nellist, C. F. et al. (2022) 'Turnip Mosaic Virus, a Virus for All Seasons', *Annals of Applied Biology*, 180: 312–27.

Nguyen, H. D. et al. (2013b) 'Turnip Mosaic Potyvirus Probably First Spread to Eurasian Brassica Crops from Wild Orchids about 1000 Years Ago', *PLoS One*, 8: e55336.

——— et al. (2013a) 'Genetic Variation of the *Turnip Mosaic Virus* Population of Vietnam: A Case Study of Founder, Regional and Local Influences', *Virus Research*, 171: 138–49.

Nishi, S. (1981) Hakuran, an Interspecific Hybrid between Chinese Cabbage and Common Cabbage. In *Chinese cabbage. Proceedings*

of the first international symposium, pp. 385–91. AVRDC, Shanhua, Taiwan.

Ohshima, K. et al. (2010) 'The Evolutionary Trajectory of Turnip Mosaic Virus Populations Adapting to a New Host', *Journal of General Virology*, 91: 788–801.

——— et al. (2007) 'Patterns of Recombination in Turnip Mosaic Virus Genomic Sequences Indicate Hotspots of Recombination', *Journal of General Virology*, 88: 298–315.

——— et al. (2002) 'Molecular Evolution of *Turnip Mosaic Virus*: Evidence of Host Adaptation, Genetic Recombination and Geographical Spread', *Journal of General Virology*, 83: 1511–21.

Posada, D., and Crandall, K. A. (2001) 'Evaluation of methods for detecting recombination from DNA sequences: Computer simulations', *Proceedings of National Academy of Sciences of the United States of America*, 98: 13757–13762.

Rakotomalala, M. et al. (2019) 'Comparing Patterns and Scales of Plant Virus Phylogeography: *Rice Yellow Mottle Virus* in Madagascar and in Continental Africa', *Virus Evolution*, 5: vez023.

Rambaut, A. et al. (2018) 'Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7', *Systematic Biology*, 67: 901–4.

Rodrigues, L. K. et al. (2021) 'Characterisation of Turnip Mosaic Virus Isolates Reveals High Genetic Variability and Occurrence of Pathotype 1 in Brazil', *European Journal of Plant Pathology*, 160: 883–900.

Sako, N. (1981) 'Virus Disease of Chinese Cabbage in Japan'. In: Talekar, N. S., and Griggs, T. D. (eds) *Chinese Cabbage. Proceedings of the First International Symposium*, pp. 129–41. Shanhua, Taiwan: AVRDC.

Salminen, M. O. et al. (1995) 'Identification of Breakpoints in Intergenotypic Recombinants of HIV Type 1 by Bootscanning', *AIDS Research and Human Retroviruses*, 11: 1423–5.

Sawyer, S. A. (1999), *GENECONV: A Computer Package for the Statistical Detection of Gene Conversion. Distributed by the author (Department of Mathematics, Washington University, St Louis, MO)*. <https://www.math.wustl.edu/~sawyer> accessed 1 June 2022.

Schierup, M. H., and Hein, J. (2000) 'Consequences of Recombination on Traditional Phylogenetic Analysis', *Genetics*, 156: 879–91.

Schluter, D. et al. (1997) 'Likelihood of Ancestor States in Adaptive Radiation', *Evolution*, 51: 1699–711.

Schultz, E. S. (1921) 'A Transmissible Mosaic Disease of Chinese Cabbage, Mustard, and Turnip', *Journal of Agricultural Research*, 22: 173–8.

Shimizu, K. (2008) 'Formation of Cabbage Production Areas and Fixation of the Cabbage Dietary Habit in Japan from the Late Meiji Era to the Early Showa Era', *Geographical Review of Japan*, 81: 1–24.

——— (2017) Spread of Chinese Cabbage in Japan during Taisho Era: Focusing on Growing Demand and Seed Supply System, Tsukuba Univ. PhD thesis. pp 220.

Smith, J. M. (1992) 'Analyzing the Mosaic Structure of Genes', *Journal of Molecular Evolution*, 34: 126–9.

Smith, K. M. (1935) 'A New Virus Disease of the Tomato', *Annals of Applied Biology*, 22: 239–42.

Suchard, M. A. et al. (2018) 'Bayesian Phylogenetic and Phylodynamic Data Integration Using BEAST 1.10', *Virus Evolution*, 4: vey016.

Takimoto, S. (1930) 'On the Mosaic Disease of Chinese Cabbage and Other Crucifers', *Japan Horticultural Magazine*, 42: 5–7.

Tan, Z. et al. (2005) 'Mutations in *Turnip Mosaic Virus* Genomes that Have Adapted *Raphanus Sativus*', *Journal of General Virology*, 86: 501–10.

——— et al. (2004) 'Inter- and Intralineage Recombinants are Common in Natural Populations of *Turnip Mosaic Virus*', *Journal of General Virology*, 85: 2683–96.

Tomimura, K. et al. (2003) 'The Phylogeny of *Turnip Mosaic Virus* ; Comparisons of 38 Genomic Sequences Reveal a Eurasian Origin and a Recent 'emergence' in East Asia', *Molecular Ecology*, 12: 2099–111.

——— et al. (2004) 'Comparisons of the Genetic Structure of Populations of *Turnip Mosaic Virus* in West and East Eurasia', *Virology*, 330: 408–23.

Tomitaka, Y., and Ohshima, K. (2006) 'A Phylogeographical Study of the *Turnip Mosaic Virus* Population in East Asia Reveals an 'Emergent' Lineage in Japan', *Molecular Ecology*, 15: 4437–57.

Tomitaka, Y., Yamashita, T., and Ohshima, K. (2007) 'The Genetic Structure of Populations of *Turnip Mosaic Virus* in Kyushu and Central Honshu, Japan', *Journal of General Plant Pathology*, 73: 197–208.

Visser, J. C., Bellstedt, D. U., and Pirie, M. D. (2012) 'The Recent Recombinant Evolution of a Major Crop Pathogen, *Potato Virus Y*', *PLoS One*, 7: e50631.

Walsh, J. A., and Jenner, C. E. (2002) 'Turnip Mosaic Virus and the Quest for Durable Resistance', *Molecular Plant Pathology*, 3: 289–300.

Wang, H.-Y. et al. (2009) 'Complete Genomic Sequence Analyses of *Turnip Mosaic Virus* Basal-BR Isolates from China', *Virus Genes*, 38: 421–8.

Weiller, G. F. (1999) *TransAlign (Version 1.0, Research School of Biological Sciences*. Canberra: Australia.

Wickham, H. (2016) *Ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Wittmann, T. J. et al. (2007) 'Isolates of Zaire Ebolavirus from Wild Apes Reveal Genetic Lineage and Recombinants', *Proceedings of the National Academy of Sciences of the United States of America*, 104: 17123–7.

Worobey, M., and Holmes, E. C. (1999) 'Evolutionary Aspects of Recombination in RNA Viruses', *Journal of General Virology*, 80: 2535–43.

Yasaka, R. et al. (2017) 'The Timescale of Emergence and Spread of Turnip Mosaic Potyvirus', *Scientific Reports*, 7: 4240.

——— et al. (2014) 'The Temporal Evolution and Global Spread of *Cauliflower Mosaic Virus*, a Plant Pararetrovirus', *PLoS One*, 9: e85641.

——— et al. (2015) 'Phylodynamic Evidence of the Migration of Turnip Mosaic Potyvirus from Europe to Australia and New Zealand', *Journal of General Virology*, 96: 701–13.

Yoshii, H. (1951) 'Mosaic Disease of Crucifers in West Japan', *Shokubutsu Byogai Kenkyu, Kyoto*, 4: 17–22.

——— (1963) 'On the Strain Distribution of Turnip Mosaic Virus', *Annals of the Phytopathological Society of Japan*, 28: 221–7.

Zhang, J. et al. (2020) 'Genetic Diversity, Phylogeography, and Evolutionary Dynamics of Highly Pathogenic Avian Influenza A (H5N6) Viruses', *Virus Evolution*, 6: veaa079.

Zhu, F. et al. (2016) 'Molecular Characterization of the Complete Genome of Three basal-BR Isolates of *Turnip Mosaic Virus* Infecting *Raphanus Sativus* in China', *International Journal of Molecular Science*, 17: 888.