



Title	系統樹駆動型研究アプローチの確立による植物ウイルス進化機構の解明
Author(s)	川久保, 修佑
Citation	北海道大学. 博士(農学) 甲第15755号
Issue Date	2024-03-25
DOI	10.14943/doctoral.k15755
Doc URL	http://hdl.handle.net/2115/91944
Type	theses (doctoral)
File Information	Kawakubo_Syusuke.pdf



[Instructions for use](#)

系統樹駆動型研究アプローチの確立による
植物ウイルス進化機構の解明
(Study on plant virus evolutionary mechanisms
by establishing phylogeny-driven research approach)

令和6年

北海道大学大学院農学院

生産フロンティアコース/植物病原学研究室 博士後期課程

川久保 修佑

目次

図表一覧	3
略語一覧	6
第 I 章 緒言	7
第 II 章 研究史	9
第 III 章 リーキ黄色条斑ウイルスの宿主適応進化	14
目的	14
材料と方法	15
結果	31
考察	53
第 IV 章 カブモザイクウイルスの日本における系統地理学的解析	57
目的	57
材料と方法	58
結果	74
考察	99
第 V 章 アレキシウイルス属の種分岐過程の解析	101
目的	101
材料と方法	101
結果	102
考察	105
第 VI 章 総合考察	107
付属資料 – 使用した主なプログラム一覧	110
摘要	114
引用文献	117
謝辞	140

図表一覧

第 III 章

Fig. 1.	Bulbs of <i>Allium</i> plants used in this study.	16
Fig. 2.	IGV capture of the reads mapped to the alliinase gene sequences of leek (MG742366).	17
Fig. 3.	RT-PCR to detect LYSV from GH garlic and leek.	32
Fig. 4.	RT-PCR to detect LYSV from LG hybrid, GH garlic and leek.	33
Fig. 5.	Scatter plots of results of DAPC using the viral nucleotide sequences of P1, HC-Pro, and CP.	34
Fig. 6.	Time-scaled MCC tree of the P1 gene.	36
Fig. 7.	Summary of the date-randomization tests.	37
Fig. 8.	Amino acid sequences of the N-terminal half of the P1 proteins.	39
Fig. 9.	Amino acid sequences of N-terminal half of P1 protein for isolates in the leek clade.	40
Fig. 10.	Time-scaled MCC trees for the P1, HC-Pro, and CP data set.	41
Fig. 11.	Time-scaled MCC tree for the P1 genes.	42
Fig. 12.	Time-scaled MCC tree for the HC-Pro genes.	43
Fig. 13.	Time-scaled MCC tree for the CP genes.	44
Fig. 14.	Time-scaled MCC tree for the NIB genes based on the filtered data set.	46
Fig. 15.	Time-scaled MCC tree for the NIB genes.	47
Fig. 16.	Diversification dynamics during LYSV evolution inferred for the P1, HC-Pro, and CP genes.	48
Fig. 17.	RSS activity of LYSV P1 and HC-Pro at 3 dpa using onion epidermis.	49
Fig. 18.	RSS activity of LYSV P1 and HC-Pro at 3 dpa using leek leaves.	50
Fig. 19.	RSS activity of LYSV P1 and HC-Pro at 5 dpa of <i>Nicotiana benthamiana</i> leaves.	51
Fig. 20.	Possible scenario of the coevolution of LYSV and <i>Allium</i> species plants.	55
Table 1.	Comparison of the ITS regions among the <i>Allium</i> plants used in this study.	18
Table 2.	Primers used in this study.	20
Table 3.	LYSV isolates used in the DAPC and phylogenetic analyses.	22

第 IV 章

Fig. 21.	Sampling map of TuMV isolates analysed in this study and classification of coded geographic regions in Japan.	65
Fig. 22.	Summary of the date-randomization tests and BETS.	69
Fig. 23.	XML code of <tipStateSwapOperator>.	71
Fig. 24.	Phylogenetic network of full genomic sequences of TuMV.	75
Fig. 25.	Phylogenetic tree of TuMV inferred by maximum-likelihood.	76
Fig. 26.	Recombination map of TuMV of Japanese isolates analysed in this study.	77
Fig. 27.	A total of 88 recombination events found in Japanese TuMV genome with recombination frequency distribution across the genome.	78
Fig. 28.	Recombination hotspots and coldspots in TuMV genome.	79
Fig. 29.	Sampling map of TuMV Japanese isolates and maximum-likelihood tree.	86
Fig. 30.	Sampling overview of Japanese TuMV isolates by year.	87
Fig. 31.	Phylogenetic tree of TuMV inferred by maximum-likelihood from polyprotein-coding sequences of nonrecombinant isolates.	88
Fig. 32.	Migration pathways of TuMV phylogenetic groups in Japan, inferred using recombination-free protein-coding regions.	90
Fig. 33.	Spatial diffusion of TuMV in Japan analysed in continuous space.	91
Fig. 34.	Dispersal history of nonrecombinants and recombinants of TuMV.	94
Fig. 35.	Major recombination events found in Japanese TuMV and inferred year of most recent common ancestor for each RTP.	35
Table 4.	TuMV Japanese isolates analysed in this study.	59
Table 5.	Model selection among clock and coalescent models.	68
Table 6.	Time to the MRCA and substitution rate inferred from recombination-free protein-coding regions.	70
Table 7.	Model selection among relaxed random walk models.	73
Table 8.	Recombination events in TuMV genome detected by recombination detecting programs.	80
Table 9.	Recombination type pattern of TuMV in Japan.	84
Table 10.	Number of recombination sites in the genomes of TuMV isolates.	85
Table 11.	Estimated dispersal routes of TuMV phylogenetic groups in Japan.	92
Table 12.	Estimated dispersal routes of TuMV by each recombination type pattern in Japan.	96
Table 13.	Mean branch dispersal speed of nonrecombinants and recombinants TuMV.	97

第 V 章

- Fig. 36. Heatmap for sequence variability matrix in allexiviruses. 103
- Fig. 37. Phylogenetic analyses of allexiviruses based on the nucleotide sequence of the CP-CRP region. 104
- Fig. 38. Phylogenetic trees based on (A) CP nucleotide sequences, (B) CRP nucleotide sequences, (C) CP amino acid sequences, and (D) CRP amino acid sequences. 106

略語一覽

AAC	<i>Allium ampeloprasum</i> complex
BAMM	Bayesian analysis of macroevolutionary mixtures
BEAST	Bayesian evolutionary analysis by sampling trees
BETS	Bayesian estimation of temporal signal
CE	common era
CI	cylindrical inclusion protein
CI	credible interval
CMV	cucumber mosaic virus
CP	coat protein
CRP	cysteine rich protein
DAPC	discriminant analysis of principal component
GH garlic	great-headed garlic
GTR	general time-reversible
HC-Pro	helper-component proteinase
IGV	Interactive Genomic Viewer
ITS	internal transcribed space
LYSV	leek yellow stripe virus
MCC	maximum clade credibility
MCMC	Markov chain Monte Carlo
MRCA	most recent common ancestor
NIa-Pro	nuclear inclusion a proteinase
NIb	nuclear inclusion b
OTU	operational taxonomic unit
OYDV	onion yellow dwarf virus
P1	protein 1
P3	protein 3
PP	posterior probability
RSS	RNA silencing suppressor
RTP	recombination type pattern
SC	strict clock
TuMV	turnip mosaic virus
UCLD	uncorrelated lognormal relaxed clock
UPGMA	unweighted pair group method with arithmetic mean
VPg	genome-linked viral protein

第 I 章 緒言

植物ウイルスは、感染した宿主植物に適応するため進化戦略を動的に変化させる (McLeish et al., 2021)。この進化機構の解明は、農業現場におけるウイルス病の防除対策や抵抗性品種の作出などに大きく寄与する (Choi et al., 2005; McLeish et al., 2019)。ウイルスは宿主細胞に依存して自己複製と翻訳を行い、伝播様式を媒介生物に依存するものも多く存在する。従ってウイルスは外界とのマルチスケールな相互作用と選択圧のもとで自身の適応度を上げる方向に進化すると考えられる。そのため、進化機構の包括的な理解には分子生物学的手法に代表されるミクロスケールの実験と分子系統学や大規模ゲノム疫学調査などによるマクロスケールの情報解析の融合研究が重要である。しかし、それぞれの手法による研究は個別の報告に留まることが多く、スケールを貫いた融合事例は未だ少ない。

植物ウイルスは動物ウイルスに比べて 1 種あたりの宿主域が広いことが多く、特定の植物種に宿主を限定している場合はスペシャリスト型、多数の植物種に感染できる場合はジェネラリスト型と呼ばれる (Elena et al., 2009)。すなわち、ウイルスの宿主決定機構には一様な法則があるのではなく、ウイルスごとに様々で多種多様な進化戦略が存在すると考えられる。ウイルスが宿主の種の壁を乗り越える異種間伝播現象には、ウイルス自身による内的要因と、外環境や宿主の応答による外的要因が複雑に絡み合っている (McLeish et al., 2019)。従って、ウイルスの進化は、宿主・媒介生物・環境要因・地理的隔離等を含む複雑な要因との相互作用を反映した結果だと考えられる (Moury et al., 2017)。

分子系統樹は塩基・アミノ酸配列情報をもとに進化の歴史を示す図だが、系統樹構築の手法はベイズ法の導入によって著しく発展した (Drummond et al., 2012; Du Plessis & Stadler, 2015)。特に、進化速度を推定する分子時計の概念と形質進化の過程を推定する系統比較法の導入により、分子レベルの遺伝子進化を生物学的解釈の得やすい表現型進化として捉えることが可能になった (Cornwell & Nakagawa, 2017; Drummond et al., 2003)。植物ウイルスはヒトの病原ウイルスと比べて、感染情報のリアルタイムな追跡が困難であることが多い。そのため、感染動態や病原性獲得過程の理解には系統

樹によって構築された進化の道筋をもとに、背後にある進化的制約について仮説を立て、それらを実験的に検証することが有効な研究アプローチとなりうる。

本研究では、系統樹を主導とした情報解析から実験までをマルチスケールに貫く系統樹駆動型研究アプローチを確立させ、複数の植物ウイルスを対象としてその進化機構を解明することを目指した。本論文ではまず第 II 章で、分子系統学の研究史とそれがウイルス進化研究に与えてきた影響について論じる。第 III 章では、リーキ黄色条斑ウイルスを対象に、分子時計と祖先状態復元解析、そして一過性タンパク質発現システムを用いた RNA サイレンシングサブレッサー活性のアッセイを行なった。ウイルスが宿主と共進化する時にどのような適応進化戦略をとってきたのかを論じる。第 IV 章では、カブモザイクウイルスを対象に行われた大規模ゲノム疫学調査のデータを用いて、系統地理学的解析を行い、日本におけるウイルス集団の時間的・空間的変遷をゲノム組換えの観点から論じる。第 V 章では、アレキシウイルス属の種間分子系統解析を行い、種分岐過程の背後にある進化の原動力について論じる。

第 II 章 研究史

1. 分子系統学の歴史と系統樹推定の手法

分子系統学は、現代の生物学において基盤的な学問領域の 1 つである。DNA や RNA そしてアミノ酸の配列データを使って系統樹を構築することで、生物の進化の歴史を理解することができる。進化論の父と呼ばれる Charles Darwin が、生物の歴史を分岐的な図で描いたのが系統樹の始まりと考えられており、1859 年に発表した『種の起源』の中で出てくる唯一の図として系統樹が誕生した。生物の進化を階層的に表現することは人間にとってはわかりやすく、画期的であったため、系統樹は広く浸透することとなり、様々な系統樹構築の手法が考案されることになった。特に、分子系統学を近年大きく発展させた 1 つの要因として、ベイズ統計学(Cheon & Liang, 2009; Mau & Newton, 1997; Rannala & Yang, 1996)による解析手法の展開がある。ベイズ統計学の概念自体は、Thomas Bayes が 1763 年に発表しているクラシカルな手法であるが、マルコフ連鎖モンテカルロ法 (Markov chain Monte Carlo 法; MCMC 法) (Hastings, 1970; Metropolis et al., 1953) の登場によりベイズ統計学が系統学的アプローチへの発展をみせ、身近なツールとなった。現在までにヒトの進化をはじめとした多くの生物種の進化について重要な知見を導いており (Eriksson & Manica, 2012; Malaspinas et al., 2016)、植物ウイルスの解析にも利用されている (Cuevas et al., 2012; Monjane et al., 2011)。そこでまずは、代表的な系統樹推定の手法について論じ、ベイズ法の導入により大きく発展した分子時計仮説について紹介する。

系統樹推定の手法は大きく分けて UPGMA 法(unweighted pair group method with arithmetic mean 法)(Sokal & Sneath, 1963)や近隣結合法 (neighbor-joining 法; NJ 法)(Saitou & Nei, 1987)を含む距離行列法と、最節約法 (maximum parsimony 法, MP 法)(Fitch, 1971; Hartigan, 1973; Swofford, 1998)、最尤法 (maximum likelihood 法, ML 法)(Felsenstein, 1981)そしてベイズ法を含む形質状態法に大別することができる(Camin & Sokal, 1965)。配列データを用いて解析する前には、配列間におけるの座位を揃えるために配列中にギャップを挿入し、配列を相互に比較するためにマルチプルシーケンスアライメントの

作成を行う必要がある。系統樹推定に用いるそれぞれの配列は遺伝子や個体もしくは種などに対応する集合であり、その操作単位を **Operational taxonomic unit (OTU)** と言う。一般的な系統樹推定においては樹根を同定することが困難なため、無根系統樹でしか推定できない。従って樹根の位置を決めるために、近縁の別種を外群として用い、無根系統樹の中で外群に繋がる枝の上に樹根があると考えられる。

1.1 距離行列法

距離行列法は OTU 間の違いを距離として表現し、その遺伝的距離が最小となる系統樹を構築する。代表的な手法としては **UPGMA 法**、近隣結合法があり、階層的クラスタリング法によりデータから最適な 1 つの系統樹を再構築する (Rzhetsky & Nei, 1993)。すなわち、配列から計算された距離行列から最も近縁な配列をペアとすることを繰り返すことにより系統樹を構築する。UPGMA 法においては 2 つのクラスター間の距離はそれぞれの距離の平均値であるため、全ての枝で進化速度が一定であると仮定している。しかし実際にはこの仮定が成り立つことは稀であり、配列の分化が進むと誤差が生じる。一方で近隣結合法においては、枝の長さを最小二乗法により計算し、枝長を最小にするものを選びクラスタリングすることによって距離行列を更新し、得られた樹形を最適とする。進化速度が一定と仮定していないものの、距離が大きい場合には妥当な推定ができないなどの欠点がある。

1.2 形質状態法

形質状態法は配列間の距離を計算せずに直接系統樹を再構築して、得られた系統樹に対してある条件を最適化するものを最適な系統樹とする。最適化する条件によって最節約法、最尤法、ベイズ法などの推定手法が用いられる。最節約法は、最小進化量を持つ系統樹が最も妥当であるという仮説に基づき、与えられたデータを説明するための塩基・アミノ酸の置換数を最小にする樹形を選抜する手法である。一方最尤法は進化を説明する塩基やアミノ酸の置換モデルが明示された上で、系統樹の対数尤度を最大にする樹形を最適とする手法である。塩基置換モデルは最も単純な **Jukes-Cantor model (Jukes & Cantor, 1969)** から transition/transversion の速度比を考慮した **Hasegawa-**

Kishino-Yano (HKY) model (Hasegawa et al., 1985)、最も複雑な General time-reversible (GTR) モデル(Tavare, 1986)があり、赤池情報量基準 (AIC)(Akaike, 1974)などの値を参考にしてモデル選択を行う。また遺伝子をコードしている領域ではコドン位置ごとに置換速度が異なるため、速度の違いをガンマ分布で近似することによって補正をかける。

形質状態法を用いて系統樹を推定する上で重要なことは、最適化するスコアを求めするために、原理的には全ての取り得る樹形を試してどれが最適なのかを比較する必要がある。OTUが増えるに伴い、可能な樹形の総数は爆発的に増加し、OTUの数が10の時点で3000万通りを越える。現実的にはこれら全ての樹形に対して最適化するスコアを計算することは不可能であるため、実際には探索空間中からランダムに初期系統樹を選び、それらの系統樹から最適化基準に基づいて系統樹を更新し続けて、最適な系統樹を選択する発見的探索法が用いられる。

1.2.1 ベイズ法

ベイズ法においては探索アルゴリズムとしてMCMC法が用いられる。ベイズ法における事後確率は、事前確率と尤度の積を正規化定数で割った値であるが、正規化定数の計算には複雑な積分を含んでおり、多くの場合は解析的に解くことが困難である。MCMC法は、この複雑な計算を行うために開発されたアルゴリズムであり、探索の過程において事後確率の比を取り、その比を採択するかどうかで探索を行うため、約分される複雑な正規化定数の計算をしなくて済む。実際には初期系統樹から出発し、樹形の変更、樹長の変更、パラメータの変更を提案し、ある棄却率のもとで採択するかどうかを判断することを1つのステップとし、この連鎖的な計算から一定頻度の反復のもとで標本抽出を行う。

1.3 分子時計

分子時計は、DNAやアミノ酸配列の置換速度が時間によらず一定である、とする仮説である(Zuckerkandl & Pauling, 1965)。また、観察される突然変異のほとんどは生存に有利でも不利でもなく、中立的とする分子進化の中立説(Kimura, 1968)も登場し、

分子時計と組み合わせて塩基配列データを用いた分岐年代の推定手法が考案されてきた。これらの年代推定においては、不確実性を適切に導入できるベイズ法が用いられることが多い。また、分岐年代を推定するためには、分子時計を較正(calibration)する必要がある、哺乳類や鳥類の場合では化石データなどが用いられる。ウイルス遺伝子の場合、塩基置換速度が速く、短い期間での標本抽出でも観察可能な進化を遂げた集団を解析できるため、それぞれの分離株が分離された年代を用いることによって較正することができる(Drummond et al., 2003; Rambaut, 2000)。

2. 動物ウイルスと分子系統学

動物ウイルスでは、特に公衆衛生上重要なヒト免疫不全ウイルス (human immunodeficiency virus; HIV) やインフルエンザウイルス(influenza virus)を用いた研究において、分子系統解析や分子時計解析により進化速度や拡散経路が推定されている(Liu et al., 2013; Nelson et al., 2009)。2019年12月に中国の武漢で発生した新型コロナウイルス (SARS-CoV-2)においても、塩基置換速度や実効再生産数の把握が感染拡大の防止に重要であると考えられ、系統樹ベースで迅速に推定された(Tian et al., 2020)。ラテンアメリカ諸国におけるジカウイルス(Zika virus; ZIKV)の感染拡大においては、ZiBRA プロジェクト(Faria et al., 2017)と呼ばれる国際組織が活躍し、オックスフォード・ナノポアテクノロジーによる MinION(Simpson et al., 2017)を利用した迅速なシーケンシング技術と、リアルタイム系統解析により導かれた疫学的情報が感染拡大の食い止めに大きく貢献した。同様の技術で2014年から2015年にかけてギニアで発生したエボラウイルス(Ebola virus)のアウトブレイクの対策にも大きく貢献するなど(Dudas & Bedford, 2019)、ウイルスの分子系統解析は公衆衛生において非常に重要な役割を果たしている。

3. 植物ウイルスと分子系統学

安定した食料生産においては植物ウイルス病の制御が非常に重要である。現在、媒

介昆虫の駆除、弱毒ウイルスの利用、抵抗性作物および遺伝子組換え作物の栽培により植物ウイルス病の防除が行われている。しかし、抵抗性を打破したウイルス系統の出現なども報告されている(Choi et al., 2005; Lee & Anderson, 1998)。新たな地域における植物ウイルスの侵入を防除するためにも、植物ウイルスの分子系統解析や系統地理学解析、分岐年代推定などにより進化機構を解明することは植物保護の観点から極めて重要である。植物 DNA ウイルスとしては、アフリカで被害を与えているジェミニウイルス科マストレウイルス属 **maize streak virus (MSV)** において、その起源や拡散経路についての詳細な研究が行われている(Harkins et al., 2009; Monjane et al., 2011)。

ポティウイルス科ポティウイルス属は 201 のウイルス種によって構成される植物 RNA ウイルス最大の分類群であり、単子葉植物から双子葉植物まで極めて多様な宿主域を持つ。ウイルスの宿主域は、種分岐に沿って系統的に変化してきたわけではなく、進化の過程で何度も独立に獲得・喪失を繰り返している(Moury & Desbiez, 2020)。ウイルス種によって宿主域の広さも様々であることから、植物ウイルスの宿主がどのように決定されるかを一概に解明することは容易ではなく、未だ明らかでないことが多い。しかし、ポティウイルスは世界中に最も広く分布するウイルスの 1 つであり、公共データベースなどから利用可能な塩基配列データも多いことから多くの分子進化的研究が進められている(Fuentes et al., 2019; Gibbs & Ohshima, 2010; Ohshima et al., 2002; Tomitaka & Ohshima, 2006)。また、地理情報と分子系統学を融合させた系統地理学的解析により従来の集団遺伝学的なアプローチから大きく発展し、祖先状態復元を用いることで起源地や拡散経路の推定も可能となっている(Gao et al., 2020; Kawakubo et al., 2021)。

第 III 章 リーキ黄色条斑ウイルスの宿主適応進化

目的

リーキ黄色条斑ウイルス (leek yellow stripe virus: LYSV) は、ポテウイルス属の 1 種であり、ニンニクに感染すると葉に黄色条斑症状を引き起こすだけでなく、鱗片の大きさを縮小させて収量低下を招く重要病原体である(Walkey & Antill, 1989)。ニンニクは栄養繁殖性作物であり、一度ウイルスが感染すると世代を超えて垂直伝播するため、ウイルスはニンニク集団内で容易に蔓延する。従って、市場で流通しているニンニクの多くは既にウイルスに感染している実態がある。LYSV 以外にも、タマネギ萎縮ウイルス(onion yellow dwarf virus: OYDV)やフレキシウイルス科アレキシウイルス属に含まれる Garlic viruses -A, -B, -C, -D, -E, -X, カルラウイルスなどもニンニクに感染し、圃場内ではこれら複数ウイルス種の混合感染として見つかることが多い(Kim et al., 2020; Sasaki et al., 2022)。LYSV はニンニク以外のネギ属植物でも、リーキやタマネギなどで感染が報告されており、感染すると葉に黄色条斑症状を引き起こす(Bos et al., 1978; Santosa & Ertunç, 2020)。リーキの葉における黄色条斑症状は、ドイツにおいて 1937 年以降頻繁に報告されていたが、当時はその病原体が明らかでなかった(Bos, 1983)。その原因病原体が LYSV であると解明されたのは、1978 年におけるオランダでの報告である(Bos et al., 1978)。また、LYSV のニンニクにおける感染初報告は、1987 年である(Lot et al., 1998)。それ以前までは、ニンニクにおける黄色条斑症状の原因ウイルスは OYDV だと考えられていた(Bos et al., 1978)。このように、複数のネギ属植物において LYSV が感染できることが明らかになっているが、異なる宿主植物への適応過程や、その背後にある分子機構は明らかにされていない。そこで、様々なネギ属植物を採集して LYSV を検出後、その遺伝子配列を決定し、分子系統解析を行った。そして、系統樹から示唆される LYSV の宿主適応進化における重要因子として、RSS 活性に着目して実験を行った。

材料と方法

1. 供試植物

北海道十勝地方、宮崎県、およびオランダで栽培されていたリーキを購入し取り寄せた。十勝地方においては、4名の異なる生産者からリーキを取り寄せた。また、佐賀県において2名の異なる生産者からジャンボニンニクを購入した。さらに、宮崎県の市場において販売されていたネギ属植物も購入した。このネギ属植物はニンニクとして販売されていたが、鱗片の大きさが明らかにニンニクとは異なるため、遺伝子解析により植物種の同定を試みた (Fig. 1)。アリナーゼ遺伝子のアミノ酸配列と Internal transcribed space (ITS) 領域の塩基配列を解析したところ、このネギ属植物はリーキ、ニンニク、およびジャンボニンニクの異種間交雑種であることが示唆された (Fig. 2, Table 1) (Ariga et al., 2002)。以降、このネギ属植物を “LG hybrid” と記す。

2. RNA 抽出

供試植物を摩砕後、TRIzol (Invitrogen, Carlsbad, CA, USA)を用いて全 RNA を抽出した。RNA-seq に供試するサンプルについては、抽出後に NucleoSpin RNA (Takara, Otsu, Japan) カラムキットを用いて精製した。

3. RT-PCR、RNA-seq、およびウイルス濃度の定量

RT-PCR は、two-step または one-step 法により行い、30-35 サイクルの条件下で、ウイルスの感染を検出した。Two-step 法においては、cDNA の合成に AMV 逆転写酵素 (Nippon Gene Tokyo, Japan) を用い、その反応液を Takara Ex Taq (Takara, Otsu, Japan) を用いて PCR 法により増幅した。One-step 法においては、Takara One Step RNA PCR kit (AMV) (Takara, Otsu, Japan) を用いた。RNA-seq においては、精製後の RNA サンプルを Novogene (Beijing, China) に外部委託して、Illumina プラットフォームを用いて反応を行った。RNA-seq によって得られたリードは、アダプター配列の除去と

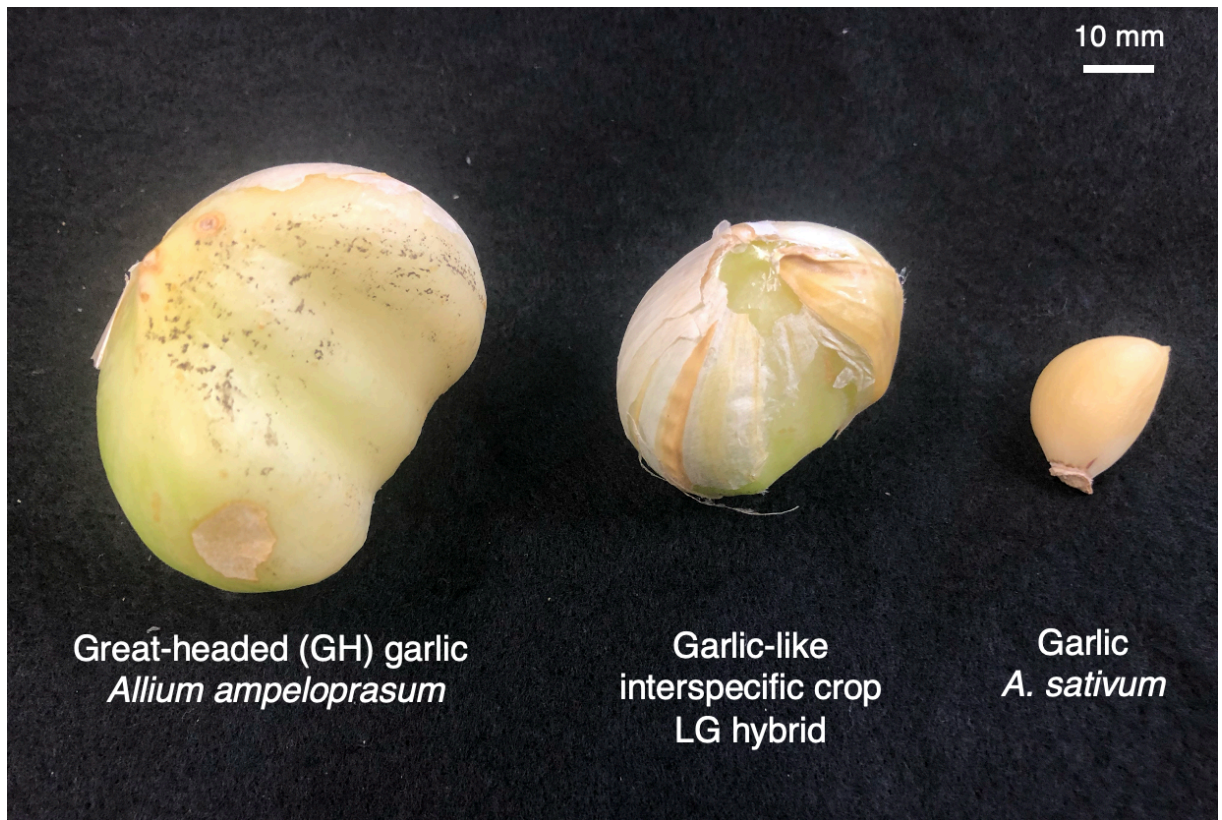


Fig. 1. Bulbs of *Allium* plants used in this study.

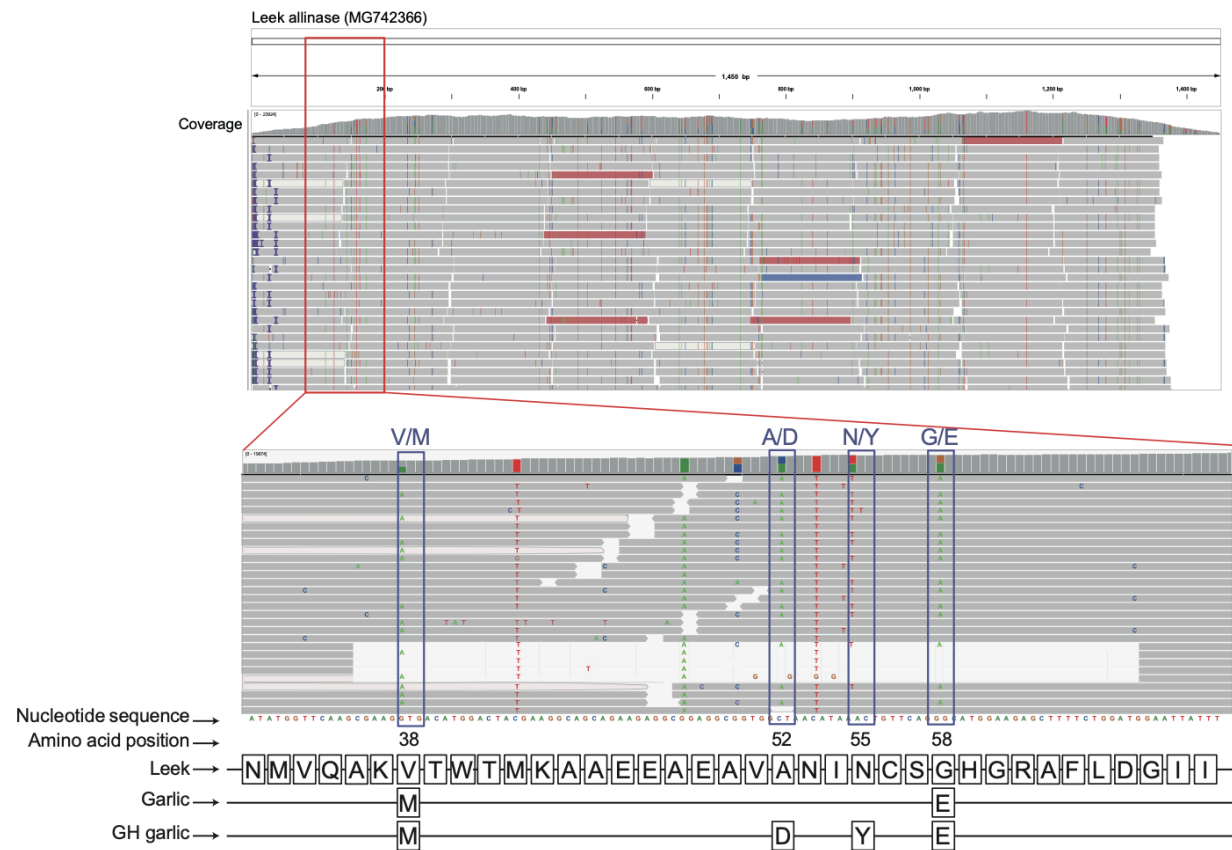


Fig. 2. IGV capture of the reads mapped to the alliinase gene sequences of leek (MG742366). The upper panel shows the overall mapped results, and the boxed section is magnified in the lower panel where differences in the unique amino acids among leek, garlic and GH garlic are indicated in blue squares. The amino acid sequences of the alliinase gene of leek, garlic and GH garlic are shown at the bottom. The amino acid positions are given relative to the methionine residue of the start codon. In the upper panel, reads in red indicate deletions; those in blue indicate insertions.

Table 1. Comparison of the ITS regions among the *Allium* plants used in this study.

Plants	Nucleotide position								
	56	71	72	95	134	142	163	178	187
Leek (FJ664307) ^a	A	T	A	T	A	C	T	A	T
GH garlic (EU626310)	A	A	C	T	A	T	T	A	T
Garlic (EU626375)	C	A	C	C	T	C	C	G	C
LG hybrid	A:37; C:34 ^b	A:78; T:1	A:1; C:84	C:53; T:46	A:24; T:102	C:111; T:19	A:19; C:63; T:52	A:63; G:60; T:1	C:102; T:20

^a GenBank accession number shown in parenthesis. GH garlic; great-headed garlic.

^b Count number of the base mapped at the corresponding site in the ITS region obtained from our RNA-seq analysis of LG hybrid.

Table 1. Continued.

Plants	Nucleotide position								
	235	256	276	476	484	490	491	492	500
Leek	C	T	G	A	T	T	G	G	G
GH garlic	C	T	G	T	T	T	G	G	G
Garlic	T	G	A	T	C	C	A	T	T
LG hybrid	C: 32; G:2; T:65	G:76; T:30	A:92; C:1; G:9	A:720; C:2; T:1	C:567; G:1 T:209	C:586; G:3; T:214	A:553; C:2; G:217; T:23	A:1; C:3; G:233; T:570	G:793; T:21

Table 1. Continued.

Plants	Nucleotide position								
	508	550	570	592	603	610	648	653	654
Leek	T	G	G	A	G	C	A	A	T
GH garlic	G	G	G	A	A	G	A	A	T
Garlic	T	A	T	G	G	G	C	G	C
LG hybrid	A:741; G:106; T:2	G:1; T:998	G:991	A:233; C:1; G:680; T:1	A:97; G:918; T:2	C:12; G:815; T:139	A:199; C:553; T:1	A:720; T:1	C:533; T:187

Table 1. Continued.

Plants	Nucleotide position								
	655	656	670	679	680	687	694	702	752
Leek	C	G	C	C	A	A	A	T	G
GH garlic	C	G	C	C	A	A	G	C	C
Garlic	T	A	T	T	C	G	G	C	C
LG hybrid	C:176; T:519	A:1; G:692	C:50; T:494	C:49; T:469	A:50; C:372; G:89	A:498	A:459; G:1	C: 386	A:7; C:7; G:53

クオリティーチェックの後、Trinity v.2.14.0 (Grabherr et al., 2011)を用いて *de novo assembly* により contig を作成し、LYSV ゲノム配列と相同性を示すものを探索した。ウイルス濃度の定量は、LYSV の P1 と CP 領域を標的とした2つのプライマーペアを用いて40サイクルのRT-PCR法によりおこなった。内部標準には、アクチン遺伝子の配列を用いた。Nested PCR法においては、RT-PCRの反応液1 μL を次のPCR反応の鋳型として供試した。実験に用いたプライマーの配列は、Table 2 に示す。

4. 塩基配列決定、データセットの作成、および遺伝子解析

RT-PCRの増幅産物を、pGEM T-easy ベクター (Promega)にライゲーションし、LYSV の P1、HC-Pro、および CP 遺伝子領域の塩基配列をサンガー法により決定した。これらのプライマー設計は、前項で記述した *de novo assembly* により得られた contig 配列をもとに行った。決定した塩基配列については、BLASTn の標準塩基配列データベースを用いた解析で、LYSV ゲノム配列との相同性があることを確認した。その後、公共データベース GenBank から利用可能な全てのLYSV ゲノム配列を取得し、本研究で配列決定した LYSV 分離株を加えてデータセットを作成した。その際、採集年代、分離宿主、採集国の情報が対応付けられていない分離株についてはデータセットから除いた。本データセットから、P1、HC-Pro、NIb、および CP 遺伝子領域に対応する塩基配列を切り出し、muscle v3.8.31(Edgar, 2004)を用いてコドンベースマルチプルシーケンスアライメントを作成した。アミノ酸の使用頻度は WebLogo (<https://weblogo.berkeley.edu>) (Crooks et al., 2004)を用いて可視化した。組換え解析には、RDP5(Martin et al., 2020)を用い、3以上の異なるプログラムで $P < 10^{-6}$ で支持される組換え信号を有意とみなし、解析に用いる領域内に組換え部位を有する分離株についてもこの時点で除いた。それぞれの遺伝子領域のデータセットにおける分離株数は、それぞれ P1 (n=179), HC-Pro (n=123), NIb (n=166), CP (n=300)であった。P1, HC-Pro, CP の3データセットについては、adegenet(Jombart, 2008)を用いて分離宿主と分離地域を形質情報とした discriminant analysis of principal component (DAPC)を行なった。

Table 2. Primers used in this study.

F/R ^a	Primer name	Sequence (5' → 3')	Purpose
F	LY-5P-L-50	ACAACTCAACTACAATCAATCG	Amplifying LYSV P1 gene
R	LY-3-L-480	TGCAGGAAATTCTTCTGTGTTGA	
F	LYSV-LCP5	GCTGGTGATGARCTAGAYGCRG	Amplifying LYSV CP gene
R	LYSV-CP-868-3	TCACTGCATATGCGCACCATC	
F	LVSV-LCP5-100	CARGCWTATATGCCAAGGTATGG	Nested-PCR to amplify PCR product of CP gene
R	LYSV-LCP3-100	GCYCTAACTGGWGTCTTGATGTG	

^a F, forward primers; R, reverse primers

主成分の数は、各解析において交差検証法により適宜選択を行なった。DAPC 解析に用いた R スクリプトは、GitHub レポジトリにて公開している (https://github.com/skawakubo/Kawakubo_et_al_LYSV-leek)。データセットに含まれる各分離株の詳細を、Table 3 に記す。

5. 分子系統解析

i. 分岐年代推定および祖先状態復元

分子系統解析には、ベイズ法の枠組みを実装した BEAST v2.6.7(Bouckaert et al., 2019) を用いた。塩基置換モデルには GTR+G+I(Tavare, 1986)を適用し、分子時計モデルには対数正規分布を用いた uncorrelated relaxed clock(Drummond et al., 2006)を適用し、系統動態モデルには、Bayesian skyline 合祖モデル(Drummond et al., 2005)を適用した。分子時計の較正には、各 LYSV 分離株の採集年代による tip-dating calibration を行なった。分子時計を較正するにあたり、各データセットに十分な時間信号が含まれているのか検討するため、クラスターランダムマイゼーション検定(Duchêne et al., 2015)を行った。検定においては TipDatingBeast(Rieux & Khatchikian, 2017)を用いて教師なし非階層的クラスタリング法により生成された擬似データセット 10 個を用いた。これらの擬似データセットから推定された塩基置換速度の 95%ベイズ信用区間が、元のデータセットと重複していなければ較正するための十分な時間信号が存在すると判断した。塩基置換速度(uclid.mean)における事前確率分布には、対数正規分布を $\mu = 7.55E-4$, $s = 0.15$ の条件下で適用した。対数正規分布のパラメータは既報のポティウイルスにおける塩基置換速度をもとに設定した(Kawakubo et al., 2021)。祖先状態復元には、各 LYSV 分離株の分離宿主情報を離散形質とした非対称置換モデルを用いた。各形質における分離株数の偏りが解析結果に与える影響を確認するため、形質データを無作為に並び替えた擬似データセットを 10 個生成し、ランダムマイゼーション解析を行なった(Edwards et al., 2011; Kawakubo et al., 2021)。元々のデータセットにおける解析結果と有意に異なったため、数の偏りが解析結果に影響していないことを確認した。事後確率分布は、マルコフ連鎖モンテカルロ (MCMC) 法により計算した。

Table 3. LYSV isolates used in the DAPC and phylogenetic analyses.^a

Index number	Accession number ^b	Protein coding region	Country	Geographic Area ^c	Host ^d	Collection of Year
LY001	D11118	Nib, CP	Japan	East Asia	garlic	1993
LY002	D28590	Nib, CP	Japan	East Asia	garlic	1990
LY003	X89711	Nib, CP	Japan	East Asia	leek	1995
LY004	AB005610	CP	China	East Asia	garlic	1998
LY005	AB005611	CP	USA	Middle East	garlic	1996
LY006	AB005612	CP	Indonesia	Southeast Asia	leek	1996
LY007	AF071525	CP	Israel	Middle East	garlic	1998
LY008	AY007693	Nib, CP	Argentina	South America	leek	2000
LY009	AF228415	CP	Brazil	South America	garlic	2001
LY010	NC_004011	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2000
LY011	AJ307032	Nib, CP	China	East Asia	garlic	2000
LY012	AJ409304	Nib, CP	China	East Asia	garlic	2000
LY013	AJ409305	Nib, CP	China	East Asia	garlic	2000
LY014	AJ409306	Nib, CP	China	East Asia	garlic	2000
LY015	AJ409307	CP	China	East Asia	garlic	2000
LY016	AJ409308	Nib, CP	China	East Asia	garlic	2000
LY017	AJ292225	Nib, CP	China	East Asia	garlic	2000
LY018	AF538950	CP	Taiwan	East Asia	garlic	2002
LY019	AY842134	Nib, CP	New Zealand	Oceania	garlic	2004
LY020	AY842136	Nib, CP	New Zealand	Oceania	garlic	2004
LY021	AY999710	Nib, CP	Argentina	South America	garlic	2005
LY022	AY999712	Nib, CP	Argentina	South America	garlic	2005
LY023	DQ296002	CP	China	East Asia	garlic	2005
LY024	AB194621	P1, HC-Pro, Nib, CP	Japan	East Asia	garlic	2006
LY025	AB194622	P1, HC-Pro, Nib, CP	Japan	East Asia	garlic	2006
LY026	AB194623	P1, HC-Pro, Nib, CP	Japan	East Asia	garlic	2006
LY027	AB194624	CP	Japan	East Asia	leek	2006
LY028	AB194625	CP	Japan	East Asia	leek	2006
LY029	AB194626	CP	Japan	East Asia	garlic	2006
LY030	AB194627	CP	Netherlands	Europe	leek	2006
LY031	AB194628	CP	Netherlands	Europe	leek	2006
LY032	AB194629	CP	Japan	East Asia	garlic	2006
LY033	AB194631	CP	Japan	East Asia	garlic	2006
LY034	AB194632	CP	Japan	East Asia	garlic	2006
LY035	AB194633	CP	South Korea	East Asia	garlic	2006
LY036	AB194634	CP	South Korea	East Asia	garlic	2006
LY037	AB194635	CP	South Korea	East Asia	garlic	2006
LY038	AB194636	CP	Japan	East Asia	garlic	2006
LY039	AB194637	CP	Japan	East Asia	garlic	2006
LY040	AB194638	CP	Japan	East Asia	garlic	2006
LY041	AB194639	CP	Japan	East Asia	garlic	2006
LY042	AB194640	CP	Japan	East Asia	garlic	2006
LY043	AB194641	CP	China	East Asia	garlic	2006
LY044	AB194642	CP	China	East Asia	garlic	2006
LY045	AB194643	P1	Netherlands	Europe	leek	2006
LY046	AB194644	P1	Japan	East Asia	leek	2006
LY047	AB194645	P1	Japan	East Asia	garlic	2006
LY048	AB194646	P1	Japan	East Asia	garlic	2006
LY049	AB194647	P1	South Korea	East Asia	garlic	2006
LY050	AB194648	P1	Japan	East Asia	garlic	2006

^a The nucleotide sequences of the LYSV isolates were obtained from GenBank after removing any isolates whose host plants, sampled country, or year of collection were not available. Any isolates that contain a recombination site were also discarded. See Materials and Methods for details.

^b For the isolates where different GenBank accession numbers were assigned in the protein regions, we show them in the order of P1/HC-Pro/CP.

^c Assigned geographic areas used in the DAPC analysis. Africa, n = 21; East Asia, n = 207; Europe, n = 48; Middle East, n = 53; North America, n = 14; Oceania, n = 7; South America, n = 16; South Asia, n = 15; Southeast Asia, n = 5.

^d Original host plants when sampled. Amaryllis, n = 5; garlic, n = 306; GH garlic, n = 4; leek, n = 49; LG hybrid, n = 1; lily, n = 3; Miami, n = 1; onion, n = 16; rakkyo, n = 1.

Table 3. Continued.

Index number	Accession number	Protein coding region	Country	Geographic Area	Host	Collection of Year
LY051	AB194648	P1	Japan	East Asia	garlic	2006
LY052	AB194649	P1	Japan	East Asia	garlic	2006
LY053	AB194650	P1	Japan	East Asia	garlic	2006
LY054	AB194651	P1	China	East Asia	garlic	2006
LY055	AB194652	P1	China	East Asia	garlic	2006
LY056	AB194653	P1	China	East Asia	garlic	2006
LY057	AB194654	P1	China	East Asia	garlic	2006
LY058	AB194655	P1	Japan	East Asia	garlic	2006
LY059	AB194656	CP	Japan	East Asia	garlic	2006
LY060	DQ402056	CP	New Zealand	Oceania	garlic	2003
LY061	DQ841554	CP	Mexico	North America	garlic	2005
LY062	DQ299380	Nib, CP	Czech Republic	Europe	garlic	2006
LY063	DQ299381	Nib, CP	Czech Republic	Europe	garlic	2006
LY064	DQ299382	Nib, CP	China	East Asia	garlic	2006
LY065	DQ925452	Nib, CP	Vietnam	Southeast Asia	leek	2004
LY066	DQ925453	Nib, CP	Vietnam	Southeast Asia	leek	2004
LY067	FJ358732	CP	Turkey	Middle East	leek	2007
LY068	GU373816	CP	China	East Asia	garlic	2007
LY069	AB551622	CP	Myanmar	Southeast Asia	garlic	2010
LY070	GQ475411	CP	Italy	Europe	garlic	2008
LY071	GQ475412	CP	Italy	Europe	garlic	2008
LY072	GQ475413	CP	Italy	Europe	garlic	2008
LY073	GQ475414	CP	Italy	Europe	garlic	2008
LY074	GQ475415	CP	Italy	Europe	garlic	2008
LY075	GQ475416	CP	Italy	Europe	garlic	2008
LY076	GQ475417	CP	Italy	Europe	garlic	2008
LY077	GQ475418	CP	Italy	Europe	garlic	2008
LY078	AB636318	P1	Japan	East Asia	garlic	2011
LY079	AB636319	P1	Japan	East Asia	garlic	2011
LY080	AB636320	P1	Japan	East Asia	garlic	2011
LY081	AB636321	P1	Japan	East Asia	garlic	2011
LY082	AB636322	P1	Japan	East Asia	garlic	2011
LY083	AB636323	P1	Japan	East Asia	garlic	2011
LY084	AB636324	P1	Japan	East Asia	garlic	2011
LY085	AB636325	P1	Japan	East Asia	garlic	2011
LY086	AB636326	P1	Japan	East Asia	garlic	2011
LY087	AB636327	P1	Japan	East Asia	garlic	2011
LY088	AB636328	P1	Japan	East Asia	garlic	2011
LY089	AB636329	P1	Japan	East Asia	garlic	2011
LY090	AB636330	P1	Japan	East Asia	garlic	2011
LY091	AB636331	P1	Spain	Europe	garlic	2011
LY092	AB636332	P1	Spain	Europe	garlic	2011
LY093	AB636333	P1	Japan	East Asia	garlic	2011
LY094	AB636334	P1	Japan	East Asia	garlic	2011
LY095	AB636335	P1	Japan	East Asia	garlic	2011
LY096	AB636336	P1	Japan	East Asia	garlic	2011
LY097	AB636337	P1	Japan	East Asia	garlic	2011
LY098	AB636338	P1	Japan	East Asia	garlic	2011
LY099	AB636339	P1	Japan	East Asia	garlic	2011
LY100	AB636340	P1	China	East Asia	garlic	2011
LY101	AB636341	P1	China	East Asia	garlic	2011
LY102	HQ918255	Nib, CP	Spain	Europe	garlic	2005
LY103	HQ873735	CP	USA	North America	garlic	2008
LY104	HQ873736	CP	USA	North America	garlic	2008
LY105	HQ873737	CP	Italy	Europe	garlic	2008
LY106	HQ873738	CP	Italy	Europe	garlic	2008
LY107	HQ873739	CP	Italy	Europe	garlic	2008
LY108	HQ873740	CP	Italy	Europe	garlic	2008

Table 3. Continued.

Index number	Accession number	Protein coding region	Country	Geographic Area	Host	Collection of Year
LY109	HQ873742	CP	Italy	Europe	garlic	2008
LY110	HQ873743	CP	Italy	Europe	garlic	2008
LY111	HQ873744	CP	Italy	Europe	garlic	2008
LY112	HQ873745	CP	Italy	Europe	garlic	2008
LY113	HQ873746	CP	USA	North America	garlic	2008
LY114	HQ873747	CP	USA	North America	garlic	2008
LY115	HQ873748	CP	USA	North America	garlic	2008
LY116	HQ873749	CP	USA	North America	garlic	2008
LY117	HQ873750	CP	USA	North America	garlic	2008
LY118	HQ873751	CP	USA	North America	garlic	2008
LY119	HQ873752	CP	USA	North America	garlic	2008
LY120	HQ873753	CP	USA	North America	garlic	2008
LY121	HQ873754	CP	USA	North America	garlic	2008
LY122	HQ873755	CP	Italy	Europe	garlic	2008
LY123	HQ873756	CP	Italy	Europe	garlic	2008
LY124	HQ873757	CP	Italy	Europe	garlic	2008
LY125	HQ873758	CP	Italy	Europe	garlic	2008
LY126	HQ873759	CP	Italy	Europe	garlic	2008
LY127	HQ873760	CP	Italy	Europe	garlic	2008
LY128	JN127339	Nib, CP	Australia	Oceania	garlic	2010
LY129	JN127340	Nib, CP	Australia	Oceania	garlic	2010
LY130	KC207712	CP	Sudan	Africa	garlic	2012
LY131	JQ899450	P1, HC-Pro, Nib, CP	Australia	Oceania	garlic	2011
LY132	JX429965	P1, HC-Pro, Nib, CP	Spain	Europe	garlic	2011
LY133	JX429967	HC-Pro, Nib	Australia	Oceania	garlic	2011
LY134	KF597283	HC-Pro, Nib, CP	Mexico	North America	garlic	2012
LY135	KF597284	CP	Argentina	South America	garlic	2012
LY136	KF597285	P1, HC-Pro, Nib, CP	Argentina	South America	garlic	2012
LY137	KF724857	CP	India	South Asia	garlic	2011
LY138	KF850539	CP	India	South Asia	garlic	2013
LY139	KM032272	CP	Poland	Europe	garlic	2012
LY140	KM032273	CP	Poland	Europe	garlic	2012
LY141	KM032274	CP	Poland	Europe	garlic	2012
LY142	KM032275	CP	Poland	Europe	garlic	2012
LY143	KM032276	CP	Poland	Europe	garlic	2012
LY144	KM032277	CP	Poland	Europe	garlic	2012
LY145	KM032278	CP	Poland	Europe	garlic	2012
LY146	KM032279	CP	Poland	Europe	garlic	2012
LY147	KM032280	CP	Poland	Europe	garlic	2012
LY148	KM032281	CP	Poland	Europe	garlic	2012
LY149	KM032282	CP	Poland	Europe	garlic	2012
LY150	KM032283	CP	Poland	Europe	garlic	2012
LY151	KP168262	CP	India	South Asia	garlic	2011
LY152	KP236094	P1	Brazil	South America	garlic	2011
LY153	KP236095	P1	Brazil	South America	garlic	2008
LY154	KP236096	P1	Brazil	South America	garlic	2007
LY155	KP236097	P1	Brazil	South America	garlic	2007
LY156	KP236098	P1	Brazil	South America	garlic	2008
LY157	KP236099	P1	Brazil	South America	garlic	2011
LY158	KP236100	P1	Brazil	South America	garlic	2008
LY159	KP236101	P1	Brazil	South America	garlic	2008
LY160	KP236102	P1	Brazil	South America	garlic	2007
LY161	KP258216	P1, HC-Pro, Nib, CP	Brazil	South America	garlic	2007
LY162	KT223098	CP	United Kingdom	Europe	Miami	2013
LY163	KP168261	P1, HC-Pro, Nib, CP	India	South Asia	garlic	2011
LY164	KR075504	Nib, CP	Serbia	Europe	leek	2013
LY165	KF850538	CP	India	South Asia	garlic	2013
LY166	MG242624	Nib, CP	Serbia	Europe	leek	2013

Table 3. Continued.

Index number	Accession number	Protein coding region	Country	Geographic Area	Host	Collection of Year
LY167	MG242625	NIb, CP	Serbia	Europe	garlic	2013
LY168	MN070122	CP	Turkey	Middle East	onion	2018
LY169	MN070123	CP	Turkey	Middle East	onion	2018
LY170	MN070124	CP	Turkey	Middle East	onion	2018
LY171	MN070125	CP	Turkey	Middle East	onion	2018
LY172	MN070126	CP	Turkey	Middle East	onion	2018
LY173	MN070127	CP	Turkey	Middle East	onion	2018
LY174	MN070128	CP	Turkey	Middle East	onion	2018
LY175	MN070129	CP	Turkey	Middle East	onion	2018
LY176	MN070130	CP	Turkey	Middle East	onion	2018
LY177	MN070131	CP	Turkey	Middle East	onion	2018
LY178	MN070132	CP	Turkey	Middle East	leek	2018
LY179	MN070133	CP	Turkey	Middle East	leek	2018
LY180	MH794571	CP	Ethiopia	Africa	garlic	2018
LY181	MH794572	CP	Ethiopia	Africa	garlic	2018
LY182	MH794573	CP	Ethiopia	Africa	garlic	2018
LY183	MH794574	CP	Ethiopia	Africa	garlic	2018
LY184	MH794575	CP	Ethiopia	Africa	garlic	2018
LY185	MH794576	CP	Ethiopia	Africa	garlic	2018
LY186	MH890554	NIb	Tunisia	Africa	garlic	2013
LY187	MH890555	NIb	Tunisia	Africa	garlic	2013
LY188	MH890556	NIb	Tunisia	Africa	garlic	2013
LY189	MH890557	NIb	Tunisia	Africa	garlic	2013
LY190	MH890558	P1	Tunisia	Africa	garlic	2013
LY191	MH890559	P1	Tunisia	Africa	garlic	2013
LY192	MH890560	P1	Tunisia	Africa	garlic	2013
LY193	MH890561	P1	Tunisia	Africa	garlic	2013
LY194	MK336972	CP	Ethiopia	Africa	garlic	2014
LY195	MK336973	CP	Ethiopia	Africa	garlic	2014
LY196	MK336974	CP	Ethiopia	Africa	garlic	2014
LY197	MK336975	CP	Ethiopia	Africa	garlic	2014
LY198	MK336976	CP	Ethiopia	Africa	garlic	2014
LY199	MK336977	CP	Ethiopia	Africa	garlic	2014
LY200	MN864794	CP	Turkey	Middle East	onion	2018
LY201	MN864795	CP	Turkey	Middle East	onion	2018
LY202	MN864796	CP	Turkey	Middle East	onion	2018
LY203	MN864797	CP	Turkey	Middle East	onion	2018
LY204	MN864798	CP	Turkey	Middle East	onion	2018
LY205	MN864799	CP	Turkey	Middle East	onion	2018
LY206	MN059436	P1, HC-Pro, NIb, CP	China	East Asia	garlic	2017
LY207	MN059437	HC-Pro, NIb	China	East Asia	garlic	2017
LY208	MN059438	P1, HC-Pro, NIb	China	East Asia	garlic	2017
LY209	MN059439	P1, HC-Pro, NIb, CP	China	East Asia	garlic	2017
LY210	MN059440	P1, HC-Pro, NIb, CP	China	East Asia	garlic	2017
LY211	MN059441	P1, HC-Pro, NIb	China	East Asia	garlic	2017
LY212	MN059442	HC-Pro, NIb, CP	China	East Asia	garlic	2017
LY213	MN059443	P1, HC-Pro, NIb, CP	China	East Asia	garlic	2017
LY214	MN059444	P1, HC-Pro, NIb, CP	China	East Asia	garlic	2017
LY215	MN059445	P1, HC-Pro, NIb, CP	China	East Asia	garlic	2017
LY216	MN059446	P1, HC-Pro, NIb, CP	China	East Asia	garlic	2017
LY217	MN059447	P1, HC-Pro, NIb, CP	China	East Asia	garlic	2017
LY218	MN059448	P1, HC-Pro, NIb, CP	China	East Asia	garlic	2017
LY219	MN059449	P1, HC-Pro, NIb, CP	China	East Asia	garlic	2017
LY220	MN059450	P1, HC-Pro, NIb, CP	China	East Asia	garlic	2017
LY221	MN059451	P1, HC-Pro, NIb, CP	China	East Asia	garlic	2017
LY222	MN059452	HC-Pro, NIb, CP	China	East Asia	garlic	2017
LY223	MN059453	P1, HC-Pro, NIb, CP	China	East Asia	garlic	2017
LY224	MN059454	P1, HC-Pro, NIb, CP	China	East Asia	garlic	2017

Table 3. Continued.

Index number	Accession number	Protein coding region	Country	Geographic Area	Host	Collection of Year
LY225	MN059455	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY226	MN059456	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY227	MN059457	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY228	MN059458	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY229	MN059459	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY230	MN059460	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY231	MN059461	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY232	MN059462	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY233	MN059463	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY234	MN059464	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY235	MN059465	Nib, CP	China	East Asia	garlic	2017
LY236	MN059466	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY237	MN059467	P1, Nib, CP	China	East Asia	garlic	2017
LY238	MN059468	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY239	MN059469	Nib	China	East Asia	garlic	2017
LY240	MN059470	P1, HC-Pro, Nib	China	East Asia	garlic	2017
LY241	MN059471	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY242	MN059472	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY243	MN059473	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY244	MN059474	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY245	MN059475	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY246	MN059476	P1, Nib, CP	China	East Asia	garlic	2017
LY247	MN059477	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY248	MN059478	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY249	MN059479	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY250	MN059480	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY251	MN059481	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY252	MN059482	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY253	MN059483	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY254	MN059484	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY255	MN059485	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY256	MN059486	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY257	MN059487	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY258	MN059488	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY259	MN059489	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY260	MN059490	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY261	MN059491	Nib, CP	China	East Asia	garlic	2017
LY262	MN059492	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY263	MN059493	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY264	MN059494	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY265	MN059495	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY266	MN059496	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY267	MN059497	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY268	MN059498	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY269	MN059499	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY270	MN059500	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY271	MN059501	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY272	MN059502	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY273	MN059503	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY274	MN059504	HC-Pro, Nib	China	East Asia	garlic	2017
LY275	MN059505	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY276	MN059506	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY277	MN059507	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY278	MN059508	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY279	MN059509	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY280	MN059510	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY281	MN059511	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY282	MN059512	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017

Table 3. Continued.

Index number	Accession number	Protein coding region	Country	Geographic Area	Host	Collection of Year
LY283	MN059513	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY284	MN059514	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY285	MN059515	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY286	MN059516	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY287	MN059517	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY288	MN059518	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY289	MN059519	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY290	MN059520	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY291	MN059521	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY292	MN059522	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY293	MN059523	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY294	MN059524	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY295	MN059525	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY296	MN059526	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY297	MN059527	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY298	MN059528	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY299	MN059529	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY300	MN059530	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY301	MN059531	P1, Nib, CP	China	East Asia	garlic	2017
LY302	MN059532	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY303	MN059533	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY304	MN059534	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY305	MN059535	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY306	MN059536	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY307	MN059537	P1, Nib, CP	China	East Asia	garlic	2017
LY308	MN059538	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY309	MN059539	HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY310	MN059540	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY311	MN059541	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY312	MN059542	Nib, CP	China	East Asia	garlic	2017
LY313	MN059543	P1, HC-Pro, Nib	China	East Asia	garlic	2017
LY314	MN059544	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY315	MN059545	P1, HC-Pro, Nib	China	East Asia	garlic	2017
LY316	MN059546	P1, HC-Pro, Nib	China	East Asia	garlic	2017
LY317	MN059547	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY318	MN059548	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY319	MN059549	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY320	MN059550	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY321	MN059551	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY322	MN059552	P1, HC-Pro, Nib, CP	China	East Asia	garlic	2017
LY323	MT038061	Nib	Turkey	Middle East	leek	2008
LY324	MT038062	Nib	Turkey	Middle East	leek	2019
LY325	MT038063	Nib	Turkey	Middle East	leek	2019
LY326	MT038064	Nib	Turkey	Middle East	leek	2008
LY327	MT038065	CP	Turkey	Middle East	leek	2008
LY328	MT038066	CP	Turkey	Middle East	leek	2019
LY329	MT038067	CP	Turkey	Middle East	leek	2019
LY330	MT038068	CP	Turkey	Middle East	leek	2008
LY331	MT731491	P1, HC-Pro, Nib, CP	India	South Asia	garlic	2020
LY332	MT731492	HC-Pro, Nib, CP	India	South Asia	garlic	2020
LY333	MW854277	P1, HC-Pro, Nib, CP	Indonesia	Southeast Asia	leek	2021
LY334	MT232838	Nib, CP	Iran	Middle East	garlic	2018
LY335	MZ076674	CP	Turkey	Middle East	leek	2014
LY336	MZ076675	CP	Turkey	Middle East	leek	2014
LY337	MZ076676	CP	Turkey	Middle East	leek	2014
LY338	MZ076677	CP	Turkey	Middle East	leek	2014
LY339	MZ076678	CP	Turkey	Middle East	leek	2014
LY340	MZ076679	CP	Turkey	Middle East	leek	2014

Table 3. Continued.

Index number	Accession number	Protein coding region	Country	Geographic Area	Host	Collection of Year
LY341	MZ076680	CP	Turkey	Middle East	leek	2015
LY342	MZ076681	CP	Turkey	Middle East	garlic	2014
LY343	MZ076682	CP	Turkey	Middle East	garlic	2015
LY344	MZ076683	CP	Turkey	Middle East	garlic	2015
LY345	MZ203474	CP	India	South Asia	amaryllis	2022
LY346	MZ203475	CP	India	South Asia	amaryllis	2022
LY347	MZ203476	CP	India	South Asia	amaryllis	2022
LY348	MZ203477	CP	India	South Asia	amaryllis	2022
LY349	MZ203478	CP	India	South Asia	amaryllis	2022
LY350	MZ004943	CP	India	South Asia	lily	2022
LY351	MZ004944	CP	India	South Asia	lily	2022
LY352	MZ004945	CP	India	South Asia	lily	2022
LY353	MT358341	CP	China	East Asia	garlic	2019
LY354	MT358342	CP	China	East Asia	garlic	2019
LY355	MT358343	P1	China	East Asia	garlic	2019
LY356	OL944631/-/OL944621	P1, CP	Turkey	Middle East	leek	2007
LY357	OL944632/-/OL944620	P1, CP	Turkey	Middle East	leek	2007
LY358	OL944633/-/OL944628	P1, CP	Turkey	Middle East	leek	2007
LY359	OL944634/-/OL944627	P1, CP	Turkey	Middle East	leek	2007
LY360	OL944635/-/OL944626	P1, CP	Turkey	Middle East	leek	2007
LY361	OL944636/-/OL944623	P1, CP	Turkey	Middle East	leek	2007
LY362	OL944637/-/OL944624	P1, CP	Turkey	Middle East	leek	2007
LY363	OL944638/-/OL944629	P1, CP	Turkey	Middle East	leek	2007
LY364	OL944639	P1	Turkey	Middle East	leek	2008
LY365	OL944640	P1	Turkey	Middle East	leek	2007
LY366	OL944641	P1	Turkey	Middle East	leek	2008
LY367	OL944622	CP	Turkey	Middle East	leek	2007
LY368	OL944625	CP	Turkey	Middle East	leek	2007
LY369	LC714765	P1	USA	North America	garlic	2022
LY370	LC714768	P1	China	East Asia	garlic	2022
LY371	LC757499	P1	Japan	East Asia	garlic	2022
LY372	LC757500	P1	Japan	East Asia	garlic	2022
LY373	LC714769	P1	Japan	East Asia	garlic	2022
LY374	LC714771	P1	Japan	East Asia	rakkyo	2022
LY375	LC714766	P1	Japan	East Asia	garlic	2022
LY376	LC714770	P1	Japan	East Asia	garlic	2022
LY377	LC714767	P1	Japan	East Asia	garlic	2022
LY378	LC757493	P1, HC-Pro, N1b, CP	Japan	East Asia	LG hybrid	2022
LY379	LC757494/LC757501/LC757505	P1, HC-Pro, CP	Japan	East Asia	leek	2022
LY380	-/LC757502/-	HC-Pro	Japan	East Asia	leek	2022
LY381	-/LC757503/LC757506	HC-Pro, CP	Japan	East Asia	leek	2022
LY382	-/LC757504/-	HC-Pro	Japan	East Asia	leek	2022
LY383	LC757495	P1	Japan	East Asia	GH garlic	2022
LY384	LC757496	P1	Japan	East Asia	GH garlic	2022
LY385	LC757497	P1	Japan	East Asia	GH garlic	2022
LY386	LC757498	P1	Japan	East Asia	GH garlic	2022

事後確率分布の最適解を確実に探索するために、独立した連鎖計算を最低 3 回解析した。それぞれの連鎖計算は 1 万世代毎抽出で 1 億回行った。各連鎖計算の収束判定には Tracer v1.7.2(Rambaut et al., 2018)を用いた。複数の独立した連鎖計算が十分な定常状態に達していることを Tracer v1.7.1 で確認した適切な範囲の burn-in 状態を除いて LogCombiner で結合し、Effective sample size (ESS) ≥ 200 であることを基準として判定した。Maximum clade credibility (MCC) 系統樹の作成には TreeAnnotator を用いた。系統樹の可視化には、ggtree(Yu et al., 2017)を用いた。

ii. Birth-death model を用いた適応度推定

系統樹の位相と枝長の情報から、系統樹枝ごとの適応度推定を行うため、BAMM(Rabosky, 2014)と BAMMtools(Rabosky et al., 2014)を用いて多様化解析を行なった。前項で記述した、BEAST の出力 MCC 系統樹を ultrametric 変換したものを本解析の入力ファイルとして用いた。MCMC の連鎖計算は 5 千世代毎抽出で 500 万回行った。事前確率分布には、*setBAMMpriors* 関数を用いて得られた結果を適用し、*globalSamplingFraction* パラメータは 0.01 を用いた。本解析における λ パラメータを、birth-death model における birth rate として適応度推定を行なった。本解析に用いた R スクリプトは GitHub レポジトリにて公開している(https://github.com/skawakubo/Kawakubo_et_al_LYSV-leek)。

6. RNA サイレンシングサプレッサー活性の定量

LYSV の P1 と HC-Pro における RSS 活性の定量には、既報の方法を参考にしたアグロインフィルトレーション法を用いた(Kim et al., 2020; Sasaki et al., 2022; Shimura et al., 2022)。タンパク質の発現クローンには、LG hybrid に感染していた LYSV 分離株(LC757493)とニンニクに感染していた LYSV 分離株(LC714765)の配列を用い、それぞれを LYSV のリーキ系統とニンニク系統の代表分離株として比較実験を行なった。HC-Pro と FLAG 付加 P1 タンパク質(P1-FLAG)をコードする DNA 断片を、XbaI と SacI の二つの制限酵素で消化した pBE2113 プラスミド

ベクターにライゲーション後、*Agrobacterium tumefaciens* KYRT1 系統に形質転換した。発現植物としては、タマネギ、リーキおよび *Nicotiana benthamiana* を用いた。緑色蛍光タンパク質(GFP)、P1、HC-Pro コンストラクトをそれぞれ形質転換させた *Agrobacterium* 懸濁液を 1:1:1 の体積比で混合させ、*N. benthamiana* とリーキの葉およびタマネギの表皮にインフィルトレーションした。接種源の懸濁液は、*N. benthamiana* には終濁度 $OD_{600}=1.0$ 、タマネギとリーキには 0.2 に調整した。GFP の蛍光強度は *N. benthamiana* については UV 光下で 5 days post-agroinfiltration (dpa)に測定し、ImageJ (<https://imagej.net/ij/index.html>)を用いて定量した。タマネギとリーキについては、3dpa において LAS AF software (Leica)を用いて測定および定量を行なった。全ての実験において、5 回の生物学的反復をとり同様の結果が再現されることを確認した。統計学的有意度検定は、 $P<0.05$ の有意水準において Tukey の多重比較検定および一元配置分散分析(One-way ANOVA)により行なった。パラメトリック検定の選択には、Shapiro-Wilk の検定により正規性を確認して行なった。タンパク質の発現は、ウエスタンブロットティング法により確認した。アグロインフィルトレーション後のサンプルを、Laemmli サンプルバッファーと 2%SDS および 0.1%の Tween 20 で調整したリン酸緩衝生理食塩水(PBST)を加えて磨砕し、10%のポリアクリルアミド電気泳動(SDS-PAGE)に供試した。抗 FLAG モノクローナル抗体(FUJIFILM Wako Pure Chemical Corporation)を用いて、P1-FLAG を検出した。

結果

1. ネギ属植物異種間交雑種に感染していた LYSV の遺伝子解析

本研究で用いた交雑種(LG hybrid)に感染していた LYSV の P1 遺伝子塩基配列をクエリー配列として用い BLASTn で解析すると、インドネシア産のリーキから分離された LYSV 分離株(MW854277)と最も高い塩基配列一致度を示したが、84%と比較的低い値であった。その後、EMBOSS-NEEDLE(Needleman & Wunsch, 1970)を用いて日本産のニンニクに感染していた LYSV 分離株の配列とペアワイズ塩基配列一致度を計算すると、51%とかなり低い値を示した。

2. 宿主植物の違いにおける LYSV 蓄積量比較

リーキ、ジャンボニンニク、および LG hybrid から RNA を抽出後、P1 と CP 遺伝子領域を標的に RT-PCR を行なった(Fig. 3-4)。ジャンボニンニクと LG hybrid においては、P1 遺伝子を標的とした 30-35 サイクルの one-step RT-PCR で標的サイズのバンドが検出された。一方、同様の条件でリーキから LYSV の検出を試みてもバンドは検出されず、40 サイクルで極めて薄いバンドが確認された (Fig. 3)。本結果は、リーキ細胞の中で LYSV は極低濃度で存在することを示唆する。また、CP 遺伝子を標的とした one-step RT-PCR、および nested PCR の結果からも、ジャンボニンニクと LG hybrid においては LYSV が高濃度で蓄積している一方で、リーキにおいては低濃度で蓄積していることが示唆された(Fig. 4)。

3. LYSV の集団遺伝学的解析

P1、HC-Pro、CP の 3 遺伝子領域について作成したデータセットを用い、分離宿主と地理情報を形質情報とした DAPC 解析を行なった(Fig. 5)。分離宿主を形質情報として用いた場合、全遺伝子領域においてニンニクとリーキが第 1 もしくは第 2 判別関数で明らかに分かれることが示された(Fig. 5A-C)。一方で、地理情報を形質情報として用いた場合、明確な集団分離は認められなかった

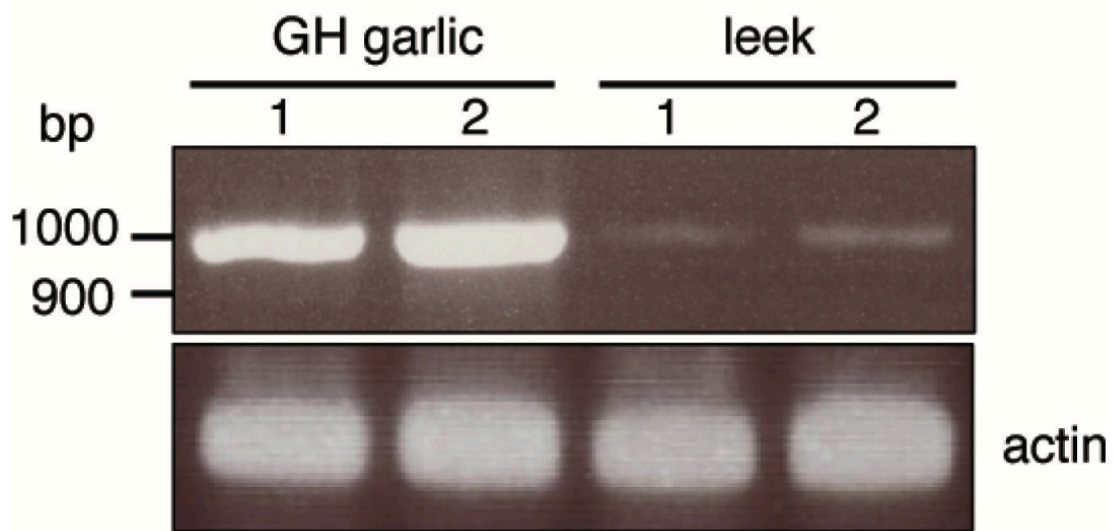


Fig. 3. RT-PCR to detect LYSV from GH garlic and leek. The PCR products were analyzed after 40 cycles. The actin gene was used as an internal control. The lane number represents each individual plant. Primers are listed in Table 2.

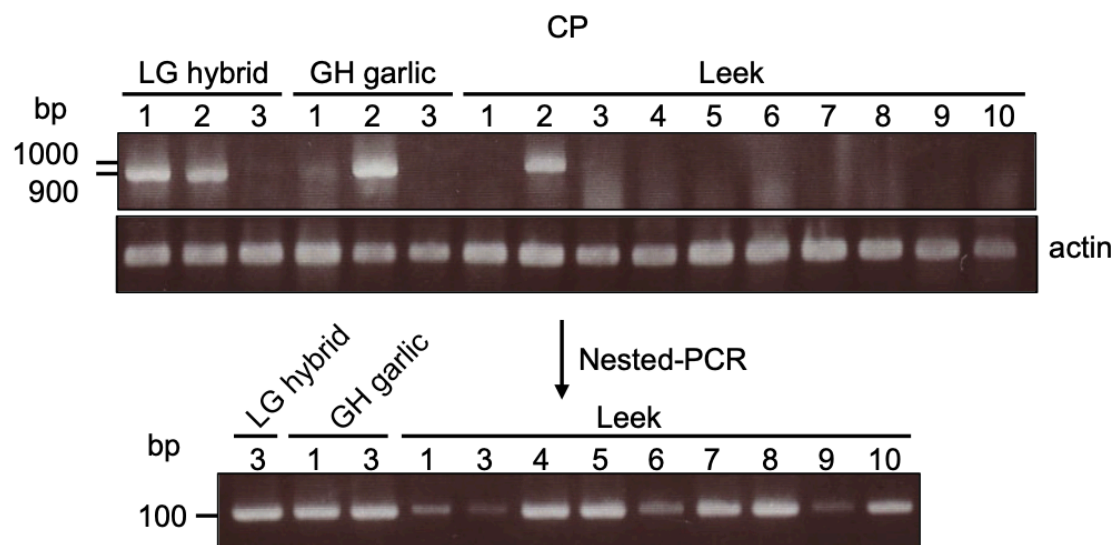


Fig. 4. RT-PCR to detect LYSV from LG hybrid, GH garlic and leek. The first-round PCR was run for 40 cycles (upper panel), then nested-PCR was run (lower panel) when no band was detected in the first-round PCR. The actin gene was used as an internal control. Lane number represents each individual plant and the number in the upper panel corresponds to those on the lower panel. PCR primers are listed in the Table 2.

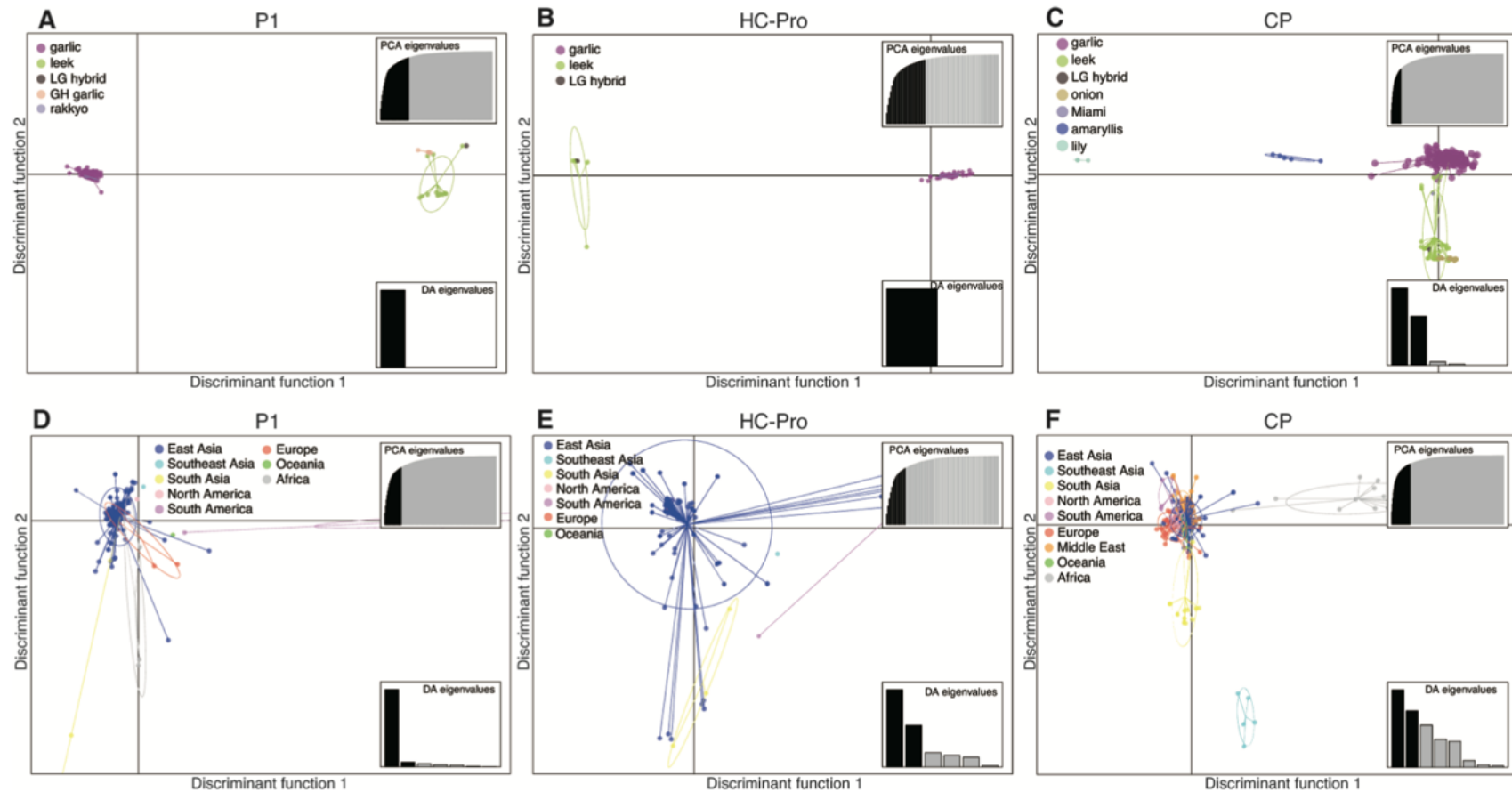


Fig. 5. Scatter plots of results of discriminant analysis of principal components using the viral nucleotide sequences of P1 (A and D), HC-Pro (B and E), and CP (C and F). Discriminant analysis of LYSV isolates with (A–C) host plants or (D–F) geographic areas used as a population trait. The circles indicate 95% inertia ellipses. Rakkyo is a common name for *Allium chinense* G. Don; Miami is an ornamental onion of *Allium* species.

(Fig. 5D-F)。これらの結果は、LYSV の適応進化は地理的分離よりも宿主の違いが原動力となっていることを示唆している。本解析で用いたリーキからの LYSV は、インドネシア、日本、オランダ、セルビア、トルコ、ベトナムから採集されている。

4. P1 遺伝子における欠失及び時間系統樹

P1 遺伝子領域の塩基配列を用いてベイズ時間系統樹を構築し、本領域内に見られる欠失の部位とともに可視化した(Fig. 6)。ランダムイゼーション検定の結果から、本解析で用いたデータセットは全て十分な時間信号を有することが示された (Fig. 7)。また、リーキに感染していた LYSV 分離株は全て一つの単系統群であり、ニンニクに感染する LYSV とは独立した系統群を形成することが明らかになった (Fig. 6)。さらに、ジャンボニンニクと LG hybrid に感染していた LYSV 分離株もこれらリーキ特異的系統群に属することが明らかになった。これらリーキ特異的系統群に属する分離株は、全て 68 アミノ酸の欠失(134_210del)を持たない N-type の分離株(Yoshida et al., 2012)であった。リーキ特異的系統群の最も近い共通祖先(most recent common ancestor: MRCA)は 1833 (95% ベイズ信用区間: 1775-1893)年に分岐したと推定された。さらに本系統群は2つのサブ系統群に分けられることが支持された (posterior probability: PP > 0.9) (Fig. 6)。ほとんど全ての N-type の分離株は 4 アミノ酸の欠失(17_21del)を持ち、その姉妹群に位置する単系統群についても同じ位置に 2 アミノ酸の欠失(18_19del)を持つことが明らかになった。この姉妹群の分岐に対応する事後確率は低かった(Fig. 6)。全ての S-type 分離株は、68 アミノ酸の欠失(134_201del)以外の欠失を持たなかった。また、S-type 分離株の共通祖先と姉妹群に位置する単系統群は 134_201del の欠失を持たなかった。全 LYSV 分離株の MRCA は 1264(95% ベイズ信用区間: 1070-1451)年に分岐したと推定された。

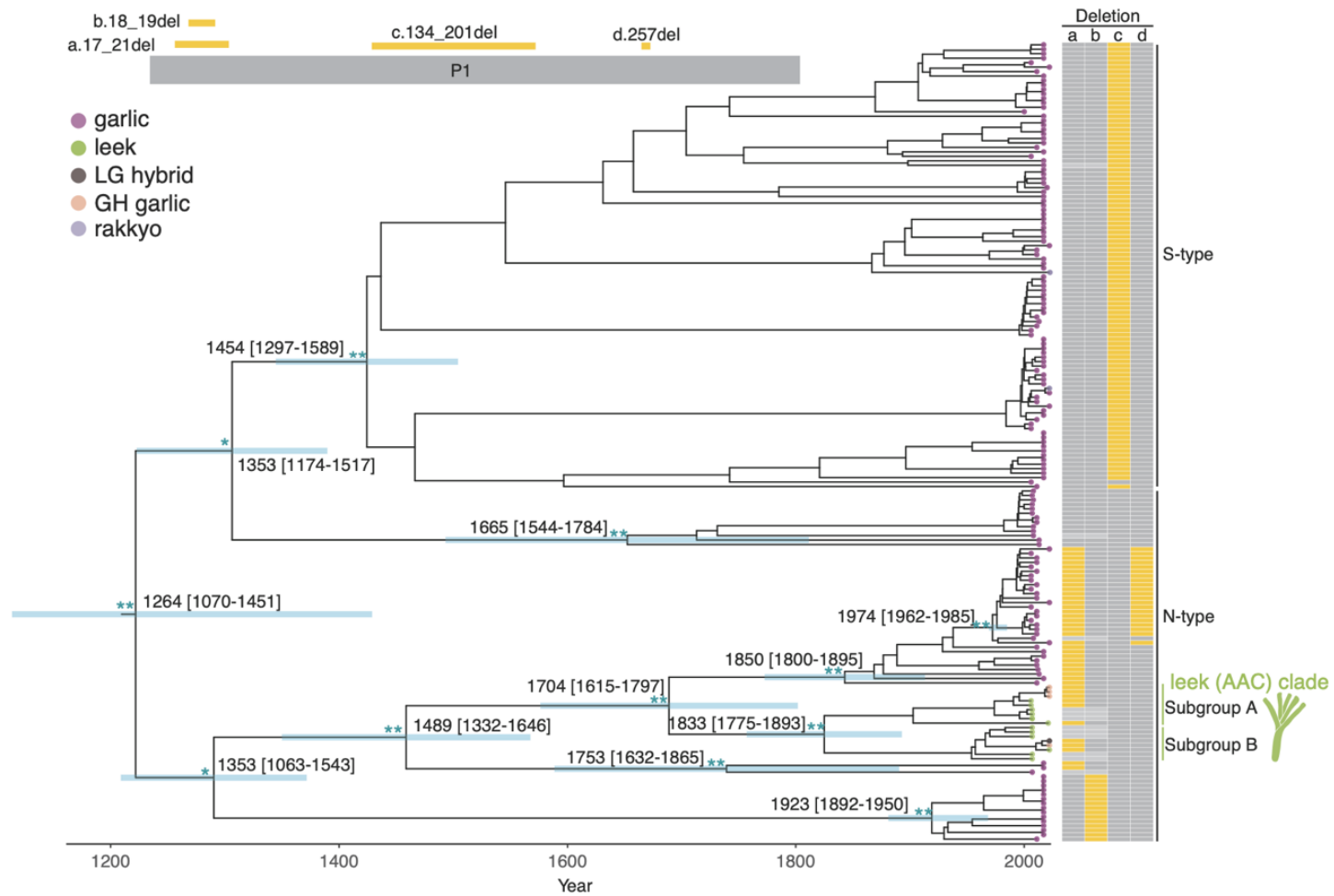
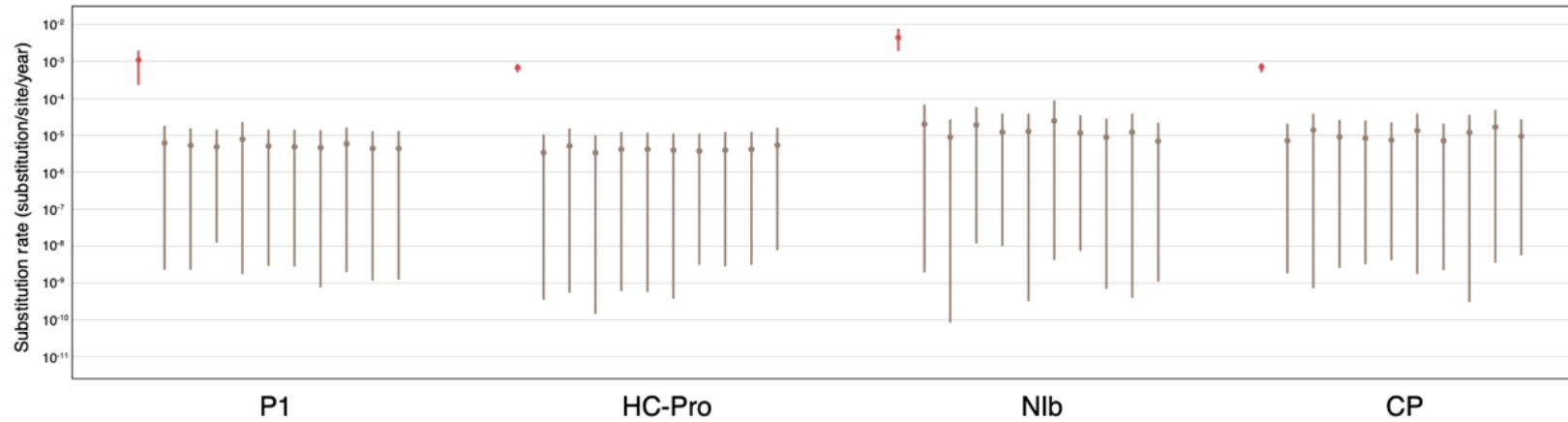


Fig. 6. Time-scaled MCC tree of the P1 gene. Colors at the tipping points represent the hosts shown in the color key. The x-axis is the timescale in years. The blue bar at each node of interest is the 95% Bayesian credible interval for the inferred time. Sequence deletions in each isolate are shown in the diagram at the top as an array of the vertical box on the right (yellow, presence of deletion; gray, absence of deletion; light gray, missing values due to partially available sequences). * posterior probability >0.7; ** posterior probability >0.9.

A. filtered data set



B. total available isolates

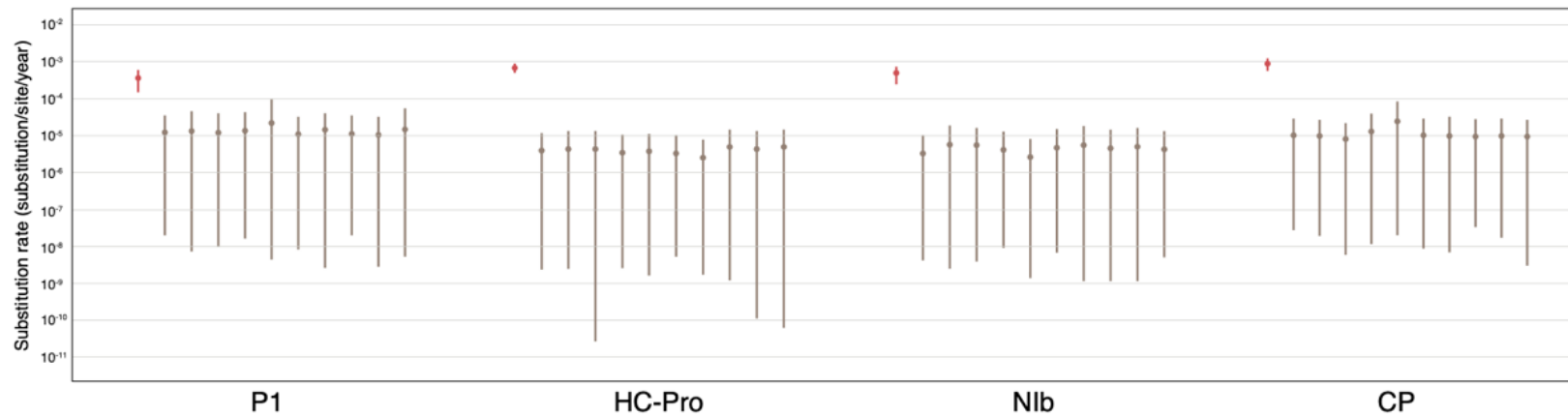


Fig. 7. Summary of the date-randomization tests. (A) Filtered data sets for four viral genes that shared the same taxon set, which were used in the Figs. 10A-C and 14. (B) Total available isolates data sets, which were used in the Figs. 10D-F, 11-13, and 15. Estimates of the nucleotides substitution rate (substitution/site/year) from the original data set and ten clustered permutations. Point symbols represent the mean rate estimate for each data set, with error bars showing the 95% credible intervals (CI). Point symbols and error bars in red color represent the estimate from the original data set, whereas those in brown color represent the estimates from the date-randomized data sets. Note that the 95% CI of the estimates from the date-randomized replicates did not overlap with the mean posterior estimate from the original data set for all the analyzed genes, indicating the presence of temporal signals.

5. P1 タンパク質のアミノ酸における反復モチーフ

LYSV の P1 タンパク質アミノ酸配列を確認すると、ポティウイルス属において保存されている IxFG モチーフ(Valli et al., 2007)が見つかったが、既報のモチーフサイズよりも長い IxFGSFETP として保存されていた(Fig. 8-9)。さらに、このモチーフは N-type においては 3 回反復して存在し、この部位が S-type における 68 アミノ酸の欠失(134_210del)と対応していることが明らかになった。S-type においては、このモチーフは 1 回反復のみで存在していた。リーキ特異的系統群においては、計 7 箇所のサイトで特異的に保存されているアミノ酸が見られた(Fig. 9)。

6. 祖先宿主の推定

P1、HC-Pro、CP の 3 遺伝子領域において祖先状態復元を行い、祖先宿主の推定を行なった。遺伝子領域ごとに含まれる分離株数が異なるため、まず全データセットに共通して含まれる分離株のみに限定して解析を行なった(n=93)。その結果 P1、HC-Pro、CP の 3 遺伝子領域全てにおいて LYSV の祖先宿主はリーキと推定され、それぞれの事後確率は 0.46、0.49、0.43 であった(Fig. 10A-C)。一方で、ニンニクを祖先宿主とする事後確率は低く、それぞれ 0.14、0.07、0.21 であった。MRCA が分岐した年代はそれぞれ 1439 (95% ベイズ信用区間: 1223-1634)年、1614 (95% ベイズ信用区間: 1477-1732)年、1680 (95% ベイズ信用区間: 1562-1785)年であった。同様の解析を、利用可能な全分離株を含めたデータセットにおいても行なった (P1, n=179; HC-Pro, n=123; CP, n=300) (Fig. 10D-F, Figs 11-13)。HC-Pro においては、リーキを祖先宿主とする事後確率が最も高く 0.55 であったが、P1 と CP においては、リーキを祖先宿主とする事後確率はそれぞれ 0.27、0.01 と低かった。P1 においては、ニンニクを祖先宿主とする事後確率が 0.44 で最も高かった。CP においても、ニンニクを祖先宿主とする事後確率が 0.37 で最も高かったが、ヒガンバナを祖先宿主とする事後確率が 0.37、ユリを祖先宿主とする事後確率が 0.17 であった(Fig. 10F)。また、これら 3 遺伝子領域に加えて、幅広い RNA ウイルス種で構造が保存されている RNA 依存性 RNA ポリメラー

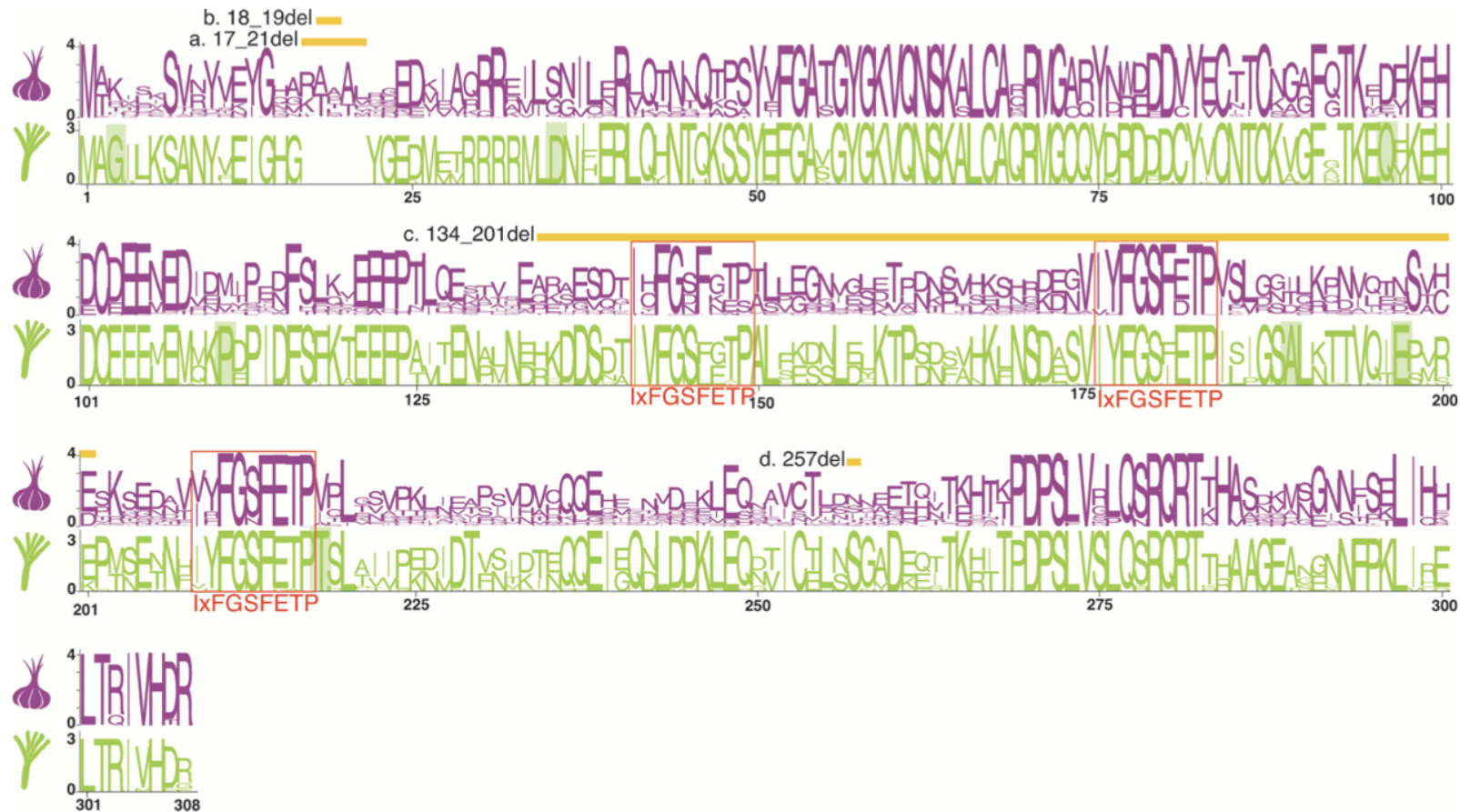


Fig. 8. Amino acid sequences of the N-terminal half of the P1 proteins. The sequences shown in purple on the top indicate LYSV isolates from garlic, and in green at the bottom from leek. The x-axis indicates translated amino acid sequence positions for the LYSV genome (AB194649). The y-axis indicates the bit score for cross-entropy at the sequence position. The conserved repetitive motifs are boxed in red. The yellow bars (a–d) correspond to the deletion positions shown in Fig. 6. Plots were created using WebLogo (<https://weblogo.berkeley.edu>).



Fig. 9. Amino acid sequences of N-terminal half of P1 protein for isolates in the leek clade. The x-axis corresponds to the amino acid sequence position of the LYSV isolate (AB194649). The y-axis indicates the bit score of cross entropy at the sequence position. The yellow bar corresponds to the deletion position shown in Fig. 6. The sequence of subgroup A at the 5' end is not shown due to the limited number of available sequences. Plots were created using WebLogo (<https://weblogo.berkeley.edu>).

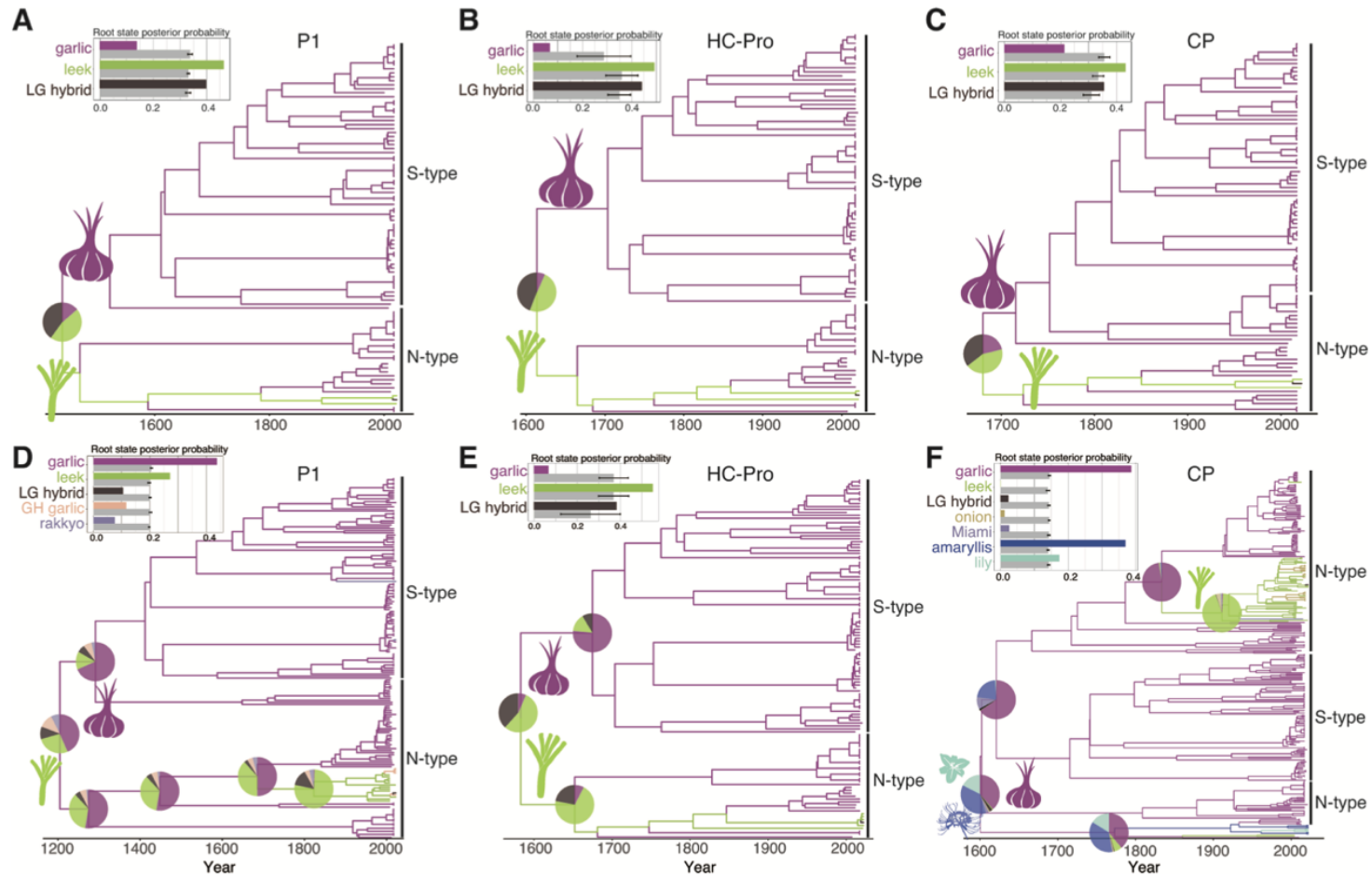


Fig. 10. Time-scaled MCC trees for the P1 (A and D), HC-Pro (B and E), and CP (C and F) data set. (A–C) is based on the filtered data set that shared the same taxon set ($n = 93$) and (D–F) is based on the total available isolates in each gene (P1, $n = 179$; HC-Pro, $n = 123$; CP, $n = 300$). Branch colors indicate the inferred results of the ancestral state reconstruction of host plants. Pie charts indicate the posterior probabilities of the host plants inferred as the ancestral state. The x-axis is scaled in years. The histogram (top left) shows the posterior probability of the root for each host plant; the gray bars indicate the posterior probability with error bars obtained from randomized replicates of the tip state.

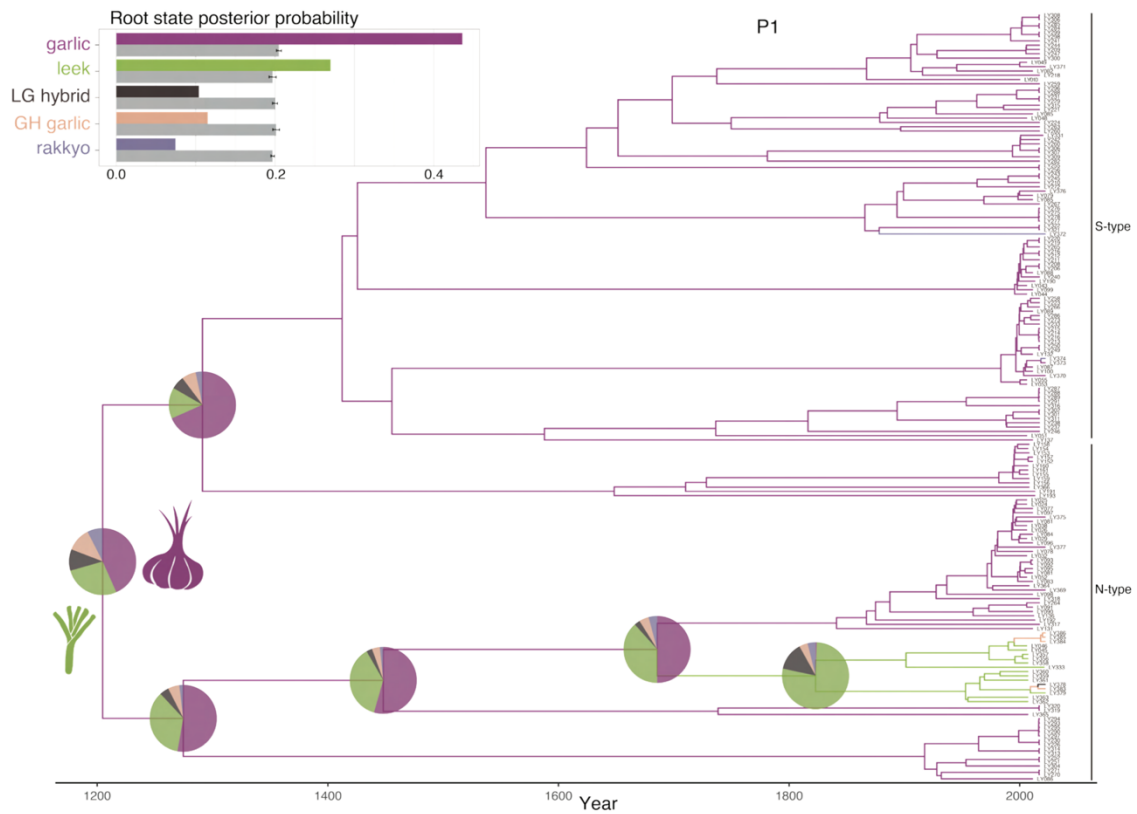


Fig. 11. Time-scaled MCC tree for the P1 genes. This is the same tree as those in Fig. 10D, but the tip labels are annotated with the isolate index number as summarized in Table 3. Branch colors indicate the host species in the inferred ancestral state reconstruction as defined in the key. Pie charts indicate the posterior probabilities of the host plants inferred as the ancestral state. The x-axis is scaled in years. The histogram (top left) shows posterior probability of the root for each host plant; the gray bars indicate the posterior probability with error bars obtained from randomized replicates of the tip state.

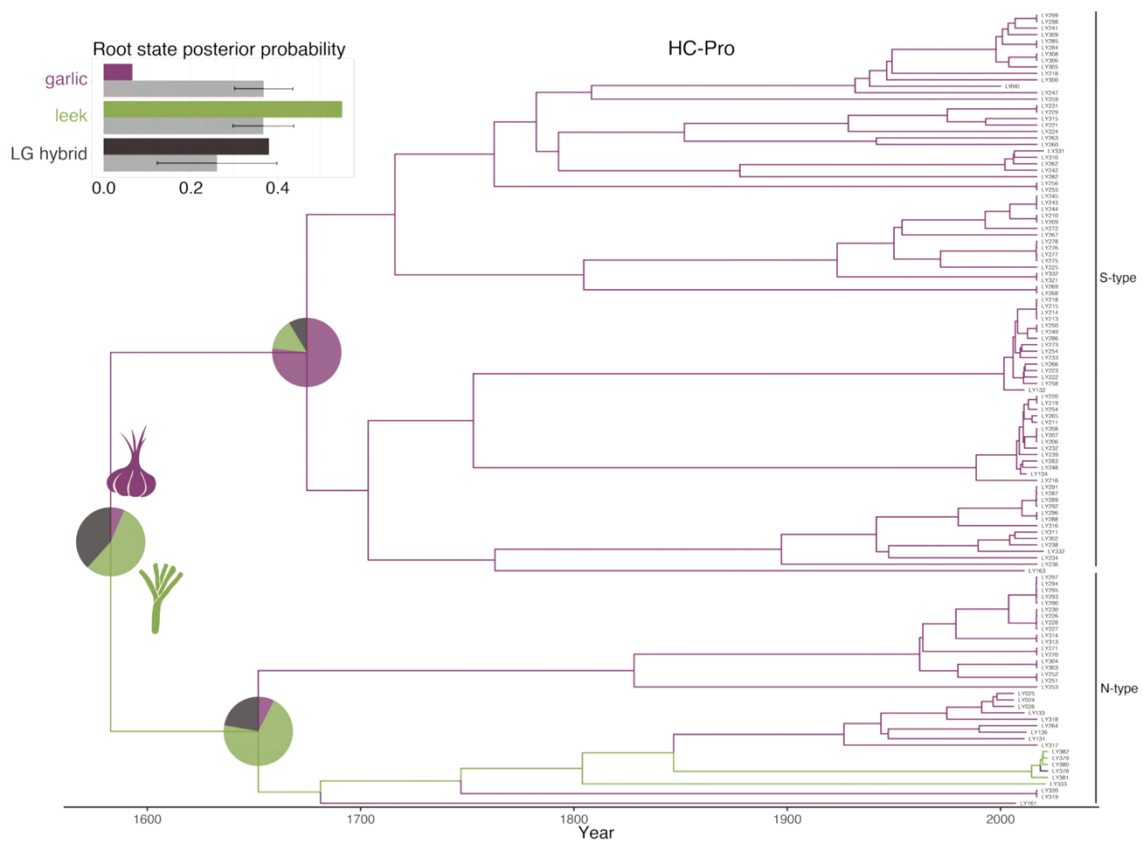


Fig. 12. Time-scaled MCC tree for the HC-Pro genes. This is the same tree as those in Fig. 10E, but the tip labels are annotated with the isolate index number as summarized in Table 3. Branch colors indicate the host species in the inferred ancestral state reconstruction as defined in the key. Pie charts indicate the posterior probabilities of the host plants inferred as the ancestral state. The x-axis is scaled in years. The histogram (top left) shows posterior probability of the root for each host plant; the gray bars indicate the posterior probability with error bars obtained from randomized replicates of the tip state.

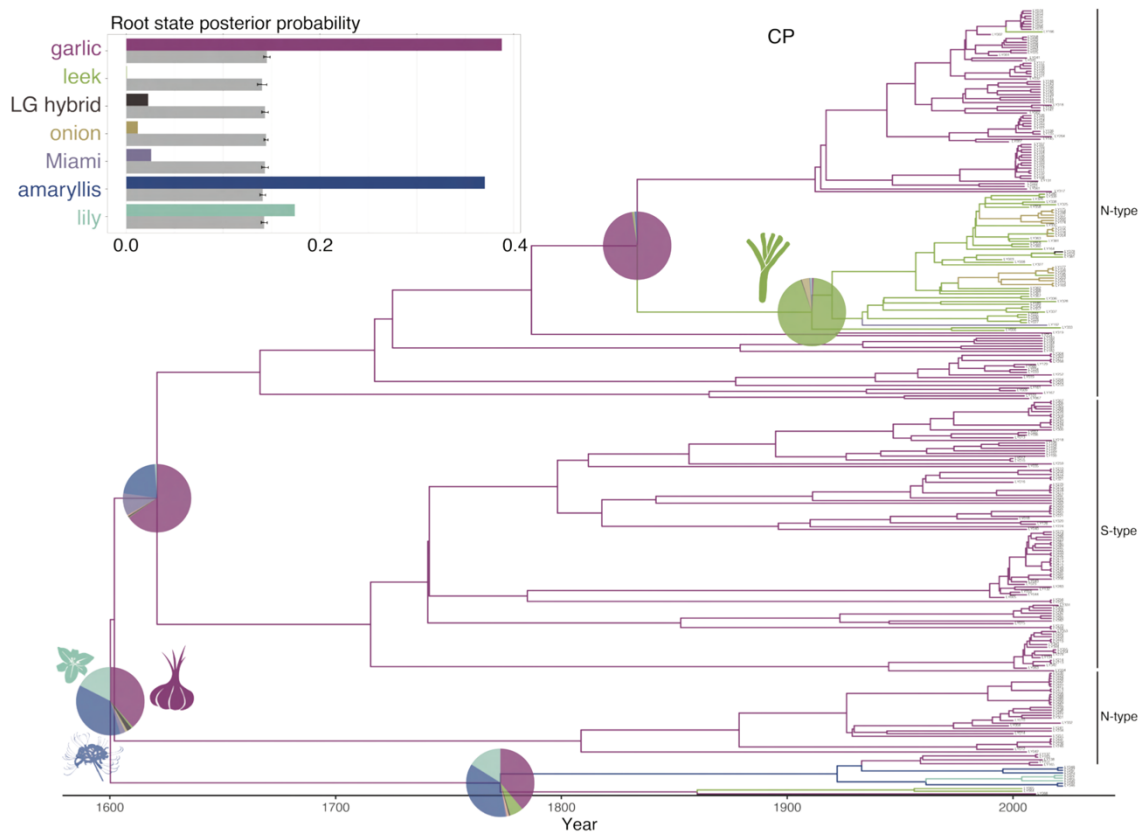


Fig. 13. Time-scaled MCC tree for the CP genes. This is the same tree as those in Fig. 10F, but the tip labels are annotated with the isolate index number as summarized in Table 3. Branch colors indicate the host species in the inferred ancestral state reconstruction as defined in the key. Pie charts indicate the posterior probabilities of the host plants inferred as the ancestral state. The x-axis is scaled in years. The histogram (top left) shows posterior probability of the root for each host plant; the gray bars indicate the posterior probability with error bars obtained from randomized replicates of the tip state.

ゼ (RdRP) である N1b 遺伝子においても同様の解析を行なった(Figs. 14-15)。N1b 遺伝子を用いた系統樹においては、全分離株と他の遺伝子領域と同様のフィルタリングを行なった両方の場合で、LYSV 祖先宿主としてニンニクが最も高い事後確率を示した(Figs. 14-15)。

7. 適応度推定

適応度推定には、前項で記述した全分離株による P1、HC-Pro、CP の 3 遺伝子領域における BEAST MCC 系統樹を入力ファイルとして用い、系統分岐上で多様化率(diversification rate)の変化がいつ起きたかを解析した。BAMM を用いて解析した結果、LYSV の進化の過程で多様化率が有意に変化したイベントは複数回起きていることが明らかになった (Fig. 16)。特に、P1 遺伝子の系統樹においては、S-type の発生時に LYSV の birth rate が大きく上昇し、高い適応度を獲得したことが明らかになった。また、リーキ特異的系統群の分岐においても、birth rate が変化したことが明らかになった。HC-Pro 遺伝子の系統樹においては、リーキ特異的系統群の分岐時に birth rate の変化は確認されなかった。一方で、CP 遺伝子の系統樹においては、リーキ特異的系統群の分岐時に birth rate が大きく上昇し、高い適応度を獲得したことが明らかになった。これらの結果は、LYSV が宿主を変える時に大きな適応度の変化が生じていることを示している。

8. RNA サイレンシングサプレッサー活性の定量

ポティウイルスの P1 タンパク質は、HC-Pro が持つ RSS 活性を促進する機能を有していることが報告されている(Kasschau & Carrington, 1998)。LYSV においては、P1 タンパク質が単独で強力な RSS 活性を示すことから(Sasaki et al., 2022)、本研究においても P1 タンパク質単独での RSS 活性を調査した。まず初めに、LYSV の自然宿主であるネギ属植物において RSS 活性を調査した。タマネギとリーキの両方において、リーキ系統とニンニク系統の両 LYSV に由来する P1 はどちらも RSS 活性を示さなかった(Figs. 17-18)。 *N. benthamiana* を宿主植物として用いた場合でも、P1 タンパク質は RSS 活性を示さなかった(Fig. 19)。

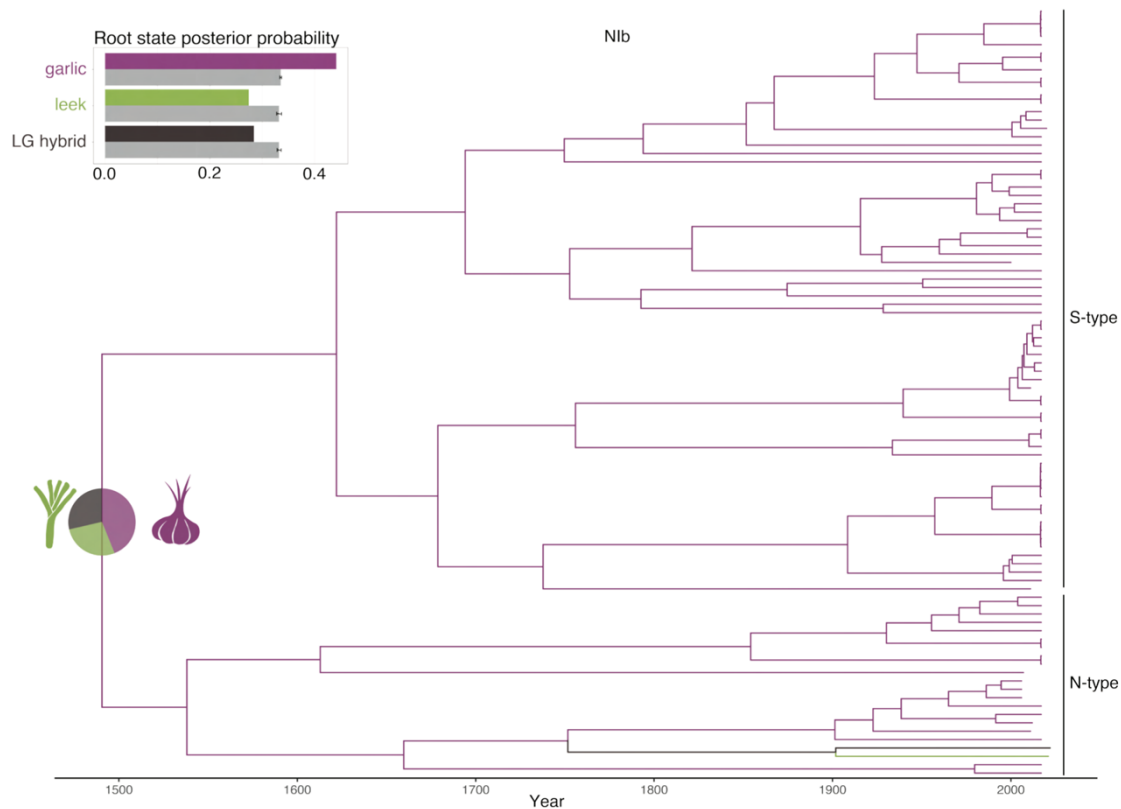


Fig. 14. Time-scaled MCC tree for the Nib genes based on the filtered data set. The taxon set were kept same with the analysis shown in the Fig. 10A-C. Branch colors indicate the host species in the inferred ancestral state reconstruction as defined in the key. Pie charts indicate the posterior probabilities of the host plants inferred as the ancestral state. The x-axis is scaled in years. The histogram (top left) shows posterior probability of the root for each host plant; the gray bars indicate the posterior probability with error bars obtained from randomized replicates of the tip state.

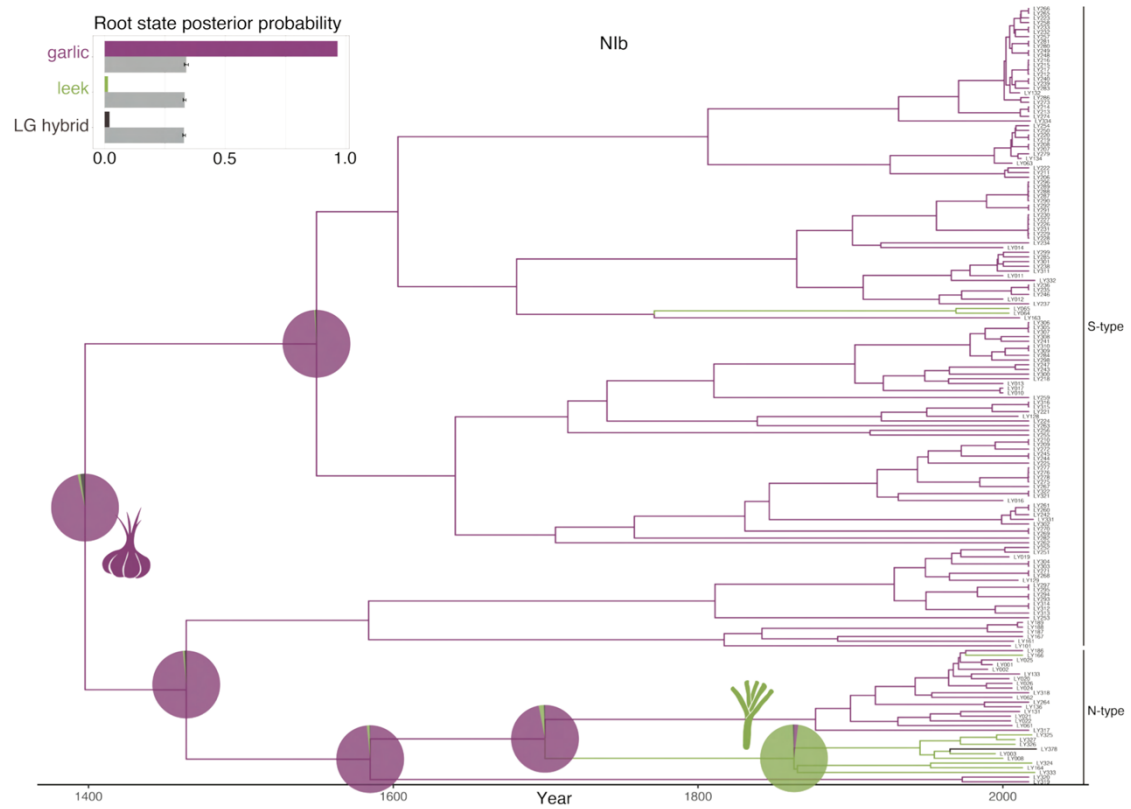


Fig. 15. Time-scaled MCC tree for the Nib genes. The tip labels are annotated with the isolate index number as summarized in Table 3. Branch colors indicate the host species in the inferred ancestral state reconstruction as defined in the key. Pie charts indicate the posterior probabilities of the host plants inferred as the ancestral state. The x-axis is scaled in years. The histogram (top left) shows posterior probability of the root for each host plant; the gray bars indicate the posterior probability with error bars obtained from randomized replicates of the tip state.

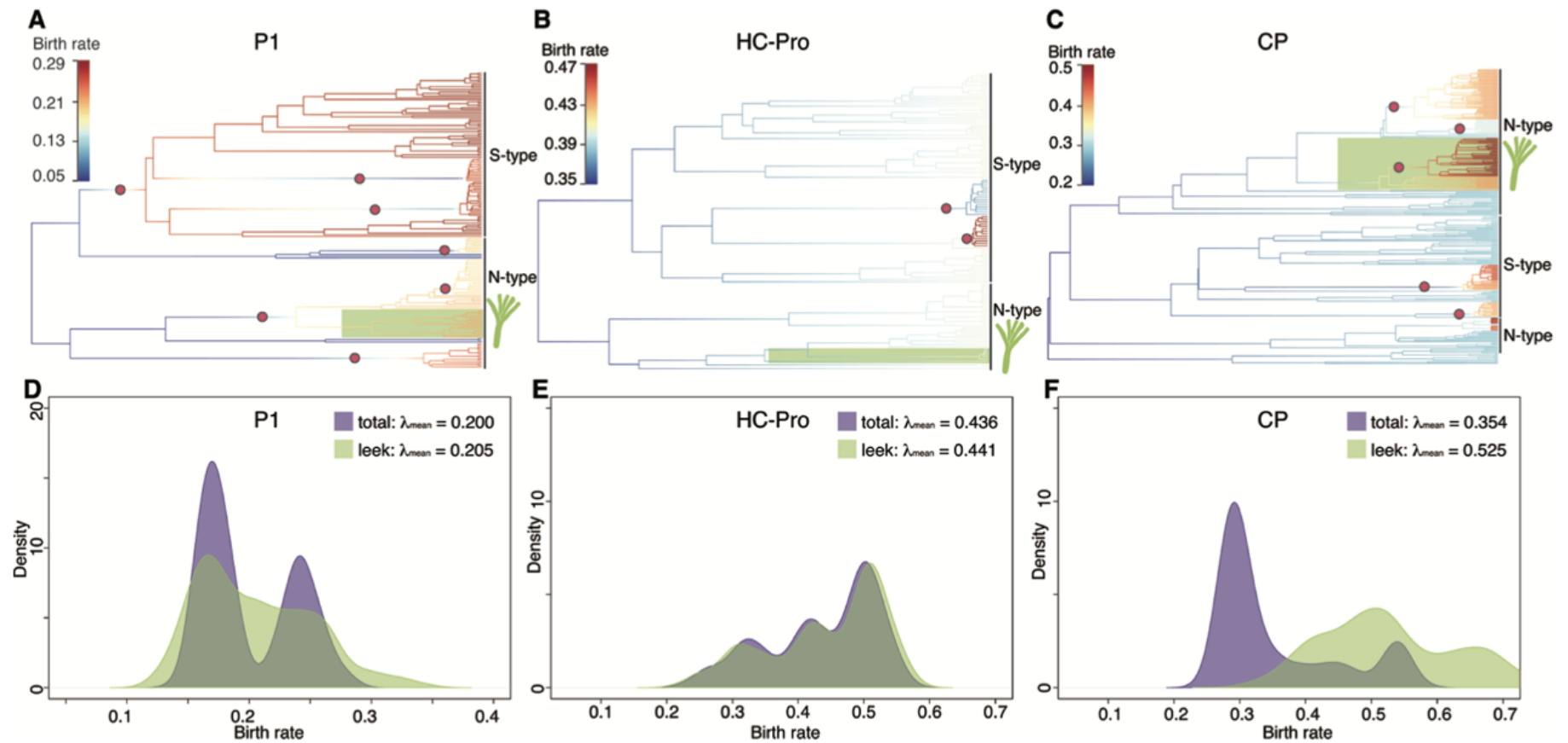


Fig. 16. Diversification dynamics during LYSV evolution inferred for the P1 (A and D), HC-Pro (B and E), and CP (C and F) genes. (A–C): Mean phylorate plots whose tree topologies were equivalent and derived from the BEAST output as shown in Fig. 10D through F. The leek (AAC) clade is shaded in green. Red circles on tree branches indicate epochs of rate shifts for each distinct shift configuration inferred from the maximum a posteriori probability. (D–F): Kernel density of birth rate

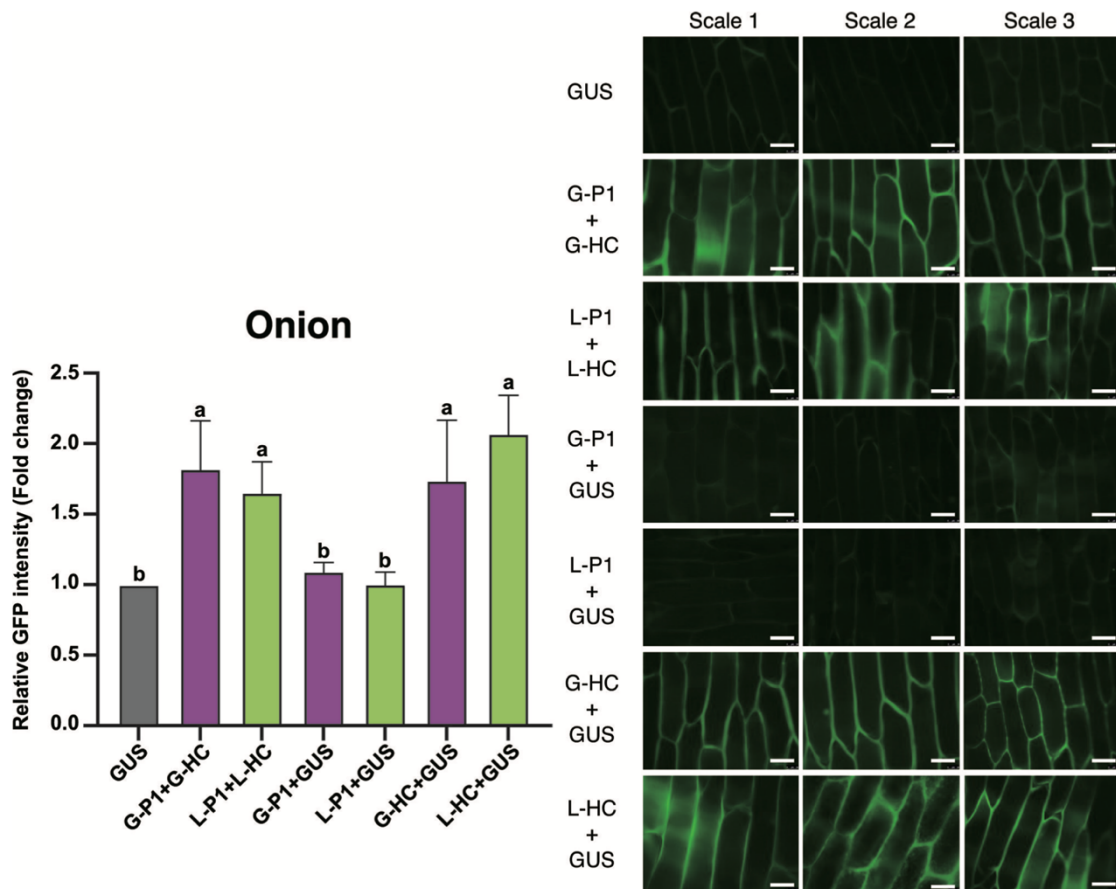


Fig. 17. RSS activity of LYSV P1 and HC-Pro at 3 dpa using onion epidermis. P1 and HC-Pro of garlic strains (G-P1 and G-HC) and leek strains (L-P1 and L-HC) were co-expressed with the GFP gene. Greater GFP fluorescence intensity indicates greater RSS activity. Three different onion scales (scales 1–3) are shown. Relative mean fold changes in GFP fluorescence intensity, calculated using the LAS AF software (Leica), are shown in the barplot with the value of GUS set to 1.0. Fill colors in the barplot represent the LYSV strains (purple, garlic strain; green, leek strain). Fold change values were log-transformed, then analyzed by one-way ANOVA ($P < 0.0001$) followed by Tukey's multiple comparison test. Different letters above the bars indicate a significant difference between means ($P < 0.05$).

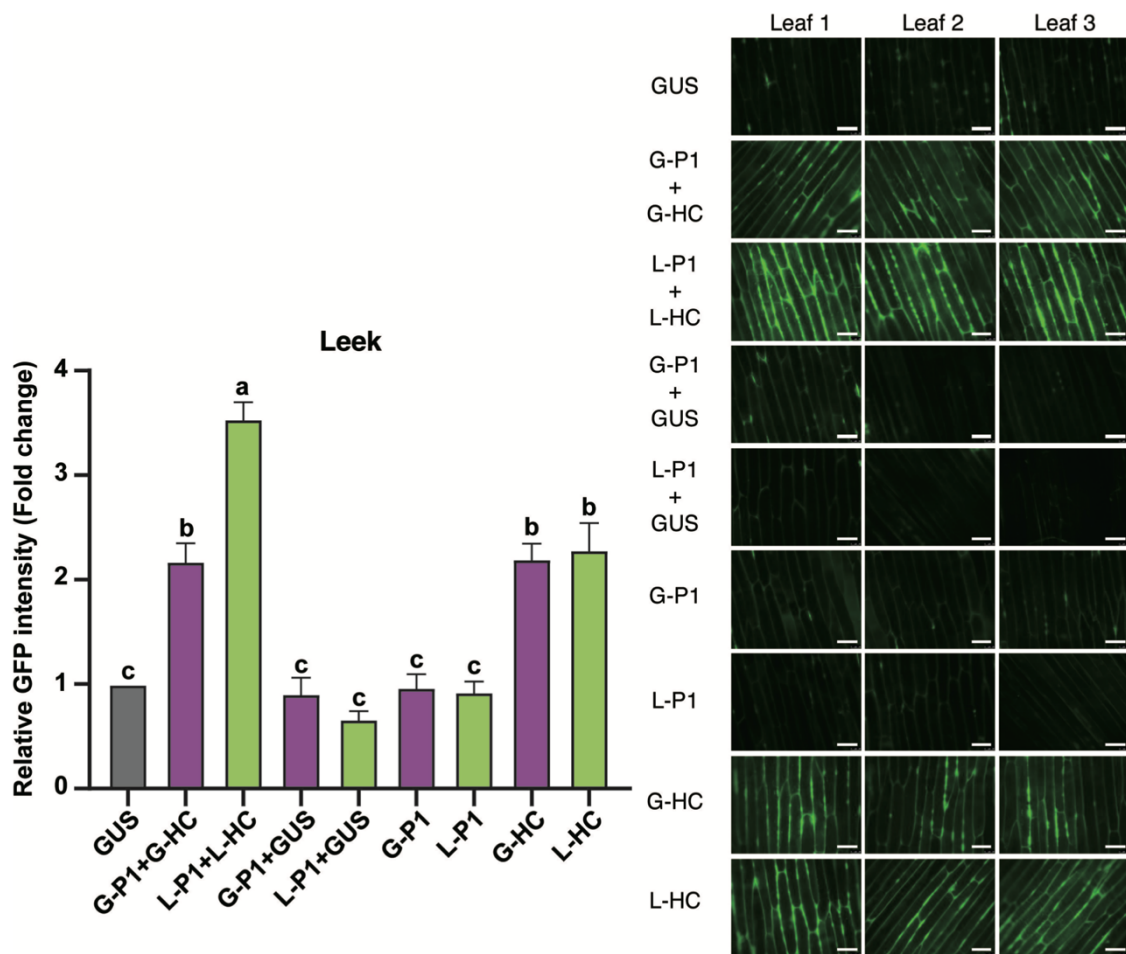


Fig. 18. RSS activity of LYSV P1 and HC-Pro at 3 dpa using leek leaves. P1 and HC-Pro of garlic strain (G-P1 and G-HC) and leek strain(L-P1 and L-HC) were co-expressed with the GFP gene. Greater GFP fluorescence intensity indicates greater RSS activity. Three different infiltrated leaves (leaves1–3) are shown. Relative mean fold changes in GFP fluorescence intensity, calculated using the LAS AF software (Leica), are shown in the barplot with the value of GUS set to 1.0. Fill colors in the barplot represent the LYSV strains (purple, garlic strain; green, leek strain). Fold change values were log-transformed, then analyzed by one-way ANOVA ($P < 0.0001$) followed by Tukey's multiple comparison test. Different letters above the bars indicate a significant difference between means ($P < 0.05$).

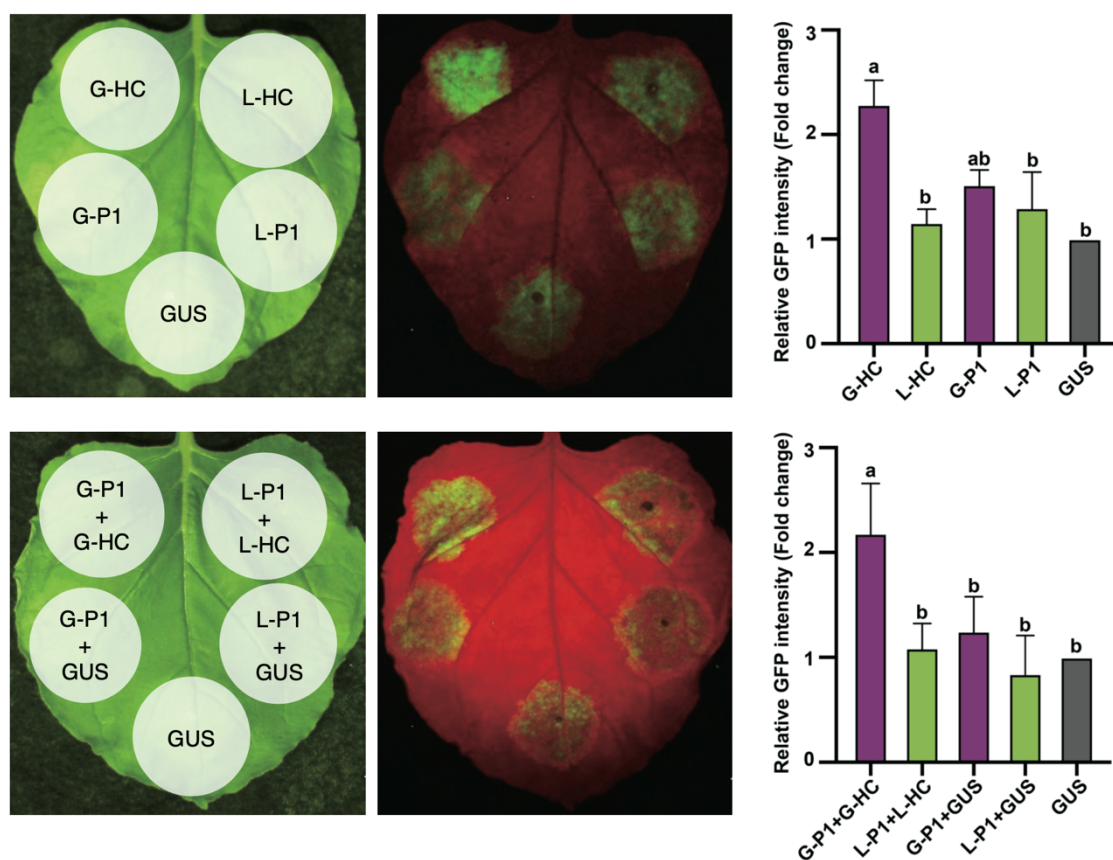


Fig. 19. RSS activity of LYSV P1 and HC-Pro at 5 dpa of *Nicotiana benthamiana* leaves. The upper panels of the left pictures show the agroinfiltrated leaves that P1 and HC-Pro proteins of garlic strain (G-P1 and G-HC) and leek strain (L-P1 and L-HC) were expressed alone with the GFP gene, whereas the lower panels show that P1 and HC-Pro were co-expressed. Greater GFP fluorescence intensity indicates greater RSS activity. The GUS gene was used as a negative control. GFP fluorescence intensity was analyzed by the ImageJ software (<https://imagej.nih.gov/ij/index.html>). Relative mean fold-change values, when the control, GUS was set to 1.0, are shown in the barplot. Fill colors in the barplot represent the LYSV strains (purple, garlic strain; green, leek strain). One-way ANOVA test was conducted (upper plot, $P = 0.0051$; lower plot, $P = 0.0002$), followed by Tukey's multiple comparison test. Different letters above the bars indicate a significant difference between means ($P < 0.05$).

次に、HC-Pro を単独発現させた場合と、P1 と共発現させた場合において、それぞれの RSS 活性を調査した。タマネギを宿主植物として用いた場合、ニンニク系統とリーキ系統の両方において HC-Pro が RSS 活性を示すことが明らかになった(Fig. 17)。また、P1 と共発現させても HC-Pro 単独発現時と同程度の RSS 活性を示した(Fig. 17)。一方リーキを宿主植物として用いた場合、HC-Pro の単独発現時にはリーキ系統とニンニク系統の両 LYSV は同程度の RSS 活性を示すのに対し、リーキ系統由来の LYSV では P1 と HC-Pro の共発現下において RSS 活性の強い相乗効果が観察された(Fig. 18)。この相乗効果現象は、ニンニク系統の LYSV では観察されなかった。*N. benthamiana* の場合では、HC-Pro の単独発現下において、ニンニク系統の LYSV がリーキ系統の LYSV よりも高い RSS 活性を示した(Fig. 19)。P1 と HC-Pro の共発現下においても、ニンニク系統の LYSV がリーキ系統の LYSV よりも高い RSS 活性を示した。以上の結果は、LYSV の RSS 活性は宿主植物によって異なり、その違いを決定する一因として、P1 による HC-Pro の RSS 活性促進機能が重要であることを示唆している。

考察

本研究は、リーキ、ニンニク、およびジャンボニンニクの異種間交雑種を発見したことで、LYSV のユニークな宿主適応進化および宿主との共進化機構の重要な知見を得ることができた。ポティウイルスにおいては、地理的分離が進化の大きな原動力となっている報告が多数示されており (Gao et al., 2020; Kawakubo et al., 2021; Ohshima et al., 2002)、特定の宿主に適応的に進化してきた報告は例がない。しかし、LYSV は他のポティウイルスとは違い、宿主適応が大きな進化の原動力になっていることが示唆された。本研究で発見した交雑種(LG hybrid)においては、LYSV の蓄積量が非常に高かったため、RNA-seq によって LYSV のゲノム配列を容易に決定することができた。この交雑種に感染していた LYSV のゲノム配列は、以前に報告されたニンニク由来の LYSV とは大きく異なり、リーキに感染する LYSV のものと高い配列類似性を示していた(Fig. 6)。その後、リーキとジャンボニンニクにおいて LYSV の蓄積量を調査すると、どちらもリーキ特異的系統群に属する LYSV の感染が確認されたにも関わらず、リーキにおける蓄積量は極めて低いのに対しジャンボニンニクには高濃度の LYSV が蓄積していることが示されたため(Figs. 3-4)、その違いを決定している因子を探るべく解析・実験を行なった。

ネギ属植物の系統関係については、トランスクリプトーム、核および葉緑体の配列解析により、ニンニクとリーキはタマネギや小ネギなどの他のネギ属植物より近縁な関係にあることが示されている(Hirscheegger et al., 2010; Jimenez et al., 2019; Sun et al., 2020; Zhu et al., 2017)。ジャンボニンニクとリーキは、*Allium ampeloprasum* complex (AAC)として、同種の変異種として分類されており、kurrat やパールオニオンも同じ AAC に含まれている(Bothmer, 1970; Fritsch & Friesen, 2002; Hirscheegger et al., 2006)。AAC の変異種間は染色体の多倍性が異なり、リーキは4倍体 ($2n = 4X = 32$) で、ジャンボニンニク6倍体 ($2n = 6X = 48$) である(Fritsch & Friesen, 2002)。また、これらのネギ属植物においては、アリナーゼ酵素のアミノ酸配列がそれぞれユニークに保存されていることが報告されている

(Ariga et al., 2002)。ジャンボニンニクではニンニクには無いユニークなアミノ酸が2箇所保存されている(Ariga et al., 2002)。本研究で用いた交雑種(LG hybrid)は、リーキとジャンボニンニクにユニークなアミノ酸を両方持っていた(Fig. 2)。LYSV がリーキ細胞内で高濃度に蓄積されない理由は明らかでないが、LYSV はリーキ細胞内では潜在感染のように弱毒化することで細胞内に保たれ続けるような進化戦略を取ったのだと考えられる。従って、ニンニクに由来する宿主因子の中に、LYSV の増殖を促進するメカニズムが存在する可能性が考えられる(Fig. 20)。

ネギ属植物は、13 世紀ごろにマルコポーロによって中央アジアから地中海沿岸地域にもたらされ、これにより現在のネギ属の種分岐が促進されたと考えられる(Pratt, 2010)。多くの植物ウイルスにおいて、このような人為的要因の介入が進化方向性を決定してきた例が報告されている(Gao et al., 2020; Kawakubo et al., 2021)。本研究において、LYSV の P1 遺伝子を用いた時間系統樹により、LYSV 共通祖先の分岐は 1264(95% ベイズ信用区間: 1070-1451)年と推定された。すなわち、LYSV の進化はネギ属植物の種分岐とほぼ同じ時間軸で起きていることが示唆される(Fig. 20)。

LYSV の多様化率が系統分岐過程で複数回起きていることは、LYSV の適応度が何度も変化していることを示唆している(Fig. 16)。Sclerotiniaceae に属する植物病原糸状菌を用いた先行研究においても同様に、宿主を変える際に糸状菌の多様化率が変化していることが示されている(Navaud et al., 2018)。

LYSV の N-type と S-type を決定する P1 タンパク質の部分欠失領域には、反復配列モチーフ(IxFGSFETP)が存在していることが明らかになった(Figs. 8-9)。この保存されているモチーフの機能は不明であるが、既報のポティウイルス共通モチーフ(IxFG)を一部含んでいる。この FG モチーフは、ストレス顆粒(stress granules: SGs)の形成に関与していると報告されている(Panas et al., 2015)。本研究では、宿主植物による RSS 活性の違いが進化方向性を決定する重要因子の一つとなっていることをしめした。しかし、SGs のようなその他の要因についても複雑に関与している可能性も十分考えられる。

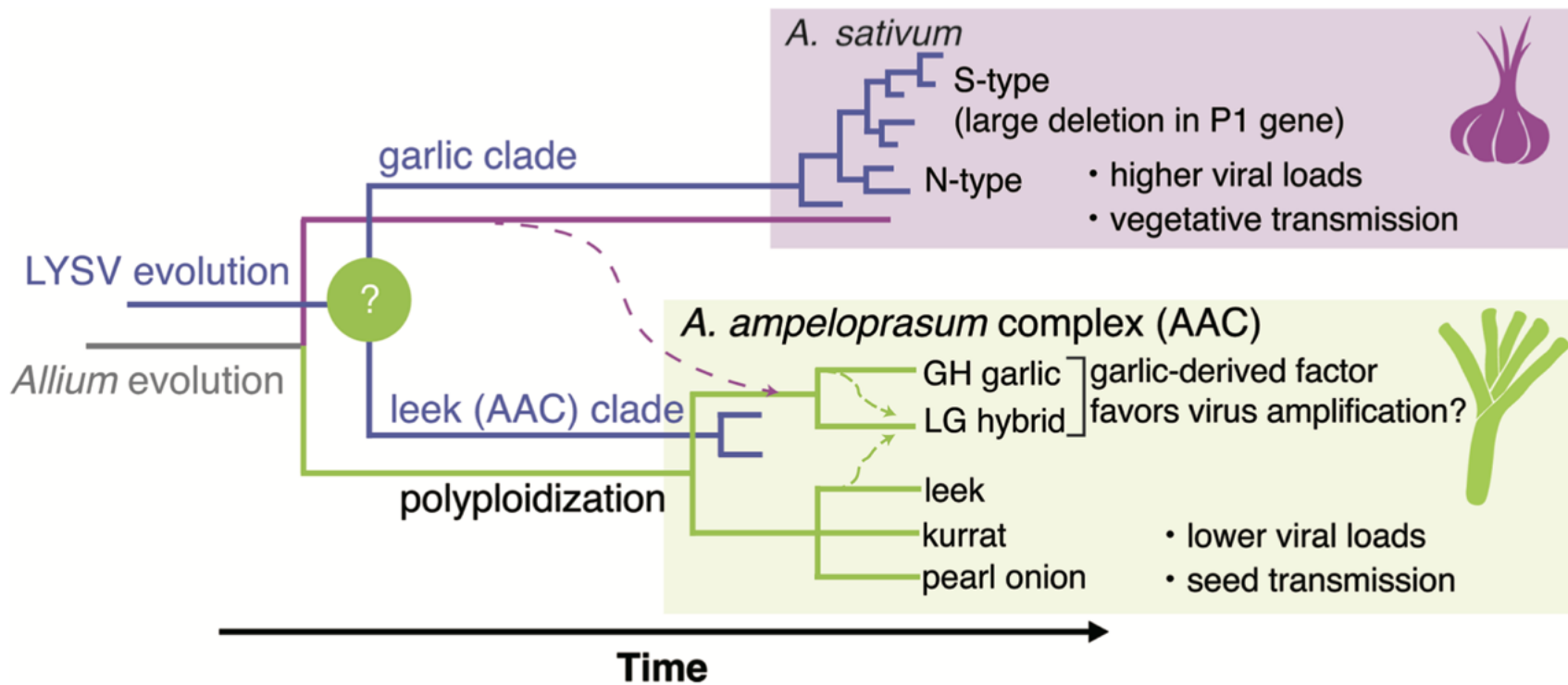


Fig. 20. Possible scenario of the coevolution of LYSV and *Allium* species plants. The dashed line indicates a potential horizontal chromosome transfer event among *Allium* species.

本章では、LYSV の進化において、P1 と HC-Pro によって決定される RSS 活性の宿主ごとの違いが重要であることを示した。そして、LYSV の共通祖先は分岐初期においてリーキを宿主としていたことが示唆され、LYSV はネギ属植物と密接な共進化関係にあることも明らかにした。

第 IV 章 カブモザイクウイルスの日本における系統地理学的解析

目的

植物病原体の時空間的な分布および拡散経路を理解することは、感染拡大を防ぎ、効果的な防除戦略を立てる上でも非常に重要である。しかし、植物病原体の分布は国や地域によっても大きく異なるのに加え、流行動態をリアルタイムに追跡することは困難である。系統地理学は、主に生物の進化の過程において集団の移動や分布を地理情報と結びつける学問領域である(Searle, 2000)。系統地理学的手法を大規模に調査されたゲノムデータに適用し、これまで数多くの重要病原体の拡散経路や感染動態が明らかにされてきた(Dudas et al., 2017; Metsky et al., 2017; Wilkinson et al., 2021)。

カブモザイクウイルス (turnip mosaic virus: TuMV) は、ポティウイルス科ポティウイルス属の 1 種である。アブラナ科をはじめとした多くの植物種に感染する植物ウイルスであり、園芸作物や農作物に発生する重要病害ウイルスとして、キュウリモザイクウイルス (cucumber mosaic virus: CMV) に次いで二番目にランクされている(Walsh & Jenner, 2002)。1921 年に米国で TuMV が初報告(Gardner & Kendrick, 1921; Schultz, 1921)された後、それらを皮切りに数々の国においてアブラナ科植物に同様のモザイク症状を引き起こす事例が報告されるようになった(Smith, 2008)。日本においては、1951 年の吉井による報告が TuMV によるモザイク病に関する最初の記述である(Yoshii, 1951)。また、TuMV の進化には、ゲノムの組換えが重要であると考えられており、ゲノムの組換えが TuMV の病原性を拡大させる 1 要因であることが報告されている(Ohshima et al., 2007)。そこで本研究では、日本全土から採集された TuMV 分離株のゲノム配列をもとに組換え解析を詳細に行なった。そして、日本における TuMV の組換え型の分布を明らかにするとともに、系統地理学的手法により時空間的拡散様式を解析した。

材料と方法

1. 塩基配列の決定およびデータセットの作成

日本全土から採集されたモザイク症状を呈するアブラナ科野菜を ELISA 法 (Clark & Adams, 1977) に供試し、TuMV 感染の有無を確認した。感染が確認されたサンプルは、0.01M カリウムリン酸緩衝液 (pH 7.0) を加えて磨砕後、*Chenopodium quinoa* に汁液接種し、単一病斑分離を 2 回繰り返した。その後、博多すわりカブ (*Brassica rapa*) もしくは *N. benthamiana* に接種してウイルスを増殖させた。接種 2-3 週間後に Isogen II (Nippon Gene, Japan) を用いて 上葉から RNA を抽出し、ダイレクトシーケンス法により TuMV のゲノム配列を決定した。cDNA の合成には、PrimeScript Molony murine leukemia virus 逆転写酵素 (Takara Bio, Japan) を用い、high-fidelity Platinum Pfx DNA 合成酵素 (Invitrogen/ThermoFisher Scientific, Japan) を用いて PCR 反応によりウイルス DNA を増幅させた。これらのゲノム配列に、公共データベースから取得した配列を合わせ、採集年代、採集地域 (都市名、経度緯度)、分離宿主、病原性型 (Ohshima et al., 2002) の情報を統合し、合計 370 分離株からなる日本産 TuMV のゲノムデータセットを作成した (Fig. 21, Table 4)。Clustal X v2.1 (Larkin et al., 2007) を用いてコドンベースのマルチプルシーケンスアライメントを作成した。

2. 組換え解析

RDP5 (Martin et al., 2020) に実装されている、RDP、BOOTSCAN、GENECONV、MAXCHI、CHIMAERA、および SISCAN を用いて TuMV のゲノムワイド組換え解析を行なった。各プログラムの統計的仮説検定において、 $P < 10^{-6}$ を有意水準とみなし、3 つ以上の異なるプログラムで指示された組換え信号を組換え部位とみなした。また、本法で検出された組換え部位については、SimPlot (Lole et al., 1999) を用いて塩基配列類似度を可視化し確認した。

Table 4. TuMV Japanese isolates analysed in this study.

Island, district and prefecture	Isolate name	Original host	Location (City or Town)	Collecting year	Host-infecting type ^a	Latitude	Longitude	Reference	Accession number
Hokkaido Island/district (n=33)									
Hokkaido Prefecture (n=33)	HOB537J	<i>Brassica juncea</i> var. <i>integrifolia</i>	Sapporo	2004	BR	43.1428	141.3109	Kawakubo et al. (2022)	LC639421
	HOB542J	<i>Brassica oleracea</i>	Sapporo	2005	BR	43.1428	141.3109	Kawakubo et al. (2022)	LC639422
	HOB546J	<i>Brassica</i> sp.	Sapporo	2006	B(R)	43.1294	141.3467	Kawakubo et al. (2022)	LC639423
	HOD500J	<i>Raphanus sativus</i>	Ohno	1998	BR	41.8345	140.6709	Kawakubo et al. (2022)	LC639424
	HOD511J	<i>R. sativus</i>	Makkari	1998	BR	42.7630	140.8037	Kawakubo et al. (2022)	LC639425
	HOD512J	<i>R. sativus</i>	Makkari	1998	BR	42.7630	140.8037	Kawakubo et al. (2022)	LC639426
	HOD517J	<i>R. sativus</i>	Kimobetsu	1998	BR	42.7955	140.9346	Tomimura et al. (2003)	AB093617
	HOD518J	<i>R. sativus</i>	Memuro	1998	BR	42.9119	143.0509	Kawakubo et al. (2022)	LC639427
	HOD519J	<i>R. sativus</i>	Obihiro	1998	BR	42.9233	143.1972	Kawakubo et al. (2022)	LC639428
	HOD523J	<i>R. sativus</i>	Fukagawa	2001	BR	43.7233	142.0533	Kawakubo et al. (2022)	LC639429
	HOD525J	<i>R. sativus</i>	Ishikari	2002	BR	43.1714	141.3155	Kawakubo et al. (2022)	LC639430
	HOD526J	<i>R. sativus</i>	Sapporo	2002	BR	43.1449	141.3586	Kawakubo et al. (2022)	LC639431
	HOD527J	<i>R. sativus</i>	Sapporo	2002	BR	43.1449	141.3586	Kawakubo et al. (2022)	LC639432
	HOD528J	<i>R. sativus</i>	Sapporo	2002	BR	43.1449	141.3586	Kawakubo et al. (2022)	LC639433
	HOD529J	<i>R. sativus</i>	Sapporo	2002	BR	43.1449	141.3586	Kawakubo et al. (2022)	LC639434
	HOD530J	<i>R. sativus</i>	Sapporo	2002	BR	43.1194	141.3994	Kawakubo et al. (2022)	LC639435
	HOD531J	<i>R. sativus</i>	Sapporo	2002	BR	43.1194	141.3994	Kawakubo et al. (2022)	LC639436
	HOD532J	<i>R. sativus</i>	Nanae	2002	BR	41.8957	140.6945	Kawakubo et al. (2022)	LC639437
	HOD533J	<i>R. sativus</i>	Nanae	2002	BR	41.8957	140.6945	Kawakubo et al. (2022)	LC639438
	HOD534J	<i>R. sativus</i>	Abuta	2002	BR	42.8467	140.8507	Kawakubo et al. (2022)	LC639439
	HOD535J	<i>R. sativus</i>	Otaru	2002	BR	43.1650	141.2389	Kawakubo et al. (2022)	LC639440
	HOD536J	<i>R. sativus</i>	Sapporo	2002	BR	43.0888	141.3754	Kawakubo et al. (2022)	LC639441
	HOD539J	<i>R. sativus</i>	Sapporo	2005	BR	43.1194	141.3994	Kawakubo et al. (2022)	LC639442
	HOD540J	<i>R. sativus</i>	Sapporo	2005	BR	43.1194	141.3994	Kawakubo et al. (2022)	LC639443
	HOD541J	<i>R. sativus</i>	Sapporo	2005	BR	43.1428	141.3109	Kawakubo et al. (2022)	LC639444
	HOD543J	<i>R. sativus</i>	Sapporo	2006	BR	43.1449	141.3586	Kawakubo et al. (2022)	LC639445
	HOD544J	<i>R. sativus</i>	Sapporo	2006	BR	43.1428	141.3109	Kawakubo et al. (2022)	LC639446
HOD545J	<i>R. sativus</i>	Sapporo	2006	BR	43.1428	141.3109	Kawakubo et al. (2022)	LC639447	
HOD547J	<i>R. sativus</i>	Sapporo	2007	BR	43.0927	141.4094	Kawakubo et al. (2022)	LC639448	
HOD548J	<i>R. sativus</i>	Sapporo	2007	BR	43.1449	141.3586	Kawakubo et al. (2022)	LC639449	
HOD549J	<i>R. sativus</i>	Sapporo	2007	BR	43.1428	141.3109	Kawakubo et al. (2022)	LC639450	
HOD551J	<i>R. sativus</i>	Sapporo	2009	BR	43.1428	141.3109	Kawakubo et al. (2022)	LC639451	
HOD553J	<i>R. sativus</i>	Sapporo	2009	BR	43.1449	141.3586	Kawakubo et al. (2022)	LC639452	
Tohoku district (n=61)									
Aomori Prefecture (n=22)	AD178J	<i>R. sativus</i>	Rokunohe	1998	BR	40.6096	141.3247	Ohshima et al. (2007)	AB252094
	AD181J	<i>R. sativus</i>	Tohoku	1998	BR	40.7280	141.2578	Ohshima et al. (2007)	AB252095
	AD851J	<i>R. sativus</i>	Ohhata	2002	BR	41.3932	141.0455	Kawakubo et al. (2022)	LC639361
	AD853J	<i>R. sativus</i>	Ohhata	2002	BR	41.3932	141.0455	Ohshima et al. (2007)	AB252096
	AD855J	<i>R. sativus</i>	Ohminato	2002	BR	41.3932	141.0455	Ohshima et al. (2007)	AB252097
	AD858J	<i>R. sativus</i>	Kawauchi	2002	BR	41.2722	140.9435	Kawakubo et al. (2022)	LC639362
	AD860J	<i>R. sativus</i>	Kawauchi	2002	BR	41.2722	140.9435	Ohshima et al. (2007)	AB252098
	AD867J	<i>R. sativus</i>	Tohoku	2002	BR	40.7280	141.2578	Kawakubo et al. (2022)	LC639363
	AD871J	<i>R. sativus</i>	Tohoku	2002	BR	40.7667	141.2503	Kawakubo et al. (2022)	LC639364
	AD873J	<i>R. sativus</i>	Tohoku	2002	BR	40.7280	141.2578	Kawakubo et al. (2022)	LC639365
	AD874J	<i>R. sativus</i>	Tohoku	2002	BR	40.7280	141.2578	Kawakubo et al. (2022)	LC639366
	AD877J	<i>R. sativus</i>	Shimoda	2002	BR	40.6020	141.4182	Kawakubo et al. (2022)	LC639367
	AD878J	<i>R. sativus</i>	Shimoda	2002	BR	40.6020	141.4182	Kawakubo et al. (2022)	LC639368
	AD880J	<i>R. sativus</i>	Rokunohe	2002	BR	40.6096	141.3247	Kawakubo et al. (2022)	LC639369
	AD886J	<i>R. sativus</i>	Fukuchi	2002	B(R)	40.4778	141.3736	Kawakubo et al. (2022)	LC639370
	AD887J	<i>R. sativus</i>	Fukuchi	2002	BR	40.4778	141.3736	Kawakubo et al. (2022)	LC639371
	AD888J	<i>R. sativus</i>	Fukuchi	2002	BR	40.4778	141.3736	Kawakubo et al. (2022)	LC639372
	AD911J	<i>R. sativus</i>	Hirosaki	2000	BR	40.5780	140.4826	Kawakubo et al. (2022)	LC639373
	AD912J	<i>R. sativus</i>	Hirosaki	2000	BR	40.5780	140.4826	Kawakubo et al. (2022)	LC639374
	AD915J	<i>R. sativus</i>	Hirosaki	2000	BR	40.5872	140.4719	Kawakubo et al. (2022)	LC639375
AT181J	<i>Eustoma grandiflorum</i>	Aomori	1998	BR	40.8222	140.7474	Ohshima et al. (2007)	AB252102	
H1J	<i>R. sativus</i>	Hirosaki	1996	BR	40.6031	140.4641	Ohshima et al. (2007)	AB252118	
Iwate Prefecture (n=9)	IWD024J	<i>R. sativus</i>	Kitakami	2000	BR	39.2868	141.1132	Kawakubo et al. (2022)	LC639464
	IWD027J	<i>R. sativus</i>	Morioka	2000	BR	39.7020	141.1543	Kawakubo et al. (2022)	LC639465
	IWD028J	<i>R. sativus</i>	Niisato	2000	BR	39.6266	141.7908	Kawakubo et al. (2022)	LC639466
	IWD032J	<i>R. sativus</i>	Iwaizumi	2000	BR	39.8432	141.7965	Ohshima et al. (2007)	AB252120
	IWD033J	<i>R. sativus</i>	Morioka	2000	BR	39.7020	141.1543	Kawakubo et al. (2022)	LC639467
	IWD037J	<i>R. sativus</i>	Senmaya	2000	BR	39.3886	141.1169	Kawakubo et al. (2022)	LC639468
	IWD038J	<i>R. sativus</i>	Yahaba	2000	BR	39.6060	141.1430	Ohshima et al. (2007)	AB252121
	IWD042J	<i>R. sativus</i>	Ohfunato	2000	BR	39.0819	141.7085	Kawakubo et al. (2022)	LC639469
	IWD043J	<i>R. sativus</i>	Rikuzentakada	2000	BR	39.0278	141.6255	Kawakubo et al. (2022)	LC639470
Miyagi Prefecture (n=5)	MYD011J	<i>R. sativus</i>	Yoneyama	2000	BR	38.6325	141.2301	Kawakubo et al. (2022)	LC639530
	MYD012J	<i>R. sativus</i>	Kannari	2000	BR	38.8424	141.0681	Kawakubo et al. (2022)	LC639531
	MYD013J	<i>R. sativus</i>	Yamamoto	2000	BR	37.9618	140.8782	Ohshima et al. (2007)	AB252128
	MYD014J	<i>R. sativus</i>	Watarai	2000	BR	38.0378	140.8525	Kawakubo et al. (2022)	LC639532
	MYD015J	<i>R. sativus</i>	Kesennuma	2000	BR	38.9076	141.5703	Ohshima et al. (2007)	AB252129
Akita Prefecture (n=6)	AKD161J	<i>R. sativus</i>	Ogachi	1998	BR	39.1797	140.6672	Ohshima et al. (2007)	AB252099
	AKD932J	<i>R. sativus</i>	Hachiryu	2000	BR	40.0996	140.0044	Kawakubo et al. (2022)	LC639376
	AKD934J	<i>R. sativus</i>	Hachiryu	2000	BR	40.0996	140.0044	Ohshima et al. (2007)	AB252100
	AKD938J	<i>R. sativus</i>	Yuzawa	2000	BR	39.1644	140.4957	Kawakubo et al. (2022)	LC639377
	AKH937J	<i>Brassica pekinensis</i>	Yuzawa	2000	BR	39.1644	140.4957	Ohshima et al. (2007)	AB252101
	AKT939J	<i>Brassica rapa</i>	Yuzawa	2000	BR	39.1644	140.4957	Kawakubo et al. (2022)	LC639378

^a Host type B; *Brassica*, isolates infected *B. rapa* cv. Hakatasuwari systemically giving mosaic symptoms. Host type (B); isolates infected *B. rapa* only occasionally. Host type BR; these isolates infected both *B. rapa* and *R. sativus* systemically giving mosaic symptoms. Host type B(R); isolates infected *B. rapa* systemically giving mosaic symptoms and infected *R. sativus* only occasionally.

Table 4. Continued.

Island, district and prefecture	Isolate name	Original host	Location (City or Town)	Collecting year	Host-infecting type ^a	Latitude	Longitude	Reference	Accession number
Yamagata Prefecture (n=12)	YAD016J	<i>R. sativus</i>	Kawanishi	2000	BR	38.0044	140.0459	Kawakubo et al. (2022)	LC639659
	YAD019J	<i>R. sativus</i>	Sagae	2000	BR	38.3805	140.2764	Kawakubo et al. (2022)	LC639660
	YAD020J	<i>R. sativus</i>	Shirataka	2000	BR	38.1831	140.0985	Ohshima et al. (2007)	AB252140
	YAD022J	<i>R. sativus</i>	Yamagata	2000	BR	38.2554	140.3396	Kawakubo et al. (2022)	LC639661
	YAD925J	<i>R. sativus</i>	Yonezawa	2000	BR	37.9223	140.1167	Kawakubo et al. (2022)	LC639662
	YAD926J	<i>R. sativus</i>	Yonezawa	2000	BR	37.9223	140.1167	Kawakubo et al. (2022)	LC639663
	YAD927J	<i>R. sativus</i>	Tsuruoka	2000	BR	38.7272	139.8267	Kawakubo et al. (2022)	LC639664
	YAD928J	<i>R. sativus</i>	Tsuruoka	2000	BR	38.7272	139.8267	Kawakubo et al. (2022)	LC639665
	YAD929J	<i>R. sativus</i>	Tsuruoka	2000	BR	38.7272	139.8267	Kawakubo et al. (2022)	LC639666
	YAD930J	<i>R. sativus</i>	Tsuruoka	2000	BR	38.7272	139.8267	Kawakubo et al. (2022)	LC639667
	YAH931J	<i>B. pekinensis</i>	Tsuruoka	2000	BR	38.7272	139.8267	Kawakubo et al. (2022)	LC639668
YAL018J	<i>Lactuca sativa</i>	Sagae	2000	BR	38.3805	140.2764	Ohshima et al. (2007)	AB252141	
Fukushima Prefecture (n=7)	FKD001J	<i>R. sativus</i>	Sukagawa	2000	BR	37.2869	140.3728	Ohshima et al. (2007)	AB252109
	FKD004J	<i>R. sativus</i>	Funehiki	2000	BR	37.4342	140.5705	Ohshima et al. (2007)	AB252110
	FKD005J	<i>R. sativus</i>	Soma	2000	BR	37.7970	140.9191	Kawakubo et al. (2022)	LC639403
	FKD006J	<i>R. sativus</i>	Ryozen	2000	BR	37.7691	140.4747	Kawakubo et al. (2022)	LC639404
	FKD008J	<i>R. sativus</i>	Aizuwakamatsu	2000	BR	37.4900	139.9277	Kawakubo et al. (2022)	LC639405
	FKD120J	<i>R. sativus</i>	Koriyama	1998	BR	37.4005	140.3597	Kawakubo et al. (2022)	LC639406
	FKH122J	<i>B. pekinensis</i>	Naraha	1998	BR	37.2827	140.9935	Ohshima et al. (2007)	AB252111
Kanto district (n=27)									
Ibaraki Prefecture (n=4)	IBD222J	<i>R. sativus</i>	Tsukuba	2011	BR	36.1164	140.0714	Kawakubo et al. (2022)	LC639454
	IBD223J	<i>R. sativus</i>	Tsukuba	2011	BR	36.1164	140.0714	Kawakubo et al. (2022)	LC639455
	IBD224J	<i>R. sativus</i>	Tsukuba	2011	BR	36.1164	140.0714	Kawakubo et al. (2022)	LC639456
	IBD226J	<i>R. sativus</i>	Tsukuba	2011	BR	36.0596	140.0984	Kawakubo et al. (2022)	LC639457
Tochigi Prefecture (n=7)	2J	<i>B. pekinensis</i>	-	1994	BR	36.5551	139.8826	Tomimura et al. (2003)	AB093622
	DMJ	<i>R. sativus</i>	-	1996	BR	36.5551	139.8826	Tomimura et al. (2003)	AB093623
	Ka1J	<i>B. pekinensis</i>	-	1994	BR	36.5551	139.8826	Tomimura et al. (2003)	AB093624
	N217	<i>B. rapa</i>	Aso	1963	B	36.3661	139.5790	Kawakubo et al. (2022)	LC639535
	TCB794J	<i>Brassica</i> sp.	Sano	2012	B	36.3292	139.5764	Kawakubo et al. (2022)	LC639642
	Tu-2R1	<i>R. sativus</i>	-	Not known	BR	36.5551	139.8826	Suehiro et al. (2004)	AB105135
	Tu-3	<i>B. oleracea</i>	-	Not known	B	36.5551	139.8826	Suehiro et al. (2004)	AB105134
Gunma Prefecture (n=3)	GMB790J	<i>Brassica</i> sp.	Ota	2012	B	36.2456	139.3133	Kawakubo et al. (2022)	LC639410
	GMB792J	<i>Brassica</i> sp.	Ota	2012	B	36.2725	139.3234	Kawakubo et al. (2022)	LC639411
	GMD791J	<i>R. sativus</i>	Ota	2012	BR	36.2456	139.3133	Kawakubo et al. (2022)	LC639412
Saitama Prefecture (n=2)	STD795J	<i>R. sativus</i>	Kumagaya	2012	BR	36.1726	139.3374	Kawakubo et al. (2022)	LC639639
	STD860J	<i>R. sativus</i>	Kawagoe	2010	BR	35.8906	139.4434	Kawakubo et al. (2022)	LC639640
Chiba Prefecture (n=4)	CD848J	<i>R. sativus</i>	Kimitsu	2010	BR	35.3168	139.9226	Kawakubo et al. (2022)	LC639379
	CI822J	<i>Papaver nudicaule</i>	Tateyama	1999	BR	34.9965	139.8700	Kawakubo et al. (2022)	LC639380
	CP845J	<i>Calendula officinalis</i>	Shirahama	1997	BR	34.9137	139.9008	Tomimura et al. (2003)	AB093614
	CS820J	<i>Matthiola incana</i>	Tateyama	1999	BR	34.9965	139.8700	Kawakubo et al. (2022)	LC639381
Tokyo Prefecture (n=1)	TD88J	<i>R. sativus</i>	Hino	1998	BR	35.6714	139.3951	Tomimura et al. (2003)	AB093615
Kanagawa Prefecture (n=6)	KND853J	<i>R. sativus</i>	Miura	2010	BR	35.1442	139.6208	Kawakubo et al. (2022)	LC639499
	KND854J	<i>R. sativus</i>	Fujisawa	2010	BR	35.3587	139.4491	Kawakubo et al. (2022)	LC639500
	KND855J	<i>R. sativus</i>	Atsugi	2010	BR	35.4981	139.3628	Kawakubo et al. (2022)	LC639501
	KND856J	<i>R. sativus</i>	Atsugi	2010	BR	35.5186	139.3399	Kawakubo et al. (2022)	LC639502
	KND857J	<i>R. sativus</i>	Atsugi	2010	BR	35.5186	139.3399	Kawakubo et al. (2022)	LC639503
	KND865J	<i>R. sativus</i>	-	2010	BR	35.3092	139.1226	Kawakubo et al. (2022)	LC639504
Chubu district (n=54)									
Niigata Prefecture (n=9)	NID046J	<i>R. sativus</i>	Kajikawa	2000	BR	37.1099	138.2446	Kawakubo et al. (2022)	LC639552
	NID048J	<i>R. sativus</i>	Niitsu	2000	BR	37.8025	139.1106	Ohshima et al. (2007)	AB252131
	NID049J	<i>R. sativus</i>	Niigata	2000	BR	37.9161	139.0364	Kawakubo et al. (2022)	LC639553
	NID078J	<i>R. sativus</i>	Joetsu	2000	BR	37.1479	138.2361	Kawakubo et al. (2022)	LC639554
	NID117J	<i>R. sativus</i>	Shiozawa	1998	BR	37.0400	138.8503	Kawakubo et al. (2022)	LC639555
	NID119J	<i>R. sativus</i>	Yuzawa	1998	BR	36.9341	138.8173	Ohshima et al. (2007)	AB252132
	NID569J	<i>R. sativus</i>	Myouko	2008	B	37.0252	138.2536	Kawakubo et al. (2022)	LC639556
	NID922J	<i>R. sativus</i>	Mitsuke	2000	BR	37.5325	138.9141	Kawakubo et al. (2022)	LC639557
	NID924J	<i>R. sativus</i>	Mitsuke	2000	BR	37.5325	138.9141	Kawakubo et al. (2022)	LC639558
Toyama Prefecture (n=3)	TYD083J	<i>R. sativus</i>	Kosugi	2000	BR	36.6493	137.2332	Kawakubo et al. (2022)	LC639653
	TYD085J	<i>R. sativus</i>	Uozu	2005	BR	36.8274	137.4091	Kawakubo et al. (2022)	LC639654
	TYD300J	<i>R. sativus</i>	Takaoka	2005	BR	36.7541	137.0258	Kawakubo et al. (2022)	LC639655
Ishikawa Prefecture (n=6)	ISD089J	<i>R. sativus</i>	Kaga	2000	BR	36.3027	136.3149	Kawakubo et al. (2022)	LC639458
	ISD091J	<i>R. sativus</i>	Kanazawa	2000	BR	36.5610	136.6566	Kawakubo et al. (2022)	LC639459
	ISD092J	<i>R. sativus</i>	Uchinada	2000	BR	36.6535	136.6451	Kawakubo et al. (2022)	LC639460
	ISD093J	<i>R. sativus</i>	Hakui	2000	BR	36.5993	136.6466	Kawakubo et al. (2022)	LC639461
	ISD420J	<i>R. sativus</i>	Komatsu	2005	BR	36.4084	136.4459	Kawakubo et al. (2022)	LC639462
	ISD440J	<i>R. sativus</i>	Negami	2005	BR	36.4366	136.4442	Kawakubo et al. (2022)	LC639463

Table 4. Continued.

Island, district and prefecture	Isolate name	Original host	Location (City or Town)	Collecting year	Host-infecting type ^a	Latitude	Longitude	Reference	Accession number
Fukui Prefecture (n=4)	FID096J	<i>R. sativus</i>	Kanazu	2000	BR	36.2081	136.2286	Kawakubo et al. (2022)	LC639399
	FID96J	<i>R. sativus</i>	Tsuruga	2005	BR	35.6459	136.0557	Kawakubo et al. (2022)	LC639402
	FID430J	<i>R. sativus</i>	Awara	2005	BR	36.0761	136.2125	Kawakubo et al. (2022)	LC639400
	FID565J	<i>R. sativus</i>	Oono	2008	BR	35.9798	136.4875	Kawakubo et al. (2022)	LC639401
Yamanashi Prefecture (n=2)	YSD442J	<i>R. sativus</i>	Fujiyoshida	2001	BR	35.4880	138.8072	Kawakubo et al. (2022)	LC639671
	YST441J	<i>Brassica juncea</i>	Fujiyoshida	2001	BR	35.4880	138.8072	Kawakubo et al. (2022)	LC639672
Nagano Prefecture (n=16)	N216	<i>B. oleracea</i>	-	1961	B	36.6485	138.1950	Kawakubo et al. (2022)	LC639534
	N224	<i>B. rapa</i>	Matsumoto	1966	B	36.2380	137.9720	Kawakubo et al. (2022)	LC639540
	N232	<i>R. sativus</i>	Nobeyama	1966	B(R)	35.9482	138.4637	Kawakubo et al. (2022)	LC639542
	NND453J	<i>R. sativus</i>	Suwa	2005	BR	36.0392	138.1141	Kawakubo et al. (2022)	LC639559
	NND456J	<i>R. sativus</i>	Shinano	1998	BR	36.8064	138.2072	Kawakubo et al. (2022)	LC639560
	NND457J	<i>R. sativus</i>	Shiojiri	2005	BR	36.1148	137.9532	Kawakubo et al. (2022)	LC639561
	NND459J	<i>R. sativus</i>	Iida	2005	BR	35.5149	137.8217	Kawakubo et al. (2022)	LC639562
	NND574J	<i>R. sativus</i>	Sakaki	2008	BR	36.4617	138.1803	Kawakubo et al. (2022)	LC639563
	NND575J	<i>R. sativus</i>	Sakaki	2008	BR	36.4617	138.1803	Kawakubo et al. (2022)	LC639564
	NND900J	<i>R. sativus</i>	Matsumoto	2005	BR	36.2380	137.9720	Kawakubo et al. (2022)	LC639565
	NND901J	<i>R. sativus</i>	Matsumoto	2000	BR	36.2380	137.9720	Kawakubo et al. (2022)	LC639566
	NND902J	<i>R. sativus</i>	Minowa	2005	BR	36.1877	138.2440	Kawakubo et al. (2022)	LC639567
	NND904J	<i>R. sativus</i>	Shiojiri	2000	BR	36.1148	137.9532	Kawakubo et al. (2022)	LC639568
	NND905J	<i>R. sativus</i>	Suwa	2000	BR	36.0392	138.1141	Kawakubo et al. (2022)	LC639569
	NND908J	<i>R. sativus</i>	Ina	2000	BR	35.8275	137.9541	Kawakubo et al. (2022)	LC639570
NNT451J	<i>B. juncea</i>	Iida	2005	BR	35.5149	137.8217	Kawakubo et al. (2022)	LC639571	
Gifu Prefecture (n=4)	GFD097J	<i>R. sativus</i>	Toki	2000	BR	35.3526	137.1834	Kawakubo et al. (2022)	LC639407
	GFD098J	<i>R. sativus</i>	Hida	2005	BR	36.3562	137.2885	Kawakubo et al. (2022)	LC639408
	GFD460J	<i>R. sativus</i>	Ena	2005	BR	35.4493	137.4128	Kawakubo et al. (2022)	LC639409
	GFD462J	<i>R. sativus</i>	Yoro	2001	BR	35.3084	136.5615	Ohshima et al. (2007)	AB252115
Shizuoka Prefecture (n=8)	SHD467J	<i>R. sativus</i>	Iwata	2001	BR	34.7179	137.8515	Kawakubo et al. (2022)	LC639619
	SHD468J	<i>R. sativus</i>	Iwata	2001	BR	34.7179	137.8515	Kawakubo et al. (2022)	LC639620
	SHD470J	<i>R. sativus</i>	Gotenba	2001	BR	35.3088	138.9348	Kawakubo et al. (2022)	LC639621
	SHD473J	<i>R. sativus</i>	Susono	2001	BR	35.1740	138.9066	Kawakubo et al. (2022)	LC639622
	SHD477J	<i>R. sativus</i>	Fuji	2001	BR	35.1615	138.6762	Kawakubo et al. (2022)	LC639623
	SHD585J	<i>R. sativus</i>	Shimada	2008	BR	34.8362	138.1760	Kawakubo et al. (2022)	LC639624
	SHH472J	<i>B. pekinensis</i>	Susono	2001	BR	35.1740	138.9066	Kawakubo et al. (2022)	LC639625
	SHT479J	<i>E. grandiflorum</i>	Shimizu	2002	BR	35.1132	138.4987	Kawakubo et al. (2022)	LC639626
Aichi Prefecture (n=2)	ACD099J	<i>R. sativus</i>	Iwakura	2000	B(R)	35.2795	136.8713	Kawakubo et al. (2022)	LC639359
	ACD483J	<i>R. sativus</i>	Toyota	2001	BR	35.0826	137.1561	Kawakubo et al. (2022)	LC639360
Kinki district (n=23)									
Mie Prefecture (n=3)	MED301J	<i>R. sativus</i>	Nabari	2001	BR	34.6275	136.1085	Kawakubo et al. (2022)	LC639527
	MED302J	<i>R. sativus</i>	Shiroyama	2001	BR	34.6763	136.5054	Ohshima et al. (2007)	AB252127
	MED303J	<i>R. sativus</i>	Tsu	2001	BR	34.7186	136.5052	Kawakubo et al. (2022)	LC639528
Shiga Prefecture (n=3)	SGB088J	<i>B. rapa</i>	Hikone	2000	BR	35.2744	136.2597	Ohshima et al. (2007)	AB252136
	SGD071J	<i>R. sativus</i>	Hikone	2000	BR	35.2744	136.2597	Kawakubo et al. (2022)	LC639618
	SGD311J	<i>R. sativus</i>	Nishiazai	1998	BR	35.4825	136.1111	Tomimura et al. (2003)	AB093619
Kyoto Prefecture (n=3)	KYD073J	<i>R. sativus</i>	Mineyama	2000	BR	35.1973	135.6924	Ohshima et al. (2007)	AB252126
	KYD074J	<i>R. sativus</i>	Miyazu	2000	BR	35.5356	135.1956	Kawakubo et al. (2022)	LC639517
	KYD81J	<i>R. sativus</i>	Joyo	1998	BR	34.8530	135.7801	Tomimura et al. (2003)	AB093613
Osaka Prefecture (n=1)	OSD189J	<i>R. sativus</i>	Osaka	1999	BR	34.6937	135.5023	Kawakubo et al. (2022)	LC639611
Hyogo Prefecture (n=7)	HGB356J	<i>Brassica sp.</i>	Itami	2012	BR	34.7674	135.4033	Kawakubo et al. (2022)	LC639414
	HGC346J	<i>Brassica oleracea</i> var. <i>capitata</i>	Taishi	2001	BR	34.6513	135.1276	Kawakubo et al. (2022)	LC639415
	HGD076J	<i>R. sativus</i>	Kasumi	2000	BR	34.7398	135.3246	Kawakubo et al. (2022)	LC639416
	HGD341J	<i>R. sativus</i>	Himeji	2005	BR	34.8154	134.6856	Kawakubo et al. (2022)	LC639417
	HGD347J	<i>R. sativus</i>	Kakogawa	2011	BR	34.6513	135.1276	Kawakubo et al. (2022)	LC639418
	HGD348J	<i>R. sativus</i>	Tatsuno	2001	BR	34.7343	135.3154	Kawakubo et al. (2022)	LC639419
HGD355J	<i>R. sativus</i>	Itami	2012	B(R)	34.7674	135.4033	Kawakubo et al. (2022)	LC639420	
Nara Prefecture (n=3)	NRD350J	<i>R. sativus</i>	Gojyo	2001	BR	34.3520	135.6935	Ohshima et al. (2007)	AB252134
	NRD351J	<i>R. sativus</i>	Gojyo	2001	BR	34.3520	135.6935	Kawakubo et al. (2022)	LC639572
	NRD353J	<i>R. sativus</i>	Kashiwara	2001	BR	34.4526	135.7628	Kawakubo et al. (2022)	LC639573
Wakayama Prefecture (n=3)	WB364J	<i>B. oleracea</i>	Katsuragi	2001	BR	34.2967	135.5037	Kawakubo et al. (2022)	LC639656
	WD361J	<i>R. sativus</i>	Iwade	2001	BR	34.2562	135.3113	Kawakubo et al. (2022)	LC639657
	WD365J	<i>R. sativus</i>	Hashimoto	2001	BR	34.3147	135.6054	Kawakubo et al. (2022)	LC639658
Chugoku district (n=28)									
Tottori Prefecture (n=9)	N219	<i>B. rapa</i>	Kurayoshi	1961	B	35.4304	133.8261	Kawakubo et al. (2022)	LC639537
	TRB203J	<i>B. rapa</i>	Iwami	2005	BR	35.5760	134.3321	Kawakubo et al. (2022)	LC639647
	TRD050J	<i>R. sativus</i>	Yonago	2000	BR	35.4281	133.3310	Kawakubo et al. (2022)	LC639648
	TRD051J	<i>R. sativus</i>	Daisen	2000	BR	35.5107	133.4963	Kawakubo et al. (2022)	LC639649
	TRD052J	<i>R. sativus</i>	Akasaki	2000	BR	35.5134	133.6399	Ohshima et al. (2007)	AB252138
	TRD053J	<i>R. sativus</i>	Tomari	2000	BR	35.5145	133.9413	Ohshima et al. (2007)	AB252139
	TRD054J	<i>R. sativus</i>	Tottori	2000	BR	35.5011	134.2351	Kawakubo et al. (2022)	LC639650
	TRD204J	<i>R. sativus</i>	Daisen	2005	BR	35.5107	133.4963	Kawakubo et al. (2022)	LC639651
	TRD811J	<i>R. sativus</i>	Sakaminato	2017	BR	35.5322	133.2171	Kawakubo et al. (2022)	LC639652

Table 4. Continued.

Island, district and prefecture	Isolate name	Original host	Location (City or Town)	Collecting year	Host-infecting type ^a	Latitude	Longitude	Reference	Accession number	
Shimane Prefecture (n=11)	SMB063J	<i>B. juncea</i>	Izumo	2000	BR	35.3670	132.7547	Kawakubo et al. (2022)	LC639627	
	SMD058J	<i>R. sativus</i>	Misumi	2000	BR	34.8947	132.0731	Kawakubo et al. (2022)	LC639628	
	SMD060J	<i>R. sativus</i>	Gotsu	2000	BR	35.0113	132.2208	Ohshima et al. (2007)	AB252137	
	SMD061J	<i>R. sativus</i>	Oda	2000	BR	35.1921	132.4993	Kawakubo et al. (2022)	LC639629	
	SMD064J	<i>R. sativus</i>	Izumo	2000	BR	35.3670	132.7547	Kawakubo et al. (2022)	LC639630	
	SMD065J	<i>R. sativus</i>	Hikawa	2000	BR	35.3782	132.8272	Kawakubo et al. (2022)	LC639631	
	SMD211J	<i>R. sativus</i>	Hirata	2005	BR	35.4411	132.8272	Kawakubo et al. (2022)	LC639632	
	SMD212J	<i>R. sativus</i>	Misumi	2005	BR	34.8947	132.0731	Kawakubo et al. (2022)	LC639633	
	SMD213J	<i>R. sativus</i>	Masuda	2005	BR	34.6726	131.8240	Kawakubo et al. (2022)	LC639634	
	SMD214J	<i>R. sativus</i>	Oda	2005	BR	35.1921	132.4993	Kawakubo et al. (2022)	LC639635	
	SMD310J	<i>R. sativus</i>	Hikawa	2005	BR	35.3782	132.8272	Kawakubo et al. (2022)	LC639636	
Okayama Prefecture (n=2)	OYD066J	<i>R. sativus</i>	Kume	2000	BR	34.6463	133.8698	Kawakubo et al. (2022)	LC639612	
	OYD221J	<i>R. sativus</i>	Mimazaka	1998	BR	35.0086	134.1486	Kawakubo et al. (2022)	LC639613	
Hiroshima Prefecture (n=2)	HD75J	<i>R. sativus</i>	Nishiku	1998	BR	34.6463	133.8698	Kawakubo et al. (2022)	LC639413	
	HRD805J	<i>R. sativus</i>	Kitahiroshima	2017	BR	35.0086	134.1486	Kawakubo et al. (2022)	LC639453	
Yamaguchi Prefecture (n=4)	YMD068J	<i>R. sativus</i>	Mine	2000	BR	34.1665	131.2061	Kawakubo et al. (2022)	LC639669	
	YMD069J	<i>R. sativus</i>	Misumi	2000	BR	34.1444	132.2313	Ohshima et al. (2007)	AB252142	
	YMD070J	<i>R. sativus</i>	Abu	2000	BR	34.5033	131.4714	Ohshima et al. (2007)	AB252143	
	YMD240J	<i>R. sativus</i>	Mine	2005	BR	34.1665	131.2061	Kawakubo et al. (2022)	LC639670	
Shikoku Island/district (n=29)										
Tokushima Prefecture (n=4)	TKD761J	<i>R. sativus</i>	Miyoshi	2002	BR	34.0260	133.8072	Kawakubo et al. (2022)	LC639643	
	TKD764J	<i>R. sativus</i>	Donari	2002	BR	34.1104	134.3378	Kawakubo et al. (2022)	LC639644	
	TKD766J	<i>R. sativus</i>	Tokushima	2002	BR	34.0704	134.5550	Kawakubo et al. (2022)	LC639645	
	TKD767J	<i>R. sativus</i>	Tokushima	2002	BR	34.0704	134.5550	Kawakubo et al. (2022)	LC639646	
Kagawa Prefecture (n=9)	KWB778J	<i>B. oleracea</i>	Takamatsu	2004	B	34.3425	134.0465	Ohshima et al. (2007)	AB252124	
	KWB779J	<i>B. rapa</i>	Takamatsu	2004	BR	34.3425	134.0465	Ohshima et al. (2007)	AB252125	
	KWB780J	<i>B. oleracea</i>	Takamatsu	2004	B	34.3425	134.0465	Kawakubo et al. (2022)	LC639510	
	KWD775J	<i>R. sativus</i>	Takamatsu	2002	BR	34.3425	134.0465	Kawakubo et al. (2022)	LC639511	
	KWD776J	<i>R. sativus</i>	Ryonan	2002	BR	34.2474	133.9455	Kawakubo et al. (2022)	LC639512	
	KWD777J	<i>R. sativus</i>	Ayauta	2002	BR	34.2303	133.9627	Kawakubo et al. (2022)	LC639513	
	KWD781J	<i>R. sativus</i>	Sakaide	2011	BR	34.3215	133.8745	Kawakubo et al. (2022)	LC639514	
	KWD785J	<i>R. sativus</i>	Takamatsu	2011	BR	34.3425	134.0465	Kawakubo et al. (2022)	LC639515	
	KWD872J	<i>R. sativus</i>	Sakaide	2011	BR	34.3215	133.8745	Kawakubo et al. (2022)	LC639516	
Ehime Prefecture (n=11)	EHD709J	<i>R. sativus</i>	Honai	2002	BR	33.4999	132.3889	Kawakubo et al. (2022)	LC639382	
	EHD717J	<i>R. sativus</i>	Honai	2002	BR	33.4999	132.3889	Kawakubo et al. (2022)	LC639383	
	EHD719J	<i>R. sativus</i>	Yawatahama	2002	BR	33.4629	132.4234	Kawakubo et al. (2022)	LC639384	
	EHD720J	<i>R. sativus</i>	Uwajima	2002	BR	33.2236	132.5604	Kawakubo et al. (2022)	LC639385	
	EHD721J	<i>R. sativus</i>	Uwajima	2002	BR	33.2236	132.5604	Kawakubo et al. (2022)	LC639386	
	EHD723J	<i>R. sativus</i>	Uwajima	2002	BR	33.2236	132.5604	Kawakubo et al. (2022)	LC639387	
	EHD727J	<i>R. sativus</i>	Misho	2002	BR	33.6025	132.7858	Kawakubo et al. (2022)	LC639388	
	EHD729J	<i>R. sativus</i>	Ipponmatsu	2002	BR	32.9610	132.6508	Kawakubo et al. (2022)	LC639389	
	EHD731J	<i>R. sativus</i>	Saijo	2002	BR	33.9195	133.1811	Kawakubo et al. (2022)	LC639390	
	EHD733J	<i>R. sativus</i>	Tanbara	2002	BR	33.9087	133.0669	Kawakubo et al. (2022)	LC639391	
	EHD739J	<i>R. sativus</i>	Honai	2002	BR	33.4999	132.3889	Kawakubo et al. (2022)	LC639392	
Kochi Prefecture (n=5)	KOD740J	<i>R. sativus</i>	Sukumo	2002	BR	32.9390	132.7262	Kawakubo et al. (2022)	LC639505	
	KOD742J	<i>R. sativus</i>	Nakamura	2002	BR	32.9846	132.9440	Kawakubo et al. (2022)	LC639506	
	KOD743J	<i>R. sativus</i>	Ohkata	2002	BR	33.0271	133.0084	Kawakubo et al. (2022)	LC639507	
	KOD752J	<i>R. sativus</i>	Tosa	2002	BR	33.4961	133.4253	Kawakubo et al. (2022)	LC639508	
	KOD753J	<i>R. sativus</i>	Tosa	2002	BR	33.4961	133.4253	Kawakubo et al. (2022)	LC639509	
Kyushu Island/district (n=83)										
Fukuoka Prefecture (n=11)	FB998J	<i>B. pekinensis</i>	Yame	2004	B(R)	33.2120	130.5579	Kawakubo et al. (2022)	LC639393	
	FB999J	<i>B. pekinensis</i>	Yame	2004	B(R)	33.2120	130.5579	Kawakubo et al. (2022)	LC639394	
	FD21J	<i>R. sativus</i>	Shime	1998	BR	33.5914	130.4790	Kawakubo et al. (2022)	LC639395	
	FD27J	<i>R. sativus</i>	Fukuoka	1998	BR	33.5902	130.4017	Tomimura et al. (2003)	AB093618	
	FD810J	<i>R. sativus</i>	Kitakyushu	2000	BR	33.8835	130.8752	Kawakubo et al. (2022)	LC639396	
	FD891J	<i>R. sativus</i>	Ukiha	2007	BR	33.3474	130.7550	Kawakubo et al. (2022)	LC639397	
	FD997J	<i>R. sativus</i>	Ukiha	2001	BR	33.3474	130.7550	Kawakubo et al. (2022)	LC639398	
	N218	<i>B. rapa</i>	Chikugo	1963	BR	33.2124	130.5022	Kawakubo et al. (2022)	LC639536	
	N220	<i>B. rapa</i>	Chikugo	1974	B	33.2124	130.5022	Kawakubo et al. (2022)	LC639538	
	N223	<i>B. rapa</i>	Chikugo	1965	B	33.2124	130.5022	Kawakubo et al. (2022)	LC639539	
	N227	<i>B. rapa</i>	Chikugo	1960	B	33.2124	130.5022	Kawakubo et al. (2022)	LC639541	
	Saga Prefecture (n=10)	1J	<i>R. sativus</i>	Saga	1977	BR	33.2631	130.3009	Ohshima, Tanaka, and Sako (1996)	D83184
		59J	<i>R. sativus</i>	Saga	1996	BR	33.2631	130.3009	Tomimura et al. (2003)	AB093620
		C42J	<i>B. rapa</i>	Saga	1993	B	33.2631	130.3009	Tomimura et al. (2003)	AB093625
		SB20J	<i>Brassica</i> sp.	Yamato	2004	BR	33.3015	130.2809	Kawakubo et al. (2022)	LC639614
SD22J		<i>R. sativus</i>	Karatsu	2004	BR	33.4501	129.9683	Kawakubo et al. (2022)	LC639615	
SE251J		<i>Gypsophila elegans</i>	Saga	2001	BR	33.2631	130.3009	Kawakubo et al. (2022)	LC639616	
SE252J		<i>G. elegans</i>	Saga	2001	BR	33.2631	130.3009	Kawakubo et al. (2022)	LC639617	
SN4J		<i>Brassica</i> sp.	Saga	1998	BR	33.2631	130.3009	Kawakubo et al. (2022)	LC639637	
ST19J		<i>B. juncea</i>	Nabeshima	1998	B	33.2819	130.2667	Kawakubo et al. (2022)	LC639638	
STD8J		<i>R. sativus</i>	Tosu	1997	BR	33.3779	130.5062	Kawakubo et al. (2022)	LC639641	

Table 4. Continued.

Island, district and prefecture	Isolate name	Original host	Location (City or Town)	Collecting year	Host-infecting type ^a	Latitude	Longitude	Reference	Accession number
Nagasaki Prefecture (n=12)	NDJ	<i>R. sativus</i>	Takaki	1997	BR	32.9051	130.1357	Tomimura et al. (2003)	AB093616
	ND10J	<i>R. sativus</i>	Hirado	1998	BR	33.3680	129.5537	Ohshima et al. (2007)	AB252130
	ND256J	<i>R. sativus</i>	Kinkai	2001	BR	32.8926	129.7704	Kawakubo et al. (2022)	LC639543
	ND257J	<i>R. sativus</i>	Takaki	2001	BR	32.9051	130.1357	Kawakubo et al. (2022)	LC639544
	ND258J	<i>R. sativus</i>	Higashisonogi	2002	BR	33.0371	129.9172	Kawakubo et al. (2022)	LC639545
	ND260J	<i>R. sativus</i>	Takaki	2001	BR	32.9051	130.1357	Kawakubo et al. (2022)	LC639546
	ND264J	<i>R. sativus</i>	Kinkai	2001	BR	32.8926	129.7704	Kawakubo et al. (2022)	LC639547
	ND271J	<i>R. sativus</i>	Tsushima	2009	BR	34.2028	129.2875	Kawakubo et al. (2022)	LC639548
	ND276J	<i>R. sativus</i>	Isahaya	2010	BR	32.8434	130.0531	Kawakubo et al. (2022)	LC639549
	ND277J	<i>Raphanus sativus</i> var. <i>longipinnatus</i>	Nagasaki	2010	BR	32.7503	129.8779	Kawakubo et al. (2022)	LC639550
	ND278J	<i>R. sativus</i> var. <i>longipinnatus</i>	Nagasaki	2010	BR	32.7503	129.8779	Kawakubo et al. (2022)	LC639551
	NTD253J	<i>R. sativus</i>	Tsushima	2001	BR	34.2696	129.3151	Kawakubo et al. (2022)	LC639574
Kumamoto Prefecture (n=7)	KD32J	<i>R. sativus</i>	Nankan	1997	BR	33.0616	130.5412	Tomimura et al. (2003)	AB093621
	KD41J	<i>R. sativus</i>	Taura	1998	BR	32.3645	130.5000	Kawakubo et al. (2022)	LC639472
	KD302J	<i>R. sativus</i>	Kikuchi	2006	BR	32.9789	130.8130	Kawakubo et al. (2022)	LC639471
	KD834J	<i>R. sativus</i>	Ichinomiya	2000	BR	32.9477	131.1163	Kawakubo et al. (2022)	LC639473
	KD941J	<i>R. sativus</i>	Ueki	2001	BR	32.9046	130.6924	Kawakubo et al. (2022)	LC639474
	KD945J	<i>R. sativus</i>	Gyokuto	2001	BR	32.9183	130.6288	Kawakubo et al. (2022)	LC639475
	KD946J	<i>R. sativus</i>	Sakamoto	2003	BR	32.4251	130.6727	Kawakubo et al. (2022)	LC639476
Oita Prefecture (n=9)	OB284J	<i>Brassica</i> sp.	Hiji	2007	BR	33.3692	131.5326	Kawakubo et al. (2022)	LC639575
	OB290J	<i>Brassica</i> sp.	Yufuin	2007	BR	33.2841	131.3511	Kawakubo et al. (2022)	LC639576
	OB995J	<i>Brassica</i> sp.	Hita	2001	BR	33.3206	130.9404	Kawakubo et al. (2022)	LC639577
	OB996AJ	<i>Brassica</i> sp.	Yamaga	2001	BR	33.4580	131.5060	Kawakubo et al. (2022)	LC639578
	OB996BJ	<i>Brassica</i> sp.	Yamaga	2001	BR	33.4580	131.5060	Kawakubo et al. (2022)	LC639579
	OD14J	<i>R. sativus</i>	Amagase	1998	BR	33.2548	131.0238	Kawakubo et al. (2022)	LC639580
	OD15J	<i>R. sativus</i>	Hita	1998	BR	33.3206	130.9404	Kawakubo et al. (2022)	LC639581
	OD985J	<i>R. sativus</i>	Tsukumi	2001	BR	33.0720	131.8611	Kawakubo et al. (2022)	LC639582
	ON16J	<i>B. napus</i>	Yufuin	1998	BR	33.2841	131.3511	Kawakubo et al. (2022)	LC639610
Miyazaki Prefecture (n=10)	MB649J	<i>B. rapa</i>	Kawaminami	2007	BR	32.1920	131.5259	Kawakubo et al. (2022)	LC639518
	MC982J	<i>B. oleracea</i>	Kitaura	2001	BR	32.6846	131.7994	Kawakubo et al. (2022)	LC639519
	MD45J	<i>R. sativus</i>	Miyazaki	1998	BR	31.9077	131.4202	Kawakubo et al. (2022)	LC639520
	MD49J	<i>R. sativus</i>	Miyakonojo	1998	BR	31.7196	131.0616	Kawakubo et al. (2022)	LC639521
	MD641J	<i>R. sativus</i>	Miyazaki	2007	BR	31.9077	131.4202	Kawakubo et al. (2022)	LC639522
	MD644J	<i>R. sativus</i>	Shintomi	2007	BR	32.0690	131.4880	Kawakubo et al. (2022)	LC639523
	MD647J	<i>R. sativus</i>	Kawaminami	2007	BR	32.1920	131.5259	Kawakubo et al. (2022)	LC639524
	MD652J	<i>R. sativus</i>	Nobeoka	2007	BR	32.5823	131.6651	Kawakubo et al. (2022)	LC639525
	MD966J	<i>R. sativus</i>	Tano	2001	BR	31.8383	131.3043	Kawakubo et al. (2022)	LC639526
	MN978J	<i>Brassica</i> sp.	Tsuno	2001	BR	31.8965	131.4068	Kawakubo et al. (2022)	LC639529
Kagoshima Prefecture (n=24)	KGB383J	<i>Brassica</i> sp.	Makurazaki	2007	BR	31.2729	130.2967	Kawakubo et al. (2022)	LC639477
	KGB394J	<i>Brassica oleracea</i> var. <i>botrytis</i>	Kanoya	2007	BR	31.3782	130.8523	Kawakubo et al. (2022)	LC639478
	KGB397J	<i>Brassica</i> sp.	Ohsaki	2007	BR	31.3824	130.9562	Kawakubo et al. (2022)	LC639479
	KGB662J	<i>Brassica chinensis</i>	Ei	2003	BR	31.2382	130.5178	Kawakubo et al. (2022)	LC639480
	KGB964J	<i>B. rapa</i>	Okinoerabu	2002	BR	27.3662	128.6005	Kawakubo et al. (2022)	LC639481
	KGD54J	<i>R. sativus</i>	Satsumasendai	1998	BR	31.8133	130.3035	Ohshima et al. (2007)	AB252123
	KGD55J	<i>R. sativus</i>	Fukiage	1998	BR	31.5126	130.3368	Kawakubo et al. (2022)	LC639483
	KGD58J	<i>R. sativus</i>	Kanoya	1998	BR	31.3782	130.8523	Kawakubo et al. (2022)	LC639484
	KGD373J	<i>R. sativus</i>	Sastumasendai	2007	BR	31.8133	130.3035	Kawakubo et al. (2022)	LC639482
	KGD661J	<i>R. sativus</i>	Ei	2003	BR	31.2382	130.5178	Kawakubo et al. (2022)	LC639485
	KGD677J	<i>R. sativus</i>	Kasari	2003	BR	28.4904	129.6905	Kawakubo et al. (2022)	LC639486
	KGD678J	<i>R. sativus</i>	Kasari	2003	BR	28.4904	129.6905	Kawakubo et al. (2022)	LC639487
	KGD680J	<i>R. sativus</i>	Tanegashima	2004	BR	30.7323	130.9970	Kawakubo et al. (2022)	LC639488
	KGD682J	<i>Brassica</i> sp.	Tanegashima	2004	BR	30.7323	130.9970	Kawakubo et al. (2022)	LC639489
	KGD687J	<i>R. sativus</i>	Tanegashima	2004	BR	30.6979	131.0612	Kawakubo et al. (2022)	LC639490
	KGD690J	<i>R. sativus</i>	Tanegashima	2004	BR	30.5631	130.9621	Kawakubo et al. (2022)	LC639491
	KGD692J	<i>R. sativus</i>	Tanegashima	2004	BR	30.4140	130.9010	Kawakubo et al. (2022)	LC639492
	KGD701J	<i>R. sativus</i>	Yakushima	2004	BR	30.3259	130.6144	Kawakubo et al. (2022)	LC639493
	KGD851J	<i>R. sativus</i>	Kagoshima	2000	BR	31.5969	130.5571	Kawakubo et al. (2022)	LC639494
	KGD946J	<i>R. sativus</i>	Kagoshima	2001	BR	31.5833	130.6500	Kawakubo et al. (2022)	LC639495
	KGD947J	<i>R. sativus</i>	Kagoshima	2001	BR	31.5833	130.6500	Kawakubo et al. (2022)	LC639496
	KGD954J	<i>R. sativus</i>	Kagoshima	2001	BR	31.5833	130.6500	Kawakubo et al. (2022)	LC639497
	KGD963J	<i>R. sativus</i>	Okinoerabu	2002	BR	27.3662	128.6005	Kawakubo et al. (2022)	LC639498
	N214	<i>B. rapa</i>	Kagoshima	1961	B	31.5969	130.5571	Kawakubo et al. (2022)	LC639533

Table 4. Continued.

Island, district and prefecture	Isolate name	Original host	Location (City or Town)	Collecting year	Host-infecting type ^a	Latitude	Longitude	Reference	Accession number
Okinawa Island/district (n=30)									
Okinawa Prefecture (n=30)	OKB613J	<i>B. chinensis</i>	Gushikami	2000	BR	26.1047	127.7266	Kawakubo et al. (2022)	LC639674
	OKB616J	<i>B. juncea</i>	Nishihara	2000	BR	26.2441	127.7348	Kawakubo et al. (2022)	LC639583
	OKB617J	<i>B. juncea</i>	Nishihara	2000	BR	26.2506	127.7650	Kawakubo et al. (2022)	LC639584
	OKB634J	<i>Nasturtium officinale</i>	Chinen	2005	B	26.1699	127.8274	Kawakubo et al. (2022)	LC639585
	OKB636J	<i>B. juncea</i>	Ginowan	2005	BR	26.2815	127.7785	Kawakubo et al. (2022)	LC639586
	OKB638J	<i>B. juncea</i>	Naha	2006	BR	26.2137	127.7256	Kawakubo et al. (2022)	LC639587
	OKB639J	<i>B. juncea</i>	Naha	2006	BR	26.2137	127.7256	Kawakubo et al. (2022)	LC639588
	OKC198J	<i>B. rapa</i>	Nakagusuku	1999	BR	26.2672	127.7910	Kawakubo et al. (2022)	LC639589
	OKD192J	<i>R. sativus</i>	Nishihara	1999	BR	26.1953	127.7211	Kawakubo et al. (2022)	LC639590
	OKD195J	<i>R. sativus</i>	Nakagusuku	1999	BR	26.2718	127.7920	Kawakubo et al. (2022)	LC639675
	OKD197J	<i>R. sativus</i>	Nakagusuku	1999	BR	26.2718	127.7920	Kawakubo et al. (2022)	LC639591
	OKD604J	<i>R. sativus</i>	Nakijin	2000	BR	26.6825	127.9721	Kawakubo et al. (2022)	LC639592
	OKD608J	<i>R. sativus</i>	Higashi	2000	BR	26.6107	128.1107	Kawakubo et al. (2022)	LC639593
	OKD609J	<i>R. sativus</i>	Higashi	2000	BR	26.6574	128.1847	Kawakubo et al. (2022)	LC639594
	OKD610J	<i>R. sativus</i>	Higashi	2000	BR	26.6357	128.1504	Kawakubo et al. (2022)	LC639673
	OKD611J	<i>R. sativus</i>	Higashi	2000	BR	26.6357	128.1504	Kawakubo et al. (2022)	LC639595
	OKD612J	<i>R. sativus</i>	Nago	2000	BR	26.5755	128.0552	Kawakubo et al. (2022)	LC639596
	OKD621J	<i>R. sativus</i>	Kumejima	2002	BR	26.3504	126.7712	Kawakubo et al. (2022)	LC639597
	OKD622J	<i>R. sativus</i>	Kumejima	2002	BR	26.3504	126.7712	Kawakubo et al. (2022)	LC639598
	OKD624J	<i>R. sativus</i>	Kumejima	2002	BR	26.3504	126.7712	Kawakubo et al. (2022)	LC639599
	OKD626J	<i>R. sativus</i>	Kumejima	2002	BR	26.3504	126.7712	Kawakubo et al. (2022)	LC639600
	OKD627J	<i>R. sativus</i>	Kumejima	2002	BR	26.3504	126.7712	Kawakubo et al. (2022)	LC639601
	OKD630J	<i>R. sativus</i>	Kumejima	2002	BR	26.3504	126.7712	Kawakubo et al. (2022)	LC639602
	OKD633J	<i>R. sativus</i>	Ishigaki	2004	BR	24.6157	124.1565	Kawakubo et al. (2022)	LC639603
	OKD637J	<i>R. sativus</i>	Nakagusuku	2005	BR	26.2718	127.7920	Kawakubo et al. (2022)	LC639604
	OKD645J	<i>R. sativus</i>	Nakagusuku	2008	BR	26.2718	127.7920	Kawakubo et al. (2022)	LC639605
	OKM62J	<i>B. juncea</i>	Hateruma	1998	BR	24.0679	123.7803	Kawakubo et al. (2022)	LC639607
	OKM191J	<i>B. juncea</i>	Nakagusuku	1999	BR	26.2672	127.7910	Kawakubo et al. (2022)	LC639606
	OKM631J	<i>B. juncea</i>	Shuri	2003	BR	26.2176	127.7211	Kawakubo et al. (2022)	LC639608
	OKM632J	<i>B. juncea</i>	Shuri	2003	BR	26.2176	127.7211	Kawakubo et al. (2022)	LC639609
Not known	JPN1	<i>R. sativus</i>	-	1995	Not known	Not known	Not known	-	KM094174
	JPN1ARA	<i>R. sativus</i>	-	Not known	Not known	Not known	Not known	Vijayan et al. (2017)	KX377967

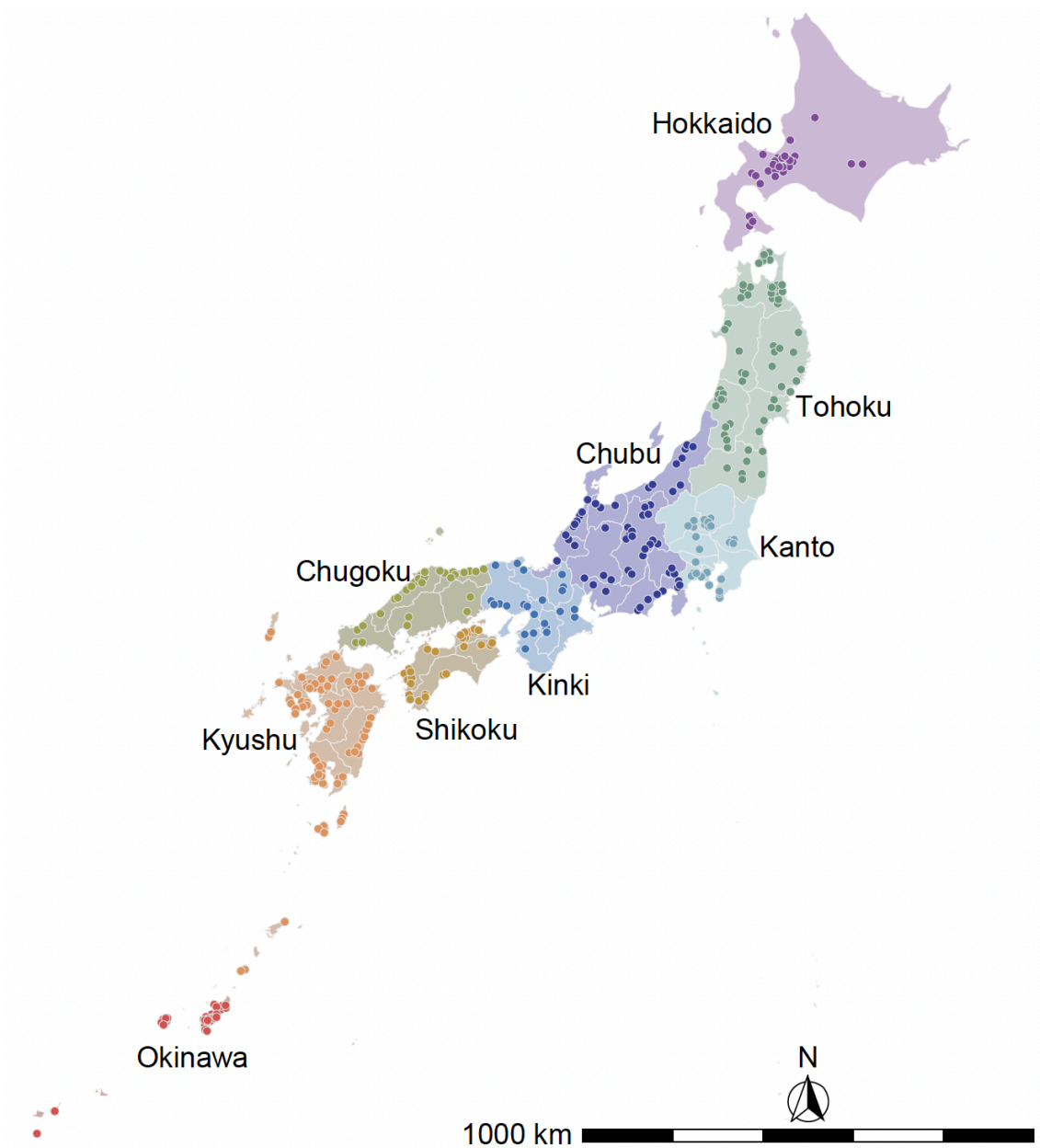


Fig. 21. Sampling map of TuMV isolates analysed in this study and classification of coded geographic regions in Japan. Each isolate was annotated by commonly acknowledged district division in Japan and used for phylogeographic analysis in discrete space.

3. 分子系統解析および分子時計解析

まず初めに、全分離株を用いて Neighbor-Net 法を実装した SplitsTree v4.11.3(Huson & Bryant, 2005)による系統ネットワーク解析、最尤法を実装した PhyML v3.0 (Guindon et al., 2010)による分子系統解析を行なった。その後、前項で記述した組換え解析の結果から、HC-Pro、P3、NIb、CP 遺伝子領域に組換え部位を持たない、非組換え配列を用いて PhyMLv3 による最尤系統樹を推定した。塩基置換モデルとして jModelTest2(Darriba et al., 2012)の結果から最適と判断された GTR+I+G モデルを用い、各ノードの分岐の支持度として 1000 回反復のブートストラップ法を用いた。推定された系統樹を FigTree (<http://tree.bio.edu.ac.uk/software/figtree/>) によって可視化した。また、推定された最尤系統樹と各分離株の採集年代を用いて TempEst v1.5.1 (Rambaut et al., 2016) による root-to-tip 線形回帰分析を行い、以降の時間解析において適切でない分離株、即ち、系統樹における枝長と年代に相関がないため、データセットにおける外れ値としてみなされる分離株がある場合それらを以降のデータセットから除いた。その後、BEAST v1.10.4(Suchard et al., 2018)を用いて分子時計を考慮したベイズ時間系統樹を構築し、TuMV の分岐年代および塩基置換速度を推定した。事後確率分布の最適解を確実に探索するために、独立した連鎖計算を最低 3 回解析した。それぞれの連鎖計算の長さに 3 万世代毎抽出の 3 億回を用いた。各連鎖計算の収束判定に Tracer v1.7.1(Rambaut et al., 2018)を用いた。複数の独立した連鎖計算が十分な定常状態に達していることを Tracer v1.7.1 で確認した適切な範囲の burn-in 状態を除いて LogCombiner v1.10.4 で結合し ESS \geq 200 であることを基準として判定した。Maximum clade credibility (MCC) 系統樹の作成には TreeAnnotator v1.10.4 を用いた。TempEst v1.5.1 による root-to-tip 線形回帰分析で示唆されたリネージ毎の変異率のばらつきを考慮するため、uncorrelated lognormal distribution relaxed clock (UCLD)(Drummond et al., 2006)を分子時計モデルとして用いた。集団動態モデルを、合祖理論に基づく constant size tree prior, exponential growth tree prior および ノンパラメトリックな Bayesian skyline plot tree prior の 3 つの中から Path sampling と Stepping-stone sampling 法(Baele et al.,

2012)により計算された周辺尤度をもとに選択した (Table 5)。また、分離株の採集年代を用いて分子時計を較正するにあたり、各データセットに十分な時間信号が含まれているのか検討するため、クラスターランダムマイゼーション検定 (Duchêne et al., 2015)および BETS 検定(Duchene et al., 2020)を行った(Fig. 22, Table 6)。検定において、R パッケージの TipDatingBeast (Rieux & Khatchikian, 2017)を用いた教師なし非階層的クラスタリング法により生成された擬似データセット 10 個を用いた。時間信号が十分でないと判断されたデータセットには、先行研究にて推定された TuMV の塩基置換速度(Kawakubo et al., 2021)を事前確率分布として用いた。

4. 時空間的拡散経路解析

系統地理学的手法を用いた時空間的拡散経路解析は、BEAST v1.10.4 を用いて、離散空間および連続空間の両方において行なった。離散空間においては、北海道、東北、関東、中部、近畿、中国、四国、九州、沖縄の計 9 地域を離散的状態とみなした。拡散経路の推定には、非対称モデルを用いたベイジアン確率的選択変数探索 (Bayesian Stochastic Search Variable Selection; BSSVS) (Lemey et al., 2009)を用いた。各 MCMC の計算には前項と同じ条件を適用した。Spread3 (Bielejec et al., 2016)を用いてベイズ因子 (BF) を計算した。BF が 3 以上かつ事後確率が 0.5 以上の拡散経路を支持されていると判断した。また、分離株の採集地の偏りが推定結果に与える影響を少なくするため、特定の地域で採集された分離株がデータセットに占める割合が 30%を超える場合、それらの分離株を無作為に抽出して、分離株がデータセットに占める割合を 30 %未満にした。データセットに特定の地域で採集された分離株が 1 つしかない場合、その分離株を解析から除いた。分離株の採集地毎の偏りが無視できる程度であることを確認するために、<tipStateSwapOperator> を追加し(Fig. 23)、推定結果の信頼性を確認した。各分離株の採集地を無作為に並べ替えた 10 反復のデータセットを用いて得られた起源地の事後確率分布を帰無分布とみなし、得られた帰無分布と元のデータセットから推定された祖先地の事後確率分布を比較し差を確認した。

Table 5. Model selection among clock and coalescent models.

Phylogenetic group or subgroup	Protein-coding region ^a	Clock and coalescent model ^b	Log marginal likelihood ^c	
			Path sampling	Stepping-stone sampling
Asian-BR	HC-Pro*	SC Constant Size	-7969.93	-7971.40
		SC Exponential Growth	-7927.83	-7929.90
		SC Bayesian Skyline	-7875.70	-7879.76
		UCLD Constant Size	-7978.18	-7980.05
		UCLD Exponential Growth	-7878.44	-7883.17
		UCLD Bayesian Skyline	-7888.92	-7892.14
	P3*	SC Constant Size	-7378.75	-7379.95
		SC Exponential Growth	-7326.35	-7328.16
		SC Bayesian Skyline	-7280.23	-7283.95
		UCLD Constant Size	-7382.73	-7383.88
		UCLD Exponential Growth	-7334.09	-7336.49
		UCLD Bayesian Skyline	-7285.61	-7288.86
	N1b*	SC Constant Size	-3663.69	-3663.75
		SC Exponential Growth	-3646.83	-3647.00
		SC Bayesian Skyline	-3604.80	-3605.10
		UCLD Constant Size	-3670.65	-3670.77
		UCLD Exponential Growth	-3653.47	-3653.71
		UCLD Bayesian Skyline	-3613.02	-3613.46
CP	SC Constant Size	-1894.72	-1894.71	
	SC Exponential Growth	-1895.05	-1895.14	
	SC Bayesian Skyline	-1872.82	-1872.84	
	UCLD Constant Size	-1900.85	-1900.85	
	UCLD Exponential Growth	-1900.94	-1900.95	
	UCLD Bayesian Skyline	-1882.20	-1882.42	
basal-BR	HC-Pro*	SC Constant Size	-3868.82	-3869.28
		SC Exponential Growth	-3867.67	-3868.25
		SC Bayesian Skyline	-3822.94	-3823.52
		UCLD Constant Size	-3868.83	-3869.63
		UCLD Exponential Growth	-3867.13	-3867.97
		UCLD Bayesian Skyline	-3825.93	-3826.33
	P3*	SC Constant Size	-4498.54	-4499.38
		SC Exponential Growth	-4497.64	-4498.51
		SC Bayesian Skyline	-4433.06	-4434.12
		UCLD Constant Size	-4501.74	-4502.78
		UCLD Exponential Growth	-4500.21	-4501.57
		UCLD Bayesian Skyline	-4436.43	-4437.56
	N1b*	SC Constant Size	-4427.43	-4428.56
		SC Exponential Growth	-4427.59	-4429.26
		SC Bayesian Skyline	-4365.34	-4365.66
		UCLD Constant Size	-4430.31	-4431.57
		UCLD Exponential Growth	-4375.87	-4377.92
		UCLD Bayesian Skyline	-4375.87	-4377.92
CP	SC Constant Size	-4158.59	-4161.48	
	SC Exponential Growth	-4153.78	-4155.83	
	SC Bayesian Skyline	-4083.40	-4083.58	
	UCLD Constant Size	-4160.21	-4161.08	
	UCLD Exponential Growth	-4156.80	-4158.74	
	UCLD Bayesian Skyline	-4094.28	-4096.91	
world-B3	HC-Pro*	SC Constant Size	-4911.96	-4912.01
		SC Exponential Growth	-4905.22	-4905.24
		SC Bayesian Skyline	-4881.61	-4881.99
		UCLD Constant Size	-4916.51	-4916.65
		UCLD Exponential Growth	-4909.94	-4909.98
		UCLD Bayesian Skyline	-4883.74	-4884.21
	P3*	SC Constant Size	-5909.81	-5909.97
		SC Exponential Growth	-5907.34	-5907.42
		SC Bayesian Skyline	-5883.52	-5883.68
		UCLD Constant Size	-5914.15	-5914.52
		UCLD Exponential Growth	-5911.67	-5911.69
		UCLD Bayesian Skyline	-5880.27	-5880.27
	N1b*	SC Constant Size	-6985.88	-6986.79
		SC Exponential Growth	-6909.75	-6911.76
		SC Bayesian Skyline	-6909.05	-6910.03
		UCLD Constant Size	-6990.94	-6992.14
		UCLD Exponential Growth	-6958.33	-6959.24
		UCLD Bayesian Skyline	-6909.20	-6911.14
CP	SC Constant Size	-5300.40	-5301.43	
	SC Exponential Growth	-5276.98	-5277.73	
	SC Bayesian Skyline	-5233.19	-5234.58	
	UCLD Constant Size	-5304.00	-5305.21	
	UCLD Exponential Growth	-5287.42	-5288.34	
	UCLD Bayesian Skyline	-5248.88	-5250.60	

^a HC-Pro*, P3* and N1b* indicate partial protein-coding regions.

^b SC: Strict clock, UCLD: Uncorrelated lognormal relaxed clock. Selected models are shown in bold.

^c Log marginal likelihood was calculated as described in Baele et al. (2012).

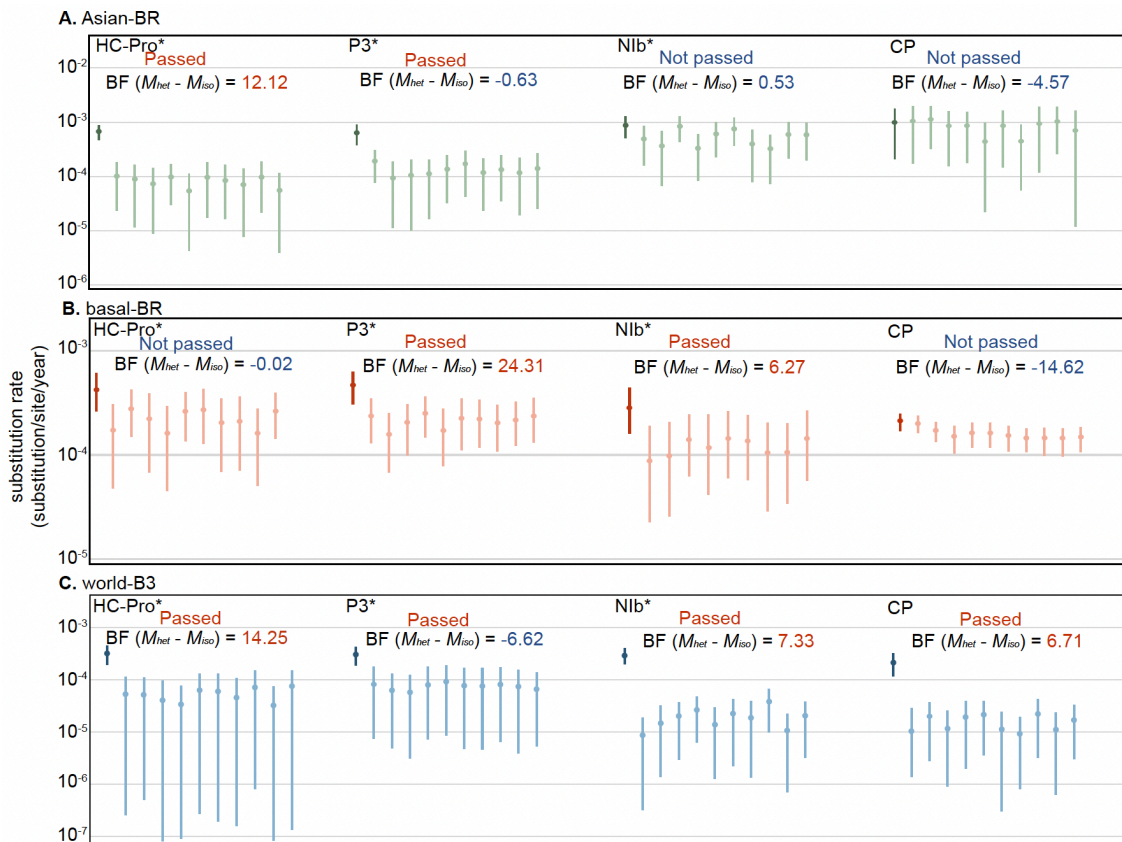


Fig. 22. Summary of the date-randomization tests and BETS of the data set used in this study. (A) Asian-BR, (B) basal-BR, and (C) world-B3. Estimates of nucleotides substitution rate (substitution/site/year) from the original data set and ten clustered permutations. Symbols represent the mean rate estimate for each data set, with error bars showing the 95% credible intervals (CI). Symbols and error bars in dark colour represent the estimate from the original data set, whereas those in light colour represent estimates from date-randomized data sets. For the data set that failed to support the presence of temporal signal by the date-randomization test, we specified an informative prior distribution for the substitution rate based on a previous estimate (Kawakubo et al., 2021). Bayes factor (BF) was calculated by log marginal likelihood of heterochronous model (M_{het}) minus isochronous model (M_{iso}), by following the Bayesian evaluation of temporal signal (Duchene et al., 2020). $BF > 3$ was considered as strong support of the presence of temporal signal indicated in red font, whereas the other in blue. Note that our Bayesian evaluation failed to support the presence of temporal signal for P3* data sets of Asian-BR and world-B3, even though they were supported by date-randomization tests.

Table 6. Time to the most recent common ancestor (MRCA) and substitution rate inferred from recombination-free protein-coding regions.

(Sub-)group	Protein-coding regions ^a	Results of date-randomization ^b	MRCA (CE) ^c	Substitution rate (substitution/site/year)
Asian-BR	HC-Pro* (n=153)	Passed	1968 (1954 – 1980) ^d	4.86×10^{-4} (3.25×10^{-4} – 6.44×10^{-4})
	P3* (n=152)	Passed	1951 (1928 – 1970)	3.93×10^{-4} (2.48×10^{-4} – 5.46×10^{-4})
	Nlb* (n=74)	Not passed	-	-
	CP (n=21)	Not passed	-	-
basal-BR	HC-Pro* (n=84)	Not passed	-	-
	P3* (n=132)	Passed	1548 (1198 – 1808)	1.41×10^{-4} (5.88×10^{-5} – 2.28×10^{-4})
	Nlb* (n=108)	Passed	1848 (1784 – 1903)	3.71×10^{-4} (2.52×10^{-4} – 4.98×10^{-4})
	CP (n=144)	Not passed	-	-
world-B3	HC-Pro* (n=55)	Passed	1817 (1711 – 1897)	2.51×10^{-4} (1.30×10^{-4} – 3.67×10^{-4})
	P3* (n=63)	Passed	1792 (1711 – 1858)	2.96×10^{-4} (1.97×10^{-4} – 3.92×10^{-4})
	Nlb* (n=140)	Passed	1860 (1817 – 1898)	2.65×10^{-4} (1.98×10^{-4} – 3.34×10^{-4})
	CP (n=138)	Passed	1812 (1749 – 1872)	1.53×10^{-4} (1.10×10^{-4} – 1.96×10^{-4})

^a HC-Pro*, P3* and Nlb* indicate partial protein-coding regions of HC-Pro, P3, and Nlb, respectively.

^b Date-randomization tests were performed with 10 clustered permutations.

^c Common Era.

^d 95% credibility intervals are shown in parentheses.

```
<operators>
<!-- randomize tip location start from this line -->
    <tipStateSwapOperator weight="10">
        <ancestralTreeLikelihood idref="location.treeLikelihood"/>
    </tipStateSwapOperator>
<!-- randomize tip location the end line -->
</operators>
```

Fig. 23. XML code of `<tipStateSwapOperator>`. This element represents an operator to swap tip traits between two random tips. This operator element can be added within operators block.

連続空間における拡散経路解析においては、relaxed random walk model の緩和分布をコーシー分布、ガンマ分布、対数正規分布の 3 つから前項と同様の方法でモデル選択を行なった(Table 7)。そして、burn-in 後の事後確率分布から無作為に系統樹標本を 100 回抽出し、それらをもとに 95%信用区間を地図上に投射した。また、拡散速度解析には seraphim(Dellicour et al., 2016)を用い、祖先地からの拡散前線を継時的に解析および可視化した。

Table 7. Model selection among relaxed random walk models.

Phylogenetic group or subgroup	Protein-coding region ^a	Relaxed random walk model ^b	Log marginal likelihood ^c	
			Path sampling	Stepping-stone sampling
Asian-BR	HC-Pro*	Cauchy	-8081.60	-8086.29
		Gamma	-8020.22	-8025.41
		Lognormal	-8127.55	-8132.11
	P3*	Cauchy	-7745.30	-7749.66
		Gamma	-7672.20	-7676.56
		Lognormal	-7790.00	-7794.65
	NIb*	Cauchy	-3907.82	-3907.90
		Gamma	-3909.82	-3910.15
		Lognormal	-3916.23	-3916.74
	CP	Cauchy	-1986.82	-1986.72
		Gamma	-1987.29	-1987.15
		Lognormal	-1987.57	-1987.29
basal-BR	HC-Pro*	Cauchy	-4141.22	-4141.59
		Gamma	-4134.28	-4135.70
		Lognormal	-4160.43	-4161.49
	P3*	Cauchy	-4870.02	-4871.33
		Gamma	-4870.97	-4872.83
		Lognormal	-4881.27	-4883.86
	NIb*	Cauchy	-4724.17	-4725.89
		Gamma	-4716.11	-4718.66
		Lognormal	-4726.89	-4728.29
	CP	Cauchy	-4540.61	-4542.68
		Gamma	-4555.19	-4558.07
		Lognormal	-4576.15	-4579.73
world-B3	HC-Pro*	Cauchy	-5026.94	-5027.95
		Gamma	-4997.53	-4998.21
		Lognormal	-5051.11	-5049.56
	P3*	Cauchy	-6111.15	-6111.26
		Gamma	-6104.36	-6104.89
		Lognormal	-6140.87	-6140.75
	NIb*	Cauchy	-7007.50	-7010.91
		Gamma	-6954.99	-6958.67
		Lognormal	-7062.58	-7065.41
	CP	Cauchy	-5469.91	-5472.40
		Gamma	-5409.26	-5412.82
		Lognormal	-5533.91	-5535.65

^a HC-Pro*, P3*, and NIb* indicate partial protein-coding regions.

^b Selected models are shown in bold.

^c Log marginal likelihood was calculated as described in Baele et al. (2012).

結果

1. 組換え解析と組換え部位の同定

まず初めに、日本産 TuMV 全 370 分離株を用いた系統ネットワーク解析と最尤法による系統樹構築により、TuMV 日本集団の中には組換え体が多く存在することが示唆された(Figs. 24-25)。その後、ゲノムワイドに組換え解析を行い、網羅的に組換え部位を同定した(Figs. 26-31, Tables 8-10)。TuMV の日本集団は遺伝学的に 1. basal-BR 系統内組換え体および非組換え体、2. Asian-BR×world-B3 の系統間組換え体、3. world-B 系統内組換え体の大きく 3 つの集団に分かれることが示された (Fig. 29)。今回解析した全日本産 TuMV 分離株の 41%が Asian-BR×world-B3 の系統間組換え体であった。また、TuMV のゲノムには組換えが起こりやすいホットスポットと組換え部位が少ないコールドスポットが存在することが明らかになった(Fig. 27-28)。組換えホットスポットは、P1、HC-Pro、NIa-Pro、CP 遺伝子領域、コールドスポットは P3 遺伝子領域に多く見られた。TuMV の日本集団は、合計で 88 の独立した組換え部位が存在することが明らかになった(Figs. 26-27, Table 8)。これらの組換え部位のうち、4 箇所の組換え部位のみが海外産の TuMV でも報告されているもので(Kawakubo et al., 2021; Ohshima et al., 2007; Yasaka et al., 2017; Yasaka et al., 2015)、残り 84 箇所の組換え部位は TuMV の日本集団のみで見られた。

2. 組換え体型の地理的分布と継時的変遷

日本産 TuMV 全 370 分離株のうち、92%に相当する 339 分離株が少なくとも 1 箇所組換え部位を持つ組換え体で、8%に相当する 31 分離株のみが非組換え体であった(Fig. 26)。これらの非組換え体は全て、basal-BR もしくは world-B3 グループに属する分離株であった。また、日本産 TuMV の組換え体は、全 82 パターンの独立した組換え体型(RTP)に分けられた(Fig. 26)。これらのうち、日本以外でも分布が報告されているのは、2 パターンのみで 1 つ(RTP71)は台湾、もう 1 つ

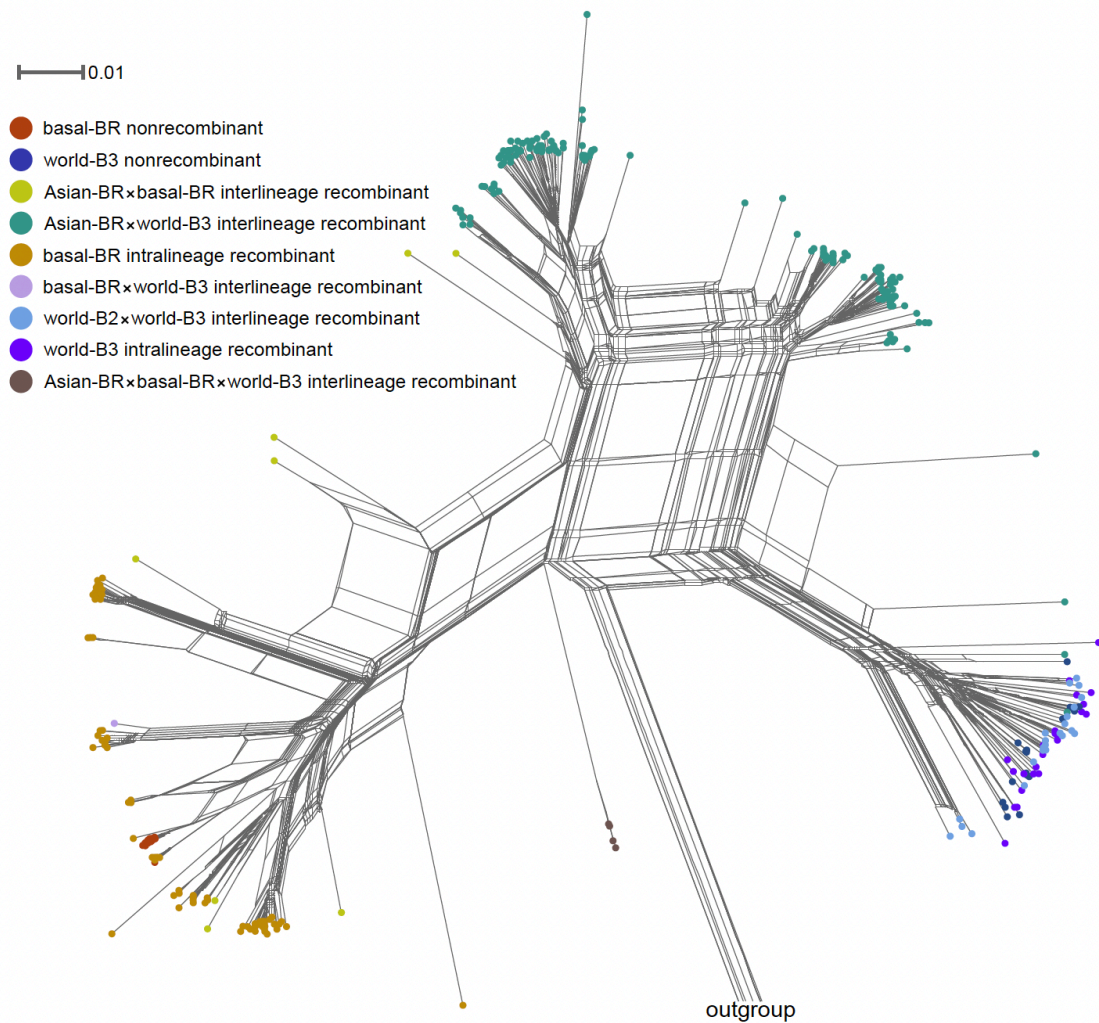


Fig. 24. Phylogenetic network of full genomic sequences of TuMV. Neighbor-Net method was used. Tip colours are shown in the inset key. The genome sequences of the isolates of narcissus late season yellows virus (NLSYV; accession numbers JQ326210, JX156421, and NC_023628), narcissus yellow stripe virus (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) were used as outgroup taxa.

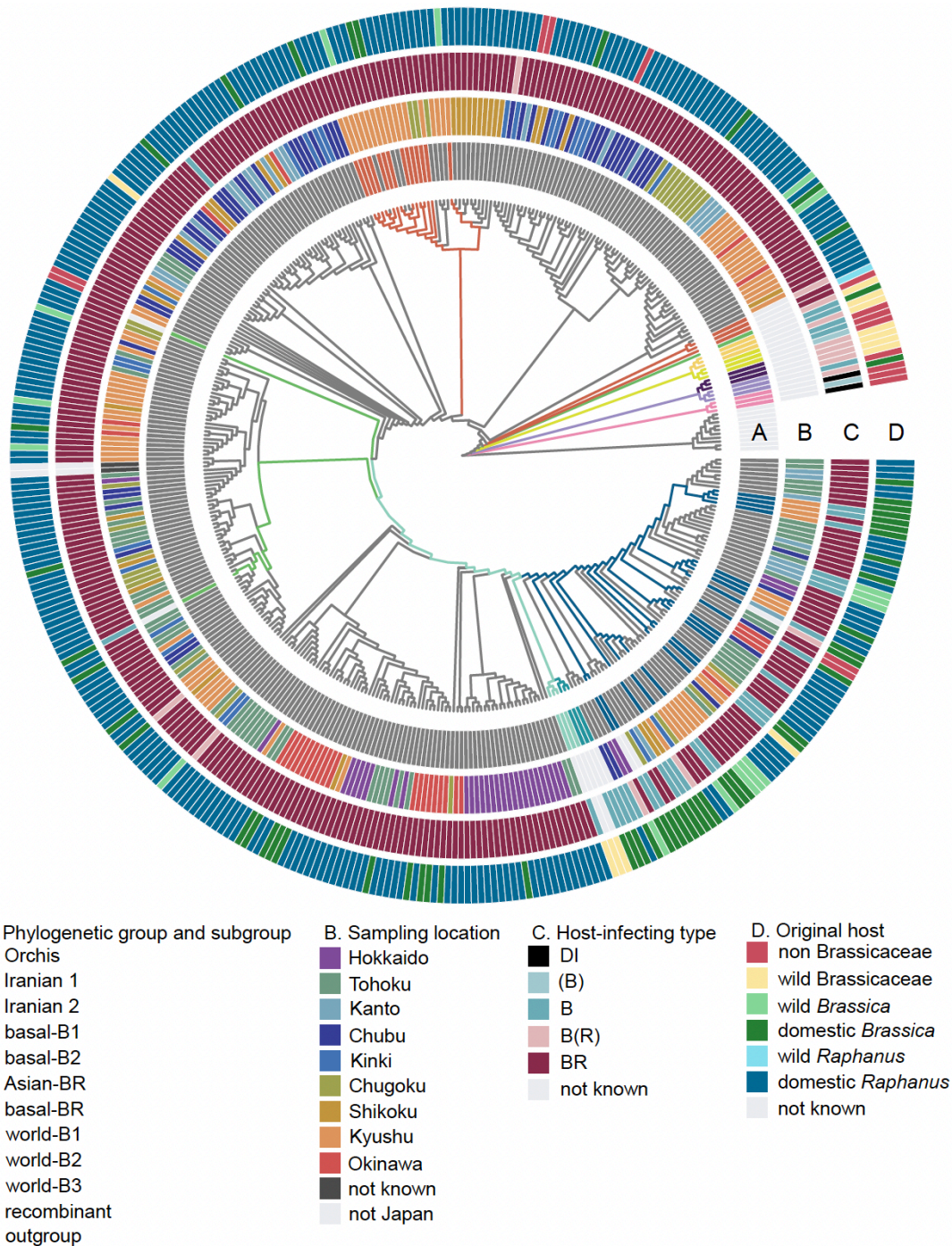


Fig. 25. Phylogenetic tree of TuMV inferred by maximum likelihood. Full genomic sequences of 370 Japanese isolates assembled with previously reported 30 nonrecombinant isolates were used. Concentric circles are coloured according to (A) phylogenetic group and subgroup, (B) sampling location, (C) host-infecting type, and (D) original host. The genome sequences of the isolates of 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) were used as outgroup taxa. The aligned genomic sequences are 8832 nt in length. Host-infecting type (B): isolates infected *Brassica rapa* only occasionally. Host-infecting type BR: isolates infected both *B. rapa* and *Raphanus sativus* systemically giving mosaic symptoms. Host-infecting type B(R): isolates infected *B. rapa* systemically giving mosaic symptoms, and infected *R. sativus* only occasionally. Host-infecting type DI: difficult to infect *Brassica* plants.

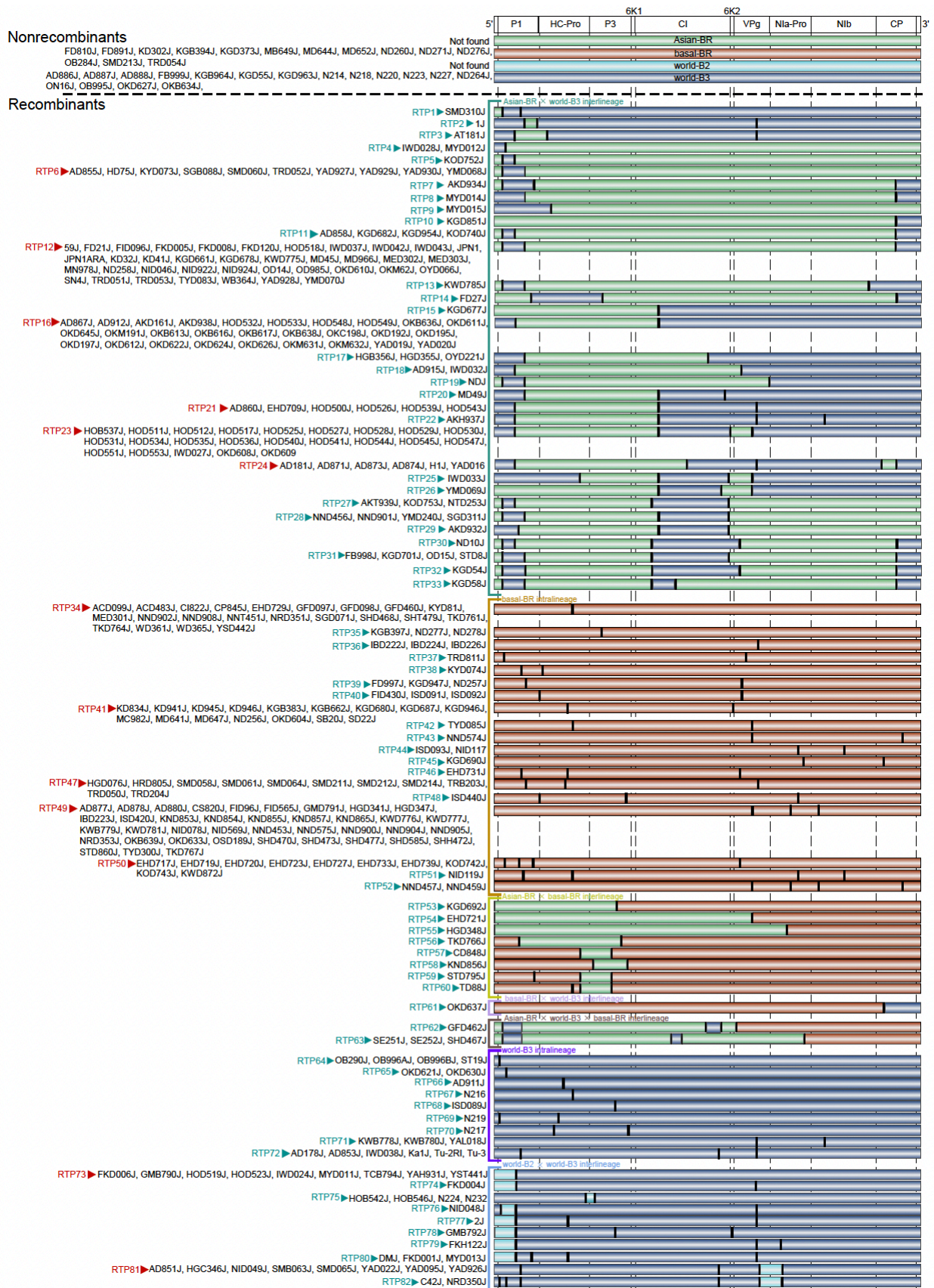


Fig. 26. Recombination map of TuMV of Japanese isolates analysed in this study. A total of 82 recombination type pattern (RTP) are listed and assigned with the number. RTP number in red font shows major recombination type pattern that were found in more than six isolates in Japan, hence used for subsequent phylogeographic analysis, whereas those in blue shows minor. Note that due to lack of sampling location information, RTP72 was not used for phylogeographic analysis even though it was found in six isolates.

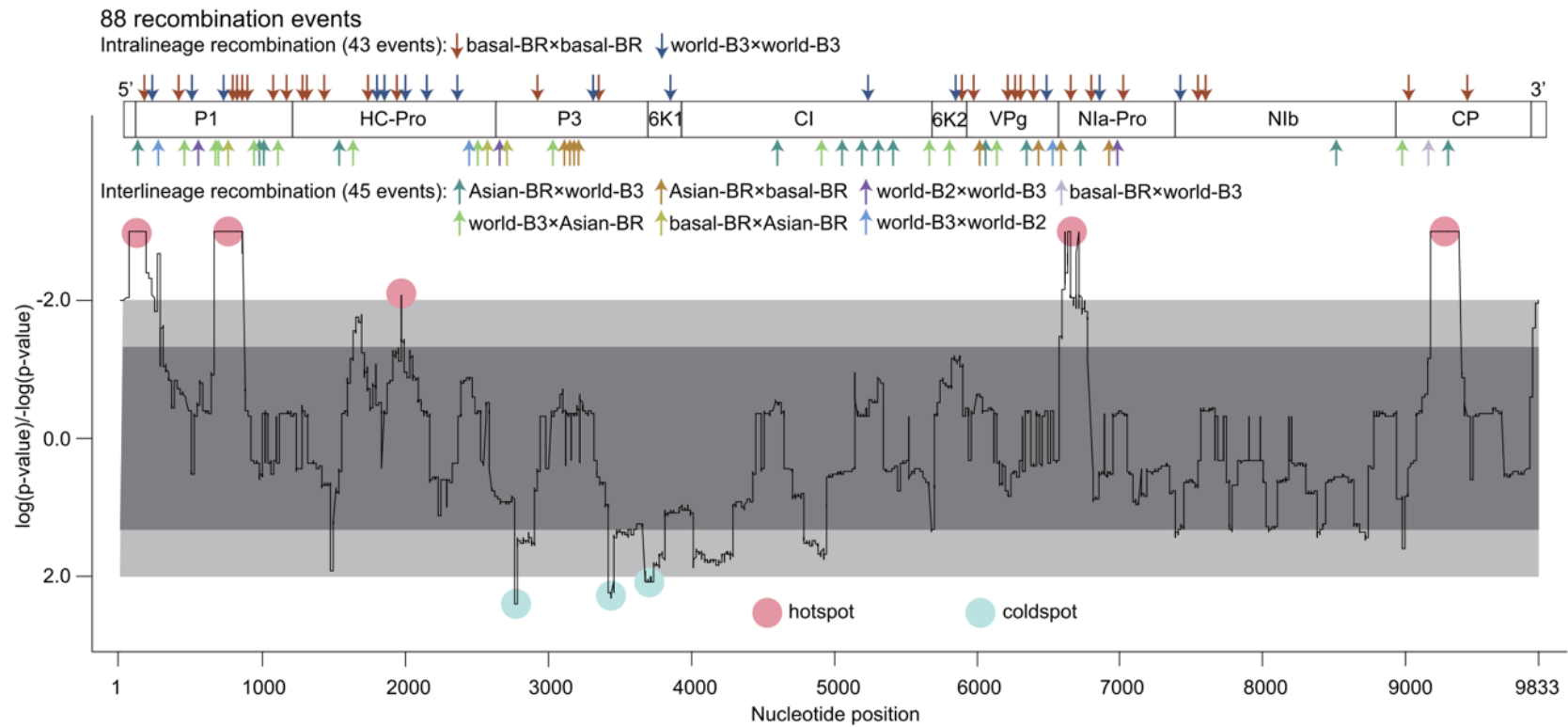


Fig. 27. A total of 88 recombination events found in Japanese TuMV genome with recombination frequency distribution across the genome. *P*-value for recombination frequency was plotted against TuMV genome corresponding to the nucleotide position of UK1 isolate(Jenner et al., 2000). Dark and light grey shadow indicates the range of 95 % and 99 % confidence intervals, respectively.

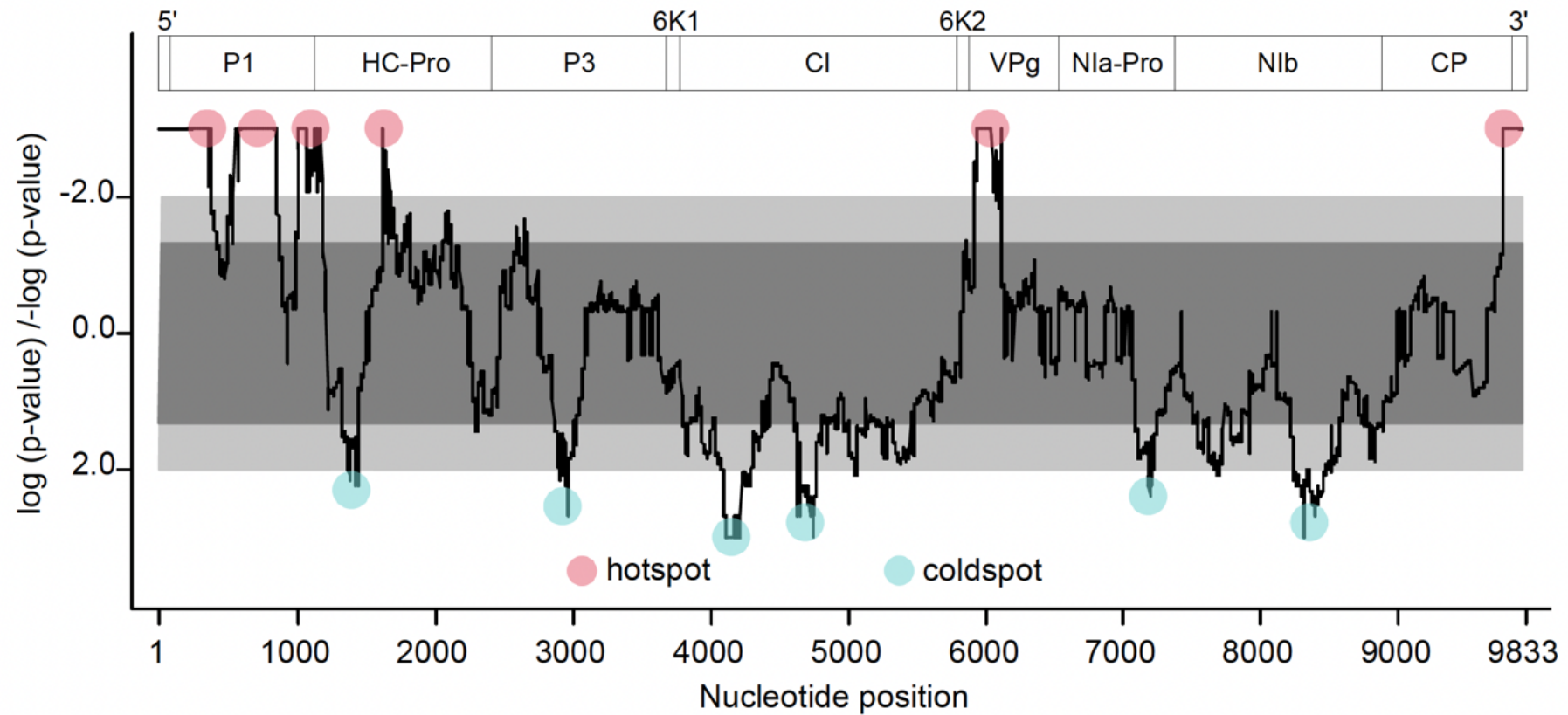


Fig. 28. Recombination hotspots and coldspots in TuMV genome. *P*-value for recombination frequency was plotted against TuMV genome corresponding to the nucleotide position of UK1 isolate (Jenner et al., 2000). Dark and light grey shadow indicate the range of 95 % and 99 % confidence intervals, respectively.

Table 8. Recombination events in TuMV genome detected by recombination detecting programs.

Protein-coding region and recombination events ^a	Isolate (recombination detecting program detected) ^b	Parental isolate ^c	Type of recombination ^d	P-value ^e
P1 (nt131-1216)				
1. P1-1 (nt179-225; n=75)	59J (<u>R</u> GBMCS), AD855J (<u>R</u> GBMCS), AD858J (<u>R</u> GBMCS), AKD934J (<u>R</u> GBMCS), AKT939J (<u>R</u> GB), FB998J (<u>R</u> GBMCS), FD21J (<u>R</u> GBMCS), FKD005J (<u>R</u> GBS), FKD008J (<u>R</u> GBS), FID096J (<u>R</u> GBS), FKD120J (<u>R</u> GBMCS), GFD462J (<u>R</u> GBMCS), HD75J (<u>R</u> GBMCS), HOD518J (<u>R</u> GBMS), IWD037J (<u>R</u> GBS), IWD042J (<u>R</u> GBS), IWD043J (<u>R</u> GBS), JPN1 (<u>R</u> GBMCS), JPN1ARA (<u>R</u> GBMCS), KD32J (<u>R</u> GBCS), KD41J (<u>R</u> GBMCS), KGD54J (<u>R</u> GBMCS), KGD58J (<u>R</u> GBMCS), KGD661J (<u>R</u> GBMCS), KGD682J (<u>R</u> GBMCS), KGD701J (<u>R</u> GBS), KGD954J (<u>R</u> GBMCS), KOD740J (<u>R</u> GBMCS), KOD752J (<u>R</u> GBMCS), KOD753J (<u>R</u> GBMCS), KWD775J (<u>R</u> GBMS), KWD785J (<u>R</u> GBMS), KGD678J (<u>R</u> GBMS), KYD073J (<u>R</u> GBMCS), MD45J (<u>R</u> GBS), MD966J (<u>R</u> GBMCS), MED302J (<u>R</u> GBS), MED303J (<u>R</u> GBS), MN978J (<u>R</u> GBMS), ND10J (<u>R</u> GBMCS), ND258J (<u>R</u> GBMS), NDJ (<u>R</u> GBMCS), NID046J (<u>R</u> GBMCS), NID922J (<u>R</u> GBS), NID924J (<u>R</u> GBS), NND456J (<u>R</u> GBMCS), NND901J (<u>R</u> GBMCS), NTD253J (<u>R</u> GBMCS), OD14J (<u>R</u> GBS), OD15J (<u>R</u> GBS), OD985J (<u>R</u> GBMCS), OKD610J (<u>R</u> GBMCS), OKM62J (<u>R</u> GBS), OYD066J (<u>R</u> GBMS), SE251J (<u>R</u> GBMCS), SE252J (<u>R</u> GBMCS), SGB088J (<u>R</u> GBMCS), SGD311J (<u>R</u> GBS), SHD467J (<u>R</u> GBMCS), SMD060J (<u>R</u> GBMCS), SMD310J (<u>R</u> GBMCS), SN4J (<u>R</u> GBS), STD8J (<u>R</u> GBMS), TRD051J (<u>R</u> GBMCS), TRD052J (<u>R</u> GBS), TRD053J (<u>R</u> GBMCS), TYD083J (<u>R</u> GBS), WB364J (<u>R</u> GBS), YAD927J (<u>R</u> GBMCS), YAD928J (<u>R</u> GBMCS), YAD929J (<u>R</u> GBMCS), YAD930J (<u>R</u> GBMCS), YMD068J (<u>R</u> GBMCS), YMD070J (<u>R</u> GBS), YMD240J (<u>R</u> GBMCS)	YAD016J/CH7 × T22/ Tu-2RI	Asian-BR × world-B3	$1.07 \times 10^{-70} - 1.49 \times 10^{-19}$
2. P1-2 (nt223; n=1)	TRD81J (<u>R</u> GBMCS)	AD886J × SMD212J	basal-BR × basal-BR	1.37×10^{-101}
3. P1-3 (nt226-279; n=7)	C42J (<u>R</u> GBMCS), N219 (<u>R</u> GBMCS), NRD350J (<u>R</u> GBMS), OB290J (<u>R</u> GB), OB996AJ (<u>R</u> GBS), OB996BJ (<u>R</u> GBS), ST19J (<u>R</u> GBS)	GBR7 × NND459J ISD089J × CHHZ3A	world-B3 × world-B3	$3.73 \times 10^{-57} - 8.24 \times 10^{-10}$
4. P1-4 (nt254; n=1)	NID048J (<u>R</u> GBMCS)	VIET180 × DNK2	world-B3 × world-B2	1.08×10^{-27}
5. P1-5 (nt388-458; n=10)	EHD717J (<u>R</u> BS), EHD719J (<u>R</u> GBMS), EHD720J (<u>R</u> BCS), EHD723J (<u>R</u> BS), EHD727J (<u>R</u> GBMS), EHD733J (<u>R</u> BS), EHD739J (<u>R</u> GBMCS), KOD742J (<u>R</u> GBMS), KOD743J (<u>R</u> GBMCS), KWD872J (<u>R</u> BCS)	ITA7 × TRD050J	basal-BR × basal-BR	$2.19 \times 10^{-14} - 4.99 \times 10^{-7}$
6. P1-6 (nt394-602; n=4)	AKD932J (<u>R</u> GBMCS), IWD028J (<u>R</u> GBMCS), MD49J (<u>R</u> GBMCS), MYD012J (<u>R</u> GBMCS)	OKD611J/ CHSE1AB × CHN342J/YAD019J	world-B3 × Asian-BR	$9.92 \times 10^{-41} - 2.13 \times 10^{-28}$
7. P1-7 (nt657-665; n=4)	C42J (<u>R</u> GBMCS), NRD350J (<u>R</u> GBMS), OKD621J (<u>R</u> GBMCS), OKD630J (<u>R</u> GBMCS)	NND459J × GBR7	world-B3 × world-B3	$1.98 \times 10^{-18} - 6.38 \times 10^{-13}$
8. P1-8 (nt660-700; n=17)	2J (<u>R</u> GBCS), DMJ (<u>R</u> GB), FKD001J (<u>R</u> GB), FKD004J (<u>R</u> BC), FKD006J (<u>R</u> GBCS), FKH122J (<u>R</u> GBCS), GMB790J (<u>R</u> GBMCS), GMB792J (<u>R</u> GBS), HOD519J (<u>R</u> GBMCS), HOD523J (<u>R</u> GBMCS), IWD024J (<u>R</u> GBCS), MYD011J (<u>R</u> GBMS), MYD013J (<u>R</u> BS), NID048J (<u>R</u> GBMCS), TCB794J (<u>R</u> GB), YAH931J (<u>R</u> GBMCS), YST441J (<u>R</u> GBS)	DNK2 × HOD533J	world-B2 × world-B3	$1.94 \times 10^{-32} - 4.51 \times 10^{-12}$
9. P1-9 (nt703-790; n=72)	1J (<u>R</u> GBMS), 59J (<u>R</u> GBMS), AD855J (<u>R</u> GBMCS), FB998J (<u>R</u> GBMS), FD21J (<u>R</u> GBMCS), FID096J (<u>R</u> GBS), FKD005J (<u>R</u> GBMS), FKD008J (<u>R</u> GBMS), FKD120J (<u>R</u> GBMCS), GFD462J (<u>R</u> GBMCS), HD75J (<u>R</u> GBMS), HGB356J (<u>R</u> GBMS), HGD355J (<u>R</u> GBMS), HOD518J (<u>R</u> GBMS), IWD037J (<u>R</u> GBS), IWD042J (<u>R</u> GBS), IWD043J (<u>R</u> GBS), JPN1 (<u>R</u> GBMS), JPN1ARA (<u>R</u> GBMS), KD32J (<u>R</u> GBCS), KD41J (<u>R</u> GBMS), KGD54J (<u>R</u> GBMS), KGD58J (<u>R</u> GBMS), KGD661J (<u>R</u> GBMS), KGD677J (<u>R</u> GBMS), KGD678J (<u>R</u> GBMS), KGD701J (<u>R</u> GBMS), KWD775J (<u>R</u> GBMS), KWD785J (<u>R</u> GBMS), KYD073J (<u>R</u> GBMCS), MD45J (<u>R</u> GBS), MD49J (<u>R</u> GBMCS), MD966J (<u>R</u> GBMS), MED302J (<u>R</u> GBS), MED303J (<u>R</u> GBS), MN978J (<u>R</u> GBMS), ND10J (<u>R</u> GBMS), ND258J (<u>R</u> GBMS), NDJ (<u>R</u> GBMCS), NID046J (<u>R</u> GBS), NID922J (<u>R</u> GBS), NID924J (<u>R</u> GBS), NND456J (<u>R</u> GBMS), NND901J (<u>R</u> GBMS), OD14J (<u>R</u> GBS), OD15J (<u>R</u> GBS), OD985J (<u>R</u> GBMS), OKD610J (<u>R</u> GBMCS), OKM62J (<u>R</u> GBS), OYD066J (<u>R</u> GBMS), OYD221J (<u>R</u> GBMS), SGB088J (<u>R</u> GBMS), SGD311J (<u>R</u> GBMS), SE251J (<u>R</u> GBMCS), SE252J (<u>R</u> GBMCS), SHD467J (<u>R</u> GBMCS), SMD060J (<u>R</u> GBMS), SN4J (<u>R</u> GBS), STD8J (<u>R</u> GBMS), TRD051J (<u>R</u> GBMS), TRD052J (<u>R</u> GBS), TRD053J (<u>R</u> GBMS), TYD083J (<u>R</u> GBS), WB364J (<u>R</u> GBS), YAD927J (<u>R</u> GBMS), YAD928J (<u>R</u> GBMS), YAD929J (<u>R</u> GBMS), YAD930J (<u>R</u> GBMS), YMD068J (<u>R</u> GBMS), YMD069J (<u>R</u> GBMS), YMD070J (<u>R</u> GBS), YMD240J (<u>R</u> GBMS)	CHZH34A × TUR10	world-B3 × Asian-BR	$2.57 \times 10^{-56} - 1.61 \times 10^{-21}$
10. P1-10 (nt722-785; n=78)	AD181J (<u>R</u> GBMCS), AD858J (<u>R</u> GBMCS), AD860J (<u>R</u> GBMCS), AD867J (<u>R</u> GBMCS), AD871J (<u>R</u> GBMCS), AD873J (<u>R</u> GBMCS), AD874J (<u>R</u> GBMCS), AD912J (<u>R</u> GBMCS), AD915J (<u>R</u> GBMCS), AKD161J (<u>R</u> GBMCS), AKD938J (<u>R</u> GBMCS), AKH937J (<u>R</u> GBMCS), AKT939J (<u>R</u> GB), AT181J (<u>R</u> GBMCS), EHD709J (<u>R</u> GBMCS), KOD752J (<u>R</u> GBMCS), H1J (<u>R</u> GBMCS), HOB537J (<u>R</u> GBMCS), HOD500J (<u>R</u> GBMCS), HOD511J (<u>R</u> GBMCS), HOD512J (<u>R</u> GBMCS), HOD517J (<u>R</u> GBMCS), HOD525J (<u>R</u> GBMCS), HOD526J (<u>R</u> GBMCS), HOD527J (<u>R</u> GBMCS), HOD528J (<u>R</u> GBMCS), HOD529J (<u>R</u> GBMCS), HOD530J (<u>R</u> GBMCS), HOD531J (<u>R</u> GBMCS), HOD532J (<u>R</u> GBMCS), HOD533J (<u>R</u> GBMCS), HOD534J (<u>R</u> GBMCS), HOD535J (<u>R</u> GBMCS), HOD536J (<u>R</u> GBMCS), HOD539J (<u>R</u> GBMCS), HOD540J (<u>R</u> GBMCS), HOD541J (<u>R</u> GBMCS), HOD543J (<u>R</u> GBMCS), HOD544J (<u>R</u> GBMCS), HOD545J (<u>R</u> GBMCS), HOD547J (<u>R</u> GBMCS), HOD548J (<u>R</u> GBMCS), HOD549J (<u>R</u> GBMCS), HOD551J (<u>R</u> GBMCS), HOD553J (<u>R</u> GBMCS), IWD027J (<u>R</u> GBMCS), IWD032J (<u>R</u> GBMCS), KOD753J (<u>R</u> GBMCS), KOD740J (<u>R</u> GBMCS), KGD682J (<u>R</u> GBMCS), KGD954J (<u>R</u> GBMCS), ND10J (<u>R</u> GBMCS), NTD253J (<u>R</u> GBMCS), OKB613J (<u>R</u> GBMCS), OKB616J (<u>R</u> GBMCS), OKB617J (<u>R</u> GBMCS), OKB636J (<u>R</u> GBMCS), OKB638J (<u>R</u> GBMCS), OKC198J (<u>R</u> GBMCS), OKD192J (<u>R</u> GBMCS), OKD195J (<u>R</u> GBMCS), OKD197J (<u>R</u> GBMCS), OKD608J (<u>R</u> GBMCS), OKD609J (<u>R</u> GBMCS), OKD611J (<u>R</u> GBMCS), OKD612J (<u>R</u> GBMCS), OKD621J (<u>R</u> GBMCS), OKD622J (<u>R</u> GBMCS), OKD624J (<u>R</u> GBMCS), OKD626J (<u>R</u> GBMCS), OKD645J (<u>R</u> GBMCS), OKM191J (<u>R</u> GBMCS), OKM631J (<u>R</u> GBMCS), OKM632J (<u>R</u> GBMCS), YAD016J (<u>R</u> GBMCS), YAD019J (<u>R</u> GBMCS), YAD020J (<u>R</u> GBMCS)	OB995J × HOD500J	world-B3 × Asian-BR	$7.18 \times 10^{-129} - 7.72 \times 10^{-20}$
11. P1-11 (nt737-751; n=17)	AD178J (<u>R</u> GBMCS), AD851J (<u>R</u> GBMCS), AD853J (<u>R</u> GBMCS), C42J (<u>R</u> GBMCS), IWD038J (<u>R</u> GBMCS), HGC346J (<u>R</u> GBMCS), Ka1J (<u>R</u> GBMCS), NID049J (<u>R</u> GBMCS), NRD350J (<u>R</u> GBMCS), SMB063J (<u>R</u> GBMCS), SMD065J (<u>R</u> GBMCS), SMD310J (<u>R</u> GBMCS), Tu-2RI (<u>R</u> GBMCS), Tu-3 (<u>R</u> GBMCS), YAD022J (<u>R</u> GBMCS), YAD925J (<u>R</u> GBMCS), YAD926J (<u>R</u> GBMCS)	IWD033J/GBR7 × YAD929J/YAD020J	world-B3 × world-B3	$6.08 \times 10^{-69} - 2.32 \times 10^{-32}$

^a Recombination events detected in the turnip mosaic virus genomes by RDP5 (Martin et al., 2020) from the aligned sequences of their genomes. Nucleotide positions are numbered corresponding to UK1 genome (Jenner et al., 2000). P1, protein 1; HC-Pro, helper component proteinase; P3, protein 3; 6K1, first 6kDa protein; CI, cylindrical inclusion protein; 6K2, second 6kDa protein; VPg, genome-linked viral protein; NIa-Pro, nuclear inclusion a proteinase; NIb, nuclear inclusion b protein; and CP, coat protein.

^b Recombination isolates identified by the recombination detecting programs; R (RDP), G (GENECONV), B (BOOTSCAN), M (MAXCHI), C (CHIMAERA) and S (SISCAN) programs in RDP5. The program which gave smallest *P*-value among the other detected programs is shown in underlined and bold font.

^c Parental isolates are shown in the order of 5' side of the recombination event to 3' side.

^d Phylogenetic (sub-)groups of parental isolates shown in the order of 5' side of the recombination event to 3' side.

^e The smallest *P*-values of the programs shown in underlined and bold font were reported in a range among all isolates which share the same recombination event.

Table 8. Continued.

Protein-coding region and recombination events	Isolate (recombination detecting program detected)	Parental isolate	Type of recombination	P-value
12. P1-12 (nt756; n=1)	TKD766J (R GBMCS)	CH7 × OD985J	basal-BR × Asian-BR	1.60×10^{-35}
13. P1-13 (nt771-1083; n=14)	FD997J (R GBMCS), HGD076J (R GBMCS), HRD805J (R GBMCS), KGD947J (R GBMCS), ND257J (R GBMCS), SMD058J (R GBMCS), SMD061J (R GBMCS), SMD064J (R GBMCS), SMD211J (R GBMCS), SMD212J (R GBMCS), SMD214J (R GBMCS), TRB203J (R GBMCS), TRD050J (R GBMCS), TRD204J (R GBMCS)	NND908J/YSD442J × KGB394J/DEU4	basal-BR × basal-BR	$8.03 \times 10^{-57} - 9.45 \times 10^{-45}$
14. P1-14 (nt857; n=10)	EHD717J (R BS), EHD719J (R GBMS), EHD720J (R BCS), EHD723J (R BS), EHD727J (R GBMS), EHD733J (R BS), EHD739J (R GBMCS), KOD742J (R GBMS), KOD743J (R BMC S), KWD872J (R BCS)	TRD050J × ITA7	basal-BR × basal-BR	$4.15 \times 10^{-14} - 8.76 \times 10^{-8}$
15. P1-15 (nt879; n=1)	EHD731J (R BS)	KGD947J × WFLB06	basal-BR × basal-BR	2.62×10^{-10}
16. P1-16 (nt880; n=2)	KYD074J (R GBCS), NID119J (R GMS)	DJ5 × IRNWRSh40	basal-BR × basal-BR	$4.56 \times 10^{-9} - 1.65 \times 10^{-8}$
17. P1-17 (nt905; n=1)	AKD934J (R GBMCS)	HOD519J × KGD851J	world-B3 × Asian-BR	5.70×10^{-41}
18. P1-18 (nt930; n=1)	FD27J (R BMC S)	MYD014J × N219	Asian-BR × world-B3	3.25×10^{-40}
19. P1-19 (nt950; n=1)	IJ (R GBCS)	JPN1 × N216	Asian-BR × world-B3	3.58×10^{-25}
20. P1-20 (nt1079; n=1)	STD795J (R GBMCS)	NND575J × INRWRSh40	basal-BR × basal-BR	3.78×10^{-37}
21. P1-21 (nt1111; n=1)	MYD014J (R GBMCS)	KGD851J × HOD519J	world-B3 × Asian-BR	1.39×10^{-80}
22. P1-22 (nt1177-1118; n=4)	FID430J (R GBCS), ISD091J (R GBMCS), ISD092J (R GBMCS), ISD440J (R GBMS)	SMD213J × YSD442J	basal-BR × basal-BR	$4.69 \times 10^{-49} - 1.09 \times 10^{-27}$
HC-Pro (nt1217-2590)				
23. HC-Pro-1 (nt1254; n=1)	KYD074J (R GBMCS)	IRNWRSh40 × DJ5	basal-BR × basal-BR	1.50×10^{-12}
24. HC-Pro-2 (nt1266; n=10)	EHD717J (R GBMCS), EHD719J (R GBMS), EHD720J (R GBMCS), EHD723J (R GBMCS), EHD727J (R GBMCS), EHD733J (R GBMCS), EHD739J (R GBMCS), KOD742J (R GBMS), KOD473J (R GBMCS), KWD872J (R GBMCS)	ITA7 × TRD050J	basal-BR × basal-BR	$4.25 \times 10^{-65} - 4.12 \times 10^{-62}$
25. HC-Pro-3 (nt1418-1466; n=10)	HGD076J (R GBMCS), HRD805J (R GBMCS), SMD058J (R GBMCS), SMD061J (R GBMCS), SMD064J (R GBMCS), SMD211J (R GBMCS), SMD212J (R GBMCS), SMD214J (R GBMCS), TRB203J (R GBMCS), TRD050J (R GBMCS), TRD204J (R GBMCS)	ISD902 × R065	basal-BR × basal-BR	$4.12 \times 10^{-38} - 4.92 \times 10^{-36}$
26. HC-Pro-4 (nt1562; n=1)	AT181J (R GBMCS)	KOD753J × N216	Asian-BR × world-B3	4.44×10^{-90}
27. HC-Pro-5 (nt1624; n=1)	MYD015J (R GBMCS)	HOD519J × CHBJ1	world-B3 × Asian-BR	1.71×10^{-133}
28. HC-Pro-6 (nt1676-1700; n=17)	EHD731J (R BS), KD834J (R GBMCS), KD941J (R GBMCS), KD945J (R GBMCS), KD946J (R GBMCS), KGB383J (R GBMCS), KGB662J (R GBMCS), KGD680J (R GBMCS), KGD687J (R GBMCS), KGD946J (R GBMCS), MC982J (R GBMCS), MD641J (R GBMCS), MD647J (R GBMCS), ND256J (R GBMCS), OKD604J (R GBMCS), SB20J (R GBMCS), SD22J (R GBMCS)	ITA7/WFLB06 × R041/KGD947J	basal-BR × basal-BR	$6.41 \times 10^{-28} - 5.31 \times 10^{-9}$
29. HC-Pro-7 (nt1900; n=1)	N217 (R BMC S)	CZE11 × CHZJ27A	world-B3 × world-B3	1.67×10^{-16}
30. HC-Pro-8 (nt1918; n=1)	N219 (R GBMCS)	CHHZ3A × ISD089J	world-B3 × world-B3	8.68×10^{-17}
31. HC-Pro-9 (nt1916-1922; n=25)	ACD099J (R GBMCS), ACD483J (R GBMCS), CI822J (R GBMCS), CP845J (R GBMCS), EHD729J (R GBMCS), GFD097J (R GBMCS), GFD098J (R GBMCS), GFD460J (R GBMCS), KYD81J (R GBMCS), MED301J (R GBMCS), NID119J (R BMS), NND902J (R GBMCS), NND908J (R GBMCS), NNT451J (R GBMCS), NRD351J (R GBMCS), SGD071J (R GBMCS), SHD468J (R GBMCS), SHH479J (R GBMCS), TD88J (R GBMCS), TKD761J (R GBMCS), TKD764J (R GBMCS), TYD085J (R GBMCS), YSD442J (R GBMCS), WD361J (R GBMCS), WD365J (R GBMCS)	STD795J × IRNWRSh40	basal-BR × basal-BR	$2.40 \times 10^{-39} - 1.62 \times 10^{-28}$
32. HC-Pro-10 (nt1974; n=1)	AD911J (R BMS)	N219 × CHZH34A	world-B3 × world-B3	2.56×10^{-15}
33. HC-Pro-11 (nt2059-2205; n=4)	2J (R GBMS), FKD001J (R GBMCS), DMJ (R GBMCS), MYD013J (R BS)	CHZH33A × N223	world-B3 × world-B3	$1.61 \times 10^{-16} - 3.77 \times 10^{-7}$
34. HC-Pro-12 (nt2306; n=1)	N216 (R GBMS)	NRD350J × CHZH33A	world-B3 × world-B3	6.37×10^{-13}
35. HC-Pro-13 (nt2385; n=4)	HOB542J (R GB), HOB546J (R GB), N224 (R GB), N232 (R GB)	ISD069J × DNK3	world-B3 × world-B2	$7.00 \times 10^{-8} - 2.94 \times 10^{-7}$
36. HC-Pro-14 (nt2401; n=1)	IWD033J (R GBMCS)	AT181J × MYD015J	world-B3 × Asian-BR	4.30×10^{-82}
37. HC-Pro-15 (nt2447-2484; n=3)	CD848J (R GBMCS), STD795J (R GBMCS), TD88J (R GBMCS)	SHH472J × NID924J	basal-BR × Asian-BR	$3.58 \times 10^{-109} - 2.03 \times 10^{-50}$
P3 (nt2591-3655)				
38. P3-1 (nt2602; n=4)	HOB542J (R GB), HOB546J (R GB), N224 (R GB), N232 (R GB)	DNK3 × ISD069J	world-B2 × world-B3	$7.00 \times 10^{-8} - 2.94 \times 10^{-7}$
39. P3-2 (nt2618; n=1)	KND856J (R GBMCS)	EHD729J × SE251J	basal-BR × Asian-BR	8.30×10^{-78}
40. P3-3 (nt2845; n=3)	ND277J (R GBMCS), ND278J (R GBMCS), KGB397J (R GBMCS)	SMD213J × KD946J	basal-BR × basal-BR	$4.79 \times 10^{-100} - 2.54 \times 10^{-36}$
41. P3-4 (nt2988; n=1)	FD27J (R BMC S)	N219 × MYD014J	world-B3 × Asian-BR	1.49×10^{-80}
42. P3-5 (nt3007-3021; n=3)	CD848J (R GBMCS), STD795J (R GBMCS), TD88J (R GBMCS)	NID924J × SHH472J	Asian-BR × basal-BR	$3.58 \times 10^{-109} - 2.63 \times 10^{-50}$
43. P3-6 (nt3144; n=1)	KGD692J (R GBMCS)	ND258J × IND33	Asian-BR × basal-BR	1.49×10^{-168}
44. P3-7 (nt3210; n=1)	TKD766J (R GBMCS)	NND456J × WD361J	Asian-BR × basal-BR	4.47×10^{-148}
45. P3-8 (nt3228; n=1)	KND856J (R GBMCS)	SE251J × EHD729J	Asian-BR × basal-BR	8.30×10^{-78}
46. P3-9 (nt3279-3282; n=2)	GMB792J (R GMS), ISD089J (R BS)	MYD013J × AD911J	world-B3 × world-B3	$1.09 \times 10^{-11} - 3.49 \times 10^{-8}$
47. P3-10 (nt3345; n=1)	ISD440J (R GBMS)	SMD213J × OKD637J	basal-BR × basal-BR	1.02×10^{-55}
6K1 (nt3656-3811)				
48. 6K1-1 (nt3688; n=1)	N217 (R BMC S)	CHZJ27A × CZE11	world-B3 × world-B3	1.67×10^{-16}
CI (nt3812-5743)				
49. CI-1 (nt4622-4644; n=13)	AKT939J (R GBMCS), FB998J (R GBMCS), KGD54J (R GBMCS), KGD58J (R GBS), KGD701J (R GBMCS), KOD753J (R GBMCS), ND10J (R GBMCS), NTD253J (R GBMCS), OD15J (R GBMCS), SE251J (R GBS), SE252J (R GBS), SHD467J (R GBS), STD8J (R GBMCS)	NDJ/NID922J/IWD033J × N216/N214/ISD089J	Asian-BR × world-B3	$2.16 \times 10^{-47} - 7.52 \times 10^{-32}$
50. CI-2 (nt4847-4917; n=4)	KGD58J (R GBS), SE251J (R GBS), SE252J (R GBS), SHD467J (R GBS)	ISD089J × IWD033J	world-B3 × Asian-BR	$3.37 \times 10^{-37} - 3.41 \times 10^{-31}$

Table 8. Continued.

Protein-coding region and recombination events	Isolate (recombination detecting program detected)	Parental isolate	Type of recombination	<i>P</i> -value
51. CI-3 (nt4982-5016; n=69)	AD860J (R GBMCS), AD867J (R GBMCS), AD912J (R GBMCS), AKD161J (R GBMCS), AKD932J (R GBMCS), AKD938J (R GBMCS), AKH937J (R GBMCS), EHD709J (R GBMCS), HOB537J (R GBMCS), HOD500J (R GBMCS), HOD511J (R GBMCS), HOD512J (R GBMCS), HOD517J (R GBMCS), HOD525J (R GBMCS), HOD526J (R GBMCS), HOD527J (R GBMCS), HOD528J (R GBMCS), HOD529J (R GBMCS), HOD530J (R GBMCS), HOD531J (R GBMCS), HOD532J (R GBMCS), HOD533J (R GBMCS), HOD534J (R GBMCS), HOD535J (R GBMCS), HOD536J (R GBMCS), HOD539J (R GBMCS), HOD540J (R GBMCS), HOD541J (R GBMCS), HOD543J (R GBMCS), HOD544J (R GBMCS), HOD545J (R GBMCS), HOD547J (R GBMCS), HOD548J (R GBMCS), HOD549J (R GBMCS), HOD551J (R GBMCS), HOD553J (R GBMCS), IWD027J (R GBMCS), IWD033J (R GBMCS), KGD677J (R GBMCS), MD49J (R GBMCS), NND456J (R GBMCS), NND901J (R GBMCS), OKB613J (R GBMCS), OKB616J (R GBMCS), OKB617J (R GBMCS), OKB636J (R GBMCS), OKB638J (R GBMCS), OKC198J (R GBMCS), OKD192J (R GBMCS), OKD195J (R GBMCS), OKD197J (R GBMCS), OKD611J (R GBMCS), OKD612J (R GBMCS), OKD621J (R GBMCS), OKD622J (R GBMCS), OKD624J (R GBMCS), OKD626J (R GBMCS), OKD645J (R GBMCS), OKD608J (R GBMCS), OKD609J (R GBMCS), OKM191J (R GBMCS), OKM631J (R GBMCS), OKM632J (R GBMCS), SGD311J (R GBMCS), YAD019J (R GBMCS), YAD020J (R GBMCS), YMD069J (R GBMCS), YMD240J (R GBMCS)	OKM62J/FD27J × N223/N214	Asian-BR × world-B3	$1.10 \times 10^{144} - 1.76 \times 10^{65}$
52. CI-4 (nt5183-5253; n=6)	AD181J (R GBMCS), AD871J (R GBMCS), AD873J (R GBMCS), AD874J (R GBMCS), H1J (R GBMCS), YAD016J (R GBMCS)	AD858J × N223	Asian-BR × world-B3	$6.57 \times 10^{129} - 2.69 \times 10^{110}$
53. CI-5 (nt5150-5151; n=16)	AD178J (R BMC S), AD851J (R BMC S), AD853J (R BMC S), C42J (R GBMCS), HGC346J (R GBMCS), IWD038J (R BMC S), Ka1J (R BMC S), NID049J (R GBMCS), NRD350J (R GBMCS), SMD063J (R GBMCS), SMD065J (R GBMCS), Tu-2R1 (R BMC S), Tu-3 (R GBMCS), YAD022J (R BMC S), YAD925J (R BMC S), YAD926J (R BMC S)	GBR7 × YAD020J	world-B3 × world-B3	$5.19 \times 10^{125} - 3.97 \times 10^{21}$
54. CI-6 (nt5261-5349; n=1)	GFD462J (R GBS)	AD858J × N223	Asian-BR × world-B3	5.38×10^{28}
55. CI-7 (nt5360-5394; n=3)	HGB356J (R GBMS), HGD355J (R GBMS), OYD221J (R GBMS)	AD858J × N223	Asian-BR × world-B3	$4.60 \times 10^{110} - 2.96 \times 10^{34}$
56. CI-8 (nt5714-5737; n=2)	GFD462J (R GB), YMD069J (R GBMCS)	N216 × FKD005J	world-B3 × Asian-BR	$1.63 \times 10^{47} - 1.67 \times 10^{22}$
6K2 (nt5744-5902)				
57. 6K2-1 (nt5730-5763; n=37)	AKD932J (R GBMCS), AKT939J (R GBMCS), FB998J (R GBMCS), HOB537J (R BMC S), HOD511J (R GBMCS), HOD512J (R GBMCS), HOD517J (R BMC S), HOD525J (R GBMCS), HOD527J (R BMC S), HOD528J (R GBMCS), HOD529J (R GBMCS), HOD530J (R GBMCS), HOD531J (R GBMCS), HOD534J (R GBMCS), HOD535J (R BMC S), HOD536J (R GBMCS), HOD540J (R GBMCS), HOD541J (R GBMCS), HOD544J (R GBMCS), HOD545J (R GBMCS), HOD547J (R GBMCS), HOD551J (R GBMCS), HOD553J (R GBMCS), IWD027J (R BMC S), IWD033J (R BMC S), KGD701J (R GBMCS), KOD753J (R GBMCS), ND16J (R GBMCS), NND456J (R GBMCS), NND901J (R GBMCS), NTD253J (R GBMCS), OD15J (R GBMCS), OKD608J (R GBMCS), OKD609J (R GBMCS), SGD311J (R GBMCS), STD8J (R GBMCS), YMD240J (R GBMCS)	1J/N216 × CHSE1AB/ NDJ	world-B3 × Asian-BR	$1.42 \times 10^{130} - 1.32 \times 10^{70}$
58. 6K2-2 (nt5864; n=1)	GMB792J (R GBMCS)	AD911J × MYD013J	world-B3 × world-B3	1.09×10^{11}
59. 6K2-3 (nt5882-5909; n=15)	KD834J (R GBMCS), KD941J (R GBMCS), KD945J (R GBMCS), KD946J (R GBMCS), KGB383J (R GBMCS), KGB662J (R GBMCS), KGD680J (R GBMCS), KGD687J (R GBMCS), KGD946J (R GBMCS), MC982J (R GBMCS), MD641J (R GBMCS), MD647J (R GBMCS), ND256J (R GBMCS), SB20J (R GBMCS), SD22J (R GBMCS)	R041 × ITA7	basal-BR × basal-BR	$8.49 \times 10^{33} - 5.61 \times 10^{32}$
VPg (nt5903-6478)				
60. VPg-1 (nt5942; n=11)	EHD717J (R GBMCS), EHD719J (R GBMS), EHD720J (R GBMCS), EHD723J (R GBMCS), EHD727J (R GBMCS), EHD731J (R BMC S), EHD733J (R GBMCS), EHD739J (R GBMCS), KOD742J (R GBMS), KOD732J (R GBMCS), KWD872J (R GBMCS)	TRD050J × ITA7	basal-BR × basal-BR	$5.46 \times 10^{62} - 1.26 \times 10^{13}$
61. VPg-2 (nt5978; n=1)	GFD462J (R GBMCS)	HOD541J × NNT451J	Asian-BR × basal-BR	1.83×10^{188}
62. VPg-3 (nt5996-6027; n=2)	AD915J (R GBMCS), IWD032J (R GBMCS)	MYD015J × YST441J	Asian-BR × world-B3	$2.44 \times 10^{139} - 3.03 \times 10^{124}$
63. VPg-4 (nt6011-6030; n=2)	KGD54J (R GBMCS), ND10J (R GBMCS)	N216 × DMJ	world-B3 × Asian-BR	$1.90 \times 10^{70} - 3.73 \times 10^{67}$
64. VPg-5 (nt6032-6055; n=6)	FD997J (R GBMCS), FID430J (R GBMCS), ISD091J (R GMCS), ISD092J (R GMCS), KGD947J (R GBMCS), ND257J (R GBMCS)	DEU4 × YSD442J	basal-BR × basal-BR	$8.24 \times 10^{54} - 4.93 \times 10^{53}$
65. VPg-6 (nt6050; n=1)	TRD811J (R BMC S)	SMD212J × AD886J	basal-BR × basal-BR	1.37×10^{101}
66. VPg-7 (nt6111-6422; n=43)	AD877J (R GBMCS), AD878J (R GBMCS), AD880J (R GBMCS), CS820J (R GBMCS), FID96J (R GBMCS), FID565J (R GBMCS), GMD791J (R GBMCS), HGD341J (R GBMCS), HGD347J (R GBMCS), IBD223J (R GBMCS), ISD420J (R GBMS), KND853J (R GBMCS), KND854J (R GBMCS), KND855J (R GBMCS), KND857J (R GBMCS), KND865J (R GBMCS), KWB779J (R GBMCS), KWD776J (R GBMCS), KWD777J (R GBMCS), KWD781J (R GBMCS), NID078J (R GBMCS), NID569J (R GBMCS), NND453J (R GBMCS), NND457J (R GBMCS), NND459J (R GBMCS), NND574J (R GBMCS), NND575J (R GBMCS), NND900J (R GBMCS), NND904J (R GBMCS), NND905J (R GBMCS), NRD353J (R GBMCS), OKB639J (R GBMCS), OKD633J (R GBMCS), OSD189J (R GBMS), SHD470J (R GBMS), SHD473J (R GBMCS), SHD477J (R GBMCS), SHD585J (R GBMCS), SHH472J (R GBMCS), STD860J (R GBMCS), TKD767J (R GBMCS), TYD085J (R GBMS), TYD300J (R GBMCS)	SMD213J × OKD637J	basal-BR × basal-BR	$1.66 \times 10^{31} - 5.45 \times 10^{26}$
67. VPg-8 (nt6110-6121; n=23)	HOB537J (R GBMCS), HOD517J (R GBMCS), HOD525J (R GBMCS), HOD527J (R GBMCS), HOD528J (R GBMCS), HOD529J (R GBMCS), HOD530J (R GBMCS), HOD531J (R GBMCS), HOD534J (R GBMCS), HOD535J (R GBMCS), HOD536J (R GBMCS), HOD540J (R GBMCS), HOD541J (R GBMCS), HOD544J (R GBMCS), HOD545J (R GBMCS), HOD547J (R GBMCS), HOD551J (R GBMCS), HOD553J (R GBMCS), IWD027J (R GBMCS), IWD033J (R GBMCS), OKD608J (R GBMCS), OKD609J (R GBMCS), YMD069J (R GBMCS)	AKT939J × ISD089J	Asian-BR × world-B3	$1.66 \times 10^{31} - 5.45 \times 10^{26}$
68. VPg-9 (nt6055-6234; n=14)	HGD076J (R GBMCS), HRD805J (R GBMCS), IBD222J (R GBMCS), IBD224J (R GBMCS), IBD226J (R GBMCS), SMD058J (R GBMCS), SMD061J (R GBMCS), SMD064J (R GBMCS), SMD211J (R GBMCS), SMD212J (R GBMCS), SMD214J (R GBMCS), TRB203J (R GBMCS), TRD050J (R GBMCS), TRD204J (R GBMCS)	KGB394J × NND908J	basal-BR × basal-BR	$1.77 \times 10^{42} - 4.36 \times 10^{29}$
69. VPg-10 (nt6272; n=1)	EHD721J (R GBMCS)	NDJ × EHD717J	Asian-BR × basal-BR	3.12×10^{171}
70. VPg-11 (nt6256-6346; n=31)	1J (R GBMCS), 2J (R BMC), AD178J (R GBMCS), AD181J (R GBMCS), AD853J (R GBMCS), AD860J (R GBMCS), AD871J (R BMC S), AD873J (R BMC S), AD874J (R BMC S), AT181J (R GBMCS), AKH937J (R B S), DMJ (R BMS), EHD709J (R GBMS), FKD001J (R BMC), FKD004J (R BMC), FKH122J (R BMC), H1J (R GMCS), HOD500J (R GBMCS), HOD5026J (R GBMCS), HOD539J (R GBMS), HOD543J (R GBMS), IWD038J (R BMS), Ka1J (R BCS), KWB778J (R GBCS), KWB780J (R GBCS), MYD013J (R BMC), NID048J (R GBMCS), Tu-2R1 (R BMC S), Tu-3 (R GBMCS), YAD016J (R GMCS), YAL018J (R GBCS)	FD810J × Q-Ca	world-B3 × world-B3	$6.69 \times 10^{15} - 1.28 \times 10^{11}$

Table 8. Continued.

Protein-coding region and recombination events	Isolate (recombination detecting program detected)	Parental isolate	Type of recombination	<i>P</i> -value
71. VPg-12 (nt6310-6333; n=10)	AD851J (R B MCS), C42J (R B MCS), HGC346J (R B MCS), NID049J (R B MCS), NRD350J (R B CS), SMB063J (R B MCS), SMD065J (R B MCS), YAD022J (R B MCS), YAD925J (R B MCS), YAD926J (R B MCS)	GBR36 × CDN1	world-B3 × world-B2	$3.33 \times 10^{-21} - 4.16 \times 10^{-19}$
Nla-Pro (nt6479-7207)				
72. Nla-Pro-1 (nt6422; n=1)	HGD348J (R B MCS)	NDJ × KND865J	Asian-BR × basal-BR	6.79×10^{-169}
73. Nla-Pro-2 (nt6423-6595; n=40)	AD877J (R B BS), AD878J (R B MCS), AD880J (R B MCS), CS820J (R B MCS), FID96J (R B MCS), FID565J (R B MCS), GMD791J (R B MCS), HGD341J (R B MCS), IBD223J (R B MCS), ISD420J (R B MCS), KND853J (R B MCS), KND854J (R B MCS), KND855J (R B MCS), KND857J (R B MCS), KND865J (R B MCS), KWB779J (R B MCS), KWD776J (R B MCS), KWD777J (R B MCS), KWD781J (R B MCS), NID078J (R B MCS), NID569J (R B MCS), NND453J (R B MCS), NND457J (R B MCS), NND459J (R B MCS), NND575J (R B MCS), NND900J (R B MCS), NND904J (R B MCS), NND905J (R B MCS), NRD353J (R B MCS), OKB639J (R B MCS), OKD633J (R B MCS), OSD189J (R B MS), SHD470J (R B MCS), SHD473J (R B MCS), SHD477J (R B MCS), SHD585J (R B MCS), SHH471J (R B MCS), STD860J (R B MCS), TYD300J (R B MCS), TKD767J (R B MCS)	KWD872J × EHD731J	basal-BR × basal-BR	$2.85 \times 10^{-29} - 5.46 \times 10^{-21}$
74. Nla-Pro-3 (nt6527; n=1)	NDJ (R B MS)	IWD037J × AD911J	Asian-BR × world-B3	1.55×10^{-174}
75. Nla-Pro-4 (nt6583-6613; n=4)	NID117J (R B MCS), NID119J (R B MCS), ISD093J (R B MCS), ISD440J (R B MCS)	ND276 × IBD226J	basal-BR × basal-BR	$1.26 \times 10^{-52} - 4.06 \times 10^{-16}$
76. Nla-Pro-5 (nt6791; n=1)	FKH122J (R B MC)	Q-Ca × FD810J	world-B3 × world-B3	6.85×10^{-10}
77. Nla-Pro-6 (nt6921; n=3)	SE251J (R B MCS), SE252J (R B MCS), SHD467J (R B MCS)	NDJ × NND908J	Asian-BR × basal-BR	$2.53 \times 10^{-168} - 3.90 \times 10^{-141}$
78. Nla-Pro-7 (nt6922; n=10)	AD851J (R B MCS), C42J (R B MCS), HGC346J (R B MCS), NID049J (R B MCS), NRD350J (R B CS), SMD063J (R B MCS), SMD065J (R B MCS), YAD022J (R B MCS), YAD095J (R B MCS), YAD926J (R B MCS)	CDN1 × GBR36	world-B2 × world-B3	$6.47 \times 10^{-26} - 5.12 \times 10^{-22}$
79. Nla-Pro-8 (nt6943; n=1)	KGD690J (R B MCS)	IRNWRSh40 × OKD604J	basal-BR × basal-BR	3.39×10^{-46}
Nlb (nt7208-8758)				
80. Nlb-1 (nt7382-7416; n=4)	AKH937J (R B S), KWB778J (R B MCS), KWB780J (R B MCS), YAL018J (R B S)	GBR36 × Q-Ca	world-B3 × world-B3	$9.92 \times 10^{-30} - 1.97 \times 10^{-17}$
81. Nlb-2 (nt7606-7712; n=40)	AD877J (R B BS), AD878J (R B MCS), AD880J (R B S), CS820J (R B MCS), FID96J (R B MCS), FID565J (R B MCS), GMD791J (R B MCS), HGD341J (R B MCS), IBD223J (R B MCS), ISD420J (R B MCS), KND853J (R B MCS), KND854J (R B MCS), KND855J (R B MCS), KND857J (R B MCS), KND865J (R B MCS), KWB779J (R B MCS), KWD776J (R B MCS), KWD777J (R B MCS), KWD781J (R B MCS), NID078J (R B MCS), NID569J (R B MCS), NND453J (R B MCS), NND457J (R B MCS), NND459J (R B MCS), NND575J (R B MCS), NND900J (R B MCS), NND904J (R B MCS), NND905J (R B MCS), NRD353J (R B MCS), OKB639J (R B MCS), OKD633J (R B MCS), OSD189J (R B MS), SHD470J (R B MCS), SHD473J (R B MCS), SHD477J (R B MCS), SHD585J (R B MCS), SHH472J (R B MCS), STD860J (R B MCS), TKD767J (R B MCS), TYD300J (R B MCS)	EHD731J × KWD872J	basal-BR × basal-BR	$1.65 \times 10^{-27} - 5.45 \times 10^{-26}$
82. Nlb-3 (nt7655-8132; n=3)	ISD093J (R B MCS), NID117J (R B MCS), NID119J (R B MCS)	ND276 × IBD226J	basal-BR × basal-BR	$1.26 \times 10^{-52} - 1.63 \times 10^{-16}$
83. Nlb-4 (nt8443; n=1)	KWD785J (R B MCS)	KGD851J × N216	Asian-BR × world-B3	1.75×10^{-49}
CP (nt8759-9622)				
84. CP-1 (nt8792-8802; n=5)	AD181J (R B MCS), AD871J (R B MCS), AD873J (R B MCS), H1J (R B MCS), YAD016J (R B MCS)	AD912J × CHN342	world-B3 × Asian-BR	$1.98 \times 10^{-44} - 6.65 \times 10^{-15}$
85. CP-2 (nt8953; n=1)	KGD690J (R B MCS)	OKD604J × IRNWRSh40	basal-BR × basal-BR	3.39×10^{-43}
86. CP-3 (nt9025; n=1)	OKD637J (R B MCS)	R007 × OB996BJ	basal-BR × world-B3	4.88×10^{-88}
87. CP-4 (nt9177-9234; n=44)	59J (R B BS), AD181J (R B MCS), AD871J (R B MCS), AD873J (R B MCS), AD874J (R B BS), FB998J (R B BS), FD21J (R B BS), FD27J (R B BS), IWD042J (R B BS), JPN1 (R B BS), JPN1ARA (R B BS), KD32J (R B BS), KD41J (R B BS), KGD54J (R B BS), KGD58J (R B BS), KGD661J (R B BS), KGD678J (R B BS), KGD701J (R B BS), KGD850J (R B BS), KGD954J (R B BS), KOD740J (R B BS), KWD775J (R B BS), MD45J (R B BS), MD966J (R B BS), MED302J (R B BS), MED303J (R B BS), MN978J (R B BS), MYD013J (R B BS), ND10J (R B BS), ND258J (R B BS), NID046J (R B BS), NID922J (R B BS), NID924J (R B BS), NTD253J (R B BS), OD14J (R B BS), OD15J (R B BS), OD985J (R B BS), OKD610J (R B BS), OKM62J (R B BS), SN4J (R B BS), STD8J (R B BS), TRD051J (R B BS), TRD053J (R B BS), TYD083J (R B BS), YMD070J (R B BS)	MYD012J × CHYK56A	Asian-BR × world-B3	$1.98 \times 10^{-44} - 2.13 \times 10^{-14}$
88. CP-5 (nt9427-9528; n=3)	NND457J (R B BS), NND459J (R B BS), NND574J (R B MCS)	TYD085J × ISD093J	basal-BR × basal-BR	$2.01 \times 10^{-21} - 7.70 \times 10^{-20}$

Table 9. Recombination type pattern of TuMV in Japan.

Recombination type pattern ^a	Phylogenetic groups and subgroups	Number of isolates	Proportion (%)
Intralineage recombinant (n=141)	basal-BR × basal-BR	29	8
	world-B3 × world-B3	9	2
	basal-BR × basal-BR × basal-BR	27	7
	world-B3 × world-B3 × world-B3	5	1
	basal-BR × basal-BR × basal-BR × basal-BR	52	14
	world-B3 × world-B3 × world-B3 × world-B3	6	2
	basal-BR × basal-BR × basal-BR × basal-BR × basal-BR	13	4
Interlineage recombinant (n=198)	Asian-BR × basal-BR	3	1
	Asian-BR × world-B3	2	1
	basal-BR × world-B3	1	0.3
	world-B2 × world-B3	9	2
	world-B3 × Asian-BR	3	1
	Asian-BR × world-B3 × Asian-BR	11	3
	Asian-BR × world-B3 × world-B3	1	0.3
	basal-BR × Asian-BR × basal-BR	3	1
	basal-BR × basal-BR × Asian-BR	2	1
	world-B2 × world-B3 × world-B3	1	1
	world-B3 × Asian-BR × world-B3	34	9
	world-B3 × world-B2 × world-B3	4	1
	Asian-BR × world-B3 × Asian-BR × world-B3	46	12
	basal-BR × basal-BR × Asian-BR × basal-BR	2	1
	world-B2 × world-B3 × world-B3 × world-B3	3	1
	world-B3 × Asian-BR × world-B3 × world-B3	9	2
	world-B3 × Asian-BR × world-B3 × Asian-BR	1	0.3
	world-B3 × world-B2 × world-B3 × world-B3	1	0.3
	Asian-BR × world-B3 × Asian-BR × world-B3 × Asian-BR	7	2
	world-B3 × Asian-BR × world-B3 × world-B3 × world-B3	1	0.3
	world-B2 × world-B3 × world-B3 × world-B3 × world-B3	3	1
	world-B3 × Asian-BR × world-B3 × Asian-BR × world-B3	24	6
	world-B3 × world-B3 × world-B3 × world-B2 × world-B3	8	2
	Asian-BR × world-B3 × Asian-BR × world-B3 × Asian-BR × basal-BR	4	1
	Asian-BR × world-B3 × Asian-BR × world-B3 × Asian-BR × world-B3	7	2
	world-B3 × Asian-BR × world-B3 × world-B3 × Asian-BR × world-B3	6	2
	world-B3 × world-B3 × world-B3 × world-B3 × world-B3 × world-B2 × world-B3	2	1
	Nonrecombinant (n=31)	Orchis	0
Iranian 1		0	0
Iranian 2		0	0
basal-B1		0	0
basal-B2		0	0
Asian-BR		0	0
basal-BR		14	4
world-B1		0	0
world-B2		0	0
world-B3		17	5
Total		370	100

^a Intralineaage recombinant; parents from same major group or subgroup lineages, interlineage recombinants; parents from different major group or subgroup lineages.

Table 10. Number of recombination sites in the genomes of TuMV isolates.

Number of recombination sites	Number of isolates	Proportion (%)
Non	31	8
Single	56	15
Double	88	24
Triple	120	32
Fourth	56	15
Fifth	17	5
Sixth	2	1
Total	370	100

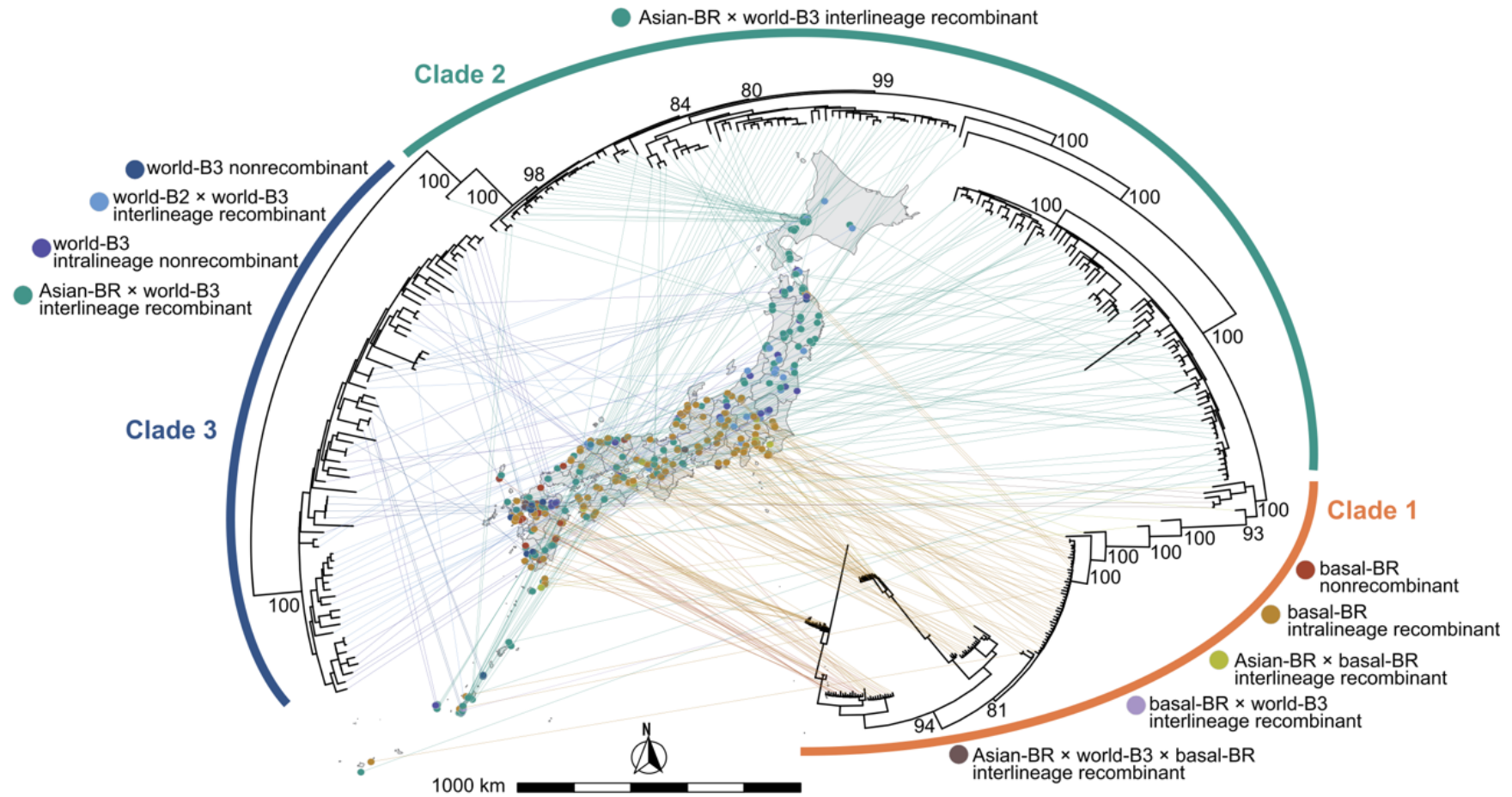


Fig. 29. Sampling map of TuMV Japanese isolates and maximum-likelihood tree. The aligned polyprotein-coding sequences are 9,486 nucleotides in length. Tree tips are connected to the sampling locations of the isolates. Outer lines show three major clades of Clade 1, 2, and 3. Colors of points on the map and connecting lines indicate recombination-type patterns, as shown in the key. Numbers at internal nodes indicate bootstrap percentages, based on 1,000 pseudoreplicates.

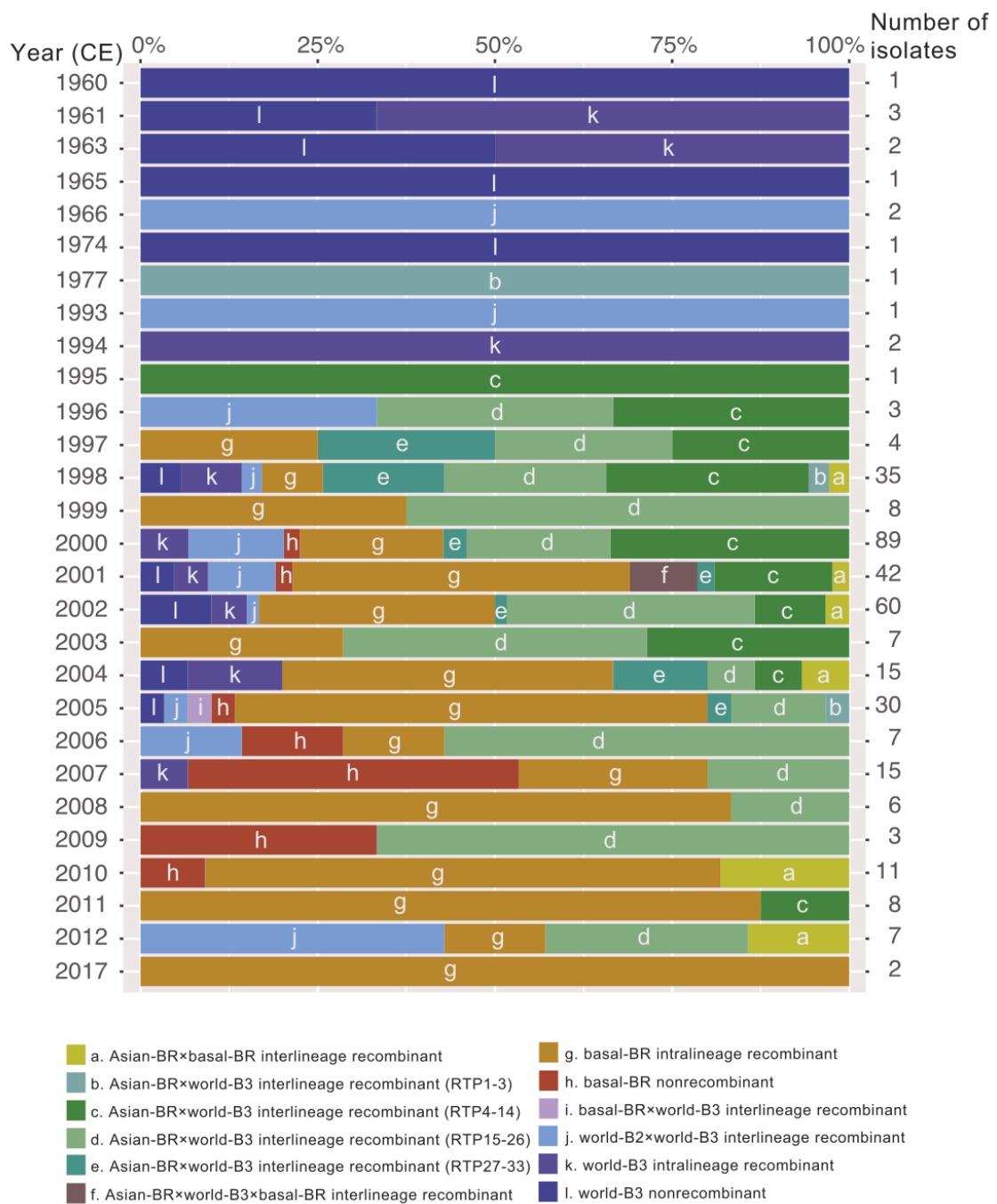


Fig. 30. Sampling overview of Japanese TuMV isolates by year. (B) Total numbers of TuMV nonrecombinant isolates and different recombination-type pattern (RTP) isolates in each sampling year. The classification of RTP is summarized in Fig. 26. The Asian-BR × world-B3 interlineage recombinants are classified 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.

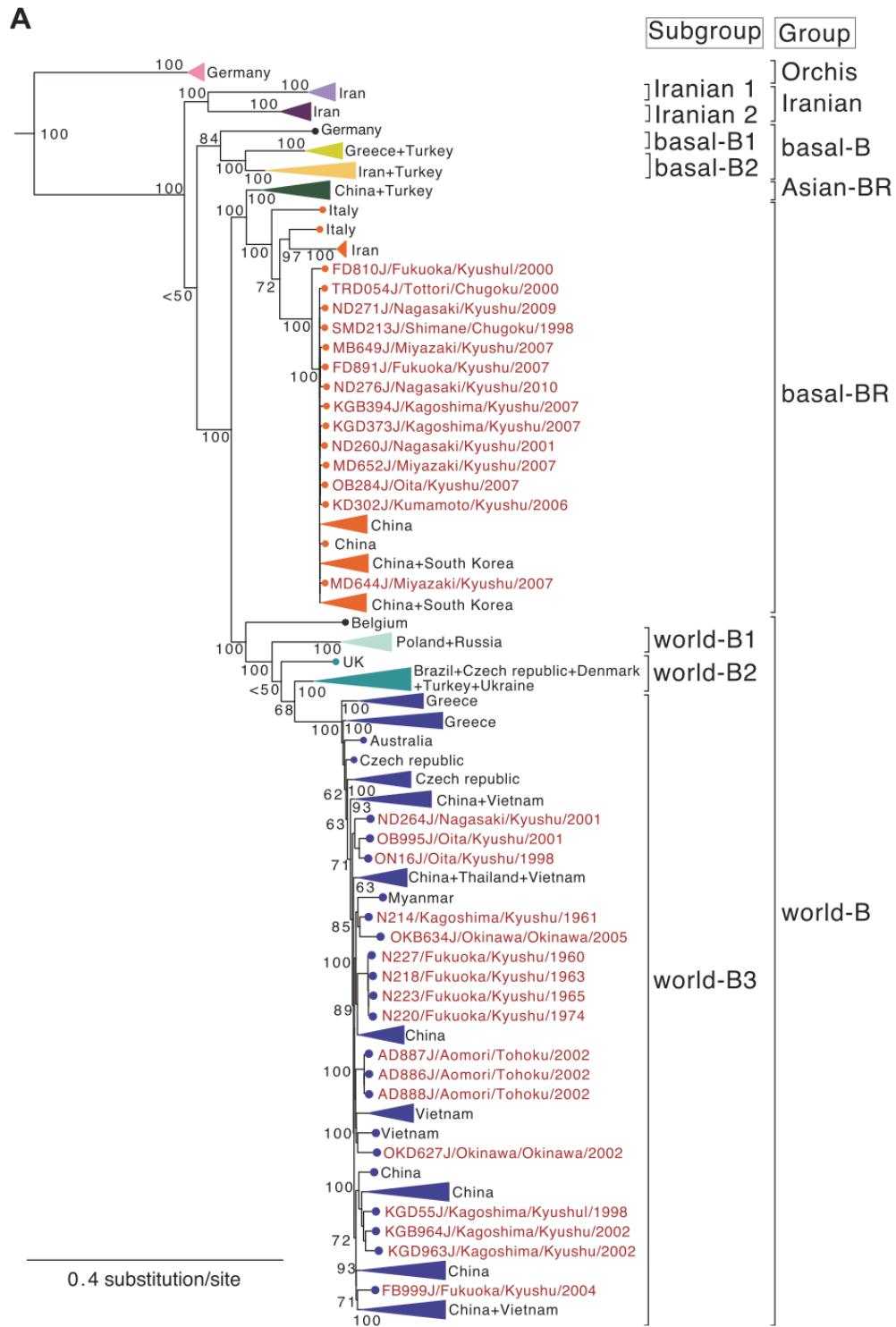


Fig. 31. Phylogenetic tree of TuMV inferred by maximum-likelihood from polyprotein-coding sequences of nonrecombinant isolates. Tree of 184 nonrecombinant isolates including thirty-one Japanese nonrecombinant isolates newly found in this study. Numbers at internal nodes indicate bootstrap percentages, 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.

(RTP81)は、東アジア、ヨーロッパ、および米国で報告されている(Fig. 26)。また、日本においては basal-BR の非組換え体および系統内組換え体は西日本に局在して分布している一方で、その他の組換え体型は日本全土に幅広く分布していることが明らかになった(Fig. 29)。また、各分離株が採集された年代を、その組換え体型ごとにプロットし可視化した(Fig. 30)。1960年から1994年までに採集された分離株は、1分離株を除いて全て world-B グループの非組換え体もしくは系統内組換え体であった。Asian-BR は、world-B との系統間組換え体として1977年に初めて採集されたが、1995年以降から急激に増加して、以前の world-B の非組換え体と系統内組換え体の集団を置き換えた。basal-BR の組換え体は1997年に初めて採集され、その後2000年代に入ると急激に増加し、以前の Asian-BR と world-B の系統間組換え体の集団を置き換えた。このように、TuMV の日本集団は、年代とともに組換え体型が変化していることが明らかになった。

3. 分子系統解析

組換え解析で同定した非組換え体については、ポリタンパク質の塩基配列を用いて最尤法による系統解析を行なった (Fig. 31)。日本における TuMV の非組換え体は、basal-BR と world-B3 グループのみであり、Asian-BR の非組換え体は分布していなかった。basal-BR グループの日本産分離株は、中国と韓国産分離株と同じ系統群内に位置した。一方で、world-B3 グループの日本産分離株は、7つの系統群に分かれ、中国産分離株と同じ系統群に位置した。ほとんどの日本産分離株の非組換え体は、九州地方に分布していた。

4. 時空間的拡散経路解析

組換え部位が比較的少なく、解析に用いることのできる分離株数の多かった、HC-Pro、P3、NIb、CP の4遺伝子領域の非組換え配列を用いて、日本における TuMV の拡散経路を離散空間および連続空間内で解析した(Figs. 32-33, Table 11)。Asian-BR グループでは、HC-Pro 遺伝子を用いた離散空間における解析の場合、東北から北海道、中部、近畿、中国、四国、九州、沖縄への7つの拡散経路が推

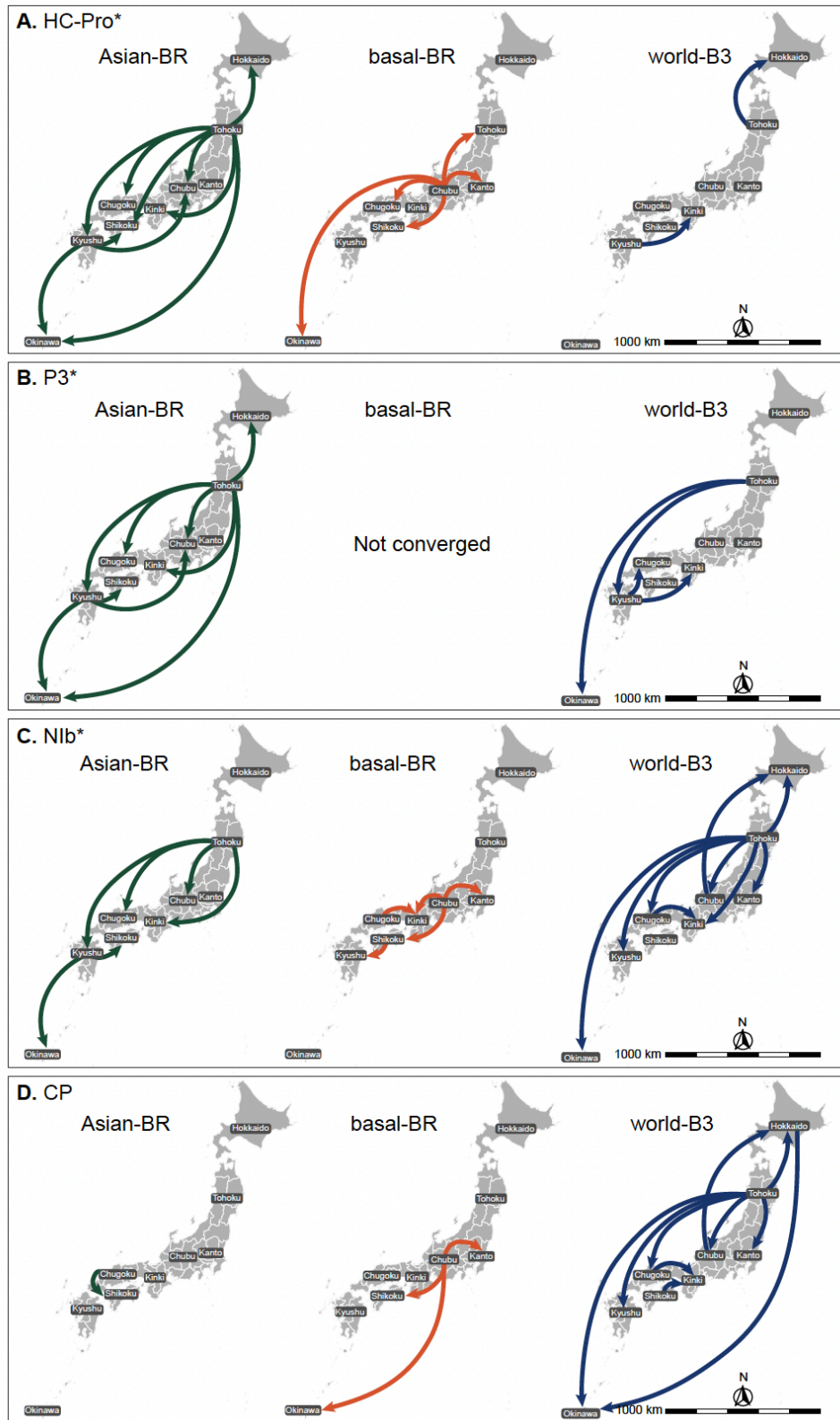


Fig. 32. Migration pathways of TuMV phylogenetic groups in Japan, inferred using recombination-free protein-coding regions. (A) Partial HC-Pro*, (B) Partial P3*, (C) Partial NIb* (D) Complete CP. The analysis of basal-BR using P3* coding regions is not shown here because the Bayesian Markov chain Monte Carlo analysis failed to achieve convergence. 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.

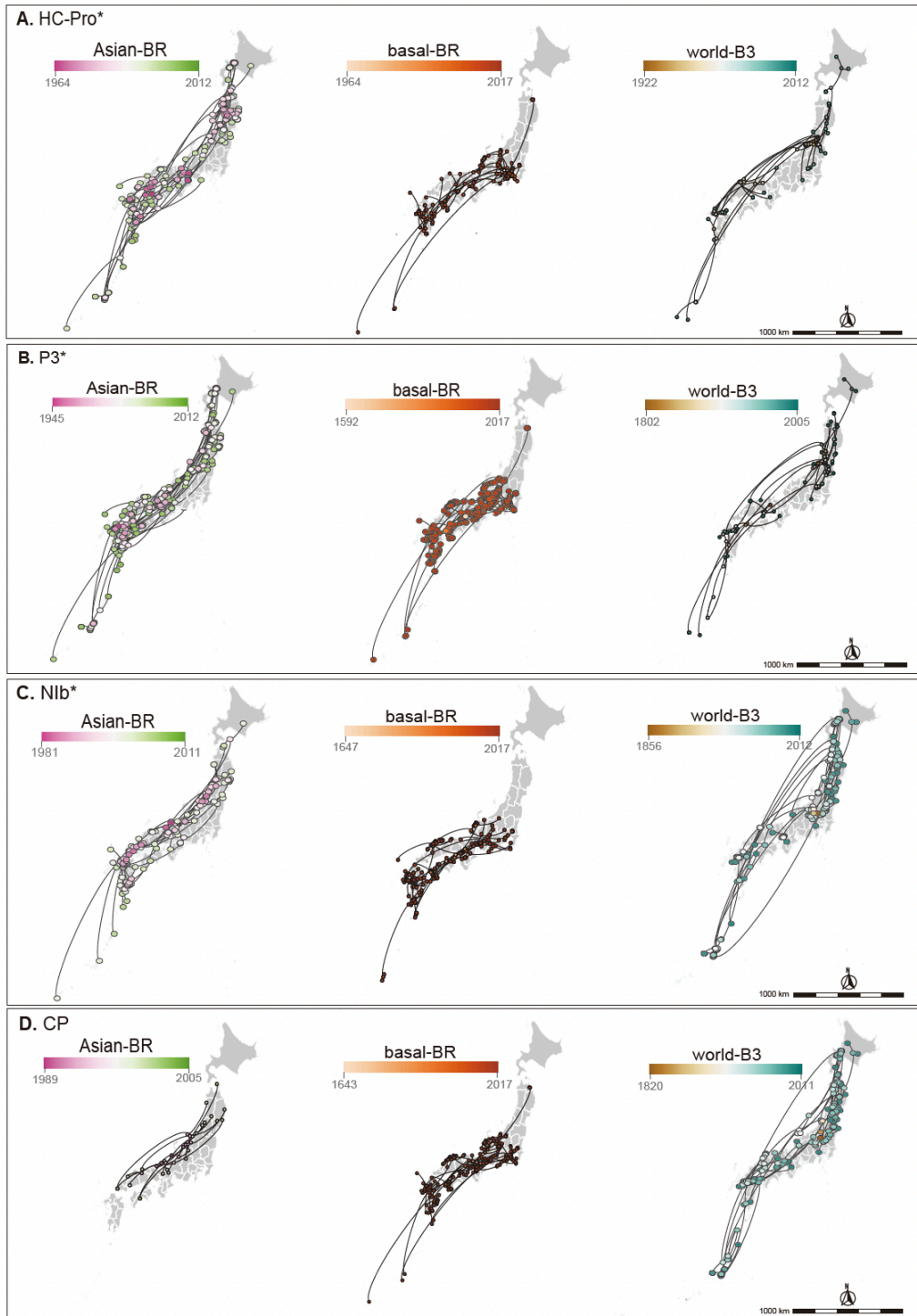


Fig. 33. Spatial diffusion of TuMV in Japan analysed in continuous space. (A) Partial HC -Pro* (B) Partial P3*(C) Partial Nib (D) Complete CP region. We annotated each isolate with its geographic coordinates (i.e., latitude and longitude), then inferred the spatial diffusion of each phylogenetic group. Cauchy, gamma, and lognormal relaxed-random-walk models were compared using marginal likelihoods estimated by path sampling and stepping-stone sampling (Table 7) (Baele et al., 2012). The reconstructed phylogeographic tree was projected onto a map using the seraphim package (Dellicour et al., 2016). Nodes are coloured according to their inferred time of occurrence.

Table 11. Estimated dispersal routes^a of TuMV phylogenetic groups in Japan.

Phylogenetic group and subgroup	Protein-coding region ^b																			
	Helper-component proteinase* (HC-Pro*)					Protein 3* (P3*)					Nuclear inclusion b* (NIb*)					Coat protein (CP)				
	From	To	Bayes factor ^c	PP ^d	Dispersal time ^e	From	To	Bayes factor	PP	Dispersal time	From	To	Bayes factor	PP	Dispersal time	From	To	Bayes factor	PP	Dispersal time
Asian-BR	Tohoku ^f	Hokkaido	457	1.0	1983-1993	Kyushu	Okinawa	1147	1.0	1990-1994	Tohoku	Kinki	1050	1.0	1992-1995	Chugoku	Shikoku	74	1.0	1991-1996
	Kyushu	Okinawa	274	1.0	1990-1993	Kyushu	Shikoku	565	1.0	1970-1992	Tohoku	Chugoku	628	1.0	1984-1994					
	Tohoku	Chugoku	203	1.0	1982-1988	Tohoku	Hokkaido	301	1.0	1970-1988	Tohoku	Chubu	408	1.0	1987-1995					
	Tohoku	Kinki	144	1.0	1979-1990	Kyushu	Chubu	46	0.9	1984-1991	Kyushu	Shikoku	390	1.0	1991-1993					
	Tohoku	Kyushu	118	1.0	1976-1989	Tohoku	Chubu	17	0.7	1972-1988	Tohoku	Kyushu	259	1.0	1985-1991					
	Kyusyu	Shikoku	110	1.0	1982-2000	Tohoku	Kinki	16	0.7	1969-1988	Kyushu	Okinawa	247	1.0	1991-1994					
	Kyushu	Chubu	67	0.9	1984-1993	Tohoku	Chugoku	13	0.7	1970-1985										
	Tohoku	Chubu	49	0.9	1979-1993	Tohoku	Kyushu	12	0.7	1955-1985										
	Tohoku	Shikoku	37	0.9	1983-1991	Tohoku	Okinawa	7	0.5	1966-1977										
	Tohoku	Okinawa	27	0.8	1972-1987															
basal-BR	Chubu	Shikoku	7058	1.0	1987-1995	Not converged					Chugoku	Kinki	394	1.0	1886-2000	Chubu	Kanto	2685	1.0	1933-1980
	Chubu	Okinawa	1147	1.0	1997-1999						Chubu	Shikoku	179	1.0	1872-1998	Chubu	Okinawa	728	1.0	2002-2002
	Chubu	Tohoku	64	1.0	1989-1996						Chubu	Kanto	106	1.0	1876-1995	Chubu	Shikoku	177	1.0	2002-2002
	Chubu	Kanto	56	1.0	1972-1985						Chubu	Kinki	32	0.9	1871-2008					
	Chubu	Chugoku	7	0.5	1991-1998						Shikoku	Kyushu	9	0.6	1805-1992					
world-B3	Tohoku	Hokkaido	588	1.0	1958-1979	Tohoku	Kyushu	28	0.8	1892-1955	Tohoku	Kyushu	28253	1.0	1924-1948	Tohoku	Kanto	14741	1.0	1939-1996
	Kyushu	Kinki	7	0.5	1969-1982	Kyushu	Kinki	26	0.8	1853-1933	Tohoku	Kanto	11298	1.0	1937-2004	Tohoku	Kyushu	368	1.0	1915-1944
						Kyushu	Chugoku	10	0.6	1931-1969	Tohoku	Okinawa	6274	1.0	1939-1980	Tohoku	Okinawa	313	1.0	1933-1978
						Tohoku	Okinawa	6	0.5	1866-1928	Tohoku	Hokkaido	5132	1.0	1948-1979	Tohoku	Hokkaido	270	1.0	1920-1971
											Tohoku	Chugoku	1706	1.0	1931-1965	Chugoku	Kinki	42	0.9	1917-1956
											Tohoku	Chubu	877	1.0	1946-1974	Shikoku	Kinki	33	0.8	1957-1988
											Chubu	Hokkaido	40	0.9	1947-2001	Chubu	Hokkaido	30	0.8	1939-1993
											Chugoku	Kinki	8	0.6	1948-2006	Tohoku	Chugoku	30	0.8	1925-1965
											Tohoku	Kinki	7	0.5	1914-1950	Tohoku	Chubu	23	0.8	1918-1961
																Hokkaido	Okinawa	11	0.6	1961-1987

^a Dispersal routes supported by three, two, and one gene region(s) are shown in yellow, green, and blue, respectively.

^b HC-Pro*, P3* and NIb* indicate partial protein-coding regions. CP (coat protein) indicates complete protein-coding region of CP.

^c Dispersal routes with Bayes factor ≥ 3 and posterior probability ≥ 0.5 were considered as supported.

^d PP; Posterior probability.

^e Dispersal time ranges are given as year Common Era (CE).

^f Classification scheme of geographic areas appears in Fig. 21.

定された。東北から他の地域への拡散経路は、P3 と N1b 遺伝子領域を用いた解析でも同様に支持された。従って、Asian-BR グループの拡散は東北地方がハブとなっていることが示唆された。basal-BR グループにおいては、中部地方から関東と四国への拡散が HC-Pro、N1b、CP の 3 遺伝子領域のデータセットで支持された。world-B3 グループにおいては、HC-Pro と P3 遺伝子領域の結果と、N1b と CP 遺伝子領域の結果で異なる傾向が見られた。N1b と CP 遺伝子領域を用いた場合、東北地方から北海道、関東、中部、中国、九州、沖縄への計 6 経路が指示され東北地方が拡散のハブであるように思えた一方で、HC-Pro と P3 遺伝子領域を用いた場合、東北地方からの拡散経路はほとんど推定されず、九州から近畿地方への拡散経路が支持された。

また、連続空間における拡散経路の解析も行なった(Fig. 33)。HC-Pro 遺伝子領域を用いた解析では、Asian-BR グループの分離株の日本への侵入は 1900 年代中期頃と推定された。侵入後は、東北、近畿、九州地方を 2000 年代にかけて循環している経路が支持された。basal-BR グループの分離株は西日本に限定した分布であるが、2000 年代急速に拡散した経路が推定された。world-B3 グループの分離株は、1800 年代後期から 1900 年代初期にかけて日本へ侵入してきたと推定された。侵入後は、東北、中国、九州地方を中心に循環している経路が支持された。

5. 組換え体の発生年代および拡散速度

前項で記述した解析のように、特定の遺伝子領域のみを用いた時に組換え部位を跨いだ遺伝子領域ごとで解析結果が異なる場合、ウイルス個体としての拡散を検討することが困難である(Martin et al., 2011)。そこで、82 パターン認められた組換え体型の中から系統地理学的解析が可能な 15 パターンを選抜し、それらのポリタンパク質遺伝子の配列を用いてそれらの発生年代と拡散経路および拡散速度を推定した(Figs. 34-35, Tables 12-13)。選抜した中で発生年代が最も古い組換え体型は、1963(95% ベイズ信用区間: 1952-1973) 年に東北地方から拡散した world-B2 × world-B3 グループ間組換え体型と思われ、平均速度 18 km/年

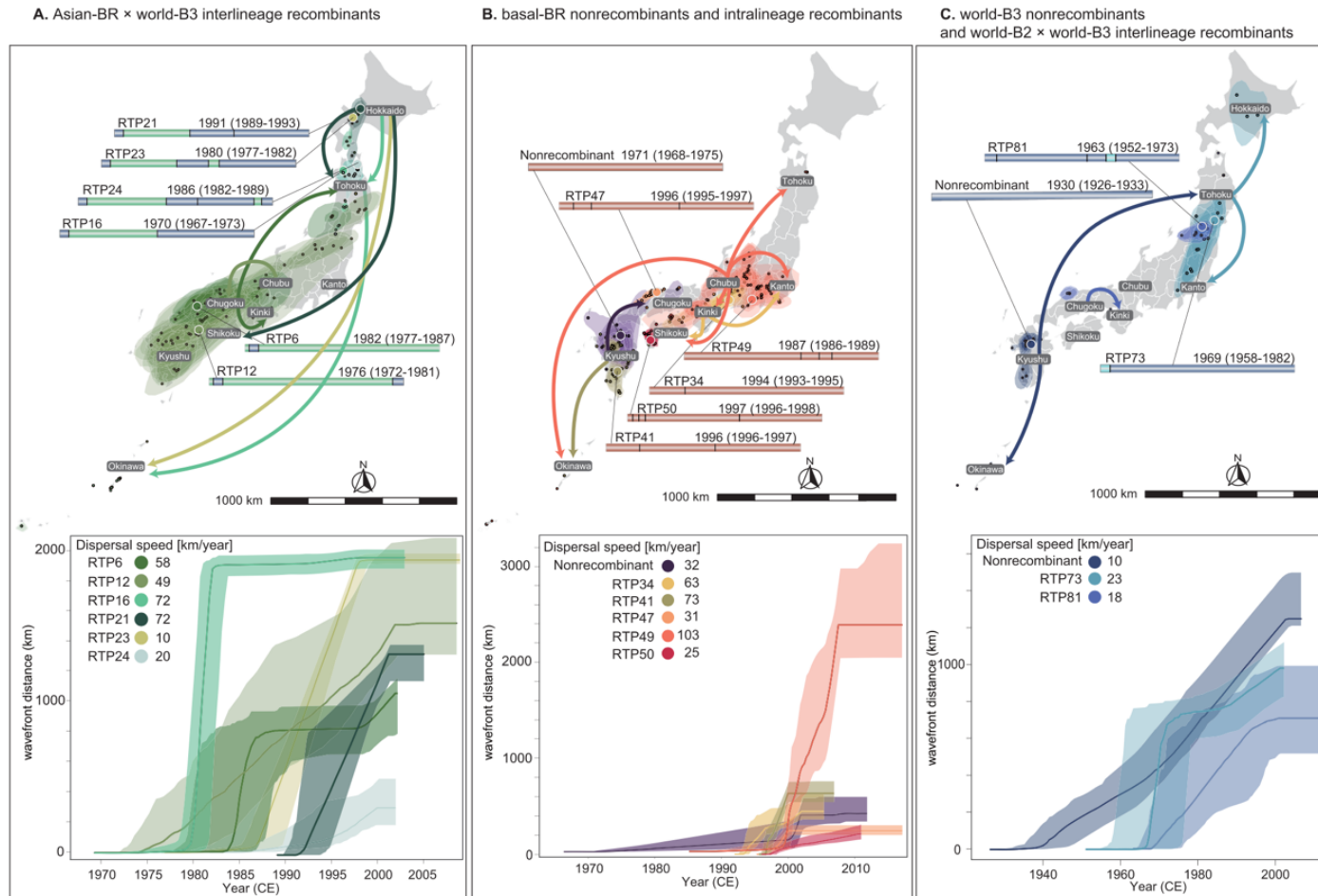


Fig. 34. Dispersal history of nonrecombinants and recombinants of TuMV. (A) Asian-BR × world-B3 interlineage recombinants; (B) basal-BR nonrecombinants and intralineage recombinants; and (C) world-B3 nonrecombinants and world-B2 × world-B3 interlineage recombinants. Upper panels show the year of the emergence and inferred dispersal routes. Inferred year of the emergence and 95 % 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 % 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 Fig. 26. 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 % credible intervals, respectively. The dispersal speed of each recombination-type pattern in every ten years is summarized in Table 13.

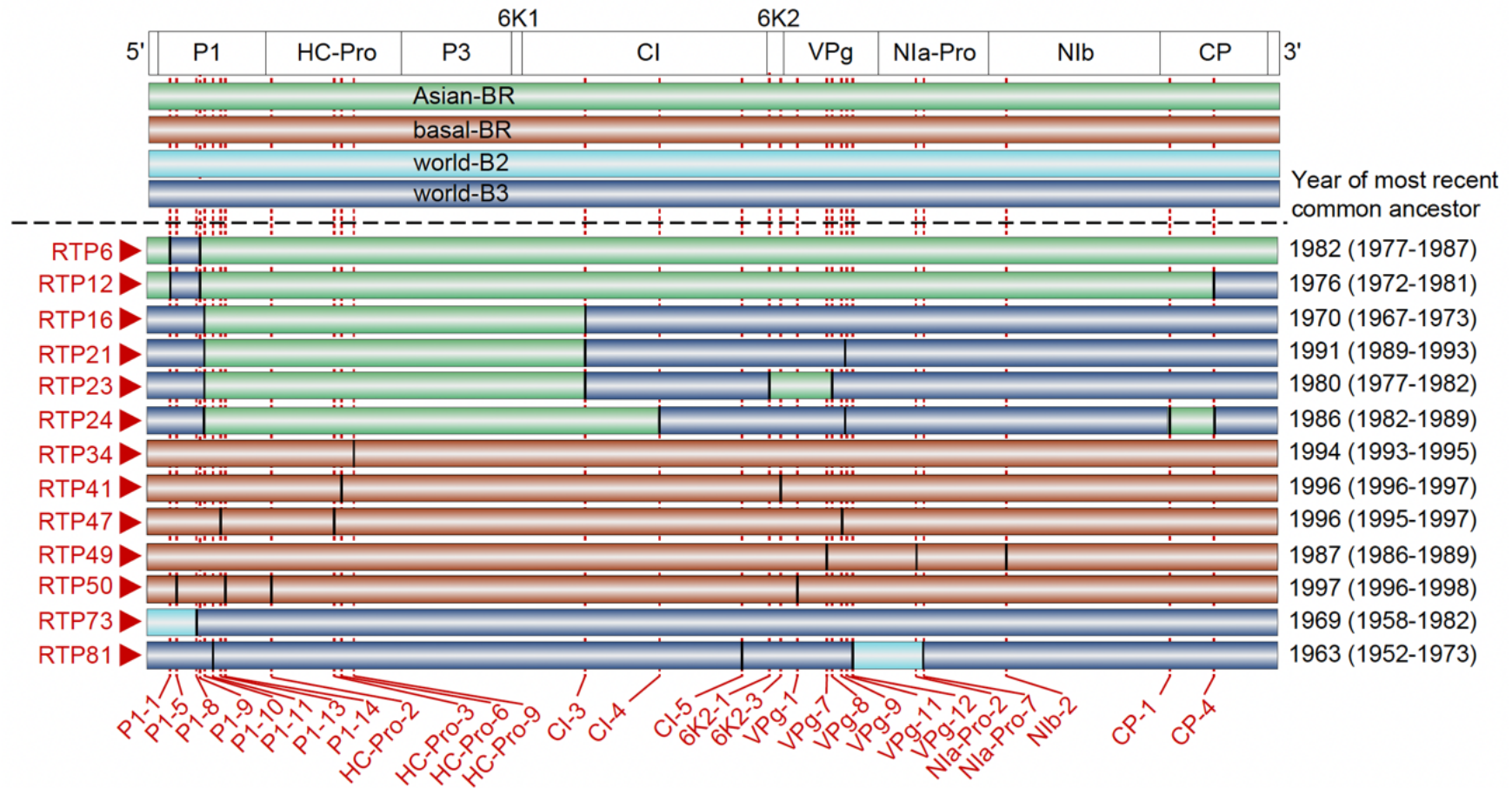


Fig. 35. Major recombination events found in Japanese TuMV and inferred year of most recent common ancestor for each RTP. The details of RTP are shown in Fig. 26 and Table 8.

Table 12. Estimated dispersal routes of TuMV by each recombination type pattern^a in Japan.

Phylogenetic group or subgroup	Recombination type pattern (RTP)		From	To	Bayes factor ^b	Posterior probability	Dispersal time ^c
	Name	Graphical representation					
Asian-BR × world-B3	RTP6		Chugoku ^d	Tohoku	6	0.8	1979 – 1988
			Chugoku	Kinki	4	0.8	1979 – 1988
	RTP12		Chubu	Chugoku	6	0.5	1978 – 1991
	RTP16		Tohoku	Okinawa	54	1.0	1976 – 1980
			Hokkaido	Tohoku	15	0.9	1988 – 1991
	RTP21		Hokkaido	Tohoku	3	0.7	1991 – 1995
			Hokkaido	Shikoku	3	0.7	1991 – 1995
	RTP23		Hokkaido	Okinawa	4	0.7	1984 – 1988
	RTP24				Not applicable ^e		
basal-BR	Nonrecombinant		Kyushu	Chugoku	1631	1.0	1998 – 1999
basal-BR intralineage	RTP34		Kinki	Kanto	8	0.7	1994 – 1995
			Kinki	Shikoku	6	0.6	1994 – 1996
	RTP41		Kyushu	Okinawa	1631	1.0	1996 – 1997
	RTP47				Not applicable		
	RTP49		Chubu	Kinki	749	1.0	1997 – 1998
			Chubu	Tohoku	658	1.0	1997 – 1998
			Chubu	Kanto	263	1.0	1998 – 2001
			Chubu	Shikoku	24	0.9	1997 – 1998
		Chubu	Okinawa	13	0.7	1998 – 2000	
	RTP50				Not applicable		
world-B3	Nonrecombinant		Kyushu	Tohoku	34	1.0	1950 – 1965
			Kyushu	Okinawa	9	0.9	1934 – 1948
world-B2 × world-B3	RTP73		Tohoku	Kanto	3	0.7	1968 – 1981
			Tohoku	Hokkaido	3	0.7	1967 – 1977
	RTP81		Chugoku	Kinki	9	0.8	1981 – 1990

^a Recombination type patterns that were found among more than six isolates in Japan were applied for this analysis.

^b Dispersal routes with a Bayes factor ≥ 3 and posterior probability ≥ 0.5 were considered as supported.

^c Dispersal time ranges are given as year Common Era (CE).

^d Classification scheme of geographic areas appears in Fig. 21.

^e All of the isolates were each sampled in a single coded location.

Table 13. Mean branch dispersal speed of nonrecombinants and recombinants TuMV.

Phylogenetic groups and subgroups	Recombination type pattern (RTP)	Time slices ^a						
		1940-1950	1950-1960	1960-1970	1970-1980	1980-1990	1990-2000	2000-2010
Asian-BR×world-B3 interlineage recombinant	RTP6	-	-	-	82 (8-269)	91 (6-319)	18 (9-36)	40 (27-57)
	RTP12	-	-	-	47 (11-135)	45 (13-112)	62 (16-171)	29 (12-49)
	RTP16	-	-	2 (1-6)	146 (71-279)	149 (2-503)	7 (1-19)	4 (3-7)
	RTP21	-	-	-	-	52 (9-181)	119 (37-280)	12 (5-27)
	RTP23	-	-	-	3 (1-5)	25 (9-64)	5 (1-15)	2 (1-5)
	RTP24	-	-	-	-	9 (4-22)	26 (8-67)	49 (12-137)
basal-BR nonrecombinant		-	-	-	-	-	38 (12-111)	15 (11-21)
basal-BR intralineaage recombinant	RTP34	-	-	-	-	-	80 (23-230)	20 (13-31)
	RTP41	-	-	-	-	-	127 (22-442)	18 (10-32)
	RTP47	-	-	-	-	-	38 (12-111)	15 (11-21)
	RTP49	-	-	-	-	-	164 (27-711)	70 (21-129)
	RTP50	-	-	-	-	-	33 (9-106)	12 (6-19)
world-B3 nonrecombinant		8 (3-15)	6 (1-20)	3 (0.4-8)	22 (1-47)	34 (21-48)	3 (1-10)	5 (3-10)
world-B2 × world-B3 interlineage recombinant	RTP73	-	-	29 (8-90)	19 (4-80)	16 (7-34)	14 (5-34)	7 (3-14)
	RTP81	-	26 (7-68)	38 (1-127)	31 (3-91)	7 (5-11)	5 (2-6)	16 (8-17)

^a Time slice range is given in Common Era (CE).

^b Dispersal speed is shown in km/year. The 95% credible intervals are in parentheses.

で北海道、関東地方へ拡散したと推定された。Asian-BR × world-B3 の系統間組換え体型の中で発生年代が最も古いと思われた組換え体型は、1970 (95 % ベイズ信用区間: 1967-1973) 年に東北地方で発生し、平均速度 72 km/年で北海道へ拡散したと推定された。また、basal-BR 系統内組換え体型の中で発生年代が最も古いと思われた組換え体型は、1994 (95 % ベイズ信用区間: 1993-1995) 年に近畿地方で発生後、平均速度 63 km/年で関東および四国地方へ拡散したと推定された。

考察

本研究は、植物ウイルスとして単一国内でのゲノム疫学調査を行なった最大規模の事例であり、最も解像度の高い網羅的な組換えおよび系統地理学的解析だと思われる。TuMVにおけるゲノムの組換えが初めて報告されたのは2002年であり(Ohshima et al., 2002)、その後オセアニア諸国(Yasaka et al., 2015)や中東(Yasaka et al., 2017)、ユーラシア大陸(Kawakubo et al., 2021)において網羅的な組換え解析が報告されきた。本研究で同定した全82パターンの組換え体型のうち、これまでに報告されてきた海外におけるTuMVの組換え体型と同じパターンは2つのみであった(Fig. 26, Table 8)。そして、日本におけるTuMVの主要な組換え体型の発生年代は1990年代から2000年代にかけてと比較的最近であることが示唆された(Figs. 34-35, Tables 12-13)。近隣諸国におけるウイルス分離株採集が不十分な可能性も否定できないが、これらのTuMV組換え体型は日本国内で発生している可能性が高いと思われた。すなわち、TuMVにおいてゲノムの組換えイベントは極めて頻繁に起きており、進化の大きな推進力となっていると思われる。

また、時空間的系統地理学解析により、TuMVへの日本への侵入は何度も独立に起きていることが示唆された(Figs. 32-33, Table 11)。今回推定された拡散経路は、world-Bグループが最も古く1800年代後半の拡散経路が複数支持された(Figs. 32-33, Table 11)。日本においては、1639年から1854年にかけてオランダなど一部の国を除いて他国からの輸入を禁止する鎖国制度を約200年間導入していたため、その間は植物病原体の侵入も制限されていたと考えられる。江戸時代後期の1854年に日米和親条約が締結されたことにより開国することとなったが、本研究で推定されたTuMVの日本への侵入時期も同時期である。開国に伴い、TuMVのような植物ウイルスは、農作物の他国からの輸入の際に感染植物として日本にもたらされたと考えられる。

さらに、拡散速度解析により、TuMVは地方間を1年当たり数十キロと拡散することが明らかとなった(Figs. 34, Table 12-13)。また、離散空間による解析によ

り、東北地方から沖縄への拡散など、離れた地域間を長距離移動する拡散が多数支持されたため(Fig. 32, Table 11)、日本国内の TuMV の拡散様式にも感染植物の人為的輸送が関与しているように思われた。

本章では、大規模なゲノム疫学調査を行うことで TuMV の進化がゲノムの組換えパターンによって特徴づけられることを示した。そして、その組換えパターンを網羅的に同定し、時空間的系統地理学解析を行うことで、TuMV の侵入時期と拡散様式を分子系統グループ、組換え体型のレベルで高精度に推定した。

第 V 章 アレキシウイルス属の種分岐過程の解析

目的

アルファフレキシウイルス科は植物ウイルスと菌類ウイルスからなる 7 つの属から構成され、そのうちの 1 つにアレキシウイルス属がある (Kreuze et al., 2020)。アレキシウイルス属は、Garlic virus A、B、C、D、E および X (GarV-A、GarV-B、GarV-C、GarV-D、GarV-E および GarV-X)、Shallot virus X (ShVX)、Garlic mite-borne filamentous virus の 8 種から構成され、これらは全てネギ属植物を宿主とするウイルスである。特に GarV-A、-B、-C、-D、E および-X は、ニンニク圃場において高頻度で感染が認められ、LYSV や OYDV と混合感染した状態で発見される (Sasaki et al., 2022; Yoshida et al., 2018)。アレキシウイルス属は他のウイルスに比べて宿主域が狭くネギ属植物に限定させているにも関わらず、8 種のウイルス種に種分化している。すなわち、異なる宿主植物への適応進化以外に重要な進化原動力が存在していることを示唆している。従って、アレキシウイルス属は植物ウイルスの種分岐過程 (speciation process) を理解する上での重要なモデルウイルスであり、その進化的戦略を明らかにすることは、ウイルス進化機構の解明において重要な知見を与えると考えられる。そこで、塩基配列の類似性解析や、ベイズ法による分子系統解析を行い、これらアレキシウイルス属の進化的解析を行なった。

材料と方法

1. データセット

GarV-A、GarV-B、GarV-C、GarV-D、GarV-E および GarV-X と ShVX の CP から CRP 遺伝子にかけての領域に相当するゲノム配列 (CP-CRP) を、公共データベース GenBank から取得した。BLAST 検索の結果、アレキシウイルス属と最も高い配列類似性を示したマンドレウイルス属の 2 種、Citrus yellow vein clearing virus (CYVCV) (MF563877) と Indian citrus ringspot virus (ICRV) (NC_003093) を外郡として用いた。マルチプルシーケンスアライメントは、muscle v3.8.31 (Edgar, 2004) を

用いて作成した。

2. 塩基配列類似性解析

MEGA v11 (Tamura et al., 2021)を用いて、CP 遺伝子領域の塩基配列ペアワイズ類似度を計算し、R v4.2.1 (R Core Team, 2022)を用いヒートマップとして可視化した。

3. 分子系統解析

分子系統樹の構築は、MrBayes v3.2.7(Ronquist & Huelsenbeck, 2003)を用いてベイズ法で行なった。また、分子系統解析には、1. CP-CRP にかけての領域、2. CP の翻訳領域、3. CRP の翻訳領域の 3 つの領域の塩基配列データを用いた。塩基置換モデルは、GTR+G+I を用いた。2. CP の翻訳領域、3. CRP の翻訳領域については、アミノ酸配列でも系統樹を構築した。アミノ酸置換モデルには、JTT モデルを用いた。MCMC の連鎖計算の長さには、3 千世代毎抽出の 3 百万回を用いた。連鎖計算の初めの 25%を burn-in として取り除いた。系統樹は、ggtree (Yu et al., 2017)を用いて可視化した。

結果

1. 塩基配列類似性解析

GarV-A、-B、-C、-D、-E の CP 遺伝子領域を用いて、塩基配列の種内類似度をヒートマップで可視化した(Fig. 36)。最も低いペアワイズ塩基類似度は、GarV-A、-B、-C、-D、-E のそれぞれにおいて、72.35%、87.61%、76.15%、77.95%、88.80%であった。GarV-A と GarV-C が他と比べて種内多様性が高かった。

2. 分子系統解析

CP-CRP 領域を用いた分子系統解析の結果、GarV-B、GarV-C、GarV-X および ShVX と GarV-A、GarV-E、GarV-D の大きく 2 系統群に分かれることが明らかになった(Fig. 37)。また、ShVX を除いて、これらの分岐は CP と CRP の翻訳

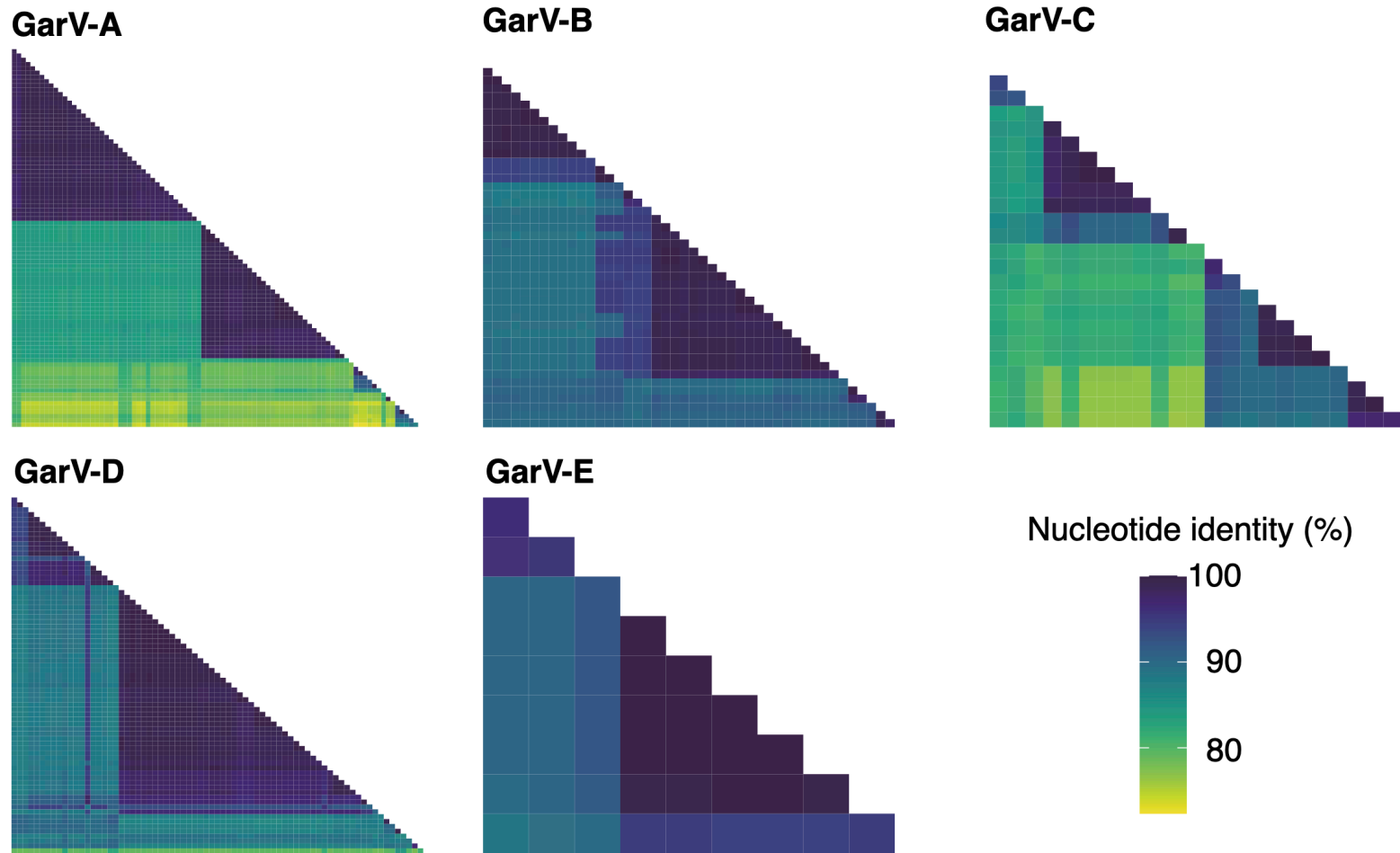


Fig. 36. Heatmap for sequence variability matrix in allelixiviruses. Garlic virus A, B, C, D, and E (GarV-A, GarV-B, GarV-C, GarV-D, and GarV-E). Each matrix represents nucleotide pairwise identity (GarV-A, n=89; GarV-B, n=45; GarV-C, n=24; GarV-D, n=74; GarV-E, n=10).

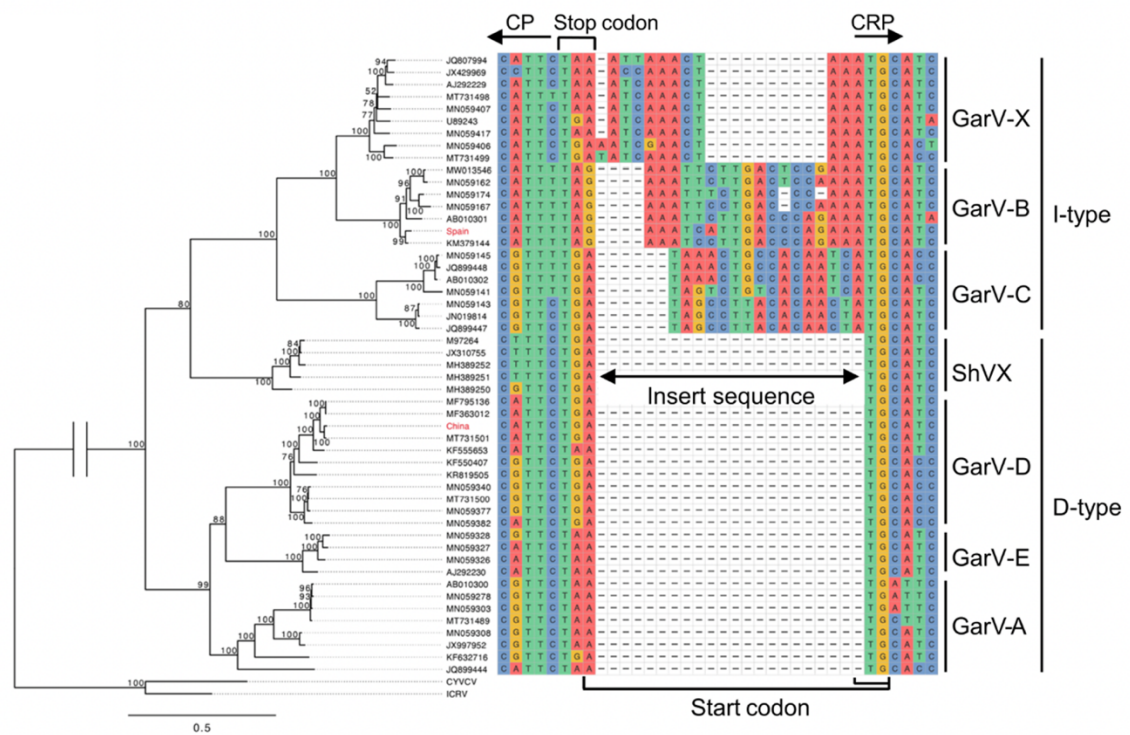


Fig. 37. Phylogenetic analyses of allelixiviruses based on the nucleotide sequence of the CP-CRP region. The phylogenetic tree (left side) was constructed using the Bayesian method as explained in the Result section. Internal node values indicate posterior probabilities (only $\geq 50\%$ are shown). CYVCV (MF563877) and ICRV (NC_003093) were used as the outgroups. Red text (Spain and China) indicates sequences obtained in this study; black text indicates GenBank accessions. The multiple sequence alignments around the insert sequences are identified on the right.

領域の間に位置する非翻訳領域の挿入がある I-type と欠失がある D-type に対応していた(Yoshida et al., 2018)。ShVX については、欠失を含む D-type であるが他の I-type と同時系統群に位置していた。CP と CRP の翻訳領域のみを用いた系統解析も行なった (Fig. 38)。CP 遺伝子の翻訳領域のみを用いた系統樹の場合、I-type である GarV-C が D-type と同じ系統群に位置し、CP-CRP の系統樹と異なる樹形を示した。CRP 遺伝子の翻訳領域のみを用いた系統樹の場合、D-type である ShVX が I-type と同じ系統群に位置し、CP-CRP の系統樹とほぼ同様の傾向を示した。

考察

本研究では、アレキシウイルス属のウイルス種を対象に種内類似性解析と種間分子系統解析を行なった。宿主域をネギ属植物に限定させているアレキシウイルス属は、異なる宿主植物への適応ではないその他の要因が原動力となって進化や種分岐が促進されていると考えられる。アレキシウイルス属は、CP と CRP の間の非翻訳領域に 10-20 塩基の挿入を持つ I-type とこの領域が欠失している D-type に分けられることが知られている(Yoshida et al., 2018)。CP-CRP 領域のベイズ系統解析の結果、D-type は系統樹の深い位置で共通祖先を有するのに対し、I-type は D-type から分岐した単系統群を構成した(Fig. 37)。すなわちこの領域は、もともと存在していたゲノムの一部を欠失させて生じたわけではなく、挿入領域を後から獲得することによって誕生した可能性が高いと考えられる。また、この挿入領域の有無が、RSS タンパク質である CRP の発現様式と細胞内局在を変化させ、宿主のウイルス感染に対する防御機構を相互排他的に抑制していることが明らかになっている(Kim et al., 2023)。短い塩基の挿入がウイルスの分子機構を変化させ、同属ウイルス内において大きく異なる進化方向性を生み出したことが明らかになった。

本章では、宿主域の狭いウイルスにおける種分岐過程を明らかにし、ウイルスがどのように進化的戦略を決めているのかの一例を示した。

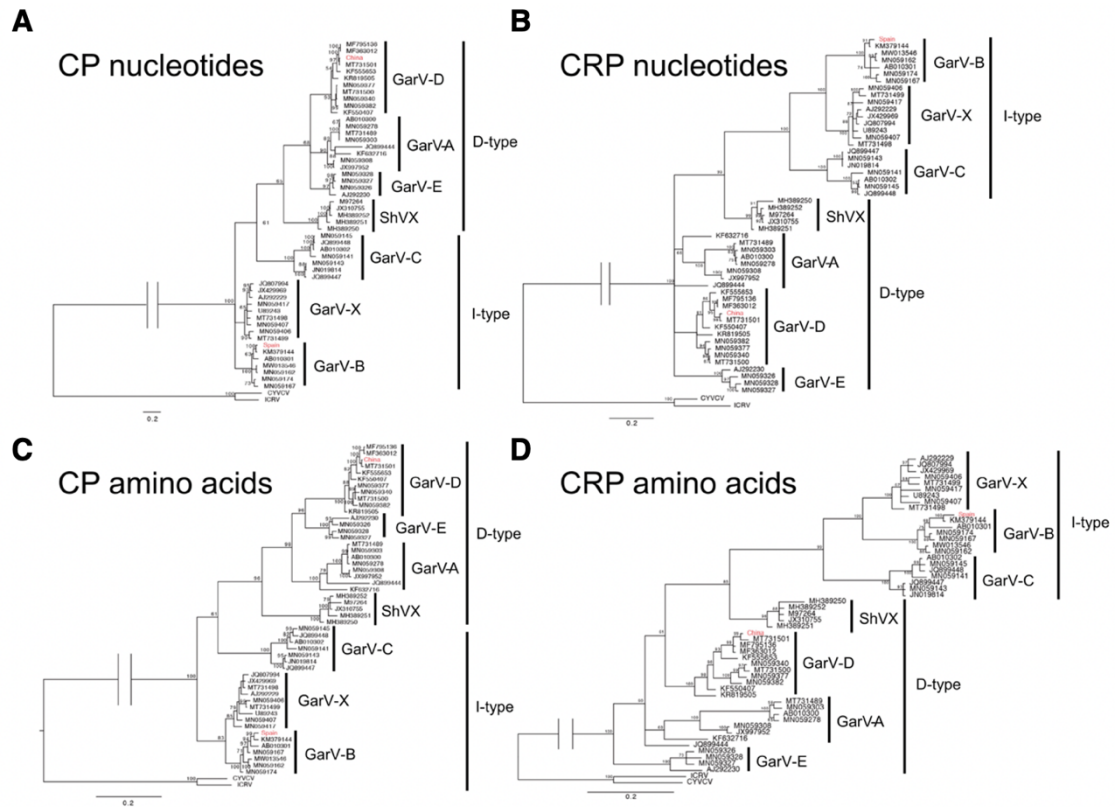


Fig. 38. Phylogenetic trees based on (A) CP nucleotide sequences, (B) CRP nucleotide sequences, (C) CP amino acid sequences, and (D) CRP amino acid sequences. The trees were constructed using the Bayesian method as explained in the Method section. Internal node values indicate posterior probability (only $\geq 50\%$ are shown). CYVCV (MF563877) and ICRV (NC_003093) were used as the outgroups. Red font isolates (Spain and China) are the sequences obtained in this study while black font isolates are the GenBank accession number of each isolate.

第 VI 章 総合考察

本研究では、分子系統解析が従来果たしてきた意義を大幅に拡張させ、系統樹を基盤とした情報解析と実験的手法のマルチスケールな融合研究のアプローチを確立した。そして、ポティウイルス属のリーキ黄色条斑ウイルス(LYSV)とカブモザイクウイルス(TuMV)、そしてアレキシウイルス属を対象に研究を行い、宿主適応過程、進化推進力、進化方向性の一端を解明することに成功し、系統樹駆動型研究アプローチが広範に有用であることを示した。

分子系統解析は生物学を支える基盤的技術であるが、特にウイルスの系統解析は感染症の制圧や感染動態を適切に把握する上で非常に効果的であることが COVID-19 パンデミック下において明確に示された(Wilkinson et al., 2021)。世界中の研究機関が SARS-CoV-2 のゲノム配列を決定し公開したことで、大規模ゲノム疫学データが急速に蓄積し、国際データベースの GISAID には約 3 年間で 1500 万分離株を超える数のゲノムデータが公開されている。そして、登録され続けるゲノムデータをリアルタイムにモニタリングすることで、拡散様式の把握や流行株超早期捕捉などが可能になっている(Hadfield et al., 2018; Ito et al., 2023; Suzuki et al., 2022)。

植物ウイルスにおいては、ゲノム疫学調査を国際的に行う体制が整っていないため、動物ウイルスと比べて感染情報をリアルタイムに追跡することが困難であることが多い。また、公共データベース上にも利用可能なゲノムデータは動物ウイルスよりも圧倒的に少ない。ゲノムデータベースの整備は植物ウイルス研究の大きな課題の 1 つであり、配列登録のインセンティブを考慮したデータベースの運営が今後必要になってくると思われる。しかし、本研究も含めていくつかの植物ウイルスでは網羅的な系統解析により、拡散様式や宿主適応進化に関する報告があり、植物保護学およびウイルス進化学上重要な知見を多く明らかにしている(Gao et al., 2020; Kawakubo et al., 2021; Kawakubo et al., 2023;

Kawakubo et al., 2022; Lefeuvre et al., 2010; Varsani et al., 2008)。植物ウイルスにおいても、このようなゲノム疫学データを用いた系統解析の事例研究を増やすことは重要である。

系統樹駆動型研究アプローチは、生物進化原理の解明において高いポテンシャルを持つが、特に生命科学の究極の目標の 1 つである進化予測技術の実現化に大きく寄与すると思われる。進化予測は、ウイルスだけでなく薬剤耐性細菌の出現予測など様々な課題に応用できる科学技術である(Han et al., 2023; Konno & Iwasaki, 2023)。進化は偶発性の強いプロセスであるため、予測は困難であると考えられていたが、データの蓄積とアルゴリズムの発展により、進化予測の実現可能性は急速に高まりつつある(Wang et al., 2023)。その可能性を高めている背景の 1 つに、情報科学分野で急速に発展している深層学習の技術でタンパク質配列を自然言語処理のように取り扱うタンパク質言語モデルの開発が挙げられる(Hie et al., 2021; Lin et al., 2023)。巨大 IT 企業も参画して開発が進められている本技術を用いて、ウイルスのタンパク質配列と **deep mutational scanning** などのハイスループット実験により計測した形質データと融合させ、配列データのみからウイルスの適応度地形(fitness landscape)を推定したり、変異の効果や進化方向性を一定の精度で予測することが可能になった(Hie et al., 2022; Kistler & Bedford, 2023; Starr et al., 2022)。これらの技術に系統樹情報を組み込んで、進化の道筋を同時に学習できれば、配列の一次情報のみではなく、その背後にある進化過程を考慮した進化予測が可能になると思われる。実際に細菌を対象にした進化予測の研究事例では、遺伝子の獲得・喪失をシンプルな離散モデルに組み込み、系統樹情報から進化過程を学習し、遺伝子獲得・喪失の予測器(predictor)を開発することに成功している(Konno & Iwasaki, 2023)。また、現実世界で実際に観測される進化は、無数に起こり得た進化の中の 1 つであるため、実験室進化のような選択圧を自在にコントロールできる条件下で得られたデータと組み合わせることができれば、進化可能性を定量することも可能になると思われる(Iwasawa et al., 2022; Maeda et al., 2020)。

このように様々な階層におけるデータを統合的に解析することで、ウイルスの進化機構や流行動態の解明と、その原動力となる分子機構を紐づけることが可能になる。系統樹駆動型研究アプローチは、これらの階層を橋渡しすることで、ウイルス進化研究における新たなパラダイムを生み出すことができると期待される。

付属資料 - 使用した主なプログラム一覧

ベイズ法による系統解析関連

* BEAST

ベイズ法による系統樹推定プログラム。時間情報を考慮した有根系統樹推定に特化して開発されている。分岐年代推定、系統地理学的解析、祖先状態復元解析などが可能。

URL: <https://beast.community>



* BEAST2

前項の BEAST から派生後、独立して開発されているプログラム。様々なアドオンパッケージが開発されており birth-death model による系統動態解析などが可能。

URL: <https://www.beast2.org>



* Tracer

MCMC の収束判定を行う GUI プログラム。推定パラメータの Effective sampling size (ESS) の計算も行う。

URL: <http://tree.bio.ed.ac.uk/software/tracer/>



* TempEst

系統樹と各 OTU の採集年代をもとに、root-to-tip 線形回帰を行うプログラム。単純なモデルにおける時間信号の推定が可能。

URL: <http://tree.bio.ed.ac.uk/software/tempest/>



* ggtree

系統樹を可視化する R パッケージ。BEAST のように、系統樹推定と同時行う他の解析結果と統合可視化が可能で拡張性が高い。

URL: <https://bioconductor.org/packages/release/bioc/html/ggtree.html>



* FigTree

系統樹を可視化する GUI プログラム。

URL: <http://tree.bio.ed.ac.uk/software/figtree/>



* Spread3

BEAST を用いた系統地理学解析結果を可視化するプログラム。

URL: <https://rega.kuleuven.be/cev/ecv/software/Spread3>

* BAMM & BAMMtools

系統樹から多様化率(diversification rate)を推定するプログラム。
R パッケージの BAMMtools と組み合わせて解析を行う。

URL: <http://bamm-project.org>



* MrBayes3

ベイズ系統樹推定プログラム。系統樹推定にベイズ法を導入した先駆けのプログラム。

URL: <https://nbisweden.github.io/MrBayes/>



最尤法による系統解析

* PhyML

最尤法による系統樹推定を行うプログラム。比較的計算に時間がかかるが、多分岐をあまり許容しないため解像度の高い系統樹を構築できる。

URL: <http://www.atgc-montpellier.fr/phyml/>

組換え解析

* RDP5

組換え解析プログラム。多数の組換えプログラムが統合して実装されており、組換え解析において最も頻繁に使用されているプログラム。

URL: <http://web.cbio.uct.ac.za/~darren/rdp.html>



* SimPlot

塩基配列の類似性を可視化するプログラム。組換え部位の検出にも使用される。

URL: <https://sray.med.som.jhmi.edu/SCRsoftware/SimPlot/>



* SplitsTree

Neighbor-Net 法で系統ネットワークを推定するプログラム。

URL: <https://uni-tuebingen.de/en/fakultaeten/mathematisch-naturwissenschaftliche-fakultaet/fachbereiche/informatik/lehrstuehle/algorithms-in-bioinformatics/software/splitstree/>



集団遺伝学的解析

* adegenet

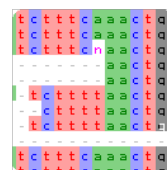
様々な集団遺伝学的解析関数を実装した R パッケージ。多変量解析や空間的遺伝子発現解析が可能。

URL: <http://adegenet.r-forge.r-project.org>

塩基配列編集およびアライメント

* AliView

塩基・アミノ酸配列編集プログラム。統合してある各種アライメントプログラムも利用可能。



URL: <https://ormbunkar.se/aliview/>

* MEGA11

塩基・アミノ酸配列編集、アラインメント作成、系統樹推定などを1つのプログラムに統合した GUI プログラム。



URL: <https://www.megasoftware.net>

* muscle

マルチプルシーケンスアライメントを作成するプログラム。Log-expectation (LE)スコアによりアライメントを行う。

URL: <https://www.ebi.ac.uk/Tools/msa/muscle/>

* ClustalX

マルチプルシーケンスアライメントを作成する GUI プログラム。ClustalWの GUI バージョン。

URL: <http://www.clustal.org>

その他の本研究に使用したスクリプトと解析データは、以下の GitHub 公開レポジトリから利用可能である。

URL: https://github.com/skawakubo/Kawakubo_et_al_LYSV-leek

摘要

植物ウイルスは、感染した宿主植物に適応するため進化戦略を動的に変化させる。この進化機構の解明は、農業現場におけるウイルス病の防除対策や抵抗性品種の作出などに大きく寄与する。ウイルスは宿主細胞に依存して自己複製と翻訳を行い、伝播様式を媒介生物に依存するものも多く存在する。従ってウイルスは外界とのマルチスケールな相互作用と選択圧のもとで自身の適応度を上げる方向に進化すると考えられる。そのため、進化機構の包括的な理解には分子生物学的手法に代表されるミクロスケールの実験と分子系統学や大規模ゲノム疫学調査などによるマクロスケールの情報解析の融合研究が重要である。しかし、それぞれの手法による研究は個別の報告に留まることが多く、スケールを貫いた融合事例は未だ少ない。本研究は、系統樹を主導とした情報解析から実験までをマルチスケールに貫く系統樹駆動型研究アプローチを確立させ、複数の植物ウイルスを対象としてその進化機構の解明を目指した。

リーキ黄色条斑ウイルス(LYSV)の宿主適応進化

- * LYSV はニンニクにおける重要病原ウイルスであるが、ニンニク以外のネギ属植物にも感染する。しかし、異なる宿主植物への適応過程や、その背後にある分子機構は明らかにされていない。
- * LYSV が感染しているネギ属植物を採集し、それら LYSV ゲノム配列の一部を決定後、公共データベース GenBank で利用可能な全ての LYSV ゲノムデータを加え、各ウイルス分離株の採集年代と地理および宿主情報を統合したデータセットを作成した。そして、**uncorrelated relaxed clock model** を用いてベイズ法による時間系統樹の構築を行なった。その結果、LYSV は P1 遺伝子に 68 アミノ酸の欠失を持つ **S-type** と欠失を持たない **N-type** に対応した 2 つの分子系統グループに 13 世紀ごろ分岐したことが明らかになった。
- * **birth-death model** を用いた適応度推定により **S-type** の方が **N-type** よりも高い適応度を持つことが示された。このことは、**S-type** が **N-type** よりも広い宿主

域を持つという本邦の野生ネギ属植物を対象とした疫学調査、および S-type の P1 タンパク質が N-type に比べて高い RNA サイレンシングサプレッサー (RSS) 活性を持つという実験結果によっても支持された。

- * N-type の分離株の中にはリーキ特異的に感染する LYSV からなる単系統グループが存在することが明らかになった。また、リーキの変異種であるジャンボニンニク、およびジャンボニンニクとニンニク、リーキの異種間交雑種を調査すると、リーキ特異的系統グループに属する LYSV 分離株の感染が確認された。
- * RT-PCR の結果からこれら LYSV 分離株はリーキ細胞内では極低濃度で存在する一方で、ジャンボニンニク、および交雑種の細胞内では、高濃度で存在することが明らかになった。
- * 祖先状態復元により祖先宿主の推定を行ったところ、LYSV の共通祖先は分岐初期においてリーキを宿主としていた可能性が高いことが示唆された。
- * リーキ特異的系統グループに属する LYSV がコードする P1 と HC-Pro を、アグロインフィルトレーション法を用いてリーキ細胞で共発現させると、ニンニクに感染する LYSV よりも高い RSS 活性を示した。この現象はタマネギと *Nicotiana benthamiana* を宿主として用いた実験では確認されなかった。

日本におけるカブモザイクウイルス(TuMV)の系統地理学的解析

- * アブラナ野菜の重要病原ウイルスである TuMV は、温帯・亜熱帯地域を中心に世界中に広く分布するが、地中海沿岸部を発生中心としシルクロードの痕跡を辿りユーラシア大陸を東に横断してきたと報告されている。
- * TuMV は 1921 年に米国で初めて発生が報告後、1951 年に日本でも報告されているが、日本国内における拡散様式や集団の年代推移などについては明らかでない。そこで、日本全土の幅広い地域で約半世紀に渡り採取された TuMV の大規模ゲノム疫学データを用い、網羅的な組換え解析および時空間的系統地理学解析を行なった。

- * 組換え解析の結果、本研究で用いた日本産 TuMV 全 370 分離株のうち、92% に相当する 339 分離株が少なくとも 1 箇所組換え部位を持つ組換え体で、8%に相当する 31 分離株のみが非組換え体であることが明らかになった。また、それらの組換え体は 82 パターンの異なる組換え体型に分けられ、そのうち 2 パターンのみが海外にも分布することが報告されており、80 パターンは日本のみに分布する組換え体型であることが明らかになった。
- * 時空間的系統地理学解析の結果、TuMV の日本への侵入は江戸時代後期頃であると推定された。

アレキシウイルス属の種分岐過程の解析

- * アレキシウイルス属は、ネギ属植物を宿主とする 8 種のウイルスから構成される。
- * ベイズ系統解析により、アレキシウイルス属共通祖先からの種分岐過程初期において、CRP 遺伝子の上流に 10~20 塩基挿入された系統群が出現したことが明らかになった。
- * この挿入領域の有無が、RSS タンパク質である CRP の発現様式と細胞内局在を変化させ、相互排他的な進化方向性を生み出したことを示した。

本研究において、分子系統解析が従来果たしてきた意義を大幅に拡張させ、系統樹を基盤とした情報解析と実験的手法のマルチスケールな融合研究のアプローチを確立した。そして、複数の植物ウイルス種を対象に、宿主適応過程、進化推進力、進化方向性の一端を解明することに成功し、系統樹駆動型研究アプローチが広範に有用であることを示した。

引用文献

- Akaike, H. (1974). A new look at the statistical model identification. 215-222. *Selected Papers of Hirotugu Akaike, Springer, New York.*
<https://doi.org/10.1109/TAC.1974.1100705>
- Ariga, T., Kumagai, H., Yoshikawa, M., Kawakami, H., Seki, T., Sakurai, H., Hasegawa, I., Etoh, T., Sumiyoshi, H., Tsuneyoshi, T., Sumi, S., & Iwai, K. (2002). Garlic-like but odorless plant *Allium ampeloprasum* 'Mushuu-ninniku'. *Journal of The Japanese Society For Horticultural Science*, 71(3), 362-369.
<https://doi.org/10.2503/jjshs.71.362>
- Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M. A., & Alekseyenko, A. V. (2012). Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution*, 29(9), 2157-2167. <https://doi.org/10.1093/molbev/mss084>
- Bielejec, F., Baele, G., Vrancken, B., Suchard, M. A., Rambaut, A., & Lemey, P. (2016). Spread3: Interactive visualization of spatiotemporal history and trait evolutionary processes. *Molecular Biology and Evolution*, 33(8), 2167-2169.
<https://doi.org/10.1093/molbev/msw082>
- Bos, L. (1983). Viruses and virus diseases of *Allium* species. *Acta Horticulturae*, 127, 11-30. <https://doi.org/10.17660/ActaHortic.1983.127.1>
- Bos, L., Huijberts, N., Huttinga, H., & Maat, D. Z. (1978). Leek yellow stripe virus and its relationships to onion yellow dwarf virus; characterization, ecology and possible control. *Netherlands Journal of Plant Pathology*, 84, 185-204.
<https://doi.org/10.1007/BF02650386>
- Bothmer, R. v. (1970). Cytological studies in *Allium*. I. chromosome numbers and

morphology in *Allium* sect. *Allium* from Greece. *Botaniska Notiser*, 518-550.

- Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F. K., Müller, N. F., Ogilvie, H. A., du Plessis, L., Poppinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., Suchard, M. A., Wu, C. H., Xie, D., Zhang, C., Stadler, T., & Drummond, A. J. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 15(4), e1006650. <https://doi.org/10.1371/journal.pcbi.1006650>
- Camin, J. H., & Sokal, R. R. (1965). A method for deducing branching sequences in phylogeny. *Evolution*, 311-326. <https://doi.org/https://doi.org/10.2307/2406441>
- Cheon, S., & Liang, F. (2009). Bayesian phylogeny analysis via stochastic approximation Monte Carlo. *Molecular Phylogenetics and Evolution*, 53(2), 394-403. <https://doi.org/https://doi.org/10.1016/j.ympev.2009.06.019>
- Choi, B. K., Koo, J. M., Ahn, H., Yum, H., Choi, C. W., Ryu, K. H., Chen, P., & Tolin, S. (2005). Emergence of Rsv-resistance breaking Soybean mosaic virus isolates from Korean soybean cultivars. *Virus Research*, 112(1-2), 42-51. <https://doi.org/https://doi.org/10.1016/j.virusres.2005.03.020>
- Clark, M. F., & 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(3), 475-483. <https://doi.org/https://doi.org/10.1099/0022-1317-34-3-475>
- Cornwell, W., & Nakagawa, S. (2017). Phylogenetic comparative methods. *Current Biology*, 27(9), R333-R336. <https://doi.org/https://doi.org/10.1016/j.cub.2017.03.049>

- Crooks, G. E., Hon, G., Chandonia, J.-M., & Brenner, S. E. (2004). WebLogo: A sequence logo generator. *Genome Research*, *14*(6), 1188-1190.
<https://doi.org/10.1101/gr.849004>
- Cuevas, J. M., Delaunay, A., Visser, J. C., Bellstedt, D. U., Jacquot, E., & Elena, S. F. (2012). Phylogeography and molecular evolution of potato virus Y. *PLoS ONE*, *7*(5), 1-10. <https://doi.org/10.1371/journal.pone.0037853>
- Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: More models, new heuristics and parallel computing. *Nature Methods*, *9*(8), 772-772.
<https://doi.org/10.1038/nmeth.2109>
- Dellicour, S., Rose, R., Faria, N. R., Lemey, P., & Pybus, O. G. (2016). SERAPHIM: Studying environmental rasters and phylogenetically informed movements. *Bioinformatics*, *32*(20), 3204-3206.
<https://doi.org/10.1093/bioinformatics/btw384>
- Drummond, A. J., Ho, S. Y. W., Phillips, M. J., & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biology*, *4*(5), 699-710.
<https://doi.org/10.1371/journal.pbio.0040088>
- Drummond, A. J., Pybus, O. G., Rambaut, A., Forsberg, R., & Rodrigo, A. G. (2003). Measurably evolving populations. *Trends in Ecology and Evolution*, *18*(9), 481-488. [https://doi.org/10.1016/S0169-5347\(03\)00216-7](https://doi.org/10.1016/S0169-5347(03)00216-7)
- Drummond, A. J., Rambaut, A., Shapiro, B., & Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, *22*(5), 1185-1192.
<https://doi.org/10.1093/molbev/msi103>
- Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian

phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, 29(8), 1969-1973. <https://doi.org/10.1093/molbev/mss075>

Du Plessis, L., & Stadler, T. (2015). Getting to the root of epidemic spread with phylodynamic analysis of genomic data. *Trends in Microbiology*, 23(7), 383-386. <https://doi.org/10.1016/j.tim.2015.04.007>

Duchêne, S., Duchêne, D., Holmes, E. C., & Ho, S. Y. W. (2015). The performance of the date-randomization test in phylogenetic analyses of time-structured virus data. *Molecular Biology and Evolution*, 32(7), 1895-1906. <https://doi.org/10.1093/molbev/msv056>

Duchene, S., Lemey, P., Stadler, T., Ho, S. Y. W., Duchene, D. A., Dhanasekaran, V., & Baele, G. (2020). Bayesian evaluation of temporal signal in measurably evolving populations. *Molecular Biology and Evolution*, 37(11), 3363-3379. <https://doi.org/10.1093/molbev/msaa163>

Dudas, G., & Bedford, T. (2019). The ability of single genes vs full genomes to resolve time and space in outbreak analysis. *BMC Evolutionary Biology*, 19(1), 232. <https://doi.org/10.1186/s12862-019-1567-0>

Dudas, G., Carvalho, L. M., Bedford, T., Tatem, A. J., Baele, G., Faria, N. R., Park, D. J., Ladner, J. T., Arias, A., Asogun, D., Bielejec, F., Caddy, S. L., Cotten, M., D'Ambrozio, J., Dellicour, S., Di Caro, A., DiClaro, J. W., Duraffour, S., Elmore, M. J., Fakoli, L. S., Faye, O., Gilbert, M. L., Gevao, S. M., Gire, S., Gladden-Young, A., Gnirke, A., Goba, A., Grant, D. S., Haagmans, B. L., Hiscox, J. A., Jah, U., Kugelman, J. R., Liu, D., Lu, J., Malboeuf, C. M., Mate, S., Matthews, D. A., Matranga, C. B., Meredith, L. W., Qu, J., Quick, J., Pas, S. D., Phan, M. V. T., Pollakis, G., Reusken, C. B., Sanchez-Lockhart, M., Schaffner, S. F., Schieffelin, J. S., Sealfon, R. S., Simon-Loriere, E., Smits, S. L., Stoecker, K.,

Thorne, L., Tobin, E. A., Vandi, M. A., Watson, S. J., West, K., Whitmer, S., Wiley, M. R., Winnicki, S. M., Wohl, S., Wölfel, R., Yozwiak, N. L., Andersen, K. G., Blyden, S. O., Bolay, F., Carroll, M. W., Dahn, B., Diallo, B., Formenty, P., Fraser, C., Gao, G. F., Garry, R. F., Goodfellow, I., Günther, S., Happi, C. T., Holmes, E. C., Kargbo, B., Keita, S., Kellam, P., Koopmans, M. P. G., Kuhn, J. H., Loman, N. J., Magassouba, N. F., Naidoo, D., Nichol, S. T., Nyenswah, T., Palacios, G., Pybus, O. G., Sabeti, P. C., Sall, A., Ströher, U., Wurie, I., Suchard, M. A., Lemey, P., & Rambaut, A. (2017). Virus genomes reveal factors that spread and sustained the Ebola epidemic. *Nature*, *544*, 309-309.

<https://doi.org/10.1038/nature22040>

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792-1797.

<https://doi.org/10.1093/nar/gkh340>

Edwards, C. J., Suchard, M. A., Lemey, P., Welch, J. J., Barnes, I., Fulton, T. L., Barnett, R., O'Connell, T. C., Coxon, P., Monaghan, N., Valdiosera, C. E., Lorenzen, E. D., Willerslev, E., Baryshnikov, G. F., Rambaut, A., Thomas, M. G., Bradley, D. G., & Shapiro, B. (2011). Ancient hybridization and an irish origin for the modern polar bear matriline. *Current Biology*, *21*(15), 1251-1258.

<https://doi.org/10.1016/j.cub.2011.05.058>

Elena, S. F., Agudelo-Romero, P., & Lalić, J. (2009). The evolution of viruses in multi-host fitness landscapes. *The Open Virology Journal*, *3*, 1-6.

<https://doi.org/10.2174/1874357900903010001>

Eriksson, A., & Manica, A. (2012). Effect of ancient population structure on the degree of polymorphism shared between modern human populations and ancient hominins. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(35), 13956-13960. <https://doi.org/10.1073/pnas.1200567109>

Faria, N. R., Quick, J., Claro, I. M., Thézé, J., De Jesus, J. G., Giovanetti, M., Kraemer, M. U. G., Hill, S. C., Black, A., Da Costa, A. C., Franco, L. C., Silva, S. P., Wu, C. H., Raghwani, J., Cauchemez, S., Du Plessis, L., Verotti, M. P., De Oliveira, W. K., Carmo, E. H., Coelho, G. E., Santelli, A. C. F. S., Vinhal, L. C., Henriques, C. M., Simpson, J. T., Loose, M., Andersen, K. G., Grubaugh, N. D., Somasekar, S., Chiu, C. Y., Muñoz-Medina, J. E., Gonzalez-Bonilla, C. R., Arias, C. F., Lewis-Ximenez, L. L., Baylis, S. A., Chieppe, A. O., Aguiar, S. F., Fernandes, C. A., Lemos, P. S., Nascimento, B. L. S., Monteiro, H. A. O., Siqueira, I. C., De Queiroz, M. G., De Souza, T. R., Bezerra, J. F., Lemos, M. R., Pereira, G. F., Loudal, D., Moura, L. C., Dhaliya, R., França, R. F., Magalhães, T., Marques, E. T., Jaenisch, T., Wallau, G. L., De Lima, M. C., Nascimento, V., De Cerqueira, E. M., De Lima, M. M., Mascarenhas, D. L., Neto, J. P. M., Levin, A. S., Tozetto-Mendoza, T. R., Fonseca, S. N., Mendes-Correa, M. C., Milagres, F. P., Segurado, A., Holmes, E. C., Rambaut, A., Bedford, T., Nunes, M. R. T., Sabino, E. C., Alcantara, L. C. J., Loman, N. J., & Pybus, O. G. (2017). Establishment and cryptic transmission of Zika virus in Brazil and the Americas. *Nature*, 546(7658), 406-410. <https://doi.org/10.1038/nature22401>

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, 17(6), 368-376. <https://doi.org/10.1007/BF01734359>

Fitch, W. M. (1971). Toward defining the course of evolution: minimum change for a specific tree topology. *Systematic Biology*, 20(4), 406-416. <https://doi.org/10.1093/sysbio/20.4.406>

Fritsch, R., & Friesen, N. (2002). Evolution, domestication and taxonomy. *Allium Crop Science: Recent Advances*, 5-30. <https://doi.org/10.1079/9780851995106.0005>

Fuentes, S., Jones, R. A. C., Matsuoka, H., Ohshima, K., Kreuze, J., & Gibbs, A. J.

- (2019). Potato virus Y; the Andean connection. *Virus Evolution*, 5(2), 1-15.
<https://doi.org/10.1093/ve/vez037>
- Gao, F., Kawakubo, S., Ho, S. Y. W., & Ohshima, K. (2020). The evolutionary history and global spatio-temporal dynamics of potato virus Y. *Virus Evolution*, 6(2), 1-10. <https://doi.org/10.1093/ve/veaa056>
- Gardner, M. W., & Kendrick, J. B. (1921). Turnip mosaic. *Journal of Agricultural Research*, 22(3).
- Gibbs, A., & Ohshima, K. (2010). Potyviruses and the digital revolution. *Annual Review of Phytopathology*, 48, 205-223. <https://doi.org/10.1146/annurev-phyto-073009-114404>
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., & Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29(7), 644-652. <https://doi.org/10.1038/nbt.1883>
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3), 307-321. <https://doi.org/10.1093/sysbio/syq010>
- Hadfield, J., Megill, C., Bell, S. M., Huddleston, J., Potter, B., Callender, C., Sagulenko, P., Bedford, T., & Neher, R. A. (2018). Nextstrain: Real-time tracking of pathogen evolution. *Bioinformatics*, 34(23), 4121-4123.
<https://doi.org/10.1093/bioinformatics/bty407>

- Han, W., Chen, N., Xu, X., Sahil, A., Zhou, J., Li, Z., Zhong, H., Gao, E., Zhang, R., Wang, Y., Sun, S., Cheung, P. P.-H., & Gao, X. (2023). Predicting the antigenic evolution of SARS-CoV-2 with deep learning. *Nature Communications*, *14*(1), 3478. <https://doi.org/10.1038/s41467-023-39199-6>
- Harkins, G. W., Martin, D. P., Duffy, S., Monjane, A. L., Shepherd, D. N., Windram, O. P., Owor, B. E., Donaldson, L., Van Antwerpen, T., & Sayed, R. A. (2009). Dating the origins of the maize-adapted strain of maize streak virus, MSV-A. *The Journal of General Virology*, *90*(12), 3066. <https://doi.org/https://doi.org/10.1099/vir.0.015537-0>
- Hartigan, J. A. (1973). Minimum mutation fits to a given tree. *Biometrics*, 53-65.
- Hasegawa, M., Kishino, H., & Yano, T.-a. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, *22*, 160-174. <https://doi.org/https://doi.org/10.1007/BF02101694>
- Hastings, W. K. (1970). Monte carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*(1), 97-109. <https://doi.org/10.2307/2529676>
- Hie, B., Zhong, E. D., Berger, B., & Bryson, B. (2021). Learning the language of viral evolution and escape. *Science*, *371*(6526), 284-288. <https://doi.org/doi:10.1126/science.abd7331>
- Hie, B. L., Yang, K. K., & Kim, P. S. (2022). Evolutionary velocity with protein language models predicts evolutionary dynamics of diverse proteins. *Cell Systems*, *13*(4), 274-285.e276. <https://doi.org/https://doi.org/10.1016/j.cels.2022.01.003>
- Hirschegger, P., Galmarini, C., & Bohanec, B. (2006). Characterization of a novel form of fertile great headed garlic (*Allium* sp.). *Plant Breeding*, *125*(6), 635-637.

<https://doi.org/https://doi.org/10.1111/j.1439-0523.2006.01279.x>

Hirschegger, P., Jakše, J., Trontelj, P., & Bohanec, B. (2010). Origins of *Allium ampeloprasum* horticultural groups and a molecular phylogeny of the section *Allium* (*Allium*: Alliaceae). *Molecular Phylogenetics and Evolution*, *54*(2), 488-497. <https://doi.org/https://doi.org/10.1016/j.ympev.2009.08.030>

Huson, D. H., & Bryant, D. (2005). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*(2), 254-267. <https://doi.org/10.1093/molbev/msj030>

Ito, J., Suzuki, R., Uriu, K., Itakura, Y., Zahradnik, J., Kimura, K. T., Deguchi, S., Wang, L., Lytras, S., Tamura, T., Kida, I., Nasser, H., Shofa, M., Begum, M. M., Tsuda, M., Oda, Y., Suzuki, T., Sasaki, J., Sasaki-Tabata, K., Fujita, S., Yoshimatsu, K., Ito, H., Nao, N., Asakura, H., Nagashima, M., Sadamasu, K., Yoshimura, K., Yamamoto, Y., Nagamoto, T., Kuramochi, J., Schreiber, G., Suzuki, S., Kato, M., Ferdous, Z., Mouri, H., Shishido, K., Misawa, N., Kimura, I., Kosugi, Y., Lin, P., Suganami, M., Chiba, M., Yoshimura, R., Yasuda, K., Iida, K., Ohsumi, N., Strange, A. P., Sauter, D., Nakagawa, S., Wu, J., Watanabe, Y., Sakamoto, A., Yasuhara, N., Nakajima, Y., Yajima, H., Shirakawa, K., Takaori-Kondo, A., Nagata, K., Kazuma, Y., Nomura, R., Horisawa, Y., Tashiro, Y., Kawa, Y., Irie, T., Kawabata, R., Shimizu, R., Takahashi, O., Ichihara, K., Motozono, C., Toyoda, M., Ueno, T., Shibatani, Y., Nishiuchi, T., Saito, A., Matsuno, K., Takayama, K., Hashiguchi, T., Tanaka, S., Fukuhara, T., Ikeda, T., Sato, K., & The Genotype to Phenotype Japan, C. (2023). Convergent evolution of SARS-CoV-2 Omicron subvariants leading to the emergence of BQ.1.1 variant. *Nature Communications*, *14*(1), 2671. <https://doi.org/10.1038/s41467-023-38188-z>

Iwasawa, J., Maeda, T., Shibai, A., Kotani, H., Kawada, M., & Furusawa, C. (2022). Analysis of the evolution of resistance to multiple antibiotics enables prediction

- of the *Escherichia coli* phenotype-based fitness landscape. *PLoS Biology*, 20(12), e3001920. <https://doi.org/10.1371/journal.pbio.3001920>
- Jenner, C. E., Sánchez, F., Nettleship, S. B., Foster, G. D., Ponz, F., & Walsh, J. A. (2000). The cylindrical inclusion gene of turnip mosaic virus encodes a pathogenic determinant to the *Brassica* resistance gene *TuRB01*. *Molecular Plant-Microbe Interactions*, 13(10), 1102-1108. <https://doi.org/10.1094/mpmi.2000.13.10.1102>
- Jimenez, H. J., Martins, L. S. S., Carvalho, R., Montarroyos, A. V. V., & Filho, R. M. M. (2019). *In silico* characterization and phylogenetic analysis of a mannose-specific lectin in *Allium* species. *Genetics and Molecular Research*, 18(2), gmr18187. <https://doi.org/10.4238/gmr18187>
- Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403-1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Jukes, T. H., & Cantor, C. R. (1969). Evolution of protein molecules. *Mammalian Protein Metabolism*, 3, 21-132. <http://dx.doi.org/10.1016/B978-1-4832-3211-9.50009-7>
- Kasschau, K. D., & Carrington, J. C. (1998). A counterdefensive strategy of plant viruses: suppression of posttranscriptional gene silencing. *Cell*, 95(4), 461-470. [https://doi.org/https://doi.org/10.1016/S0092-8674\(00\)81614-1](https://doi.org/https://doi.org/10.1016/S0092-8674(00)81614-1)
- Kawakubo, S., Gao, F., Li, S., Tan, Z., Huang, Y. K., Adkar-Purushothama, C. R., Gurikar, C., Maneechoat, P., Chiemsombat, P., Aye, S. S., Furuya, N., Shevchenko, O., Spak, J., Skoric, D., Ho, S. Y. W., & Ohshima, K. (2021). Genomic analysis of the brassica pathogen turnip mosaic potyvirus reveals its spread along the former trade routes of the Silk Road. *The Proceedings of the*

National Academy of Sciences of the United States of America, 118(12), e2021221118. <https://doi.org/10.1073/pnas.2021221118>

Kawakubo, S., Kim, H., Takeshita, M., & Masuta, C. (2023). Host-specific adaptation drove the coevolution of leek yellow stripe virus and *Allium* plants. *Microbiology Spectrum*, 11(5), e02340-02323. <https://doi.org/10.1128/spectrum.02340-23>

Kawakubo, S., Tomitaka, Y., Tomimura, K., Koga, R., Matsuoka, H., Uematsu, S., Yamashita, K., Ho, S. Y. W., & Ohshima, K. (2022). The recombinogenic history of turnip mosaic potyvirus reveals its introduction to Japan in the 19th century. *Virus Evolution*, 8(2), 1-11. <https://doi.org/10.1093/ve/veac060>

Kim, H., Aoki, N., Takahashi, H., Yoshida, N., Shimura, H., & Masuta, C. (2020). Reduced RNA silencing suppressor activity of onion yellow dwarf virus HC-Pro with N-terminal deletion may be complemented in mixed infection with another potyvirus in garlic. *Journal of General Plant Pathology*, 86(4), 300-309. <https://doi.org/10.1007/s10327-020-00926-2>

Kim, H., Kawakubo, S., Takahashi, H., & Masuta, C. (2023). Two mutually exclusive evolutionary scenarios for alexiviruses that overcome host RNA silencing and autophagy by regulating viral CRP expression. *PLoS Pathogens*, 19(6), e1011457. <https://doi.org/10.1371/journal.ppat.1011457>

Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature*, 217(5129), 624-626. <https://doi.org/10.1038/217624a0>

Kistler, K. E., & Bedford, T. (2023). An atlas of continuous adaptive evolution in endemic human viruses. *Cell Host and Microbe*. <https://doi.org/10.1016/j.chom.2023.09.012>

- Konno, N., & Iwasaki, W. (2023). Machine learning enables prediction of metabolic system evolution in bacteria. *Science Advances*, 9(2), eadc9130.
<https://doi.org/doi:10.1126/sciadv.adc9130>
- Kreuze, J. F., Vaira, A. M., Menzel, W., Candresse, T., Zavriev, S. K., Hammond, J., Hyun Ryu, K., & Report Consortium, I. (2020). ICTV virus taxonomy profile: Alphaflexiviridae. *Journal of General Virology*, 101(7), 699-700.
<https://doi.org/https://doi.org/10.1099/jgv.0.001436>
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21), 2947-2948.
<https://doi.org/10.1093/bioinformatics/btm404>
- Lee, L., & Anderson, E. (1998). Nucleotide sequence of a resistance breaking mutant of southern bean mosaic virus. *Archives of Virology*, 143, 2189-2201.
<https://doi.org/https://doi.org/10.1007/s007050050451>
- Lefevre, P., Martin, D. P., Harkins, G., Lemey, P., Gray, A. J. A., Meredith, S., Lakay, F., Monjane, A., Lett, J. M., Varsani, A., & Heydarnejad, J. (2010). The spread of tomato yellow leaf curl virus from the middle east to the world. *PLoS Pathogens*, 6(10). <https://doi.org/10.1371/journal.ppat.1001164>
- Lemey, P., Rambaut, A., Drummond, A. J., & Suchard, M. A. (2009). Bayesian phylogeography finds its roots. *PLoS Computational Biology*, 5(9).
<https://doi.org/10.1371/journal.pcbi.1000520>
- Lin, Z., Akin, H., Rao, R., Hie, B., Zhu, Z., Lu, W., Smetanin, N., Verkuil, R., Kabeli, O., Shmueli, Y., dos Santos Costa, A., Fazel-Zarandi, M., Sercu, T., Candido, S., & Rives, A. (2023). Evolutionary-scale prediction of atomic-level protein

structure with a language model. *Science*, 379(6637), 1123-1130.

<https://doi.org/doi:10.1126/science.ade2574>

Liu, D., Shi, W., Shi, Y., Wang, D., Xiao, H., Li, W., Bi, Y., Wu, Y., Li, X., & Yan, J. (2013). Origin and diversity of novel avian influenza A H7N9 viruses causing human infection: phylogenetic, structural, and coalescent analyses. *The Lancet*, 381(9881), 1926-1932. [https://doi.org/10.1016/S0140-6736\(13\)60938-1](https://doi.org/10.1016/S0140-6736(13)60938-1)

Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., Ingersoll, R., Sheppard, H. W., & Ray, S. C. (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(1), 152-160. <https://doi.org/doi:10.1128/jvi.73.1.152-160.1999>

Lot, H., Chovelon, V., Souche, S., & Delecalle, B. (1998). Effects of onion yellow dwarf and leek yellow stripe viruses on symptomatology and yield loss of three French garlic cultivars. *Plant Disease*, 82(12), 1381-1385. <https://doi.org/10.1094/PDIS.1998.82.12.1381>

Maeda, T., Iwasawa, J., Kotani, H., Sakata, N., Kawada, M., Horinouchi, T., Sakai, A., Tanabe, K., & Furusawa, C. (2020). High-throughput laboratory evolution reveals evolutionary constraints in *Escherichia coli*. *Nature Communications*, 11(1), 5970. <https://doi.org/10.1038/s41467-020-19713-w>

Malaspinas, A. S., Westaway, M. C., Muller, C., Sousa, V. C., Lao, O., Alves, I., Bergström, A., Athanasiadis, G., Cheng, J. Y., Crawford, J. E., Heupink, T. H., MacHoldt, E., Peischl, S., Rasmussen, S., Schiffels, S., Subramanian, S., Wright, J. L., Albrechtsen, A., Barbieri, C., Dupanloup, I., Eriksson, A., Margaryan, A., Moltke, I., Pugach, I., Korneliussen, T. S., Levkivskyi, I. P., Moreno-Mayar, J. V., Ni, S., Racimo, F., Sikora, M., Xue, Y., Aghakhanian, F. A., Brucato, N.,

Brunak, S., Campos, P. F., Clark, W., Ellingvåg, S., Fourmile, G., Gerbault, P., Injie, D., Koki, G., Leavesley, M., Logan, B., Lynch, A., Matisoo-Smith, E. A., McAllister, P. J., Mentzer, A. J., Metspalu, M., Migliano, A. B., Murgha, L., Phipps, M. E., Pomat, W., Reynolds, D., Ricaut, F. X., Siba, P., Thomas, M. G., Wales, T., Wall, C. M. R., Oppenheimer, S. J., Tyler-Smith, C., Durbin, R., Dortch, J., Manica, A., Schierup, M. H., Foley, R. A., Lahr, M. M., Bowern, C., Wall, J. D., Mailund, T., Stoneking, M., Nielsen, R., Sandhu, M. S., Excoffier, L., Lambert, D. M., & Willerslev, E. (2016). A genomic history of Aboriginal Australia. *Nature*, 538(7624), 207-214. <https://doi.org/10.1038/nature18299>

Martin, D. P., Lemey, P., & Posada, D. (2011). Analysing recombination in nucleotide sequences. *Molecular Ecology Resources*, 11(6), 943-955. <https://doi.org/https://doi.org/10.1111/j.1755-0998.2011.03026.x>

Martin, D. P., Varsani, A., Roumagnac, P., Botha, G., Maslamoney, S., Schwab, T., Kelz, Z., Kumar, V., & Murrell, B. (2020). RDP5: A computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. *Virus Evolution*, 7(1), 5-7. <https://doi.org/10.1093/ve/veaa087>

Mau, B., & Newton, M. A. (1997). Phylogenetic inference for binary data on dendrograms using Markov chain Monte carlo. *Journal of Computational and Graphical Statistics*, 6(1), 122-131. <https://doi.org/10.1080/10618600.1997.10474731>

McLeish, M. J., Fraile, A., & García-Arenal, F. (2019). Evolution of plant–virus interactions: host range and virus emergence. *Current Opinion in Virology*, 34, 50-55. <https://doi.org/10.1016/j.coviro.2018.12.003>

McLeish, M. J., Fraile, A., & García-Arenal, F. (2021). Population genomics of plant viruses: The ecology and evolution of virus emergence. *Phytopathology*, 111(1),

32-39. <https://doi.org/10.1094/PHYTO-08-20-0355-FI>

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6), 1087-1092. <https://doi.org/10.1063/1.1699114>

Metsky, H. C., Matranga, C. B., Wohl, S., Schaffner, S. F., Freije, C. A., Winnicki, S. M., West, K., Qu, J., Baniecki, M. L., Gladden-Young, A., Lin, A. E., Tomkins-Tinch, C. H., Ye, S. H., Park, D. J., Luo, C. Y., Barnes, K. G., Shah, R. R., Chak, B., Barbosa-Lima, G., Delatorre, E., Vieira, Y. R., Paul, L. M., Tan, A. L., Barcellona, C. M., Porcelli, M. C., Vasquez, C., Cannons, A. C., Cone, M. R., Hogan, K. N., Kopp, E. W., Anzinger, J. J., Garcia, K. F., Parham, L. A., Ramírez, R. M. G., Montoya, M. C. M., Rojas, D. P., Brown, C. M., Hennigan, S., Sabina, B., Scotland, S., Gangavarapu, K., Grubaugh, N. D., Oliveira, G., Robles-Sikisaka, R., Rambaut, A., Gehrke, L., Smole, S., Halloran, M. E., Villar, L., Mattar, S., Lorenzana, I., Cerbino-Neto, J., Valim, C., Degraeve, W., Bozza, P. T., Gnirke, A., Andersen, K. G., Isern, S., Michael, S. F., Bozza, F. A., Souza, T. M. L., Bosch, I., Yozwiak, N. L., Macinnis, B. L., & Sabeti, P. C. (2017). Zika virus evolution and spread in the Americas. *Nature*, 546(7658), 411-415. <https://doi.org/10.1038/nature22402>

Monjane, A. L., Harkins, G. W., Martin, D. P., Lemey, P., Lefevre, P., Shepherd, D. N., Oluwafemi, S., Simuyandi, M., Zinga, I., Komba, E. K., Lakoutene, D. P., Mandakombo, N., Mboukoulida, J., Semballa, S., Tagne, A., Tiendrebeogo, F., Erdmann, J. B., van Antwerpen, T., Owor, B. E., Flett, B., Ramusi, M., Windram, O. P., Syed, R., Lett, J. M., Briddon, R. W., Markham, P. G., Rybicki, E. P., & Varsani, A. (2011). Reconstructing the history of maize streak virus strain A dispersal to reveal diversification hot spots and its origin in Southern Africa. *Journal of Virology*, 85(18), 9623-9636.

<https://doi.org/10.1128/jvi.00640-11>

- Moury, B., & Desbiez, C. (2020). Host range evolution of potyviruses: A global phylogenetic analysis. *Viruses*, *12*(1), 111. <https://doi.org/10.3390/v12010111>
- Moury, B., Fabre, F., Hebrard, E., & Froissart, R. (2017). Determinants of host species range in plant viruses. *Journal of General Virology*, *98*(4), 862-873. <https://doi.org/10.1099/jgv.0.000742>
- Navaud, O., Barbacci, A., Taylor, A., Clarkson, J. P., & Raffaele, S. (2018). Shifts in diversification rates and host jump frequencies shaped the diversity of host range among Sclerotiniaceae fungal plant pathogens. *Molecular Ecology*, *27*(5), 1309-1323. <https://doi.org/https://doi.org/10.1111/mec.14523>
- Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, *48*(3), 443-453. [https://doi.org/https://doi.org/10.1016/0022-2836\(70\)90057-4](https://doi.org/https://doi.org/10.1016/0022-2836(70)90057-4)
- Nelson, M., Spiro, D., Wentworth, D., Fan, J., Beck, E., George, K. S., Ghedin, E., Halpin, R., Bera, J., & Hine, E. (2009). The early diversification of influenza A/H1N1pdm. *PLoS Currents*, *1*. <https://doi.org/10.1371/currents.RRN1126>
- Ohshima, K., Tomitaka, Y., Wood, J. T., Minematsu, Y., Kajiyama, H., Tomimura, K., & Gibbs, A. J. (2007). Patterns of recombination in turnip mosaic virus genomic sequences indicate hotspots of recombination. *Journal of General Virology*, *88*(1), 298-315. <https://doi.org/10.1099/vir.0.82335-0>
- Ohshima, K., Yamaguchi, Y., Hirota, R., Hamamoto, T., Tomimura, K., Tan, Z., Sano, T., Azuhata, F., Walsh, J. A., Fletcher, J., Chen, J., Gera, A., & Gibbs, A. (2002). Molecular evolution of *Turnip mosaic virus*: Evidence of host adaptation,

- genetic recombination and geographical spread. *Journal of General Virology*, 83(6), 1511-1521. <https://doi.org/10.1099/0022-1317-83-6-1511>
- Panas, M. D., Schulte, T., Thaa, B., Sandalova, T., Kedersha, N., Achour, A., & McInerney, G. M. (2015). Viral and cellular proteins containing FGDF motifs bind G3BP to block stress granule formation. *PLoS Pathogens*, 11(2), e1004659. <https://doi.org/10.1371/journal.ppat.1004659>
- Pratt, D. A. (2010). Garlic and other *Alliums*. The Lore and the Science. By Eric Block. *Angewandte Chemie International Edition*, 49(40), 7162-7162. <https://doi.org/https://doi.org/10.1002/anie.201004351>
- Rabosky, D. L. (2014). Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE*, 9(2), e89543. <https://doi.org/10.1371/journal.pone.0089543>
- Rabosky, D. L., Grudler, M., Anderson, C., Title, P., Shi, J. J., Brown, J. W., Huang, H., Larson, J. G., & Kembel, S. (2014). BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution*, 5(7), 701-707. <https://doi.org/10.1111/2041-210x.12199>
- Rambaut, A. (2000). Estimating the rate of molecular evolution: incorporating non-contemporaneous sequences into maximum likelihood phylogenies. *Bioinformatics*, 16(4), 395-399. <https://doi.org/10.1093/bioinformatics/16.4.395>
- Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 67(5), 901-904. <https://doi.org/10.1093/sysbio/syy032>
- Rambaut, A., Lam, T. T., Max Carvalho, L., & Pybus, O. G. (2016). Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-

- O-Gen). *Virus Evolution*, 2(1), vew007. <https://doi.org/10.1093/ve/vew007>
- Rannala, B., & Yang, Z. (1996). Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. *Journal of Molecular Evolution*, 43(3), 304-311. <https://doi.org/10.1007/BF02338839>
- Rieux, A., & Khatchikian, C. E. (2017). Tipdatingbeast: an R package to assist the implementation of phylogenetic tip-dating tests using beast. *Molecular Ecology Resources*, 17(4), 608-613. <https://doi.org/10.1111/1755-0998.12603>
- Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19(12), 1572-1574. <https://doi.org/10.1093/bioinformatics/btg180>
- Rzhetsky, A., & Nei, M. (1993). Theoretical foundation of the minimum-evolution method of phylogenetic inference. *Molecular Biology and Evolution*, 10(5), 1073-1095. <https://doi.org/10.1093/oxfordjournals.molbev.a040056>
- Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4), 406-425. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>
- Santosa, A., & Ertunç, F. (2020). Serological and molecular detection of *Leek yellow stripe virus* infecting onion (*Allium cepa* L.) and leek (*Allium ampeloprasum* L.) in Ankara, Turkey. *Bitki Koruma Bülteni*, 60(1), 49-56. <https://doi.org/10.16955/bitkorb.622150>
- Sasaki, J., Kawakubo, S., Kim, H., Kim, O. K., Yamashita, K., Shimura, H., & Masuta, C. (2022). Leek yellow stripe virus can adjust for host adaptation by trimming the N-terminal domain to allow the P1 protein to function as an RNA silencing suppressor. *The Plant Pathology Journal*, 38(4), 383-394.

<https://doi.org/10.5423/PPJ.FT.06.2022.0077>

- Schultz, E. (1921). A transmissible mosaic disease of Chinese cabbage, mustard and turnip. *Journal of Agricultural Research*, 22(3).
- Searle, J. B. (2000). Phylogeography — The history and formation of species. *Heredity*, 85(2), 201-201. <https://doi.org/10.1046/j.1365-2540.2000.0765d.x>
- Shimura, H., Kim, H., Matsuzawa, A., Akino, S., & Masuta, C. (2022). Coat protein of partitiviruses isolated from mycorrhizal fungi functions as an RNA silencing suppressor in plants and fungi. *Scientific Reports*, 12(1), 7855. <https://doi.org/10.1038/s41598-022-11403-5>
- Simpson, J. T., Workman, R. E., Zuzarte, P., David, M., Dursi, L., & Timp, W. (2017). Detecting DNA cytosine methylation using nanopore sequencing. *Nature Methods*, 14(4), 407-410. <https://doi.org/10.1038/nmeth.4184>
- Smith, K. (2008). A new virus disease of the tomato. *Annals of Applied Biology*, 22, 731-741. <https://doi.org/10.1111/j.1744-7348.1935.tb07180.x>
- Sokal, R. R., & Sneath, P. H. (1963). Principles of numerical taxonomy. *Taxon*. <https://doi.org/10.2307/1217562>
- Starr, T. N., Greaney, A. J., Hannon, W. W., Loes, A. N., Hauser, K., Dillen, J. R., Ferri, E., Farrell, A. G., Dadonaite, B., McCallum, M., Matreyek, K. A., Corti, D., Veessler, D., Snell, G., & Bloom, J. D. (2022). Shifting mutational constraints in the SARS-CoV-2 receptor-binding domain during viral evolution. *Science*, 377(6604), 420-424. <https://doi.org/doi:10.1126/science.abo7896>
- Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., & Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1 . 10. *Virus Evolution*, 4(1), 1-5. <https://doi.org/10.1093/ve/vey016>

Sun, X., Zhu, S., Li, N., Cheng, Y., Zhao, J., Qiao, X., Lu, L., Liu, S., Wang, Y., Liu, C., Li, B., Guo, W., Gao, S., Yang, Z., Li, F., Zeng, Z., Tang, Q., Pan, Y., Guan, M., Zhao, J., Lu, X., Meng, H., Han, Z., Gao, C., Jiang, W., Zhao, X., Tian, S., Su, J., Cheng, Z., & Liu, T. (2020). A chromosome-level genome assembly of garlic (*Allium sativum*) provides insights into genome evolution and alliin biosynthesis. *Molecular Plant*, *13*(9), 1328-1339.

<https://doi.org/10.1016/j.molp.2020.07.019>

Suzuki, R., Yamasoba, D., Kimura, I., Wang, L., Kishimoto, M., Ito, J., Morioka, Y., Nao, N., Nasser, H., Uriu, K., Kosugi, Y., Tsuda, M., Orba, Y., Sasaki, M., Shimizu, R., Kawabata, R., Yoshimatsu, K., Asakura, H., Nagashima, M., Sadamasu, K., Yoshimura, K., Suganami, M., Oide, A., Chiba, M., Ito, H., Tamura, T., Tsushima, K., Kubo, H., Ferdous, Z., Mouri, H., Iida, M., Kasahara, K., Tabata, K., Ishizuka, M., Shigeno, A., Tokunaga, K., Ozono, S., Yoshida, I., Nakagawa, S., Wu, J., Takahashi, M., Kaneda, A., Seki, M., Fujiki, R., Nawai, B. R., Suzuki, Y., Kashima, Y., Abe, K., Imamura, K., Shirakawa, K., Takaori-Kondo, A., Kazuma, Y., Nomura, R., Horisawa, Y., Nagata, K., Kawai, Y., Yanagida, Y., Tashiro, Y., Takahashi, O., Kitazato, K., Hasebe, H., Motozono, C., Toyoda, M., Tan, T. S., Ngare, I., Ueno, T., Saito, A., Butlertanaka, E. P., Tanaka, Y. L., Morizako, N., Sawa, H., Ikeda, T., Irie, T., Matsuno, K., Tanaka, S., Fukuhara, T., & Sato, K. (2022). Attenuated fusogenicity and pathogenicity of SARS-CoV-2 Omicron variant. *Nature*, *603*. <https://doi.org/10.1038/s41586-022-04462-1>

Swofford, D. L. (1998). Phylogenetic analysis using parsimony.

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, *38*(7), 3022-3027.

<https://doi.org/10.1093/molbev/msab120>

- Tavare, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequence. *Some Mathematical Questions in Biology-DNA Sequence Analysis*, 17, 57. <https://cir.nii.ac.jp/crid/1571417125456710400>
- Tian, H., Liu, Y., Li, Y., Wu, C.-H., Chen, B., Kraemer, M. U., Li, B., Cai, J., Xu, B., & Yang, Q. (2020). An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science*, 368(6491), 638-642. <https://doi.org/10.1126/science.abb6105>
- Tomitaka, Y., & Ohshima, K. (2006). A phylogeographical study of the *Turnip mosaic virus* population in East Asia reveals an 'emergent' lineage in Japan. *Molecular Ecology*, 15(14), 4437-4457. <https://doi.org/10.1111/j.1365-294X.2006.03094.x>
- Valli, A., Lopez-Moya, J. J., & Garcia, J. A. (2007). Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family *Potyviridae*. *Journal of General Virology*, 88(3), 1016-1028. <https://doi.org/10.1099/vir.0.82402-0>
- Varsani, A., Shepherd, D. N., Monjane, A. L., Owor, B. E., Erdmann, J. B., Rybicki, E. P., Peterschmitt, M., Briddon, R. W., Markham, P. G., Oluwafemi, S., Windram, O. P., Lefeuvre, P., Lett, J. M., & Martin, D. P. (2008). Recombination, decreased host specificity and increased mobility may have driven the emergence of maize streak virus as an agricultural pathogen. *Journal of General Virology*, 89(9), 2063-2074. <https://doi.org/10.1099/vir.0.2008/003590-0>
- Walkey, D. G. A., & Antill, D. N. (1989). Agronomic evaluation of virus-free and virus-infected garlic (*Allium sativum* L.). *Journal of Horticultural Science*, 64(1), 53-60. <https://doi.org/10.1080/14620316.1989.11515927>
- Walsh, J. A., & Jenner, C. E. (2002). Turnip mosaic virus and the quest for durable resistance. *Molecular Plant Pathology*, 3(5), 289-300.

<https://doi.org/https://doi.org/10.1046/j.1364-3703.2002.00132.x>

Wang, H., Fu, T., Du, Y., Gao, W., Huang, K., Liu, Z., Chandak, P., Liu, S., Van Katwyk, P., Deac, A., Anandkumar, A., Bergen, K., Gomes, C. P., Ho, S., Kohli, P., Lasenby, J., Leskovec, J., Liu, T.-Y., Manrai, A., Marks, D., Ramsundar, B., Song, L., Sun, J., Tang, J., Veličković, P., Welling, M., Zhang, L., Coley, C. W., Bengio, Y., & Zitnik, M. (2023). Scientific discovery in the age of artificial intelligence. *Nature*, 620(7972), 47-60. <https://doi.org/10.1038/s41586-023-06221-2>

Wilkinson, E., Giovanetti, M., Tegally, H., & San, J. E. (2021). A year of genomic surveillance reveals how the SARS-CoV-2 pandemic. *Science*, 4336, 27-27. <https://doi.org/10.1126/science.abj4336>

Yasaka, R., Fukagawa, H., Ikematsu, M., Soda, H., Korkmaz, S., Golnaraghi, A., Katis, N., Ho, S. Y. W., Gibbs, A. J., & Ohshima, K. (2017). The timescale of emergence and spread of turnip mosaic potyvirus. *Scientific Reports*, 7(1), 1-14. <https://doi.org/10.1038/s41598-017-01934-7>

Yasaka, R., Ohba, K., Schwinghame, M. W., Fletcher, J., Ochoa-Corona, F. M., Thomas, J. E., Ho, S. Y. W., Gibbs, A. J., & Ohshima, K. (2015). Phylodynamic evidence of the migration of turnip mosaic potyvirus from Europe to Australia and New Zealand. *Journal of General Virology*, 96(3), 701-713. <https://doi.org/10.1099/jgv.0.000007>

Yoshida, N., Shimura, H., & Masuta, C. (2018). Allexiviruses may have acquired inserted sequences between the CP and CRP genes to change the translation reinitiation strategy of CRP. *Archives of Virology*, 163(6), 1419-1427. <https://doi.org/10.1007/s00705-018-3749-2>

Yoshida, N., Shimura, H., Yamashita, K., Suzuki, M., & Masuta, C. (2012). Variability

in the P1 gene helps to refine phylogenetic relationships among leek yellow stripe virus isolates from garlic. *Archives of Virology*, 157(1), 147-153.

<https://doi.org/10.1007/s00705-011-1132-7>

Yoshii, H. (1951). Mosaic disease in crucifers in West Japan. *Shokubutsu Byogai Kenkyu*, 4, 17-22.

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T. Y. (2017). ggtree: an R Package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, 8(1), 28-36.

<https://doi.org/10.1111/2041-210X.12628>

Zhu, S., Tang, S., Tan, Z., Yu, Y., Dai, Q., & Liu, T. (2017). Comparative transcriptomics provide insight into the morphogenesis and evolution of fistular leaves in *Allium*. *BMC Genomics*, 18(1), 60. <https://doi.org/10.1186/s12864-016-3474-8>

Zuckermandl, E., & Pauling, L. (1965). Evolutionary divergence and convergence in proteins. In *Evolving Genes and Proteins* (pp. 97-166). Elsevier.

<https://doi.org/10.1016/B978-1-4832-2734-4.50017-6>

謝辞

本研究を行うにあたり、終始適切なご指導をくださった北海道大学大学院農学研究院植物病原学研究室 増田税教授にこの場を借りて深く感謝します。また、ことあるごとに適切な助言をくださった同研究室 畑谷達児講師、中原健二講師をはじめ、実験等丁寧に細かなところまでご指導くださった Hangil Kim 博士、松永航博士に深く感謝いたします。本論文の作成にあたり、快くご校閲を引き受けてくださり、副査として適切なご助言を賜りました同研究院植物育種学研究室 貴島祐治教授、作物生理学研究室 志村華子講師に心からの感謝をここに記します。また、日頃の研究生活の様々な場面で支えてくれた、同大学大学院農学院植物病原学、植物遺伝資源学、細胞工学研究室の皆様にも深く感謝いたします。

また、本研究で使用したネギ属植物を分譲してくださった宮崎大学農学部 竹下稔教授、情報解析について熱心かつ的確なご助言をくださった東京大学医科学研究所システムウイルス学分野 伊東潤平助教に厚く御礼申し上げます。

そして、常に的確な助言と激励をくださった、佐賀大学 大島一里名誉教授への感謝は言葉に尽くすことはできません。

本研究の一部は日本学術振興会科学研究費助成事業特別研究員奨励費 22J22297 の助成を受けています。また、情報・システム研究機構 国立遺伝学研究所が有する遺伝研スーパーコンピュータシステムを利用しました。

最後に、ここに至るまでの間、私を見守り支え続け、無償の愛を与えてくれた、父 孝広と母 理栄、妹 直美に心から感謝します。

令和6年1月

川久保修佑