

## Article

# Genetic Diversity among Rose Rosette Virus Isolates: A Roadmap towards Studies of Gene Function and Pathogenicity

Jeanmarie Verchot <sup>1,\*</sup> , Venura Herath <sup>2</sup> , Ramon Jordan <sup>3</sup>  and John Hammond <sup>3</sup><sup>1</sup> Department of Plant Pathology & Microbiology, Texas A&M University, College Station, TX 77845, USA<sup>2</sup> Department of Agriculture Biology, Faculty of Agriculture, University of Peradeniya, Peradeniya 20400, Sri Lanka<sup>3</sup> Floral and Nursery Plants Research Unit, US National Arboretum, United States Department of Agriculture, Agriculture Research Service, Beltsville, MD 20705, USA

\* Correspondence: jeanmarie.verchot@ag.tamu.edu; Tel.: +1-979-458-6369

**Abstract:** The phylogenetic relationships of ninety-five rose rosette virus (RRV) isolates with full-length genomic sequences were analyzed. These isolates were recovered mostly from commercial roses that are vegetatively propagated rather than grown from seed. First, the genome segments were concatenated, and the maximum likelihood (ML) tree shows that the branches arrange independent of their geographic origination. There were six major groups of isolates, with 54 isolates in group 6 and distributed in two subgroups. An analysis of nucleotide diversity across the concatenated isolates showed lower genetic differences among RNAs encoding the core proteins required for encapsidation than the latter genome segments. Recombination breakpoints were identified near the junctions of several genome segments, suggesting that the genetic exchange of segments contributes to differences among isolates. The ML analysis of individual RNA segments revealed different relationship patterns among isolates, which supports the notion of genome reassortment. We tracked the branch positions of two newly sequenced isolates to highlight how genome segments relate to segments of other isolates. RNA6 has an interesting pattern of single-nucleotide mutations that appear to influence amino acid changes in the protein products derived from ORF6a and ORF6b. The P6a proteins were typically 61 residues, although three isolates encoded P6a proteins truncated to 29 residues, and four proteins extended 76–94 residues. Homologous P5 and P7 proteins appear to be evolving independently. These results suggest greater diversity among RRV isolates than previously recognized.

**Keywords:** emaravirus; rose rosette virus; negative-strand RNA virus; rose; plant bunyavirus

check for updates

**Citation:** Verchot, J.; Herath, V.; Jordan, R.; Hammond, J. Genetic Diversity among Rose Rosette Virus Isolates: A Roadmap towards Studies of Gene Function and Pathogenicity. *Pathogens* **2023**, *12*, 707. <https://doi.org/10.3390/pathogens12050707>

Academic Editors: Jozef J. Bujarski and Andrea Luvisi

Received: 1 February 2023

Revised: 11 April 2023

Accepted: 10 May 2023

Published: 12 May 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Rose rosette virus (RRV) is an enveloped virus with a negative-strand RNA genome and is a member of the genus *Emaravirus* within the family *Fimoviridae* [1]. RRV has seven genome segments [2,3], and all are monocistronic except RNA6, which is bicistronic. RNA1 encodes the putative RNA-dependent RNA polymerase (RdRp), RNA2 the glycoprotein precursor (GPP), RNA3 the nucleoprotein (NP), and RNA4 the movement protein (MP) [4,5]. RNAs 1–3 alone are sufficient for local replication and encapsidation [6]. The RRV genes encoded by RNAs 5 and 7 have unknown functions [7]. RNA6 encodes two proteins known as P6a and P6b. The P6bs of other emaraviruses have been suggested to confer pathogenicity, and it is referred to as an ‘ABC’ protein [8]. The P6b homolog of High Plains wheat mosaic virus (HPWMOV) is reported to have silencing suppressor activity [9,10].

The mutation spectrum within a virus species impacts virus fitness, adaptive responses, antigenic drift, and natural selection. Mutations can offer biological advantages or disadvantages, influence the behavior of a virus population, and can also lead to the accumulation of defective interfering (DI) RNAs [11]. Recent investigations of the population of RRV across North America [12] characterized the population structure of RRV by

examining individual genome segments and reported nucleotide identities ranging from 93 to 99%. The analysis of individual genome segments did not reveal genetic changes that could be useful molecular markers to differentiate strains or isolates of RRV across North America [12]. Identifying molecular markers is important to describe the dissemination of virus strains and differential host interactions. Furthermore, identifying various subtypes and subgroups based on amino acid sequence variations is important for characterizing virulence determinants that are linked to different disease etiologies. For segmented RNA viruses, genetic changes derive from mutations, recombination, or segment reassortment. The primary source of mutations is the misincorporation of nucleotides during template copying by an error-prone viral RdRp [13,14]. Transitions are the exchange between purines or between pyrimidines and are more common than transversions, which are the exchange of a purine for a pyrimidine or vice versa [15]. Insertions or deletions, known as indel mutations, are another form of misincorporation of nucleotides. Indels happen during replication by misalignment of the RdRp along the template and most often occur at short repeated sequences or homopolymeric tracts [16]. Recombination and reassortment often occur when virus strains co-mingle in a common host or vector. Within emaraviruses, both recombination and segment reassortment have been reported for fig mosaic virus (FMV) [17] and pigeon pea sterility mosaic virus 1 and 2 (PPSMV-1 and PPSMV-2) [18]. In addition, evidence from HPWMOV supports probable reassortment [19–21].

RRV causes considerable economic losses and is a threat to rose production across the USA and Canada. Breeding priorities for rosebushes are to incorporate genetic resistance to RRV alongside other desirable traits. The general descriptions of virus strains or species subgroups are rooted in the categorization of genetic variation that influences the foliar disease pattern, or the nature of its engagement with host genetic resistance. For many virus diseases, there are known strains that elicit mild to severe symptoms, from mild leaf chlorosis or vein clearing to necrosis, and interact with host genetic resistance genes that confer protection against infection. Host resistance is described as extreme (ER) resistance where no symptoms develop, hypersensitive (HR) resistance where local lesions appear in inoculated leaves (usually without systemic infection), or immunity when plants are insusceptible to infection [22]. Plants encode virus resistance (R) genes that can recognize viral protein elicitors to trigger host defenses. The durability of resistance is determined by mutations occurring in the region of the viral elicitor that interacts with the R protein [23]. Virus resistance breaking (RB) strains typically have acquired mutations within the elicitor domains that eliminate R protein recognition to enable successful infection [24,25]. The development of durable control strategies will require understanding regions of the virus genome that show significant genetic variation that can drive virus–host interactions and virus evolution [26].

Common commercial roses are the products of interspecific hybrids or the products of hybrid crossings. For example, tea roses are themselves hybrids: the floribunda rose bush is a cross between the hybrid tea rose and the polyantha rose. Windham et al. (2023), in this special issue of *Pathogens* [27], explain that the major commercial rose cultivars are bred from genetically diverse North American, European, and Asian species or hybrids derived from several *Rosa* taxonomic sections [28–30]. Since most commercial varieties are susceptible to rose rosette disease (RRD), breeders are performing replicated trials to identify new genetic sources of RRV resistance, and this requires using a consistent inoculum and introducing good disease pressure to produce symptoms in young rose plants [28,31,32]. Breeding trials depend upon the recognition of common disease characteristics among isolates or strains, accurate and robust diagnostic detection in infected plant samples, and an understanding of the genetic diversity and population structures that influence the emergence of RB strains. This study was undertaken to gain insight into the genetic diversity and the nature of the genetic changes among the complete genomes of sequenced RRV isolates.

## 2. Materials and Methods

### 2.1. RRV Sequence Database Construction

All previously available RRV sequences were fetched from the National Center for Biotechnology Information (NCBI) Database GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>, accessed on 5 August 2021) to Geneious Prime software (<http://www.geneious.com>; accessed on 1 January 2022). Geneious Prime v. 2022.0 and v. 2023.0.4 (Dotmatics Software Co., Bishop's Stortford, UK) was used to build a database of all RRV genomes and a parallel database of accessions and common names for all reported genome segments was prepared using Microsoft® Excel® (v. 2304) for Microsoft 365. Incomplete genome sequences were discarded from both databases and only the full genome sequences with seven segments were maintained for further analysis. Sequences were aligned using translation align and then translated to validate the start and stop codons for all coding sequences and surrounding untranslated regions (UTRs). All individual genome segment nucleotide sequences were aligned using MUSCLE and trimmed down to the same length represented by the 91 GenBank sequences derived from multiplex amplicon sequencing [12], which did not contain the 16–61 nts from both the 5' and 3' terminal ends of each genome segment. This also helped to check for and trim extra bases that have been erroneously added during high-throughput sequencing (HTS). All sequences were named in the Microsoft® Excel® and Geneious Prime databases according to the isolate name reported in GenBank and their reported originating USA state.

Two new genomes were also included in this study. Samples of rose plant shoots showing severe phyllody from RRD-affected *Rosa × fortuniana* plants growing in the US National Arboretum (USNA) National Herb Garden's "Historic and Species Rose Garden" in Washington DC (isolate named RF-DC) and those from *Rosa* sp. 'Stormy Weather' shoots exhibiting typical RRD symptoms in an RRD-resistance field trial in Wilmington, DE (isolated named SW-DE), were collected in 2017 for HTS analysis. Total RNAs were isolated by the CKC (CTAB, KOAc, silica column) protocol (RF-DC sample) [33] or using the MagMAX™ Plant RNA Isolation Kit (ThermoFisher, Waltham, MA, USA) protocol (SW-DE sample) and prepared for HTS on the NextSeq® 500 platform (Illumina, San Diego, CA, USA) as described in [34]. Assembly and analysis of the resulting HTS data resulted in contigs representing nearly complete sequences (>99%) of each of the seven RRV genome segments from both isolates (RF\_DC and SW\_DE, respectively).

### 2.2. Performance of ML Analysis and MAFFT Alignments and Analysis of Sequence Variability and Mutations

Geneious Prime was used to concatenate the nucleotide and amino acid sequences of all seven segments for each virus isolate. Maximum likelihood (ML) analyses were performed using MAFFT alignment and the E-INS-i algorithm; the nucleotide scoring matrix was 200PAM/k = 2, the amino acid scoring matrix was BLOSUM62, and both analyses had a gap open penalty of 1.53. The best-fitting nucleotide and amino acid substitution models with the lowest BIC score were determined using iQTREE v. 1.6.12 [35,36] and then ML trees were imported into FigTree v.1.4.4. (both softwares from <http://github.com>, accessed on 1 January 2022) and rooted using manually selected outgroups. PHYML trees were constructed using the amino acid sequences and the VT model. The rate category and gamma distribution parameters were set to 4.

DnaSP v6.12.03 × 64 ([ub.edu/dnasp](http://ub.edu/dnasp)) was used to perform a comprehensive analysis of nucleotide polymorphisms, haplotype diversity, nucleotide diversity, and neutrality tests [37]. Recombination breakpoints were analyzed using RDP5.0, SplitsTree, and Sim-Plot++ [38–40]. The RDP5.0 package includes the inbuilt methods RDP, GENECONV, BOOTSCAN, MAXIMUM CHI SQUARE, CHIMAERA, SISTER SCAN, and 3SEQ. The analyses were conducted with default settings and events were considered significant if the *p*-values were less than  $1 \times 10^{-5}$  in at least five methods in RDP5.0. Then, parsimony-informative amino acid mutations were identified using MAFFT alignments imported into

MEGA X software v. 10.1.8 [41]. Trees were annotated in iTOL v6 to display informative mutations [42].

### 3. Results

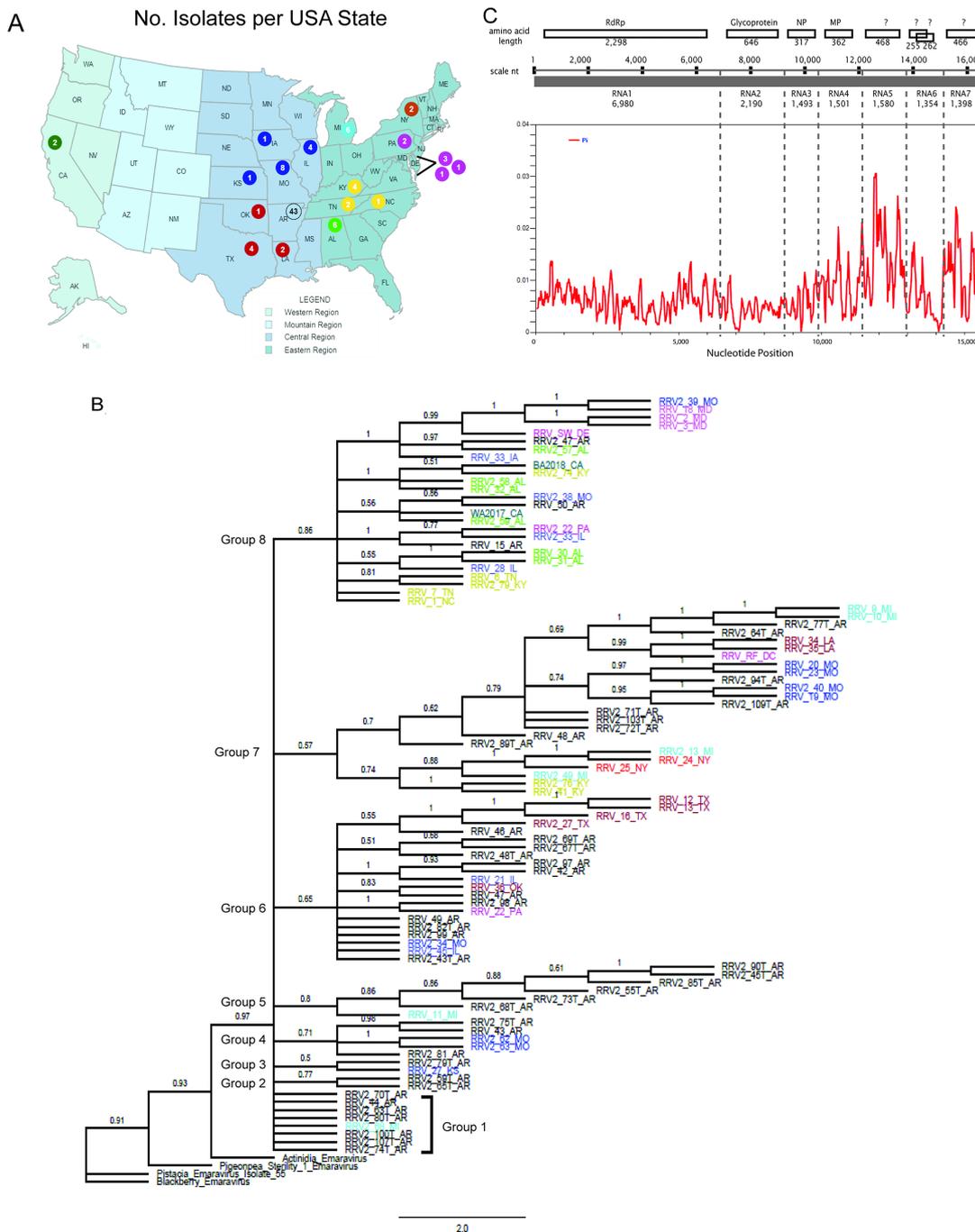
#### 3.1. Virus Isolates and Genetic Variants

Ninety-three complete genome sequences of RRV were downloaded from GenBank, and two complete genome sequences were obtained by direct sequencing. The locations reported in GenBank are distributed within seventeen USA states (Table S1 and Figure 1A) [12]. There were 43 genomes reported from Arkansas, but we were unable to independently confirm whether the designated location referred to the state location of the research team that deposited the sequences or if this referred to the garden locations that were sampled [43]. The GenBank database reports nine isolates from Alabama and Missouri and between one and four isolates for all other states. The locations of the two isolates from California were reported [44] and their geolocations are noted (BA2018: 35.41228, −119.09160; WA2017: 35.59325, −119.29216). Two additional isolates were from Wilmington, Delaware (39.66507, −75.74429), and Washington, District of Columbia (38.91166, −76.96934) (Figure 1A). These isolates are referred to as RRV\_SW\_DE and RRV\_RF\_DC. In this study, we added a suffix USA state abbreviation to each named isolate, and from here onwards, isolate names will be described without the initial ‘RRV’. The isolates’ full names (including ‘RRV’) are given in Supplementary Table S1 (which also includes the GenBank accession numbers for all RRV genome segments) and all figures. Given the low numbers of isolates per state, except for Arkansas, there are unfortunately insufficient data to determine whether the geographic distribution is a factor in the phylogenetic relatedness of these isolates [12].

The source host plants for the 91 isolates reported by Katsiani et al. (2020) [12] and the BA2018 isolate from CA [40] were only listed as ‘*Rosa* sp.’. The WA2017\_CA isolate was reported from *Rosa* sp. ‘Veteran’s Honor’ (a hybrid tea rose) [40]. The RF\_DC was taken from a sample of ‘*Rosa* × *fortuniana*’, which is believed to be a *R. banksiae* × *R. laevigata* hybrid that has nearly thornless canes and is known as a climbing shrub rose grown on fences or walls. The RF\_DC isolate was obtained from a plant in the USNA’s National Herb Garden’s “Historic and Species Rose Garden” that was demonstrated by high throughput sequence (HTS) analysis to be coinfecting with blackberry chlorotic ringspot virus and rose spring dwarf-associated virus (other viruses to be published elsewhere). The SW\_DE isolate was obtained from another climbing rose ‘Stormy Weather’ that was only infected with RRV and was obtained from an RRD-resistance research field. Notably, the reference genome reported in NCBI GenBank is derived from two different isolates and was not included in this study [12]. Another isolate from Oklahoma (NVWA6166361/OK-1) was reported in GenBank in 2021 but was not included in this study because of differences among the noncoding sequences for each segment that led us to be uncertain whether these were complete genome sequences.

ML analysis was performed using concatenated genome segments (Figure 1B). A prior phylogenetic study [8] of the genus *Emaravirus* identified four clades (A, B, C, and D). RRV was grouped into clade A along with actinidia chlorotic ringspot associated virus (AcCRaV), pigeon pea sterility mosaic virus 1 (PPSMV-1), pistacia virus B (PVB), and blackberry leaf mottle-associated virus (BLMaV). The concatenated genomes of these four virus species were used as outgroups. We identified six major groups of isolates based on their relationship to the first node above the root and identified two subgroups, A and B, within the largest group, group 6 (Figure 1B). For example, group 1 has seven RRV isolates from Arkansas and Missouri, which are in the Central Region of the USA (Figure 1C). Group 2 has eight isolates and includes the newly sequenced isolate RF\_DE alongside isolates from Arkansas, Louisiana, and Michigan. Group 3 has nine isolates from New York, Michigan, and Kentucky. Group 4 has seven isolates from Arkansas and Missouri. Group 5 has nine isolates, mainly from Arkansas, with one from Michigan. Groups 2 through 5 include isolated genomes from the Eastern and Central Regions of the USA. Group 6

has 54 isolates and is divided into subgroup A from Arkansas, Kansas, Illinois, Missouri, Oklahoma, Pennsylvania, and Texas and subgroup B, which includes isolates from across the USA, including the Eastern, Central, and Western regions. Since commercial roses are typically grown from rooted cuttings and shipped nationally, the distribution of related isolates across geography suggests that some plants may have derived from the same source materials before shipping to different states or become infected by viruliferous mites in the gardens where they were planted.



**Figure 1.** RRV isolates and nucleotide diversity. (A) RRV isolates were obtained from 16 states, and one isolate is from Washington DC. The colored circles serve as a color legend for panel C to easily compare the geographical location in the map and color-coded isolates in the phylogenetic tree. The number in each colored circle indicates the number of isolates from each state location. Notably, the

Delaware, Washington DC, and Maryland isolates are featured on the right side of the map. (B) Maximum likelihood (ML) analysis of concatenated genome segments with four emaravirus species as an outgroup. The branch length scale bar is presented at the bottom and branches with values >0.5 are labelled. Tip labels are colored according to the circles in panel (A). Groups were assigned to assist the reader. (C) Top illustration of the concatenated RRV genome segments. Open boxes represent the open reading frame (ORF), and the amino acid lengths of the encoded proteins are presented. The viral RNA-dependent RNA polymerase (RdRp), glycoprotein (GP), nucleocapsid protein (NP), and movement protein (MP) are identified above the box. Question marks identify genes of unknown functions. The scale below the boxes provides a reference for the RNA segment lengths that are concatenated and identified by the thick gray bar. The nucleotide lengths of each RNA are provided. Dotted lines represent the termini separating each segment. The graph features the nucleotide diversity along the length of the concatenated sequence.

To examine the genetic variation across isolates, genome segments were concatenated, extending a length of 15,524 nt. The linked protein sequences consisted of 4847 amino acids (Figure 1C). The statistical estimate of nucleotide diversity ( $P_i$ ) for RNA segments 1, 2, and 3 ranged from 0.000 to 0.014, which is a small variance among the segments that encode the core replicase and structural proteins (Figure 1C). The nucleotide diversity of RNA segments 4 through 7 ranged from 0.000 to 0.031, representing a broader range of variability, with RNA5 and RNA7 showing the most diversity. The genetic differences among the concatenated genomes were analyzed using DnaSP (Table 1). There are 1648 polymorphic sites (S) and 1703 mutations ( $\eta$ ), indicating that there were multiple changes among isolates, and some changes occur at the same nucleotide locations. DnaSP calculated the average pairwise differences among concatenated sequences to be 113 [12]. The overall nucleotide diversity ( $\pi$ ) was 0.0073, which is >0.005, indicating low genetic diversity among RRV isolates.

**Table 1.** Analysis of polymorphic sites and mutations for concatenated and individual RRV segments (all isolates).

Sequence	N	S	$\eta$	$\pi$	K	Tajima's D *	Fu and Li's D **†	Fu and Li's F **
Concatenated	95	1648	1703	0.0073	113	−2.2586	−4.4978	−4.0345
RNA1	95	769	792	0.0060	42	−2.4870	−4.2207	−4.1514
RNA2	95	218	221	0.0043	9	−2.6536	−4.4239	−4.3939
RNA3	95	79	83	0.0073	7	−1.8977	−4.6653	−4.2174
RNA4	95	95	98	0.0080	9	−1.8117	−4.4080	−3.9841
RNA5	95	209	223	0.0149	239	−1.5531	−3.6656	−3.3067
RNA6	95	146	154	0.0069	9	−2.3190	−4.9792	−4.6142
RNA7	95	133	134	0.0104	14	−1.4900	−3.8783	−3.4378

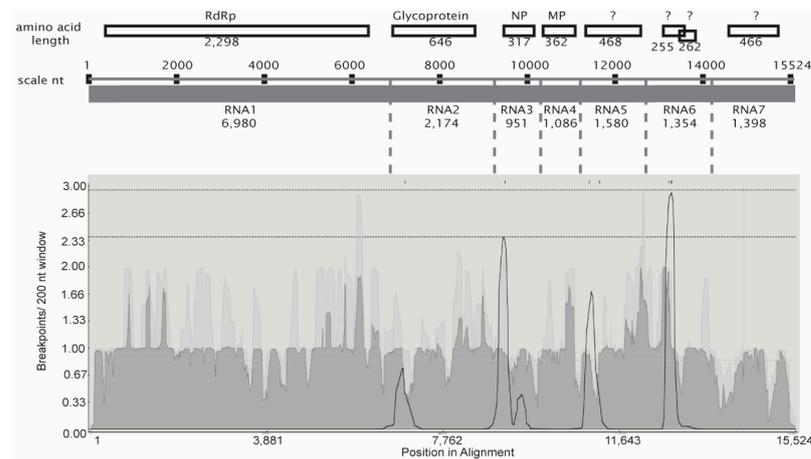
N, number of genomes; S, total number of polymorphic sites;  $\eta$  (Eta), total number of mutations;  $\pi$ , nucleotide diversity estimated by the average pairwise differences between all sequences and based on all sites; K, average pairwise nucleotide differences; \*  $p < 0.01$ ; \*\*  $p < 0.02$ ; † test using a sliding window length of 100 and step size of 25.

The uneven distribution of polymorphic sites, mutations, overall nucleotide diversity, and nucleotide differences suggests that the mutations are independent. Given that the genome segments vary in size from 6980 nt for RNA1 to 1086 nt for RNA4, the polymerase error frequency among segments may vary and there may be different selection pressures acting on viral gene products. For RNA1 through RNA6, the total number of mutations ( $\eta$ ) is greater than the number of polymorphic sites (S), i.e., some mutation sites have different substitutions in different isolates. RNA1 has the greatest number of mutations, 792. RNA3 and 4 have the lowest number of mutations, 83 and 98, respectively. RNA7 has only one more mutation than polymorphic sites (Table 1). The nucleotide diversity ( $\pi$ ) values for RNA5 and RNA7 stand out as the highest, 0.0149 and 0.0104, respectively,

while RNA2 is the lowest, 0.0060. The molecular diversity patterns were calculated using Tajima's *D*, as well as Fu and Li's *D* and *F* statistical tests (Table 1). For the concatenated sequences and each segment, tests were performed using the 'sliding window' option, and no positive values were found. The neutrality tests were significantly negative, indicating that the mutation rate is not constant and may be best explained by recent population expansion.

### 3.2. Evidence for RRV Segment Reassortment

To assess the potential for reassortment of genome segments, the concatenated sequences were analyzed for recombination breakpoints using RDP5.0, and events were validated using SplitsTree and SimPlot++. The recombination breakpoint distribution plot (Figure 2) also identifies potential breakpoints near the junctions of RNA segments. The major and minor parents, as well as the breakpoint positions and the statistical significance for the RDP tests, are provided in Table 2. Three breakpoint positions, 6928, 6982, and 7184, are neighboring the junction of RNA1 and RNA2 in the concatenated sequences. According to this plot, the breakpoints adjoining RNA1 and 2 are not represented with high confidence. The breakpoint position 9134 is at the junction of RNA2 and RNA3. The breakpoint position 12,801 lies near the junction of RNA5 and 6.



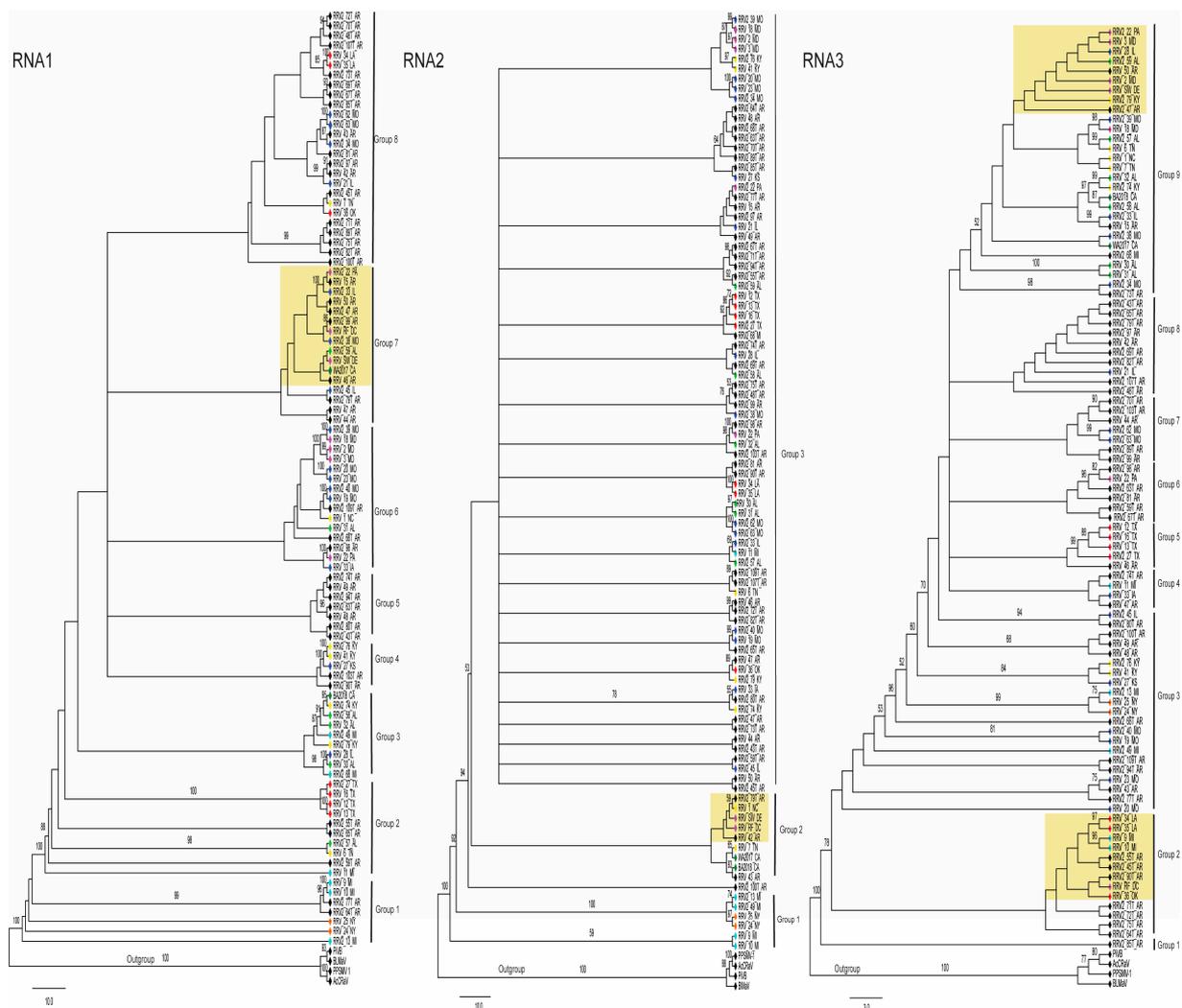
**Figure 2.** Breakpoint distribution plot presented below the illustrated concatenated genome. The dotted lines across the top of the plot identify the 95% and 99% confidence thresholds. Diagrammatic representation of the concatenated sequences, *p* values, nucleotide lengths, and termini of each segment as in Figure 1, but used here as a reference to understand the breakpoint locations in the outputs.

**Table 2.** RRV reassortment detected with RDP and SimPlot++ \*.

Minor Parent	Major Parent	Breakpoint Positions	RDP Analysis (RGCMB SPL3) **
RRV_12_TX and RRV_13_TX	RRV_9_MI and RRV_10_MI	1–9136	$6.58 \times 10^{-13}$ to $1.36 \times 10^{-54}$
RRV2_13_MI, RRV_24_NY, and RRV_25_NY,	RRV_44_AR and RRV_48_AR	6982–12,801	$5.68 \times 10^{-5}$ to $2.09 \times 10^{-45}$
RRV_27_KS, RRV_42_AR, and RRV2_69T_AR	RRV_9_MI and RRV_10_MI	6928–9134	$3.97 \times 10^{-9}$ to $5.82 \times 10^{-20}$
RRV_42_AR	RRV2_45T and RRV2_90T	7184–9134	$1.18 \times 10^{-5}$ to $3.91 \times 10^{-12}$

\* RDP output was validated by SplitsTree, SimPlot++, and BOOTSCAN using the Kimura 2-parameter distance model with a window length of 1000 and step of 200. \*\* Analysis was performed on the concatenated alignment of all seven segments using the RDP5 v. 5.34 software. Methods supporting significant recombination or reassortment signals for five or more tests are indicated with the range of significance (*p* values) obtained: R (RDP), G(GENECONV), C (Chimaera), M (MaxChi), B (BOOTSCAN), S (SISCAN), *p* (PhylPro), L (LARD), and 3 (3SEQ).

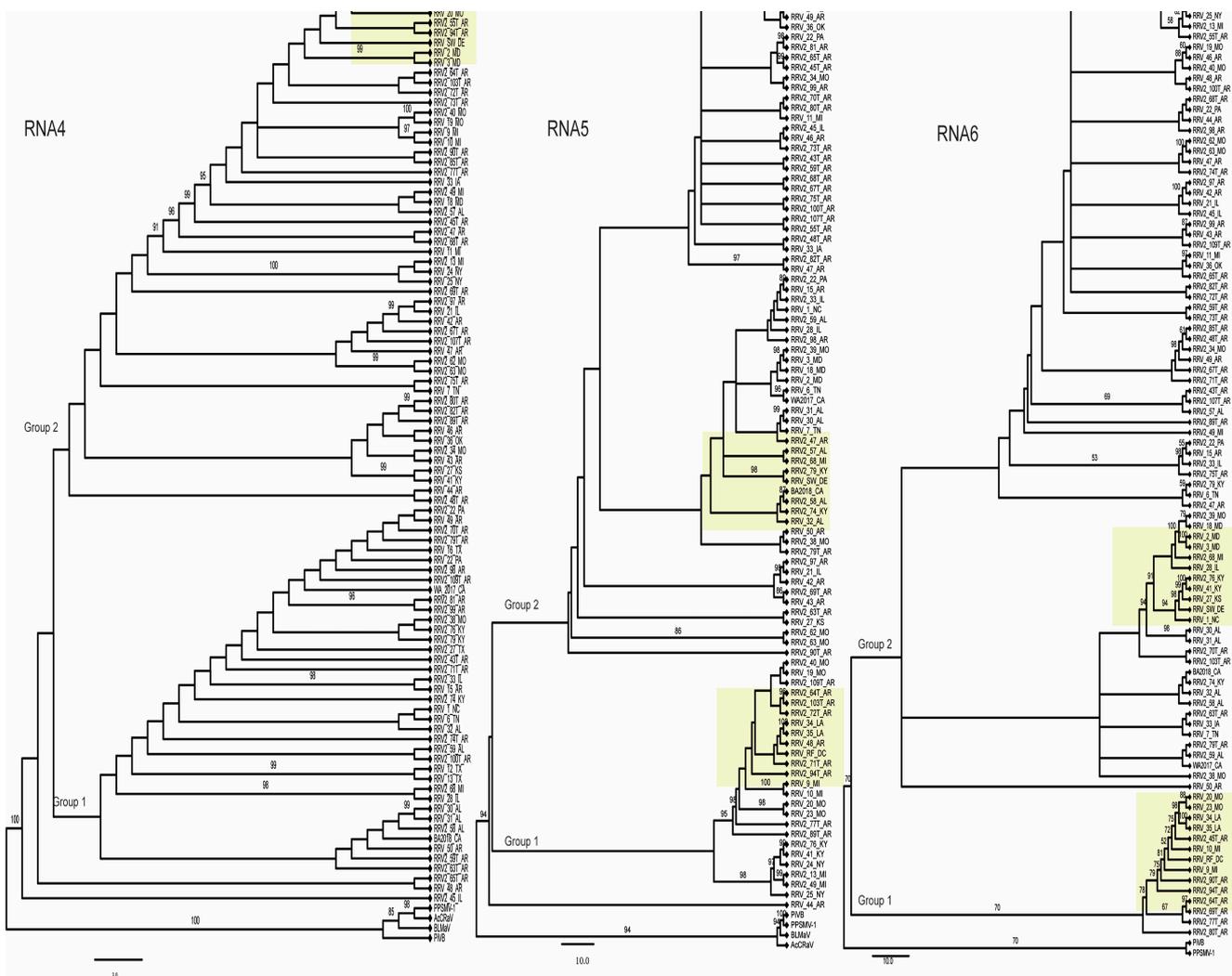
To assess the potential for genetic exchanges involving RNA2 and RNA3, we examined the topology of the phylogenetic relationships between RNA1, 2, and 3. Trees were rooted with the corresponding genome segments of the same virus species used as outgroups in the prior analysis (Figure 3). The branch patterns of these three trees reflect different evolutionary pressures acting on these genome segments. Certain groups of isolates consistently cluster across these trees, such as the four isolates from Texas, 12\_TX, 13\_TX, 16\_TX, and 2\_27\_TX. Another example is three isolates, 2\_13\_MI, 24\_NY, and 25\_NY, which are identified as minor parents in Table 2, meaning that these likely did not arise by segment reassortment. Their RNA1, RNA2, and RNA3 segments are consistently represented as closely related branches, which supports the notion that these isolates did not arise by reassortment of segments involving other isolates. This is contrasted by the three isolates from Maryland known as 2\_MD, 3\_MD, and 18\_MD, which RDP5 predicts arose by segment reassortment. The RNA1s of these three Maryland isolates are in group 6 in the tree (Figure 3) and cluster with three isolates from Missouri, 20\_MO, 23\_MO, and 2\_39\_MO. The RNA2s for the Maryland isolates cluster in a subclade at the top of the tree in group 3 with 20\_MO, 23\_MO, and 39\_MO, and two isolates from Kentucky, 41\_KY and 76\_KY. The branch positions of RNA3 for these Maryland isolates is more distinct with 2\_MD located between SW\_DE and 50\_AR, while 3\_MD is between 28\_IL and 2\_22\_PA, and 18\_MD is near 2\_39\_MO, 6\_TN, 7\_TN, 1\_NC, and 2\_57\_AL (Figure 3).



**Figure 3.** Phylogenetic relationships of RRV RNAs 1, 2, and 3 segments (left to right). The phylogenetic relationships were inferred using the maximum likelihood method (see Section 2). Bootstrap values

>50% are shown (10,000 replicates) next to the branches. Branches leading to significant clusters of isolates were assigned a group number to assist in the explanation of the results. The colored diamonds at the branch tips identify the locations as in Figure 1. The trees are drawn to scale. The yellow box surrounds the clusters of isolates that include the two new isolates, SW\_DE and RF\_DC.

We report the sequence for two new isolates, RF\_DC and SW\_DE, which are from the Eastern Region of the USA. RDP5 analysis and the ML trees in Figures 3 and 4 suggest these may have arisen by segment reassortment. To feature these isolates more prominently in the trees the subgroups of branches containing these two isolates are highlighted in yellow in Figures 3 and 4. For RNA1, both new isolates appear in group 7, but for RNA2 at different positions in the tree, in group 2. For RNA3, RF\_DC is in group 2, and SW\_DE is in group 9. (Figure 3). For RNA4, both fall closer together within group 2 at the top of the tree, while for RNA5 and RNA6, RF\_DC is in group 1 and SW\_DE in group 2 (Figure 4). Their nearest neighboring isolates also differ between all of the RNAs (Figures 3 and 4).



**Figure 4.** Phylogenetic relationships of RRV RNAs 4, 5, and 6 segments. The phylogenetic relationships were inferred using the maximum likelihood method (see Section 2). Bootstrap values >50 are shown (10,000 replicates) next to the branches. Branches leading to significant clusters of isolates were assigned a group number for reader orientation. The trees are drawn to scale. The yellow boxes surround the clusters of isolates that include the two new isolates, SW\_DE and RF\_DC. Trees were rooted as before except for RNA6, where there are only two species that can serve as outgroups. The branch patterns of these trees reflect different evolutionary pressures acting on these genome segments.



in green and the goldenrod color indicates the locations of substitution or indel mutations in the sequence. The legend represents the colors ascribed to different nucleotide substitutions in the alignment. These changes are also shown in the 1, 2, 3, and 4 boxes above the consensus bar. (B) The indel changes in regions 1 through 4 are elaborated. Each box contains the identity of the (non)recombinant genomes associated with each mutation. Some sequences have more than one change in the same region. Sequences with mutations in more than one region are highlighted in orange.

Substitutions and indel mutations throughout the four RNA6 regions of all 95 isolates were analyzed, and 46 sequences showed changes from the consensus (Figure 5B). In total, we identified single indels in region 1 of seven isolates, which is bracketed by nt positions 218 and 254. There were 10 isolates with changes in region 2 (nt positions 347–367), 3 isolates with changes in region 3 (nt positions 411–445), and 27 isolates with changes in region 4 (nt positions 445–515). Compensatory mutations may exist in other areas of RNA6 outside of these four featured regions. Across all regions, the most obvious changes were in the addition or deletion of a U along a poly U stretch (Figure 5B). Only one sequence, 2\_55T\_AR, does not have a compensatory mutation in any of the four indel regions, and its P6a and P6b ORFs are the typical lengths. Indel region 2 extends from nt position 347 to 367 and has a single U(6) tract that has an extra U or UAAA in 10 isolates. Region 4 is particularly interesting because it overlaps the translation stop codon for the ‘typical length’ P6a.

Alignment of the 95 P6a sequences showed that the majority are 61 amino acids in length. Figure 6 presents a snapshot of a larger alignment featuring isolates that deviate from the typical amino acid length of P6a. Three RRV isolates, named 2\_59\_AL, 7\_TN, and 2\_82T\_AR, encode truncated proteins that are 29 amino acids in length. Four isolates (RF\_DC, 27\_KS, 41\_KY, and 76\_KY) encode proteins that extend 75 to 94 amino acids in length.



**Figure 6.** Analysis of twenty-five P6a proteins. Alignment demonstrates that most P6a proteins are 61 amino acids in length and identifies truncated and extended sequences. Amino acid letters are colored to identify similar residues. The \* identifies the translation stop codons.

### 3.4. Amino Acid Changes within RRV Coding Sequences and Proteins

Reassortment between isolates can lead to the exchange of gene segments between viruses that can result in progeny viruses that are genetically distinct and can be associated with antigenic shift and the emergence of new pandemic strains. Understanding the patterns of amino acid changes is important to the identification of virus isolates that may be distinguishable by serological methods or virulence on different hosts. The core virion component proteins are RdRp, GPP, and NP, and these contain 31, 20, and

2 parsimony-informative residues, respectively. MP, P5, P6a, P6b, and P7 have between 3 and 29 informative changes (Table 3).

**Table 3.** Amino acid sequence variation across coding regions of 95 isolates.

Sequence	Amino Acid Length	Parsimony-Informative Amino Acids
RdRp	2297	31 (1%)
GPP	645	20 (3%)
NP	316	2 (0.6%)
MP	361	9 (2%)
P5	467	29 (6%)
P6a	61–94	3 ( $\leq 0.05\%$ )
P6b	233	7 (3%)
P7	465	23 (5%)

We mapped single amino acid variations to the phylogenetic trees derived from ML analysis of 95 isolates and used color coding of the tree branches to represent the most common changes from the consensus. Branches represented as black lines are identical to the consensus sequence. Various colored dots were used to identify additional amino acid changes that could not be represented by the colored branches (Figure 7A–D). As expected, the trees show clustering of amino acid positions with the same variant residues among closely related isolates. The largest number of amino acid changes (expected for the NP) are distributed at the base of each tree, while more distant branches mainly represent the consensus sequence.

Amino acid changes that represent a clade (and potentially with a few outliers which share the same residue but fall into other clades) are defined here as a ‘group’. Isolates sharing parsimony-informative sites but not predominantly in a single clade are not defined as a ‘group’ but are mentioned in other terms. For example, seven isolates at the base of the RdRp ML tree share a common T<sub>2036</sub>K change (magenta branches). Added mutations divide the branches in this group in a manner that suggests there are three diverging subgroups. The first subgroup has only the single T<sub>2036</sub>K mutation, the second subgroup of two PA and AR isolates has an additional V<sub>90</sub>A/I, and the third subgroup has two MI isolates with additional T<sub>848</sub>A, N<sub>851</sub>S, and F<sub>1814</sub>S (Figure 7A). There is a neighboring clade of four isolates that appear to be defined by I<sub>2088</sub>V (green branches), and there are similar groups of two to five isolates with at least one common amino acid change. One large group of 21 isolates (red branches) shares a common A<sub>1742</sub>S substitution (Figure 7A). This S<sub>1742</sub> group could be further characterized as having four subgroups: two MO isolates that have a G<sub>2210</sub>A, four AR isolates that have an A<sub>2251</sub>V/T mutation, two LA isolates with Y<sub>514</sub>H, and three IL and AR isolates with M<sub>1170</sub>I mutations, of which the 42\_AR isolate has an additional D<sub>309</sub>N mutation.

This ML tree further demonstrates that isolates from the same state may not be closely related. This is exemplified by the distinct grouping of MO isolates based on four amino acid changes: V<sub>1659</sub>I, S<sub>2033</sub>N, A<sub>1742</sub>S/G<sub>2210</sub>A, and R<sub>2173</sub>K. The 20\_MO, 23\_MO, and 2\_48T\_AR share V<sub>1659</sub>I; 34\_MO, 38\_MO, and 2\_100T\_AR share S<sub>2033</sub>N; and 19\_MO and 2\_40\_MO share R<sub>2173</sub>K. Another example is the four MI isolates. The 9\_MI and 10\_MI isolates pair together with four common changes mentioned previously, while the 11\_MI isolate clusters with four TX isolates sharing a common V<sub>139</sub>I. Isolate 13\_MI clusters with two NY isolates and 59T\_AR with a common pair of mutations, D<sub>2082</sub>E and I<sub>2242</sub>V. The 13\_MI and two NY isolates also share another two mutations, H<sub>69</sub>Y and E<sub>1462</sub>D, but only 13\_MI and 25\_NY share the V<sub>2215</sub>I mutation.

The ML tree of the GPP sequences shows clustering of amino acid changes among isolates at the base of the tree, but the ordering of branches does not match the RdRp ML

tree (Figure 7B). There are twenty parsimony-informative amino acid changes occurring in 27 isolates (Table 3), and 8 of these isolates have a V<sub>419</sub>I/A change highlighted as red branches in the tree (Figure 7B). The next most abundant mutations are M<sub>639</sub>V/I and N<sub>478</sub>D, each occurring in six isolates. Isolates with the V<sub>419</sub>I/A change have between zero and six other substitutions, including S<sub>152</sub>G, I<sub>182</sub>V, T<sub>190</sub>I, I<sub>326</sub>V, N<sub>478</sub>D, and/or M<sub>639</sub>V/I. Two isolates from MO have the V<sub>354</sub>I and N<sub>478</sub>D changes. One isolate named 47\_AR has the M<sub>639</sub>V/I mutation alone, and another isolate, 33\_IA, has two additional mutations, T<sub>21</sub>K and V<sub>189</sub>I. Isolates retaining the V<sub>419</sub>I/A substitution are primarily from NY and MI, with only one isolate from NC. Most amino acid changes occur at the base of the tree, with only six substitutions occurring in pairs of distant isolates (Figure 7B). The NP tree can be described based on two amino acid positions creating three type groups, of which 66 isolates have V<sub>61</sub> and V<sub>93</sub>, 26 isolates have V<sub>61</sub> and I<sub>93</sub>, and 3 isolates have I<sub>61</sub> and V<sub>93</sub> (Figure 7C).

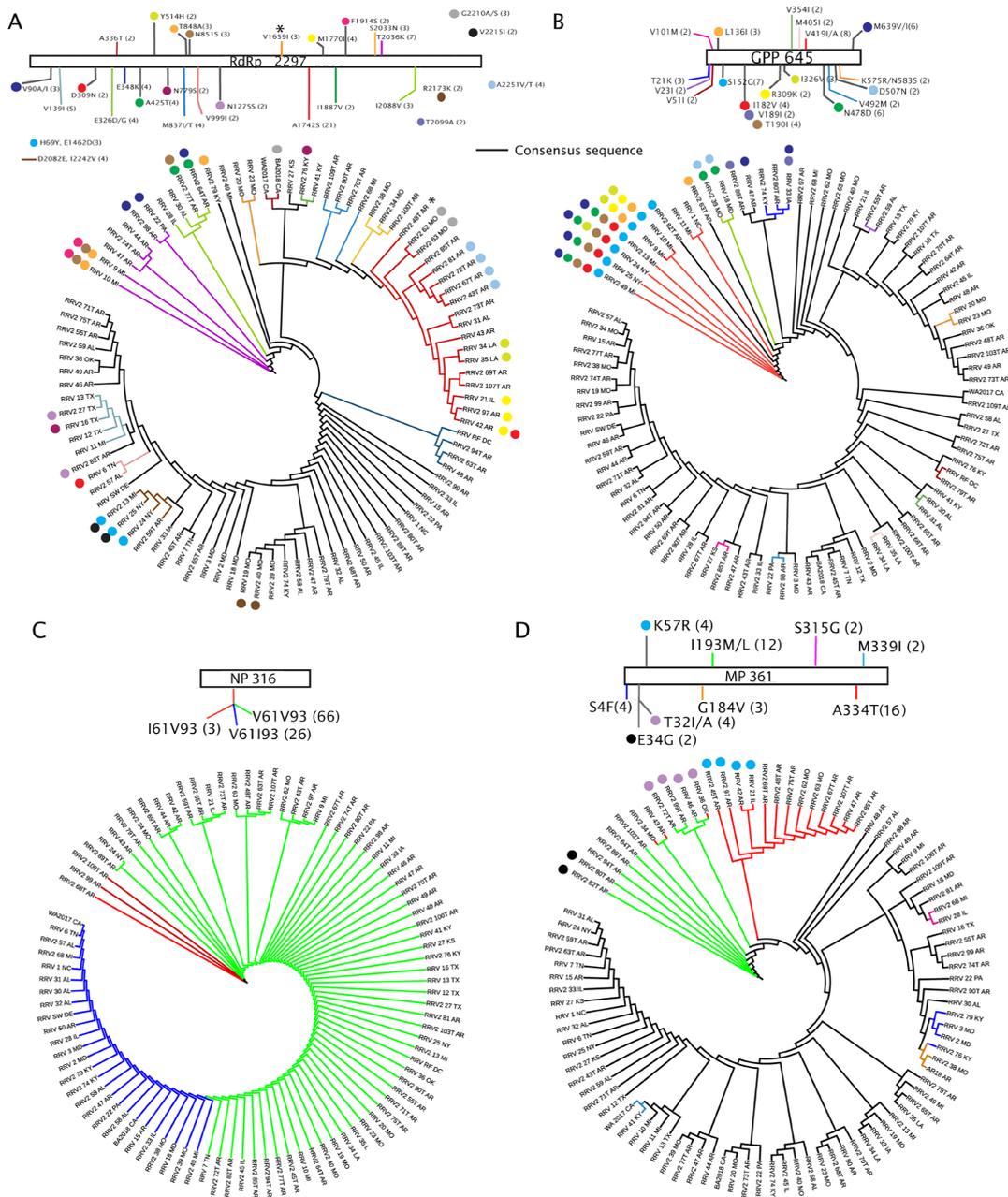


Figure 7. ML trees of sequences encoded by RNA1 through RNA4 with parsimony-informative amino acid changes. Open bars at the top of each panel provide a legend for each tree, the name of

the protein, and its amino acid length. (A) RdRp, (B) GPP, (C) NP, and (D) MP. The approximate locations of amino acid changes are represented along the bar as colored lines or colored circles. The consensus amino acid and its position are listed followed by the substituted amino acid. The numbers in parentheses identify the number of isolates with this amino acid change from the consensus. The colored branches in the phylogenetic trees correspond to the colored lines in the legends. Colored dots along the ML trees also identify amino acid changes represented by the same dots in the legends. Black branches without associating dots represent the isolates that conform to the consensus sequences.

The only described MP is encoded by RNA4, and it is not currently known if there are other movement-associated proteins encoded by other viral genome segments. Across isolates, the viral MP has nine parsimony-informative amino acid changes (Table 3). As for the GPP, most informative changes occur in isolates that reside at the base of the ML tree. There are twelve isolates with the I<sub>193</sub>M/L change (green branches) and sixteen isolates with the A<sub>334</sub>T change (red branches) (Figure 7D). Notably, there are three isolates represented by green and red branches with both I<sub>193</sub>M/L and A<sub>334</sub>T changes (2\_35\_MO, 43\_AR, and 36\_OK). These M/L<sub>193</sub> and T<sub>334</sub> groups have deeply rooted branches and are adjacent in the ML tree. Among the I<sub>193</sub>M/L isolates are two isolates with an E<sub>34</sub>G mutation and a cluster of four isolates with a T<sub>321</sub>A mutation (Figure 7D). Among the T<sub>334</sub> group are four isolates with a K<sub>57</sub>R mutation.

P5 and P7 are 467 and 465 amino acids in length, respectively, and were recently reported to be homologs of unknown function [8]. P5 and P7 have accumulated 29 and 23 parsimony-informative changes, respectively, which are an abundance of changes compared to other viral proteins and amino acid changes (Table 3). The ML tree composed of P5 sequences (Figure 8A) has four potential isolate groups. The first group of 24 isolates (blue branches) is defined by the combination of R<sub>280</sub>K, L<sub>452</sub>I, and Q<sub>467</sub>H substitutions. A second group of 36 isolates is represented by L<sub>302</sub>I (red branches). These two groups share six isolates, which are represented in Figure 8A by blue and red branches indicating that they share the R<sub>280</sub>K, L<sub>452</sub>I, Q<sub>467</sub>H, and L<sub>302</sub>I changes. A subgroup of the K<sub>280</sub>/I<sub>452</sub>/H<sub>467</sub> group also has the L<sub>302</sub>I change (blue plus red branches). Isolates lacking these mutations are represented by black branches. Eighteen members of the K<sub>280</sub>/I<sub>452</sub>/H<sub>467</sub> group have the combination of R<sub>14</sub>K/I, I<sub>297</sub>V, N<sub>440</sub>S (represented by deep red dots), and K<sub>250</sub>N/D (royal blue dot) substitutions (Figure 8A). This K<sub>280</sub>/I<sub>452</sub>/H<sub>467</sub> group also has between four and eight additional changes. Six isolates representing the combination of R<sub>280</sub>K, L<sub>302</sub>I, L<sub>452</sub>I, and Q<sub>467</sub>H substitutions (blue plus red branches) have two additional amino acid changes, K<sub>250</sub>N/D and K<sub>384</sub>R. These six isolates are from NY, MI, and KY. K<sub>250</sub>N/D is conserved with the isolate group represented by the blue branches. Regarding the I<sub>302</sub> group, there are nine isolates with additional amino acid changes other than the previously mentioned K<sub>250</sub>N/D and K<sub>384</sub>R (Figure 8A). Interestingly, there are 20 isolates with N<sub>65</sub>S, which are distributed among isolates in each of the four type groups represented by colored or black branches in the tree (Figure 8A). Four unique amino acid changes that occur only in isolates represented by black branches include N<sub>283</sub>S, F<sub>460</sub>C/Y, N<sub>345</sub>D, and S<sub>346</sub>N (Figure 8A). The N<sub>283</sub>S is also present in two isolates represented by red branches which have the L<sub>302</sub>I change: 2\_79\_KY and SW\_DE. The N<sub>345</sub>D change also occurs in 2\_58\_AR, which is also represented by a red branch identifying the L<sub>302</sub>I change.

Interestingly RNA6 encodes two overlapping proteins, 6a and 6b, and the order of branch isolates in these trees is different. P6a and P6b have three and seven informative amino acid changes (Table 3 and Figure 8). The accumulation of certain mutations in clusters along the phylogeny tree are also different, suggesting different host-selective pressures are acting on these proteins (Figure 8B,C). P6b has no parsimony-informative changes from the consensus sequence in the first 100 amino acids. All changes appear closer to the C-terminus of the protein (Figure 8B). The P6b ML tree shows seventeen isolates with a T<sub>199</sub>A mutation, which are highlighted as red branches. Interestingly, twenty-six isolates have mutations at L<sub>178</sub>. Of these 26 isolates, there are 13 with a W/Q change and 13 with a C/G change, and these occur at the base of the tree. Two deeply rooted branches

representing isolates with the T<sub>199</sub>I/A (red branches) include six isolates with the L<sub>178</sub>W/Q mutation. Two isolates at the exact base of the tree have T<sub>199</sub>I/A, L<sub>178</sub>W/Q, D<sub>218</sub>N, or H<sub>231</sub>Q (Figure 8B). The D<sub>218</sub>N and H<sub>231</sub>Q mutations also occur in a cluster of four isolates from TX, which are more distant in the tree. Five isolates with the T<sub>199</sub>I/A mutation also have the L<sub>178</sub>C/G mutation, and two have an R<sub>121</sub>K mutation with or without the C/G change at position 178 (Figure 8B). The P6a proteins in this study are more often 61 amino acids in length but seven isolates are either truncated or extended. Three isolates have three neighboring amino acid changes, TKT<sub>28–30</sub>PRH, followed by a stop codon (\*), and these are the shortest 6a proteins among the ninety-five isolates (Figures 8C and 6). There are three isolates with a single change at T<sub>28</sub> that is P/N/I and is not followed by a premature stop codon. The most common change is I<sub>24</sub>T, occurring in 10 isolates that have shallow branches. There is one isolate with R<sub>29</sub>M, two with H<sub>5</sub>Y, and three isolates from MI and TX with I<sub>24</sub>K mutations. Four isolates extend 15 to 33 amino acids beyond the most common 61 amino acid length (Figures 8C and 6).

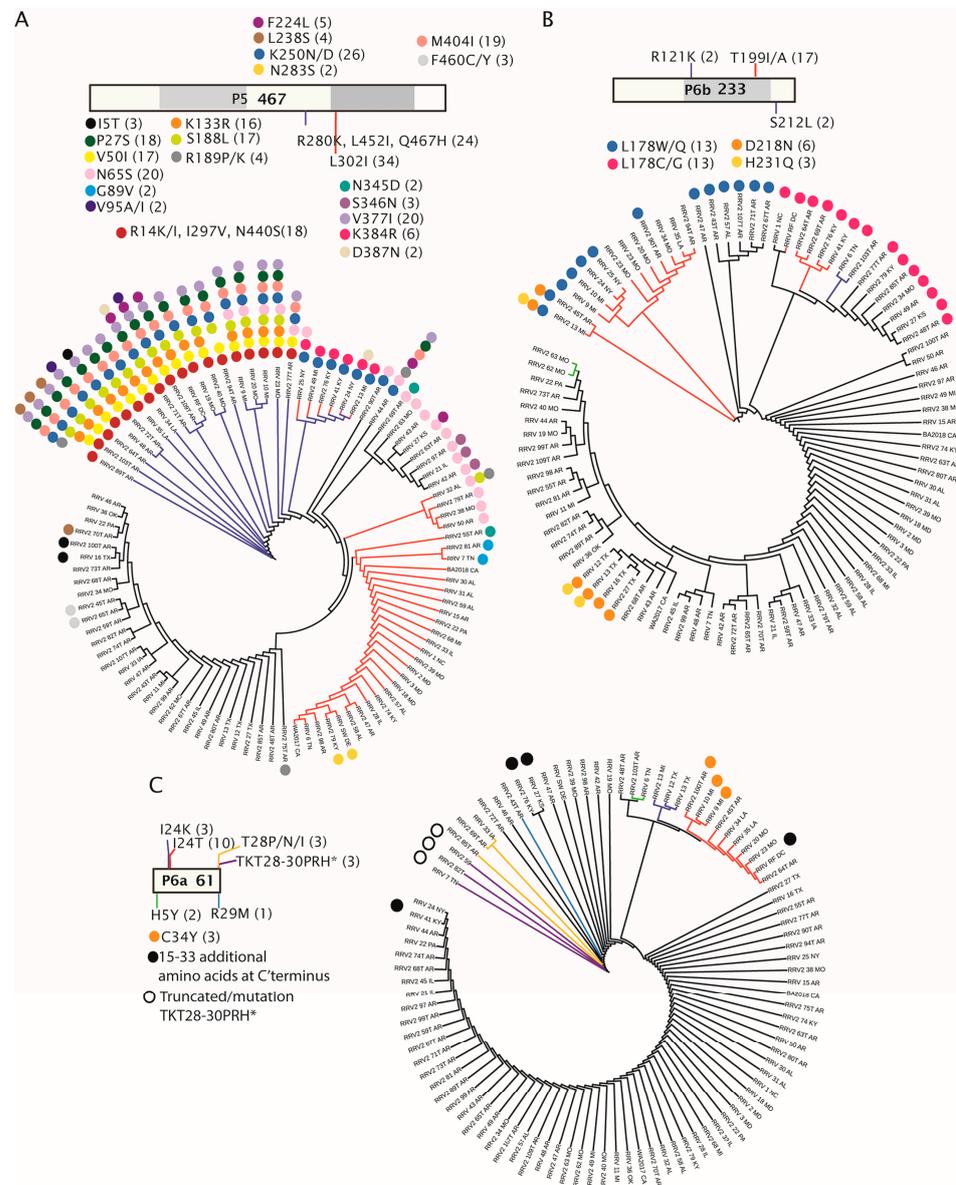


Figure 8. ML trees of sequences encoded by RNA5 and RNA6 with parsimony-informative amino acid changes. Open bars at the top of each panel provide a legend for each tree, the name of the protein,

and the amino acid length of the protein. These open bars are divided with grey boxes to represent subdivision along the sequence of 100 amino acids. (A) P5, (B) P6b, (C) P6a. The amino acid changes are listed in order below each subdivision to feature their locations along the sequence. The estimated locations of amino acid changes are represented along the bar as colored lines or colored circles. The numbers in parentheses identify the number of isolates with this amino acid change from the consensus. The colored branches in the phylogenetic trees correspond to the colored lines in the legends. Some isolates are represented by two color branches, indicating that two mutations define that cluster. Colored dots along the ML trees also identify amino acid changes represented by the same dots in the legends. Black branches without associating dots represent the isolates that conform to the consensus sequences.

P7 has a complex set of variant residues that are distinct from the changes reported in P5 (compare Figure 8A, Figure 9), suggesting that they evolved separately. Three major divisions of the P7 tree are influenced by 10 informative amino acid changes that occur in 29 to 41 isolates (Figure 9). The most notable substitutions in the ML tree are the R at position 235 in 54 isolates and K at the same position in 41 isolates. The K<sub>235</sub> group, which is the first major group at the base of the tree, includes 36 isolates that also have changes represented by Y<sub>101</sub>C, I<sub>316</sub>V, R<sub>350</sub>K, S<sub>351</sub>N, and N<sub>413</sub>D. Two isolates at the base of the tree, 41\_KY and 76\_KY, have two additional mutations that separate them from other members of the K<sub>235</sub> group, which are D<sub>342</sub>N and N<sub>448</sub>S (Figure 9). There are nine K<sub>235</sub> isolates that extend from deep branches but lack one or more of the Y<sub>101</sub>C, I<sub>316</sub>V, R<sub>350</sub>K, S<sub>351</sub>N, and N<sub>413</sub>D changes (Figure 9). A second major division of the tree is represented by blue branches in Figure 9, and we identify these isolates for the N at position 81. The N<sub>81</sub> group has few other changes. It is interesting to note that Y<sub>101</sub> only co-occurs with T<sub>81</sub> and that N<sub>81</sub> never occurs with C<sub>101</sub> in the same isolate, suggesting that these amino acid positions influence each other. Only eight isolates of this N<sub>81</sub> group have additional differences from the consensus sequence, and 2\_48T\_AR has a K<sub>350</sub>, which is more commonly seen in the K<sub>235</sub> group. The most distant branches of the N<sub>81</sub> group include two isolates from MD, four isolates from TX, and one isolate from AR (46\_AR). The MD isolates have two amino acid changes, R<sub>277</sub>K and T<sub>315</sub>I. The other five isolates from TX and AR have K<sub>183</sub>R, and two TX isolates have E<sub>191</sub>V.

There are two isolates with K<sub>235</sub> at the base of the next cluster group that lack any other change, 2\_71T\_AR and 2\_68\_MI. This next cluster group is largely defined by K<sub>39</sub>. The first member of this group, 11\_MI, has both K<sub>235</sub> and K<sub>39</sub> changes. Two isolates in this K<sub>39</sub> group also have K at position 235 alongside two additional changes known as V<sub>383</sub>A/T or V<sub>185</sub>M. Most isolates in this K<sub>39</sub> group have one or both V<sub>383</sub>A/T or V<sub>185</sub>M, further defining this group. It is interesting to note that K<sub>39</sub> can appear in isolates with either N<sub>81</sub> or K<sub>235</sub>, while these two changes do not occur in the same isolates (Figure 9).



While all RNA segments had examples of transition and transversion mutations [15], RNA6 presented a significant number of indel mutations occurring in the 3' genomic noncoding region. RNA6 encodes in the antigenomic sense two short ORFs, P6a and P6b, which overlap by 76 nt in different reading frames. The 3' genomic noncoding region is unusually long, extending ~500 nt, and has several poly(U) sequences. Several studies have reported that homopolymeric tracts can cause mutations to occur during replication caused by slippage of the replication machinery [16,49,50]. The genomic 3' noncoding region of RNA6 might be a hotspot for variation resulting from polymerase errors due to misalignment, slipping, or stuttering of the RdRp during replication [47]. Perhaps RNA secondary structure in this region may influence the mutation rate relative to other genomic regions. Figure 5 shows that some indels occur alongside other substitution mutations or other indels, suggesting that compensatory mutations are occurring as a molecular rescue mechanism. It is worth speculating that indels may be associated with a fitness cost and that secondary mutations during evolution provide solutions contributing to adaptation. Remarkably, seven P6a proteins among the 95 sequences have truncations or extensions at the C-terminus. Since the functions of P6a or P6b are currently unknown, we cannot yet speculate on the biological role of these proteins or the impacts of these amino acid changes. However, it is known that the environment and the host genetic background often shape adaptive mutations in viral genomes. The acquisition of new hosts can impact virus fitness and mutations may arise to improve survival in heterogeneous environments [51–53]. Most of the RRV genome sequences used in this study were reported in GenBank and the rose cultivars were not identified in the publication or online materials, therefore, little is known about the host genetic background [12]. Given that commercial roses can be diploid, triploid, or tetraploid, it is reasonable to assume that the RRV isolates reported in GenBank were recovered from rose varieties whose diverse genetic backgrounds likely encourage virus adaptation and host range expansion [54–57]. Until now, there is no information concerning quasi-species population structure or real estimates of heterogeneity in natural populations for RRV. Such studies will be necessary to better understand the selection pressures that specifically act on RNA6 sequences.

It is important to address the possibility that some mutations might represent errors in Illumina sequencing [58,59]. Regarding indels in RNA6, the occurrence of compensatory mutations that ensure the integrity of most RNA6 sequences suggests that these are naturally occurring mutations. Unfortunately, mutations can also be induced due to sequencing errors associated with the Illumina technology, which was used for generating the RRV reference genome in 2011 [2]. For example, the NCBI reference genome for RRV RNA1 (HQ871942.1) has a single adenine inserted at position 53, which is adjacent to a natural poly(A) sequence, and a T(10)CT(2) sequence at nucleotide position 1133–1146 that are not found in any of the other 95 isolates that were used in this current study [2,6]. The single adenine insertion significantly alters the position of the translational start codon, causing a 21 amino acid truncation near the endonuclease domain [6], and is another reason NCBI sequences HQ871942–HQ871945 were not used in this study. Ninety-one of the other sequences used in this study were collected in 2016 and were generated using Illumina NextSeq500 or HiSeq3000 devices, which have been reported in test studies to have low insertion and deletion error rates [58]. Two original sequences, RF\_DC and SW\_DE, were reported here and obtained in 2021, also using this improved Illumina technology (see Section 2).

The ML phylogenetic tree of the concatenated genomes of all 95 isolates shown in Figure 1 represents the diversity due to all three factors contributing to the evolution of RRV, namely mutation, recombination, and genome segment reassortment. In contrast, the RDP5-identified apparent recombination breakpoints for some isolates approximately at the junctions of RNA1 and RNA2, between RNA2 and RNA3, and between RNA5 and RNA6, suggest genome segment assortment as previously reported for FMV, PPSMV-1, PPSMV-2, and HPWMOV [17–21]. Other breakpoints identified by RDP5 indicate recombination within individual genome segments in various isolates.

The variations in numbers and constituent members of groups in the ML phylogenetic trees of the individual RNA segments (Figures 3 and 4) more clearly show the multiple impacts of genome segment reassortment and recombination within genome segments, exemplified by the varying positions of new isolates RF\_DC and SW\_DE in the topology of the trees in Figures 3 and 4. It has been demonstrated that identifying and removing recombinant isolates of turnip mosaic potyvirus (TuMV) prior to phylogenetic analysis yielded similar tree topologies from three distinct phylogenetic methods (ML, maximum parsimony, and neighbor joining), thus separating the results of hierarchical evolution from the results of reticulate evolution and correlating well with host adaptation of the non-recombinant isolates [60]. Notably, TuMV has a monopartite genome, so unlike RRV, TuMV is not subject to the effects of genome segment reassortment—and in RRV the combination of recombination and reassortment affects a high proportion of the isolates examined.

Most recombination breakpoints that were identified in the concatenated sequences were located near the junctions of noncoding regions between segments. Recombination and reassortment often occur when virus strains co-mingle in a common host or vector. Recombination and segment reassortment have been reported previously for FMV [17], PPSMV-1, and PPSMV-2 [18]. In addition, the occurrence of two distinct copies of RNA3 in some but not all isolates of HPWMoV, and three phylogenetically distinct clusters of RNA3, are strongly supportive of reassortment [20,21]. The HPWMoV RNAs 1 to 3 are most closely related to those of Palo Verde broom virus, and its RNA4 more closely related to those of Ti ringspot-associated and common oak ringspot-associated viruses [19]. Recombination breakpoints were detected inside RNA1, RNA2, RNA4, and RNA5 of PPSMV-1 and -2, revealing the potential recombination among isolates as well as interspecies recombination. Reassortment of RNA4 and RNA6 was also shown to occur between these two species [18].

Figures 7–9 explore the amino acid changes occurring in viral-encoded proteins in a virus population infecting a single host, roses. Genetic variation, resulting in amino acid changes, is important for adaptation to a new host or vector genotypes. Unfortunately, when these sequences were reported there was no information about host genotypes or vector populations. Future work will need to examine the interaction of virus and host/vector genotypes to understand the influence of host genotype on virus fitness. Among all the viral proteins, the NP has the least number of changes, suggesting that it might be subjected to a stringent bottleneck or have few interactions with host proteins to require adaptive mutations. The NP most likely has less influence than other viral proteins on host range expansion [61]. It is worth speculating that a combination of changes in the GPP and NP might influence particle dimensions. It is also reasonable to consider that the diversity of GPP variants may be important in influencing receptor interactions involving plant cells and insect vectors, playing a key role in horizontal virus transmission [62,63].

Although rose rosette disease (RRD) was first reported in Canada in 1940 [64], and then in Wyoming in 1941 and California in 1943 [65], and the vector *Phyllocoptes fructiphilus* was reported from the wild *R. californicus* in California [66], for many years RRD was primarily reported from populations of the introduced invasive weedy species *R. multiflora* (multiflora rose) [67,68]. As early as the late 1950s, RRD had been reported to cause significant problems in a rose breeding program in Nebraska in proximity to thickets of wild roses including *R. woodsii* and *R. suffulta* (syn. of *R. arkansana* var. *suffulta*), and also plants of the naturalized species *R. rubrifolia* (syn. of *R. glauca*) and *R. eglantheria* (syn. of *R. rubiginosa*), from all of which *P. fructiphilus* was also obtained and demonstrated to transmit RRV [69]. Despite this early report of the effects of RRV on a rose breeding program, the emergence of RRV as a significant problem in cultivated roses broadly coincided with the spread of RRD from the western to the eastern United States primarily through naturalized stands of *R. multiflora* [70], and subsequently the increased popularity of landscape roses planted in home gardens, as well as in large groupings on commercial properties and in highway medians [71]. It is probable that the expansion into commercial cultivars occurred via the wind-blown distribution of mites from wild populations of *R. multiflora* followed by the significant early adaptation to different commercial rose hybrids over a relatively short

period, and by a slower rate of further adaptation as the virus became further distributed among additional cultivated rose hybrids with different genotypes. Interestingly, different relationships among many of the isolates occur across the phylogenetic trees. Both sequence adaptation and genome segment reassortment appear to be important in the evolution of RRV.

The absence of any clear phylogenetic relationship between the geographic distribution of the sequences examined is consistent with the observations of Katsiani et al. (2020) [12] and may reflect the production of different rose cultivars by growers in multiple states and subsequent distribution within the trade. While local spread among varieties is likely to occur after planting in the landscape, some isolates (and *P. fructiphilus*) have probably been distributed to several states in dormant asymptomatic plants. No clear correlations can be drawn in the absence of cultivar and nursery origin data for most of the isolates for which sequences are currently available.

Virus pathotypes were described long before they could be associated with specific sequence differences, e.g., different soybean varieties are used to classify soybean mosaic virus pathotypes [72], or common bean varieties to differentiate bean common mosaic/bean common mosaic necrosis virus pathotypes [73]. Such pathotype differences are attributed to amino acid variations affecting the disease phenotype. Other amino acid changes have been used to describe virus pathotypes that can overcome disease resistance conferred by R genes in a gene-for-gene model or resistance conferred by quantitative trait loci (QTL) [25,74]. Key to the gene-for-gene model of resistance is that the pathogen has the capacity to evolve resistance-breaking pathotypes in a plant–virus coevolution scheme. Examples among positive-strand RNA viruses include tobamoviruses in pepper, potyviruses and potexviruses in potato, cucumoviruses in cucurbits, and other virus–host pathosystems where mutations in viral proteins impair their recognition by R proteins resulting in resistance-breaking strains [25,74–77]. Quantitative resistance to virus infection relies on QTLs decreasing pathogen virulence by reducing infection efficiency or modulating major resistance genes. Surprisingly, P5 has 29 and P7 has 23 parsimony-informative sites, which are comparable to the RdRp and GPP. P5 and P7 have been suggested to be homologs that arose by gene duplication, and it is important to note that the parsimony-informative amino acids in these proteins are quite different within an isolate, suggesting that they may be subjected to different selection pressures. Given the number and complexity of mutations in P5 and P7, it is reasonable to assume that the viral populations undergoing plant/vector bottlenecks do not influence the accumulating variants among these two proteins of unknown function. Since RRV isolates derive from commercial roses that are subjected to breeding processes to incorporate simple genetic traits or QTLs, it is worth considering the possibility that the genetic changes in the P5 and P7 proteins may be influenced by the host genetic background [25]. It is arguable that the heritable diversity of the P5 and P7 proteins may also be important for virus adaptation to changing environments. Given that accumulating mutations are often associated with resistance-breaking (RB) variants or increasing virus–host compatibility with host susceptibility factors, it is worth considering that P5 and P7 proteins are important virulence determinants.

RRV P6b has been identified as an ‘ABC protein’ with homologs in three of four proposed clades of emaraviruses [8], which may have RNA-silencing suppression activity and be involved in pathogenicity [8,9,78]. It will therefore be of considerable interest to compare the pathogenicity of RRV isolates differing in the length and amino acid sequence of P6a, or by substituting RNA6 variants of P6a or P6b length into the infectious clone [5,6].

Researchers who study molecular factors influencing R protein accumulation and the expression of disease resistance have shown that molecular chaperones, the ubiquitin-mediated proteasome systems, or post-translational modifications influence the accumulation and function of proteins and therefore influence the efficiency of resistance responses. It is well documented that molecular chaperones, the proteasome systems, and post-translational modifications can also influence virus accumulation. Thus, the host genetic background that influences genetic resistance can also drive the evolution and emergence

of RB isolates. The current analysis of mutations in RRV isolates suggests that there may be a clustering of mutations that could be determinants of subtypes or pathotypes. Unfortunately, since the host source genotypes are unknown for the 92 isolates studied here, determining potential virus pathotypes based on genomic sequence variance would be premature. Further research is needed to examine the biological properties of both the host genotype and virus isolates relative to disease severity, RNA accumulation, or resistance breaking to clearly categorize isolates into pathotypes.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/pathogens12050707/s1>, Supplementary Table S1: List of RRV accessions used in this study.

**Author Contributions:** Conceptualization, J.V., V.H., R.J. and J.H.; methodology, J.V., V.H. and R.J.; validation, V.H., R.J. and J.H.; formal analysis, J.V., V.H. and R.J.; investigation, J.V., V.H. and R.J.; resources, J.V., V.H. and R.J.; data curation, J.V., V.H. and R.J.; writing—original draft preparation, J.V. and V.H.; writing—review and editing, R.J. and J.H.; visualization, J.V. and V.H.; supervision, J.V.; project administration, J.V.; funding acquisition, J.V., R.J. and J.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by Star<sup>®</sup> Roses and Plants (grant number M2200455), through USDA funding, Agricultural Research Service (Research Project 8020-22000-052-00D), and by USDA-NIFA (grant number 2014-51181-22644) as part of the Specialty Crops Research Initiative (SCRI).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We thank Michele Scheiber, Christina King, and Kevin Ong for important discussions throughout this research and preparation of this manuscript. We also acknowledge the collaborative sequencing help from USDA-ARS colleagues Sam Grinstead and Dimitre Mollov. We thank Tom Evans (University of Delaware) for sharing RRV-infected ‘Stormy Weather’ rose material.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Kormelink, R.; Verchot, J.; Tao, X.; Desbiez, C. The Bunyavirales: The plant-infecting counterparts. *Viruses* **2021**, *13*, 842. [[CrossRef](#)] [[PubMed](#)]
2. Laney, A.G.; Keller, K.E.; Martin, R.R.; Tzanetakis, I.E. A discovery 70 years in the making: Characterization of the Rose rosette virus. *J. Gen. Virol.* **2011**, *92*, 1727–1732. [[CrossRef](#)] [[PubMed](#)]
3. Di Bello, P.L.; Ho, T.; Tzanetakis, I.E. The evolution of emaraviruses is becoming more complex: Seven segments identified in the causal agent of Rose rosette disease. *Virus Res.* **2015**, *210*, 241–244. [[CrossRef](#)] [[PubMed](#)]
4. Ishikawa, K.; Maejima, K.; Komatsu, K.; Netsu, O.; Keima, T.; Shiraiishi, T.; Okano, Y.; Hashimoto, M.; Yamaji, Y.; Namba, S. Fig mosaic emaravirus p4 protein is involved in cell-to-cell movement. *J. Gen. Virol.* **2013**, *94*, 682–686. [[CrossRef](#)]
5. Verchot, J.; Herath, V.; Urrutia, C.D.; Gayral, M.; Lyle, K.; Shires, M.K.; Ong, K.; Byrne, D. Development of a reverse genetic system for studying rose rosette virus in whole plants. *Mol. Plant-Microbe Interact.* **2020**, *33*, 1209–1221. [[CrossRef](#)]
6. Urrutia, C.D.; Romay, G.; Shaw, B.D.; Verchot, J. Advancing the rose rosette virus minireplicon and encapsidation system by incorporating GFP, mutations, and the CMV 2b silencing suppressor. *Viruses* **2022**, *14*, 836. [[CrossRef](#)]
7. Babu, B.; Washburn, B.K.; Poduch, K.; Knox, G.W.; Paret, M.L. Identification and characterization of two novel genomic RNA segments RNA5 and RNA6 in rose rosette virus infecting roses. *Acta Virol.* **2016**, *60*, 156–165. [[CrossRef](#)]
8. Rehanek, M.; Karlin, D.G.; Bandte, M.; Al Kubrusli, R.; Nourinejad Zarghani, S.; Candresse, T.; Büttner, C.; von Bargen, S. The complex world of emaraviruses—Challenges, insights, and prospects. *Forests* **2022**, *13*, 1868. [[CrossRef](#)]
9. Gupta, A.K.; Hein, G.L.; Graybosch, R.A.; Tatineni, S. Octapartite negative-sense RNA genome of High Plains wheat mosaic virus encodes two suppressors of RNA silencing. *Virology* **2018**, *518*, 152–162. [[CrossRef](#)]
10. Gupta, A.K.; Hein, G.L.; Tatineni, S. P7 and P8 proteins of High plains wheat mosaic virus, a negative-strand RNA virus, employ distinct mechanisms of RNA silencing suppression. *Virology* **2019**, *535*, 20–31. [[CrossRef](#)]
11. Domingo, E. *Virus as Populations*; Elsevier: Waltham, MA, USA, 2016; ISBN 978-0-12-800837-9.
12. Katsiani, A.; Stainton, D.; Lamour, K.; Tzanetakis, I.E. The population structure of Rose rosette virus in the USA. *J. Gen. Virol.* **2020**, *101*, 676–684. [[CrossRef](#)]
13. Domingo, E.; Holland, J.J. RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* **1997**, *51*, 151–178. [[CrossRef](#)]

14. Domingo, E. Viruses at the edge of adaptation. *Virology* **2000**, *270*, 251–253. [[CrossRef](#)]
15. Lyons, D.M.; Lauring, A.S. Evidence for the selective basis of transition-to-transversion substitution bias in two RNA viruses. *Mol. Biol. Evol.* **2017**, *34*, 3205–3215. [[CrossRef](#)]
16. Ivancic-Jelecki, J.; Slovic, A.; Santak, M.; Tešović, G.; Forcic, D. Common position of indels that cause deviations from canonical genome organization in different measles virus strains. *Viol. J.* **2016**, *13*, 134. [[CrossRef](#)]
17. Walia, J.J.; Willemsen, A.; Elci, E.; Caglayan, K.; Falk, B.W.; Rubio, L. Genetic variation and possible mechanisms driving the evolution of worldwide fig mosaic virus isolates. *Phytopathology* **2014**, *104*, 108–114. [[CrossRef](#)]
18. Patil, B.L.; Dangwal, M.; Mishra, R. Variability of emaravirus species associated with sterility mosaic disease of pigeonpea in India provides evidence of segment reassortment. *Viruses* **2017**, *9*, 183. [[CrossRef](#)]
19. Jones, R.A.C.; Vazquez-Iglesias, I.; McGreig, S.; Fox, A.; Gibbs, A.J. Genomic High Plains wheat mosaic virus sequences from Australia: Their phylogenetics and evidence for emaravirus recombination and reassortment. *Viruses* **2023**, *15*, 401. [[CrossRef](#)]
20. Tatineni, S.; McMechan, A.J.; Wosula, E.N.; Wegulo, S.N.; Graybosch, R.A.; French, R.; Hein, G.L. An Eriophyid Mite-Transmitted Plant Virus Contains Eight Genomic RNA Segments with Unusual Heterogeneity in the Nucleocapsid Protein. *J. Virol.* **2014**, *88*, 11834–11845. [[CrossRef](#)]
21. Stewart, L.R. Sequence diversity of wheat mosaic virus isolates. *Virus Res.* **2016**, *213*, 299–303. [[CrossRef](#)]
22. Valkonen, J.P.T.; Jones, R.A.C.; Slack, S.A.; Watanabe, K.N. Resistance specificities to viruses in potato: Standardization of nomenclature. *Plant Breed.* **1996**, *115*, 433–438. [[CrossRef](#)]
23. García-Arenal, F.; McDonald, B.A. An analysis of the durability of resistance to plant viruses. *Phytopathology* **2003**, *93*, 941–952. [[CrossRef](#)] [[PubMed](#)]
24. Malcuit, I.; De Jong, W.; Baulcombe, D.C.; Shields, D.C.; Kavanagh, T.A. Acquisition of multiple virulence/avirulence determinants by potato virus X (PVX) has occurred through convergent evolution rather than through recombination. *Virus Genes* **2000**, *20*, 165–172. [[CrossRef](#)] [[PubMed](#)]
25. Gallois, J.L.; Moury, B.; German-Retana, S. Role of the genetic background in resistance to plant viruses. *Int. J. Mol. Sci.* **2018**, *19*, 2856. [[CrossRef](#)] [[PubMed](#)]
26. Solomon-Blackburn, R.M.; Barker, H. Breeding virus resistant potatoes (*Solanum tuberosum*): A review of traditional and molecular approaches. *Heredity* **2001**, *86*, 17–35. [[CrossRef](#)]
27. Windham, M.T.; Evans, T.; Collins, S.; Lake, J.A.; Lau, J.; Riera-Lizarazu, O.; Byrne, D.H. Field resistance to rose rosette disease as determined by multi-year evaluations in Tennessee and Delaware. *Pathogens* **2023**, *12*, 439. [[CrossRef](#)]
28. Byrne, D.H.; Klein, P.; Yan, M.; Young, E.; Lau, J.; Ong, K.; Shires, M.; Olson, J.; Windham, M.; Evans, T.; et al. Challenges of breeding rose rosette-resistant roses. *HortScience* **2018**, *53*, 604–608. [[CrossRef](#)]
29. Zlesak, D. Pollen diameter and guard cell length as predictors of ploidy in diverse rose cultivars, species, and breeding lines. *Flor. Orn. Biotechnol.* **2009**, *3*, 53–70.
30. Raymond, O.; Gouzy, J.; Just, J.; Badouin, H.; Verdenaud, M.; Lemainque, A.; Vergne, P.; Moja, S.; Choisine, N.; Pont, C.; et al. The *Rosa* genome provides new insights into the domestication of modern roses. *Nat. Genet.* **2018**, *50*, 772–777. [[CrossRef](#)]
31. Young, E.L.; Lau, J.; Bentley, N.B.; Rawandoozi, Z.; Collins, S.; Windham, M.T.; Klein, P.E.; Byrne, D.H.; Riera-Lizarazu, O. Identification of QTLs for reduced susceptibility to rose rosette disease in diploid roses. *Pathogens* **2022**, *11*, 660. [[CrossRef](#)]
32. Lau, J.; Young, E.L.; Collins, S.; Windham, M.T.; Klein, P.E.; Byrne, D.H.; Riera-Lizarazu, O. Rose rosette disease resistance loci detected in two interconnected tetraploid garden rose populations. *Front. Plant Sci.* **2022**, *13*, 2036. [[CrossRef](#)]
33. Henderson, D.C.; Hammond, J. CKC: Isolation of Nucleic Acids from a Diversity of Plants Using CTAB and Silica Columns. *Mol. Biotechnol.* **2013**, *53*, 109–117. [[CrossRef](#)]
34. Jordan, R.; Korolev, E.; Grinstead, S.; Molloy, D. First complete genome sequence of carnation latent virus, the type member of the genus *Carlavirus*. *Arch. Virol.* **2021**, *166*, 1501–1505. [[CrossRef](#)]
35. Nguyen, L.T.; Schmidt, H.A.; Von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [[CrossRef](#)]
36. Trifinopoulos, J.; Nguyen, L.T.; von Haeseler, A.; Minh, B.Q. W-IQ-TREE: A fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* **2016**, *44*, W232–W235. [[CrossRef](#)]
37. Rozas, J.; Ferrer-Mata, A.; Sanchez-DelBarrio, J.C.; Guirao-Rico, S.; Librado, P.; Ramos-Onsins, S.E.; Sanchez-Gracia, A. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **2017**, *34*, 3299–3302. [[CrossRef](#)]
38. Martin, D.P.; Murrell, B.; Golden, M.; Khoosal, A.; Muhire, B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol.* **2015**, *1*, 1–5. [[CrossRef](#)]
39. Forslund, K.; Huson, D.H.; Moulton, V. VisRD—Visual recombination detection. *Bioinformatics* **2004**, *20*, 3654–3655. [[CrossRef](#)]
40. Samson, S.; Lord, É.; Makarenkov, V. Sequence analysis SimPlot++: A Python application for representing sequence similarity and detecting recombination. *arXiv* **2021**, arXiv:2112.09755.
41. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)]
42. Letunic, I.; Bork, P. Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [[CrossRef](#)] [[PubMed](#)]
43. Di Bello, P.L.; Thekke-Veetil, T.; Druciarek, T.; Tzanetakis, I.E. Transmission attributes and resistance to rose rosette virus. *Plant Pathol.* **2018**, *67*, 499–504. [[CrossRef](#)]

44. Al Rwahnih, M.; Karlik, J.; Diaz-Lara, A.; Ong, K.; Mollov, D.; Haviland, D.; Golino, D. First report of rose rosette virus associated with rose rosette disease affecting roses in California. *Plant Dis.* **2019**, *103*, 380. [\[CrossRef\]](#)
45. Ali, A.; Roossinck, M.J. Genetic Bottlenecks. In *Plant Virus Evolution*; Roossinck, M.J., Ed.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 123–131.
46. Suchard, M.A.; Lemey, P.; Baele, G.; Ayres, D.L.; Drummond, A.J.; Rambaut, A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **2018**, *4*, vey016. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Duffy, S.; Shackelton, L.A.; Holmes, E.C. Rates of evolutionary change in viruses: Patterns and determinants. *Nat. Rev. Genet.* **2008**, *9*, 267–276. [\[CrossRef\]](#)
48. Schirmer, M.; Sloan, W.T.; Quince, C. Benchmarking of viral haplotype reconstruction programmes: An overview of the capacities and limitations of currently available programmes. *Brief. Bioinform.* **2014**, *15*, 431–442. [\[CrossRef\]](#)
49. Agol, V.I.; Gmyl, A.P. Emergency services of viral RNAs: Repair and remodeling. *Microbiol. Mol. Biol. Rev.* **2018**, *82*, e00067-17. [\[CrossRef\]](#)
50. Alisoltani, A.; Jaroszewski, L.; Iyer, M.; Iranzadeh, A.; Godzik, A. Increased frequency of indels in hypervariable regions of SARS-CoV-2 proteins—A possible signature of adaptive selection. *Front. Genet.* **2022**, *13*, 875406. [\[CrossRef\]](#)
51. Acevedo, A.; Brodsky, L.; Andino, R. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature* **2014**, *505*, 686–690. [\[CrossRef\]](#)
52. Hily, J.M.; Poulicard, N.; Mora, M.Á.; Pagán, I.; García-Arenal, F. Environment and host genotype determine the outcome of a plant-virus interaction: From antagonism to mutualism. *New Phytol.* **2016**, *209*, 812–822. [\[CrossRef\]](#)
53. Simmonds, P.; Aiweksakun, P.; Katzourakis, A. Prisoners of war—Host adaptation and its constraints on virus evolution. *Nat. Rev. Microbiol.* **2019**, *17*, 321–328. [\[CrossRef\]](#)
54. Elena, S.F.; Fraile, A.; García-Arenal, F. Evolution and Emergence of Plant Viruses. *Adv. Virus Res.* **2014**, *88*, 161–191. [\[CrossRef\]](#)
55. Moury, B.; Fabre, F.; Hébrard, E.; Froissart, R. Determinants of host species range in plant viruses. *J. Gen. Virol.* **2017**, *98*, 862–873. [\[CrossRef\]](#)
56. García-Arenal, F.; Fraile, A. Trade-offs in host range evolution of plant viruses. *Plant Pathol.* **2013**, *62*, 2–9. [\[CrossRef\]](#)
57. Duffy, S.; Burch, C.L.; Turner, P.E. Evolution of host specificity drives reproductive isolation among RNA viruses. *Evolution* **2007**, *61*, 2614–2622. [\[CrossRef\]](#)
58. Schirmer, M.; D’Amore, R.; Ijaz, U.Z.; Hall, N.; Quince, C. Illumina error profiles: Resolving fine-scale variation in metagenomic sequencing data. *BMC Bioinform.* **2016**, *17*, 125. [\[CrossRef\]](#)
59. Stoler, N.; Nekrutenko, A. Sequencing error profiles of Illumina sequencing instruments. *NAR Genomics Bioinform.* **2021**, *3*, lqab019. [\[CrossRef\]](#)
60. Tomimura, K.; Gibbs, A.J.; Jenner, C.E.; Walsh, J.A.; Ohshima, K. The phylogeny of Turnip mosaic virus; comparisons of 38 genomic sequences reveal a Eurasian origin and a recent “emergence” in east Asia. *Mol. Ecol.* **2003**, *12*, 2099–2111. [\[CrossRef\]](#)
61. Fraile, A.; McLeish, M.J.; Pagán, I.; González-Jara, P.; Piñero, D.; García-Arenal, F. Environmental heterogeneity and the evolution of plant-virus interactions: Viruses in wild pepper populations. *Virus Res.* **2017**, *241*, 68–76. [\[CrossRef\]](#)
62. Ali, A.; Li, H.; Schneider, W.L.; Sherman, D.J.; Gray, S.; Smith, D.; Roossinck, M.J. Analysis of genetic bottlenecks during horizontal transmission of cucumber mosaic virus. *J. Virol.* **2006**, *80*, 8345–8350. [\[CrossRef\]](#)
63. McLeish, M.J.; Fraile, A.; García-Arenal, F. Ecological complexity in plant virus host range evolution. *Adv. Virus Res.* **2018**, *101*, 293–339. [\[CrossRef\]](#) [\[PubMed\]](#)
64. Conners, I.L. *Twentieth Annual Report of the Canadian Plant Disease Survey*; Department of Agriculture: Ottawa, ON, Canada, 1941; p. 98.
65. Thomas, E.A.; Scott, C.E. Rosette of rose. *Phytopathology* **1953**, *43*, 218–219.
66. Keifer, H.H. Eriophyid studies VIII. *Bull. Cali. Dept. Agric.* **1940**, *29*, 21–46.
67. Amrine, J.W., Jr.; Hindal, D.F.; Williams, R.; Appel, J.; Stasny, T.; Kassir, A. Rose rosette as a biocontrol of multiflora rose, 1987–1989. In Proceedings of the 43rd Annual Meeting of the Southern Weed Science Society, Atlanta, GA, USA, 15–17 January 1990; Volume 43, pp. 316–319.
68. Hindal, D.F.; Amrine, J.W.; Williams, R.L.; Stasny, T.A. Rose rosette disease on multiflora rose (*Rosa multiflora*) in Indiana and Kentucky. *Weed Technol.* **1988**, *2*, 442–444. [\[CrossRef\]](#)
69. Allington, W.B.; Staples, R.; Viehmeyer, G. Transmission of rose rosette virus by the eriophyid mite *Phyllocoptes fructiphilus*. *J. Econ. Entomol.* **1968**, *61*, 1137–1140. [\[CrossRef\]](#)
70. Pemberton, H.B.; Ong, K.; Windham, M.; Olson, J.; Byrne, D.H. What is rose rosette disease? *HortScience* **2018**, *53*, 592–595. [\[CrossRef\]](#)
71. Pemberton, H.B.; Karlik, J.F. A recent history of changing trends in USA garden rose plant sales, types, and production methods. *Acta Hortic.* **2015**, *1064*, 223–234. [\[CrossRef\]](#)
72. Cho, E.-K. Strains of soybean mosaic virus: Classification based on virulence in resistant soybean cultivars. *Phytopathology* **1979**, *69*, 467. [\[CrossRef\]](#)
73. Drijfhout, E.; Silbernagel, M.H.; Burke, D.W. Differentiation of strains of bean common mosaic virus. *Netherlands J. Plant Pathol.* **1978**, *84*, 13–26. [\[CrossRef\]](#)
74. Fraile, A.; García-Arenal, F. The coevolution of plants and viruses: Resistance and pathogenicity. *Adv. Virus Res.* **2010**, *76*, 1–32. [\[CrossRef\]](#)

75. Cox, B.A.; Jones, R.A.C. Genetic variability in the coat protein gene of Potato virus X and the current relationship between phylogenetic placement and resistance groupings. *Arch. Virol.* **2010**, *155*, 1349–1356. [[CrossRef](#)] [[PubMed](#)]
76. Grech-Baran, M.; Witek, K.; Szajko, K.; Witek, A.I.; Morgiewicz, K.; Wasilewicz-Flis, I.; Jakuczun, H.; Marczewski, W.; Jones, J.D.G.; Hennig, J. Extreme resistance to potato virus Y in potato carrying the *Rysto* gene is mediated by a TIR-NLR immune receptor. *Plant Biotechnol. J.* **2020**, *18*, 655–667. [[CrossRef](#)] [[PubMed](#)]
77. Moreno-Pérez, M.G.; García-Luque, I.; Fraile, A.; García-Arenal, F. Mutations that determine resistance breaking in a plant RNA virus have pleiotropic effects on its fitness that depend on the host environment and on the type, single or mixed, of infection. *J. Virol.* **2016**, *90*, 9128–9137. [[CrossRef](#)] [[PubMed](#)]
78. Lu, Y.; McGavin, W.; Cock, P.J.A.; Schnettler, E.; Yan, F.; Chen, J.; MacFarlane, S. Newly identified RNAs of raspberry leaf blotch virus encoding a related group of proteins. *J. Gen. Virol.* **2015**, *96*, 3432–3439. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.