

Article

Variability among the Isolates of Broad Bean Mottle Virus and Encapsidation of Host RNAs

Nipin Shrestha *, Melvin R. Duvall and Jozef J. Bujarski *

Department of Biological Sciences and Plant Molecular and Bioinformatics Center, Northern Illinois University, DeKalb, IL 60115, USA; duvallmel@gmail.com

* Correspondence: nipin.shrestha@gmail.com (N.S.); jbjarski@niu.edu (J.J.B.); Tel.: +1-305-684-2589 (N.S.); +1-815-753-0601 (J.J.B.); Fax: +1-815-753-7855 (J.J.B.)

Abstract: Broad bean mottle bromovirus infects legume plants and is transmissible by insects. Several broad bean mottle virus (BBMV) isolates have been identified, including one in England (isolate Ba) and five in the Mediterranean countries: Libya (LyV), Morocco (MV), Syria (SV), Sudan (TU) and Tunisia (TV). Previously, we analyzed the nucleotide sequence of the Ba RNA and here we report on and compare it with another five Mediterranean variants. The RNA segments in the latter ones were extensively homologous, with some SNPs, single nucleotide deletions and insertions, while the number of mutations was higher in isolate Ba. Both the 5' and 3' untranslated terminal regions (UTRs) among the corresponding RNAs are highly conserved, reflecting their functionality in virus replication. The AUG initiation codons are within suboptimal contexts, possibly to adjust/regulate translation. The proteins 1a, 2a, 3a and coat protein (CP) are almost identical among the five isolates, but in Ba they have more amino acid (aa) substitutions. Phylogenetic analysis revealed the isolates from Morocco and Syria clustering with the isolate from England, while the variants from Libya, Tunisia and Sudan created a different clade. The BBMV isolates encapsidate a high content of host (ribosomal and messenger) RNAs. Our studies present BBMV as a useful model for bromoviruses infecting legumes.



Citation: Shrestha, N.; Duvall, M.R.; Bujarski, J.J. Variability among the Isolates of Broad Bean Mottle Virus and Encapsidation of Host RNAs. *Pathogens* **2022**, *11*, 817. <https://doi.org/10.3390/pathogens11070817>

Academic Editor: Lawrence S. Young

Received: 25 May 2022

Accepted: 18 July 2022

Published: 21 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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/>).

Keywords: plant RNA viruses; bromoviruses; RNA genome variability; cellular RNA encapsidation; phylogeny

1. Introduction

The broad bean mottle virus is a member of the genus *Bromovirus* and infects broad bean (*Vicia faba*) and other eudicotyledonous plants, such as *Chenopodium* spp., *Nicotiana benthamiana* or *Pisum sativum* [1,2]. BBMV infection shows systemic mottling symptoms and is transmissible by insects [2,3]. There are several BBMV strains (isolates) known, one reported in England [4] while the others are in Europe, Africa and Asia. Specifically, several isolates of varying severity were found in Mediterranean countries. However, their serological properties are indistinguishable [2,3].

In our laboratory, the genome of the isolate Ba was previously sequenced, displaying three genomic RNAs and a subgenomic RNA (sgRNA4) [5,6]. Interestingly, unlike in the other *Bromoviridae*, the 5' terminal nucleotide in all of the Ba RNAs is adenosine, with its biological significance not being known [5,6]. By using the full-length cDNA clones, the RNA2-derived defective interfering (DI) RNAs were characterized in Ba; that interfered with the virus accumulation and the symptoms of BBMV infection [7]. Despite the initial molecular characterization of the BBMV, the *Brome mosaic bromovirus* (BMV) dominated plant molecular virology as a cosmopolitan positive-sense model RNA virus. However, biological features, such as high multiplicity in host cells, ease of propagation and the host range (including *Arabidopsis thaliana*), can make BBMV a more convenient model, especially for viruses infecting legumes [3].

The purpose of the study is to sequence the full genome of five BBMV isolates collected from Libya, Morocco, Syria, Sudan and Tunisia, and characterize the genetic variability among these five isolates, based on their full genome sequences, comparing them with the previously sequenced isolate from England.

2. Results and Discussion

2.1. Propagation and Purification of BBMV RNAs

The BBMV isolates were grown in *Vicia faba* seedlings, and all induced the systemic mottling symptoms, with the isolate TU being the most severe, whereas the isolates LyV, MV and TV caused mild mottling, and SV was the least symptomatic. The visual inspection of the symptoms indicated that for some of the isolates, these observations differed from previous studies, which showed that isolates MV and TV were among the most severe and that isolate TU was the least severe [2,3]. The virions were extracted by cesium chloride (CsCl) gradient centrifugation, and the potentially occluding cellular nucleic acids were removed by nucleases (see Materials and Methods). The presence of the purified virions was confirmed by transmission electron microscopy, as shown in Figure S1 (Supplementary Materials). Subsequently, the viral RNAs were extracted from the purified virions, and denaturing gel electrophoresis showed the presence of the intact BBMV RNA segments, RNAs1+2, RNA3 and sgRNA4 (Figure 1).

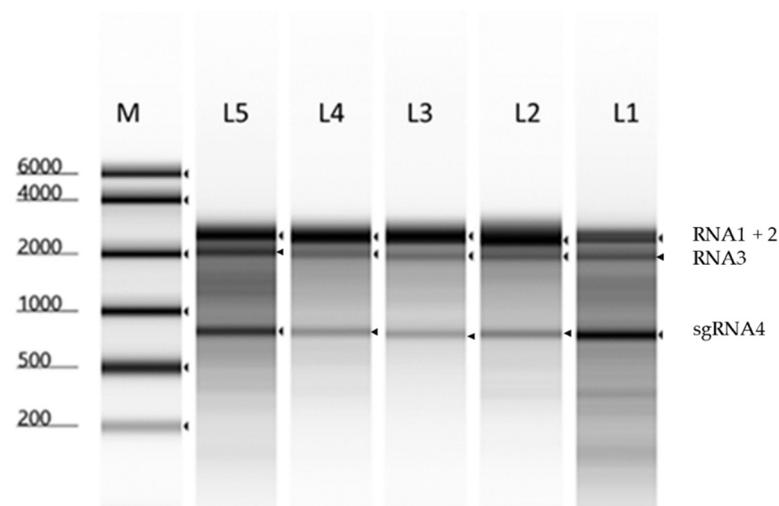


Figure 1. Denaturing gel of the BBMV RNAs isolated from five BBMV isolates. The images were taken prior to sequencing. L1, L2, L3, L4 and L5, which are the lanes loaded with the RNA from MV, LyV, SV, TU and TV isolates, respectively. Lane M is the size marker. The three major bands, characteristic for Bromoviruses, representing RNA1 + 2 (~2900–3100 bp), RNA3 (~2200 bp) and sgRNA4 (~800 bp) are visible in the gel (indicated on the right side).

2.2. Sequencing with the NGS and Sanger Protocols

The extracted BBMV RNAs were sequenced as described in Section 4. The Sanger sequencing of the 5' RACE product corroborated the 5' terminal A residues in all of the BBMV RNAs, previously reported for isolate Ba [5,6]. The other members of the bromoviruses are known to carry the 5' terminal G nucleotide [5,8,9]. Other examples of the 5' A (or U) among plant RNA viruses involved members in several families [5]. The difference in the 5' nucleotide, the substrate for the guanidyl transferase, might hint towards specific requirements for the capping mechanism in BBMV [5,6]. The sequences of RNAs 1, 2 and 3 from the five BBMV isolates were deposited in the Gene Bank, under accession numbers, OM287444 to OM287458.

2.2.1. The Particle Heterogeneity

The particle heterogeneity was analyzed, based on the number of reads (Table S1, Supplementary Materials) that were mapped to RNA1, RNA2 and RNA3, assuming that in bromoviruses the respective RNA components are packaged into separate virions [10]. The RNA1-bearing particles were the most abundant in the four isolates, LyV, MV, SV and TU, followed by RNA2 and RNA3, whereas in isolate TV, the RNA2 segments were the most abundant, followed by RNA1 and RNA3. In isolate SV, the difference in the RNA1 to RNA2 ratio was the highest among the five isolates (Figure 2; Table S2). Such a heterogeneous distribution of particles in the infected plants might influence the disease progression, including the differences in symptom severity among the different isolates shown by Makkouk et al. (1989) [3]. In brome mosaic virus (BMV), the host–RNA interactions led to particle heterogeneity that affected the timing of infection, due to an altered dissociation of CP to release the viral RNAs [10,11].

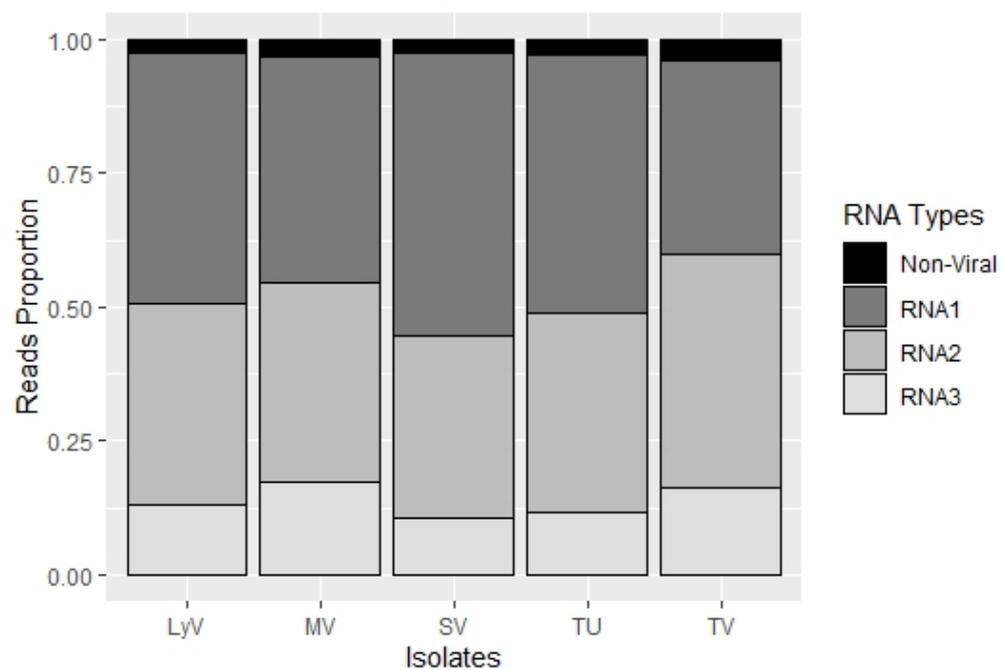


Figure 2. A stacked plot showing proportion of reads that mapped to RNA1, RNA2 and RNA3 among the five BBMV isolates.

2.2.2. Polymorphism at the Nucleotide Level

The sequence data confirmed the general organization of the BBMV RNA genome in all of the isolates, with both RNAs 1 and 2 being monocistronic whereas RNA3 encoded two proteins. The five Mediterranean isolates were similar, at 98% for RNA1 and RNA2, while RNA3 was at 93–95% (Figure 3). However, after aligning to the isolate Ba, the RNAs appeared to be different, with increased SNPs. In particular, the alignments revealed single-nucleotide polymorphisms (SNPs) and micro-indels (1–3 nts), with RNA1 carrying 72 SNPs and a single nucleotide deletion, RNA2 carrying 64 SNPs and a single nucleotide deletion, while RNA3 was carrying 90 SNPs and 28 indels. The total SNP numbers increased to 182, 376 and 230 for RNA1, RNA2 and RNA3, respectively, after comparing with the isolate Ba.

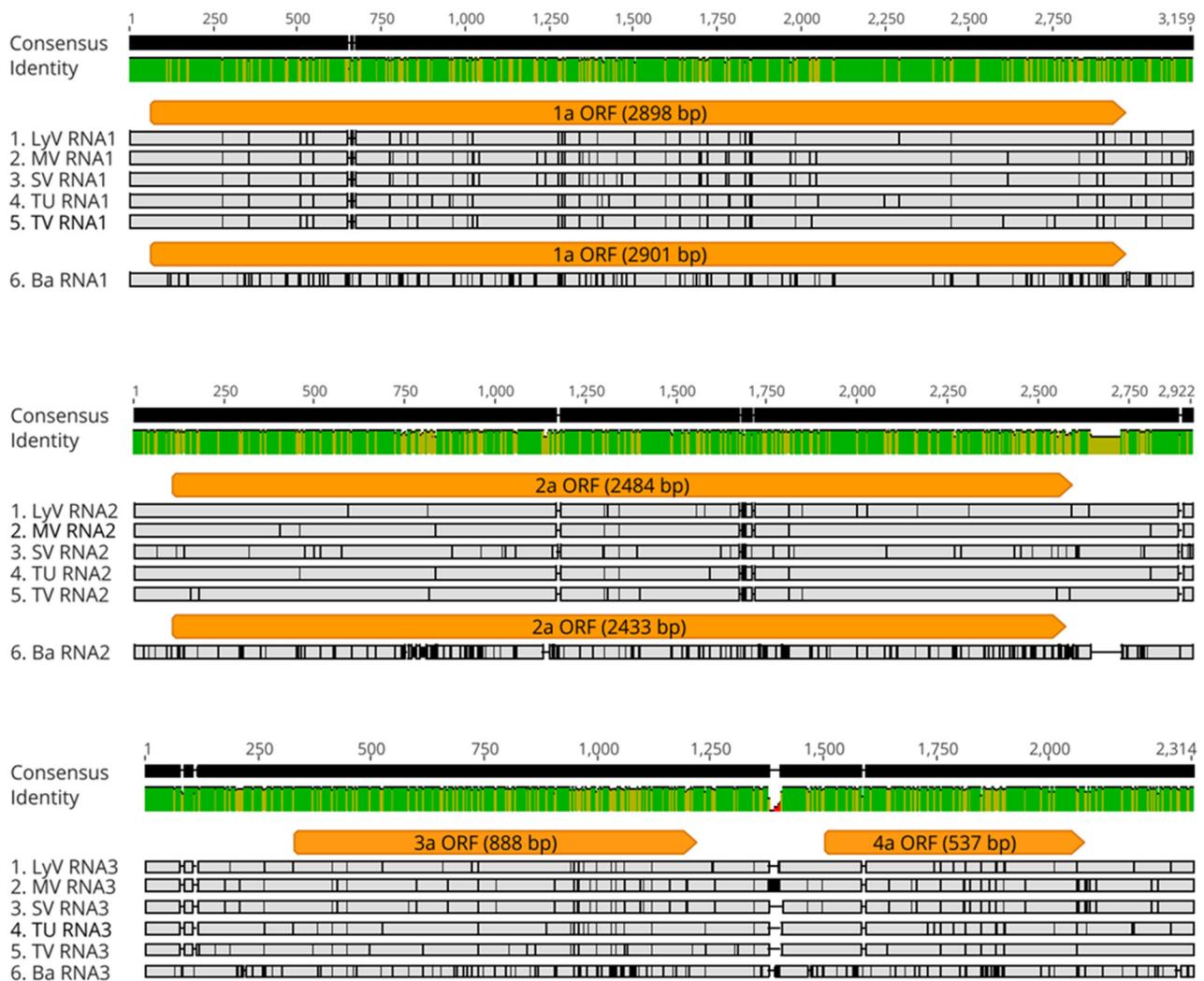


Figure 3. Multiple sequence alignment (MSA) of nucleotide sequences of BBMV RNAs 1, 2, and 3 (top to bottom) for five BBMV isolates as compared to Ba ones. The orange bars represent the ORF regions (with the length shown in parentheses). The gray regions signify identical nucleotides, while the nucleotide substitutions are depicted by the black vertical lines and the deletions by the horizontal black lines. The green bars illustrate the regions of identity.

As in Ba, the AUG initiation codons in all of the RNAs were flanked by sequences differing from the optimal consensus motifs, CCACCAUGG or AACCAUGG. Similar trends were observed for the AUG codons in BMV and Cowpea chlorotic mottle bromovirus (CCMV) RNAs [5,6]. The suboptimal AUG contexts might self-limit/regulate the viral protein synthesis, and/or allow for parallel translation from downstream AUG codons, as observed in the BMV CP open-reading frame (ORF) [12].

2.2.3. The Noncoding Regions

The intergenic regions, 5' and 3' UTR, were conserved on a particular RNA segment among the different isolates. The isolates (including Ba) shared 65 nts of 100% identity in 5' UTR of RNA1; the 5' UTRs in RNA1 were similar in length among BMV, BBMV and CCMV. In RNA2, the 5' UTR was 109 nucleotides (nts) long, 100% identical in the four isolates, but in SV there was a mutation at the 64th position; in isolate Ba, this region was 110 nts long and had five SNPs. The first 30 bases in RNA2 were similar among the three bromoviruses, BMV, BBMV and CCMV [6], likely as part of the RNA replication signal [13,14].

The 3' UTR in RNA1 was 193 nts in all five of the BBMV isolates and was not identical among any of the isolates. The least identity was between the LyV and MV isolates (97.41%), with four SNPs and one indel. This non-coding region was 272 nts for BMV and 223 nts for CCMV. These differences are due to deletions of the largely pseudo-knotted [15,16] sequences (62 and 53 nts, respectively) immediately after the 1a termination codon of BBMV and CCMV, revealing some tolerance to the mutations in the promoters of the minus-strand synthesis [5]. The RNA2 had the longest 3' UTR of 320 nts. The isolate SV had the most variation in this region, with six SNPs and a deletion.

The three non-coding regions of BBMV RNA3 included the 5' UTR, 3' UTR and the intergenic region (IGR). In contrast to RNA2, the 5' UTR of RNA3 contained 328 nts, compared to 91 nts in BMV and 238 nts in CCMV. Their first 116 nts were identical in the five BBMV isolates, but the rest carried seven SNPs. Regarding the 3' UTR, which was 238 nts long, the SV and MV isolates shared 100% identity, but the other isolates had 13 SNPs. The most polymorphic region was the IGR, which was attributed to the varied length of the homopolymeric poly A tract, and averaged between 68 nts for isolate Ba [6], and 61, 80, 51, 54 and 55 nts for the isolates LyV, MV, SV, TU and TV RNA3s, respectively. In BMV and CCMV, the poly A tracts averaged 20 and 40 nts, respectively. The molecular function(s) of the poly A tracts have not been verified in these viruses [8,9].

Besides the poly A variation, the rest of the IGR was largely conserved in BBMV, with only 10 SNPs among the five isolates. In BMV and CCMV, the sg promoter consisted of a "core" sequence that preceded the sgRNA4 initiation site. The corresponding initiation site in the five new isolates aligned with the previously mapped site in Ba [6], but with some variation within the core (Figure 4), that might affect the transcription of sgRNA4.

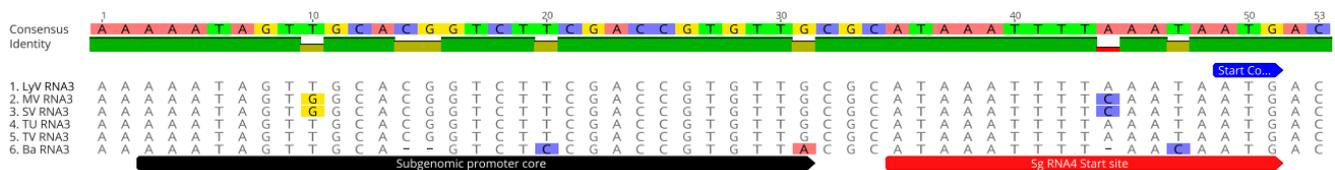


Figure 4. The alignment of the nt sequences of the intergenic noncoding region in the RNA3 of the BBMV variants covering the subgenomic promoter core (black bar) and the start site of the sg RNA4 transcription (see also Romero et al., 1992 [6]). The subgenomic RNA start site are marked by red bar, and the CP translation start codon ATG is indicated by the blue bar on top of the alignment.

2.2.4. Polymorphism at the Amino Acid Level

The open reading frames in respective RNAs were translated and aligned as shown in Figures 5 and S2–S5. The SNPs in the ORF were mostly synonymous, having little effect on the amino acid composition in the five new isolates, but in isolate Ba there were more mutations. The conserved domains were found in both the 1a and 2a proteins (Figures 5, S2 and S3). Two domains in 1a included the N- (amino acids 1 to 432) and C-terminal (amino acids 572 to 959) regions, comprising methyl transferase and helicase domains, respectively, as in the other bromoviruses [15]. For the 2a protein, the sequences of the viral RNA-dependent RNA polymerase (RdRp) were 100% identical between the isolates MV and TU, whereas in the isolates LyV, SV and TV there were one, two and five amino acid substitutions, respectively. The central domain carried the conserved Glycine–Aspartic acid–Aspartic acid (GDD) motif (Figure S3), a catalytic site of the RdRp enzyme [16,17], implying similar interactions, revealed in the crystal structures of BMV and CCMV [18,19].

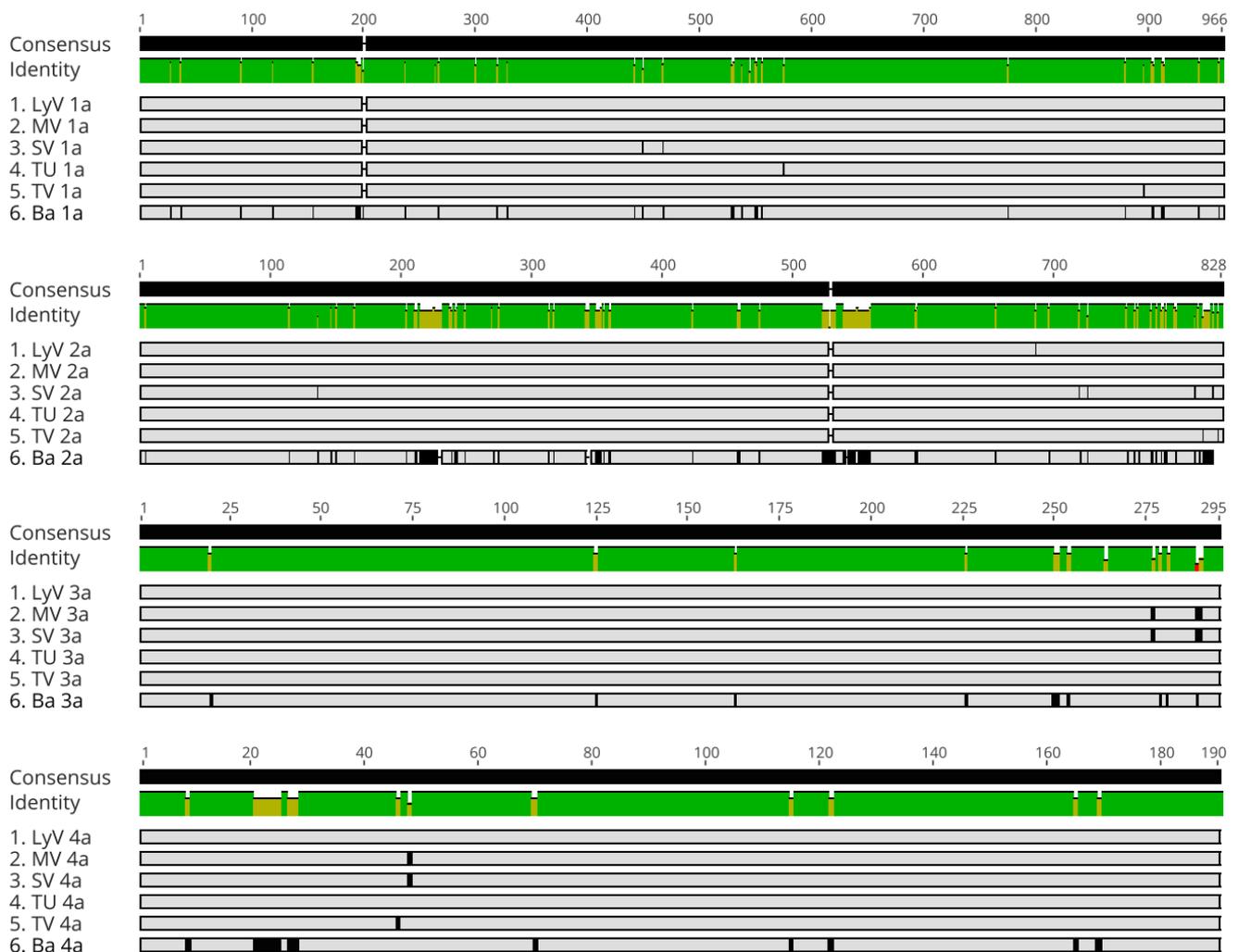


Figure 5. MSA of amino acid sequences of proteins 1a, 2a, 3a and 4a (CP) for six BBMV isolates. The gray regions represent identical nucleotides, while the aa substitutions are depicted by the black vertical lines and the deletions by small gaps. The green bars illustrate the regions of identity. The numerical scale on top marks the positions of amino acids.

Among the five isolates, the least homology was found for 3a and CP (Figures S4 and S5). The polymorphism in 3a was related to symptom severity by different isolates [20–22]. A 100% identity was found among the N-terminal LAGLI-like domains (positions 147–151) [6], that maintain the proper conformation of 3a [23]. The similarity among the 3a proteins varied between 65% with BMV vs. CCMV and 59% with BMV vs. BBMV; possibly a factor for their host range. The similarities for the 1a and for 2a proteins between BMV and CCMV also varied [5,6,9], which together might specify varied host ranges in the three bromoviruses. Along these lines, the swapping of 1a and 2a ORFs between BMV and CCMV confirmed their best fit for multiplication in these viruses' cognate hosts [15,24].

The CP was highly similar among the five BBMV isolates, with 100% similarity between the isolates LyV and TU. The isolate TV had a substitution 46-T to A, and MV and SV shared 48-T to the S substitution, whereas isolate Ba carried 12 aa substitutions (Figure S5). The well-studied arginine-rich N-terminal domain was conserved, except in the isolate Ba, where a 9-T was replaced with A and 21–25 RQLAL was replaced with SNRLR. The RQLAL motif plays a role in the selective packaging of sgRNA4 and during RNA replication in BMV and CCMV [16,17]. The observed changes within CP might affect the electrostatic balance in the virions and the infectivity [3], as shown for BMV [25,26]. In numerous viruses, the N-ARM extends inwards and directly interacts with the RNA motifs [26], affecting the

infection process. The overall homologies among the bromovirus CP sequences covered mostly the hydrophobic amino acids.

2.3. Phylogeny of BBMV Variants

The presence of the subgenomic promoter-like sequences only within 5' UTRs of BBMV and CCMV RNA3 suggested that BMV RNA3 was derived by deletion from the predecessors carrying an undivided genome [6]. Among the three bromoviruses, the higher similarity between the BMV and CCMV proteins, as well as the unique 5' A residue in BBMV, indicated the existence of an intermediate ancestor [5,6,9]. In fact, both BBMV and CCMV can infect eudicotyledonous plants, whereas BMV prefers monocots. To analyze their evolutionary relationships, the consensus cladogram was generated (Figure 6). The RNA sequences were concatenated as one sequence per isolate and aligned together, from which a cladogram was inferred by the maximum likelihood bootstrap method in Geneious, with 100 bootstrap replicates. The concatenated BBMV lineages gave bootstrap values >90. The resulting tree (Figure 6) revealed two sister clades, each with a different ancestor, but not correlating to the geographical distribution, e.g., the Morocco and Syria isolates were geographically separated, but were more closely related.

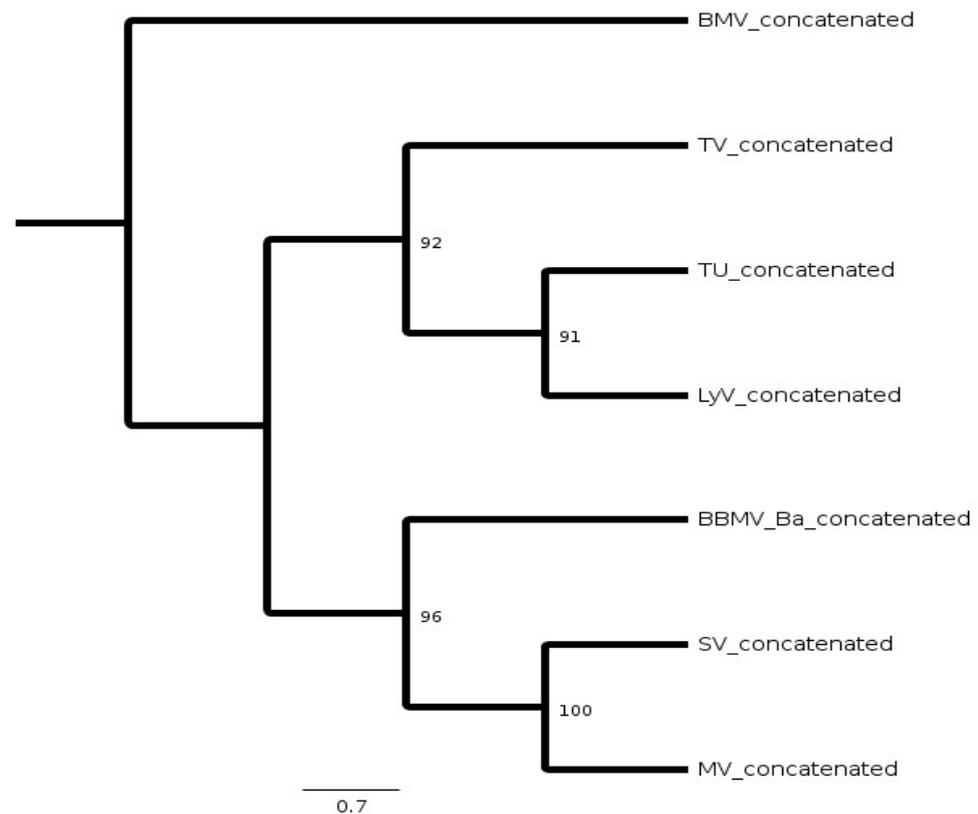


Figure 6. Maximum likelihood (ML) bootstrap consensus cladogram for five Mediterranean and the England isolates of BBMV. The sequences of RNAs 1, 2 and 3 were concatenated and aligned by MSA followed by the consensus tree reconstruction by the ML bootstrap method. The BMV RNA sequences were used as the outgroup to infer the rooted cladogram. Numbers along the branches are ML bootstrap values. The cladogram has been calculated at the Geneious default values.

2.4. BBMV Encapsidate Host RNAs

Previously, we demonstrated that BMV can encapsidate nearly 0.1% of host (cellular) RNAs, including rRNAs and mRNAs, nuclear and organellar, transposable elements and retro-transposons [27]. Here we find that the BBMV packages a 10 to 20 times (between 1.14 to 1.95%) higher content of the host RNAs. Table S2 (Supplementary Materials) summarizes the numbers of reads mapped to various categories, whereas Table S3 (Supplementary Mate-

rials) shows that over 99% of the host reads were of nuclear origin; in BMV, 50% of the reads were from nuclear RNAs and the rest from mitochondria and plastids (Figure 7; Table S3). Among the most mapped total 17 transcripts (Table S4, Supplementary Materials), although there was some variability in their concentration (based on the number of reads), five isolates encapsidated the photosystem repair protein and ribosomal protein S25, followed by aldolase, heat shock and GTP-binding proteins; the transposable elements (TEs) included retrotransposons Class 1 (LINE and LTR), and some of the Class 2 Res (Table S5). One can speculate that these differences originate either due to the distinct affinities of viral CPs with the RNA types, due to different replication sites, or because of host-specific factors. It seems that the encapsidation in the different bromoviruses might use different specificity-based mechanisms of RNA packaging. The study of the host RNA encapsidation by BBMV will be extended further to address the subject in more detail. Future analyses, including the isolate Ba, will involve the validation of host RNAs by Northern blotting, and the possibility of artifacts from residual RNAs after nuclease treatment.

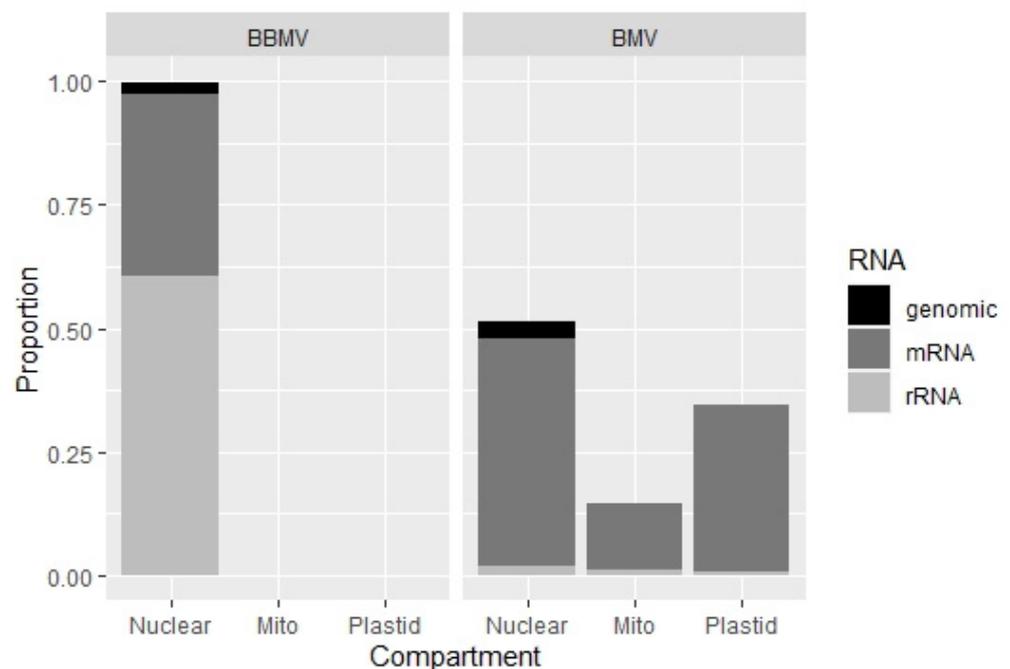


Figure 7. The stacked plot illustration profiling the proportions of different types of encapsidated host RNAs in BBMV vs. BMV, originating from three different cellular compartments. Apparently in BBMV, more than 99% of the host RNAs were nuclear RNAs with ribosomal RNAs being present more in comparison to the mRNA transcripts, whereas in BMV over 50% of the reads mapped to the nuclear RNAs followed by the organellar RNAs (mitochondrial and plastid). The proportion of the mRNAs being mapped dominantly in comparison to the rRNAs also differed between BBMV and BMV.

3. Summary and Conclusions

Our data added new information about the genus *Bromovirus*, with five new sets of BBMV RNA genomes sequenced (OM287444 to OM287458). Assuming the genome size of BBMV as nearly 8600 nts, each nucleotide was covered at the depth of 7925 reads, or 1.2×10^6 nts of reads/per reference nt. This, however, may vary, as the NGS coverage depends on multiple factors [28]. In this work, the accuracy of the NGS data was confirmed by Sanger sequencing.

These results provide the complete genome and characterization of the BBMV variants, which can be used to study the biology of bromoviruses. All of the BBMV RNAs carry the 5' terminal A, which might reflect some distinct function(s) of the 5' cap structure in BBMV; an interesting subject for future studies. Secondly, a high similarity among both the

nts and protein sequences of BBMV suggests that the previously described variations in symptom severity among the isolates [3] might be due to changes in the selected domains, working as symptom determinants, to be mapped. Third, in the same way as BMV, the BBMV isolates encapsidate the host RNAs, but their range differs, which may reflect the distinctive packaging mechanism(s). Finally, the sequences among the known BMV strains differ to a higher degree than those among the five BBMV variants. Based on the criteria defining viral strains [29], we defined the Mediterranean variants as isolates, whereas Ba BBMV should be considered as a separate strain. BMV and CCMV have been studied for host–virus interactions [16,30–33], and used in nano-biological applications [34,35]. Our findings on the genome of BBMV for different variants will contribute towards the studies of the different molecular aspects of the virus, and establish BBMV as another model Bromovirus with which to study plus-sense RNA viruses

4. Materials and Methods

4.1. Virus Material and Purification

The Mediterranean isolates of BBMV were provided by Dr. Safaa Kumari (International Center for Agricultural Research in the Dry Areas (ICARDA), Beirut, Lebanon). The viruses were propagated in the broad bean leaves, followed by extraction and purification, as described before [27,36]. Briefly, the cell extract was ultracentrifuged on a sucrose cushion and cleaned by nuclease treatment to remove any occluding RNAs/DNAs [27,37]. The nucleases were removed on Amicon Ultra-4 centrifugal filters, Ultracel-100K (UFC810008), and the washed virus suspension was further purified in 40% (w/v) CsCl gradient. The virion bands were collected and dialyzed against the virus buffer. The quality of the purified BBMV virions (Figure S1, Supplementary Materials) was examined in a transmission electron microscope, as described previously for BMV [27].

4.2. RNA Extraction and NGS Sequencing

The virions were lysed with 0.5% (w/v) sodium dodecyl sulfate (SDS), and the RNA purified by phenol-chloroform extraction in a buffer 0.5 M glycine, 0.5 M NaCl, 0.1 M EDTA, pH 9.5., followed by ethanol precipitation. The RNA integrity was analyzed in 2100 Agilent Bioanalyzer, Santa Clara, CA, USA (Figure 1). The NGS RNA-Seq sequencing was completed at the University of Illinois at Chicago, Research Resource Center—DNA Services Facility. The sequencing libraries were prepared by the Illumina Truseq stranded library preparation kit. The libraries were sequenced on an Illumina NextSeq 500 instrument, paired end 2 × 150 (Illumina, San Diego, CA, USA).

4.3. Rapid Amplification of cDNA Ends (RACE) and Sanger Sequencing

The 3' and 5' termini were determined by PCR amplification of the poly-A enriched RNAs into full-length cDNAs with the SMARTER 5'/3' RACE kit, Takara Bio USA Inc., CA, USA, using the designed RNA-specific primers. The sequences were confirmed by Sanger sequencing at Eurofins genomics, Louisville, KY, USA. The remaining regions were also sequenced with the Sanger protocol, with pairs of primers overlapping over at least 30 nts [38]. This confirmed the NGS assembly by over 99%. The exact length of the poly A stretches in RNA3 were not determined, due to the limitations of the Sanger protocol [38,39].

4.4. De Novo Genome Assembly and Sequence Comparisons

The NGS reads were quality trimmed to the minimum length of 25 bp, using the DynamicTrim and LengthSort programs from SolexaQA [27,40], and were assembled into a representative sequence of the population [41,42] using the Iterative virus assembler (IVA) pipeline [43]. The IVA has an option of the Trimmomatic program [44] that trims off the adapters and primers before the assembly process. The contigs were quality-checked by the IVA QC in a default setting, then queried against the nucleotide database, using BLASTn (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>, accessed on 24 May 2022). The verified sequences

were subjected to further comparisons, using the Geneious 10.2.3 (Biomatters, Ltd., Auckland, NZ, USA). The contigs that mapped to the previously sequenced BBMV RNAs from the Ba isolate were extracted to the length of the reference by using bedtools2 [45], and aligned to the GenBank references for BBMV RNA1 (NC_004008), RNA2 (NC_004007), and RNA3 (NC_004006) using Geneious 10.2.3.

The open reading frames for the respective RNAs were identified by the ORF finder in Geneious and translated by using standard codons. The sequence similarities and conserved motifs in the translated proteins were aligned with the MAFFT and MUSCLE sequence alignment tools.

4.5. Phylogeny Analyses

The BBMV RNA1, RNA2 and RNA3 were concatenated to generate a single RNA molecule of 8394 nts per each of the five Mediterranean isolates, whereas 8250 nts were generated for the England isolate. BMV was used as an outgroup to infer the rooted cladogram. The RNA1 (NC_002026.1), RNA2 (NC_002026.1) and RNA3 (NC_002026.1) from BMV were also concatenated into a single RNA molecule of 8201 nts. The maximum likelihood bootstrap method in Geneious, with 100 bootstrap replicates, was used to reconstruct the consensus tree.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/pathogens11070817/s1>, Figure S1: Transmission electron microscopy of extensively purified preparations of five BBMV isolates from the broad bean leaves. (A) LyV, (B) MV, (C) SV, (D) TU, (E) TV. The characteristic spherical virions with an average diameter of 27–30 nm ($n = 20$, scale μm) can be observed; Figure S2: The MSA of amino acid sequences of proteins 1a of five new BBMV isolates as compared to the Ba isolate. The identical aa are marked with the green color whereas the polymorphic regions with the olive color. The numerical scale on top marks the positions of amino acid; Figure S3: The MSA of amino acid sequences of proteins 2a of five new BBMV isolates as compared to Ba isolate. The identical aa are marked with the green color whereas the polymorphic regions with the olive color. The numerical scale on top marks the positions of amino acids. The conserved GDD motif is highlighted by red box; Figure S4: The MSA of amino acid sequences of proteins 3a of five new BBMV isolates as compared to Ba isolate. The identical aa are marked with the green color whereas the polymorphic regions with the olive color. The numerical scale on top marks the positions of amino acids; Figure S5: The MSA of amino acid sequences of 4a protein (CP) of five new BBMV isolates as compared to Ba isolate. The identical aa are marked with the green color whereas the polymorphic regions with the olive color. The numerical scale on top marks the positions of amino acids. The N-terminal RQLAL and SNLRL motifs are highlighted by the red box; Table S1: Summary of distribution of the RNA segments encapsidated in the BBMV isolates; Table S2: Summary of reads that mapped to BBMV and Broad bean references; Table S3: Summary of reads mapped to host RNAs encoded in different cellular compartments.; Table S4: The major host mRNAs encapsidated in five BBMV isolates. Table S5: Summary of the transposable elements among the encapsidated host mRNA transcripts; Table S6: The comparison of the average number of reads mapped to host RNAs originating from three different host cellular compartments in BMV vs. BBMV.

Author Contributions: Conceptualization, N.S. and J.J.B.; methodology, N.S.; software, N.S. and M.R.D.; investigation, N.S.; resources, M.R.D. and J.J.B.; writing—original draft preparation, N.S. and J.J.B.; writing—review and editing, N.S., M.R.D. and J.J.B.; funding acquisition, N.S., M.R.D. and J.J.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The genome assembled for the five isolates of BBMV in this project are deposited in the NCBI GenBank under accession: OM287444 to OM287458.

Acknowledgments: We thank Safaa Kumari, ICARDA, Lebanon and Javier Romero, CIT INIA, Spain, for providing samples of BBMV isolates that were sequenced in this work. We thank the University of Illinois at Chicago for all of the advice regarding the Illumina Next Generation Sequencing, and Paige Bothwell at Core Microscopy Facility, Department of Biological Science, Northern Illinois University for helping us with electron microscopy. This work was supported through the Plant Molecular and Bioinformatics Center, the Department of Biological Sciences (via the Elwood Ruth and Briles fellowship), and by the Graduate School, all at Northern Illinois University.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- Fortass, M.; Bos, L. Broad bean mottle virus in Morocco; variability, interaction with food legume species, and seed transmission in faba bean, pea, and chickpea. *Eur. J. Plant Pathol.* **1992**, *98*, 329–342. [[CrossRef](#)]
- Fortass, M.; Diallo, S. Broad bean mottle virus in Morocco; curculionid vectors, and natural occurrence in food legumes other than faba bean (*Vicia faba*). *Eur. J. Plant Pathol.* **1993**, *99*, 219–226. [[CrossRef](#)]
- Makkouk, K.M.; Bos, L.; Rizkallah, A.; Azzam, O.I.; Katul, L. Broad bean mottle virus: Identification, host range, serology, and occurrence on faba bean (*Vicia faba*) in West Asia and North Africa. *Eur. J. Plant Pathol.* **1989**, *94*, 195–212. [[CrossRef](#)]
- Bawden, F.C.; Chaudhuri, R.P.; Kassanis, B. Some properties of broad-bean mottle virus. *Ann. Appl. Biol.* **1951**, *38*, 774–784. [[CrossRef](#)]
- Dzianott, A.M.; Bujarski, J.I. The nucleotide sequence and genome organization of the RNA-1 Segment in two bromoviruses: Broad bean mottle virus and cowpea chlorotic mottle virus. *Virology* **1991**, *185*, 553–562. [[CrossRef](#)]
- Romero, J.; Dzianott, A.M.; Bujarski, J.J. The nucleotide sequence and genome organization of the RNA2 and RNA3 segments in broad bean mottle virus. *Virology* **1992**, *187*, 671–681. [[CrossRef](#)]
- Romero, J.; Huang, Q.; Pogany, J.; Bujarski, J.J. Characterization of Defective Interfering RNA Components That Increase Symptom Severity of Broad Bean Mottle Virus Infections. *Virology* **1993**, *194*, 576–584. [[CrossRef](#)]
- Ahlquist, P.; Dasgupta, R.; Kaesberg, P. Nucleotide sequence of the brome mosaic virus genome and its implications for viral replication. *J. Mol. Biol.* **1984**, *172*, 369–383. [[CrossRef](#)]
- Allison, R.F.; Janda, M.; Ahlquist, P. Sequence of cowpea chlorotic mottle virus RNAs 2 and 3 and evidence of a recombination event during bromovirus evolution. *Virology* **1989**, *172*, 321–330. [[CrossRef](#)]
- Ni, P.; Vaughan, R.C.; Tragesser, B.; Hoover, H.; Kao, C.C. The Plant Host Can Affect the Encapsidation of Brome Mosaic Virus (BMV) RNA: BMV Virions Are Surprisingly Heterogeneous. *J. Mol. Biol.* **2013**, *426*, 1061–1076. [[CrossRef](#)]
- Hoover, H.S.; Wang, J.C.-Y.; Middleton, S.; Ni, P.; Zlotnick, A.; Vaughan, R.C.; Kao, C.C. Phosphorylation of the Brome Mosaic Virus Capsid Regulates the Timing of Viral Infection. *J. Virol.* **2016**, *90*, 7748–7760. [[CrossRef](#)] [[PubMed](#)]
- Sacher, R.; Ahlquist, P. Effects of deletions in the N-terminal basic arm of brome mosaic virus coat protein on RNA packaging and systemic infection. *J. Virol.* **1989**, *63*, 4545–4552. [[CrossRef](#)] [[PubMed](#)]
- Marsh, L.; Hall, T. Evidence Implicating a tRNA Heritage for the Promoters of Positive-strand RNA Synthesis in Brome Mosaic and Related Viruses. *Cold Spring Harb. Symp. Quant. Biol.* **1987**, *52*, 331–341. [[CrossRef](#)] [[PubMed](#)]
- Marsh, L.E.; Huntley, C.C.; Pogue, G.P.; Connell, J.P.; Hall, T.C. Regulation of (+):(-)-strand asymmetry in replication of brome mosaic virus RNA. *Virology* **1991**, *182*, 76–83. [[CrossRef](#)]
- Dinant, S.; Janda, M.; Kroner, P.A.; Ahlquist, P. Bromovirus RNA replication and transcription require compatibility between the polymerase- and helicase-like viral RNA synthesis proteins. *J. Virol.* **1993**, *67*, 7181–7189. [[CrossRef](#)] [[PubMed](#)]
- Jacobs, A.; Hoover, H.; Smith, E.; Clemmer, D.E.; Kim, C.-H.; Kao, C.C. The intrinsically disordered N-terminal arm of the brome mosaic virus coat protein specifically recognizes the RNA motif that directs the initiation of viral RNA replication. *Nucleic Acids Res.* **2017**, *46*, 324–335. [[CrossRef](#)]
- Rao, A.; Grantham, G.L. Molecular studies on bromovirus capsid protein: II. Functional analysis of the amino-terminal arginine-rich motif and its role in encapsidation, movement, and pathology. *Virology* **1996**, *226*, 294–305. [[CrossRef](#)]
- Lucas, R.W.; Larson, S.B.; McPherson, A. The crystallographic structure of brome mosaic virus. *J. Mol. Biol.* **2002**, *317*, 95–108. [[CrossRef](#)]
- Speir, J.A.; Munshi, S.; Wang, G.; Baker, T.S.; Johnson, J.E. Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. *Structure* **1995**, *3*, 63–78. [[CrossRef](#)]
- Fujita, Y.; Mise, K.; Okuno, T.; Ahlquist, P.; Furusawa, I. A Single Codon Change in a Conserved Motif of a Bromovirus Movement Protein Gene Confers Compatibility with a New Host. *Virology* **1996**, *223*, 283–291. [[CrossRef](#)]
- Mise, K.; Allison, R.F.; Janda, M.; Ahlquist, P. Bromovirus movement protein genes play a crucial role in host specificity. *J. Virol.* **1993**, *67*, 2815–2823. [[CrossRef](#)] [[PubMed](#)]
- Takeda, A.; Kaido, M.; Okuno, T.; Mise, K. The C terminus of the movement protein of Brome mosaic virus controls the requirement for coat protein in cell-to-cell movement and plays a role in long-distance movement. *J. Gen. Virol.* **2004**, *85*, 1751–1761. [[CrossRef](#)] [[PubMed](#)]

23. Citovsky, V.; Knorr, D.; Schuster, G.; Zambryski, P. The P30 movement protein of tobacco mosaic virus is a single-strand nucleic acid binding protein. *Cell* **1990**, *60*, 637–647. [[CrossRef](#)]
24. Traynor, P.; Ahlquist, P. Use of bromovirus RNA2 hybrids to map cis- and trans-acting functions in a conserved RNA replication gene. *J. Virol.* **1990**, *64*, 69–77. [[CrossRef](#)]
25. Flasiński, S.; Dzianott, A.; Pratt, S.; Bujarski, J.J. Mutational analysis of the coat protein gene of brome mosaic virus: Effects on replication and movement in barley and in *Chenopodium hybridum*. *Molecular plant-microbe interactions. Mol. Plant-Microbe Interact.* **1995**, *8*, 23–31. [[CrossRef](#)]
26. Ni, P.; Wang, Z.; Ma, X.; Das, N.C.; Sokol, P.; Chiu, W.; Dragnea, B.; Hagan, M.; Kao, C.C. An Examination of the Electrostatic Interactions between the N-Terminal Tail of the Brome Mosaic Virus Coat Protein and Encapsidated RNAs. *J. Mol. Biol.* **2012**, *419*, 284–300. [[CrossRef](#)]
27. Shrestha, N.; Weber, P.H.; Burke, S.V.; Wysocki, W.P.; Duvall, M.R.; Bujarski, J.J. Next generation sequencing reveals packaging of host RNAs by brome mosaic virus. *Virus Res.* **2018**, *252*, 82–90. [[CrossRef](#)]
28. Sims, D.; Sudbery, I.; Ilott, N.E.; Heger, A.; Ponting, C.P. Sequencing depth and coverage: Key considerations in genomic analyses. *Nat. Rev. Genet.* **2014**, *15*, 121–132. [[CrossRef](#)]
29. Kuhn, J.H.; Bao, Y.; Bavari, S.; Becker, S.; Bradfute, S.; Brister, J.R.; Bukreyev, A.A.; Chandran, K.; Davey, R.A.; Dolnik, O.; et al. Virus nomenclature below the species level: A standardized nomenclature for natural variants of viruses assigned to the family Filoviridae. *Arch. Virol.* **2013**, *158*, 301–311. [[CrossRef](#)]
30. Takeda, A.; Nakamura, W.; Sasaki, N.; Goto, K.; Kaido, M.; Okuno, T.; Mise, K. Natural isolates of Brome mosaic virus with the ability to move from cell to cell independently of coat protein. *J. Gen. Virol.* **2005**, *86*, 1201–1211. [[CrossRef](#)]
31. Bujarski, J.J.; Kaesberg, P. Genetic recombination between RNA components of a multipartite plant virus. *Nature* **1986**, *321*, 528–531. [[CrossRef](#)] [[PubMed](#)]
32. Dzianott, A.; Rauffer-Bruyere, N.; Bujarski, J.J. Studies on Functional Interaction between Brome Mosaic Virus Replicase Proteins during RNA Recombination, Using Combined Mutants in Vivo and in Vitro. *Virology* **2001**, *289*, 137–149. [[CrossRef](#)]
33. Kolondam, B.; Rao, P.; Sztuba-Solinska, J.; Weber, P.H.; Dzianott, A.; Johns, M.A.; Bujarski, J.J. Co-infection with two strains of Brome mosaic bromovirus reveals common RNA recombination sites in different hosts. *Virus Evol.* **2015**, *1*, vev021. [[CrossRef](#)] [[PubMed](#)]
34. Shahgolzari, M.; Pazhouhandeh, M.; Milani, M.; Khosroushahi, A.Y.; Fiering, S. Plant viral nanoparticles for packaging and in vivo delivery of bioactive cargos. *WIREs Nanomed. Nanobiotechnol.* **2020**, *12*, e1629. [[CrossRef](#)] [[PubMed](#)]
35. Liu, A.; de Ruiter, M.V.; Maassen, S.J.; Cornelissen, J.J.L.M. Nanoreactors via Encapsulation of Catalytic Gold Nanoparticles within Cowpea Chlorotic Mottle Virus Protein Cages. In *Protein Scaffolds 2018*; Humana Press: New York, NY, USA, 2018; Volume 1798, pp. 1–9. [[CrossRef](#)]
36. Vaughan, R.; Tragesser, B.; Ni, P.; Ma, X.; Dragnea, B.; Kao, C.C. The Tripartite Virions of the Brome Mosaic Virus Have Distinct Physical Properties That Affect the Timing of the Infection Process. *J. Virol.* **2014**, *88*, 6483–6491. [[CrossRef](#)]
37. Routh, A.; Domitrovic, T.; Johnson, J.E. Host RNAs, including transposons, are encapsidated by a eukaryotic single-stranded RNA virus. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 1907–1912. [[CrossRef](#)]
38. Gochhait, S.; Malhotra, D.; Rai, E.; Bamezai, R.N.K. Automated Fluorescence Sequencing and Troubleshooting. In *Advanced Techniques in Soil Microbiology 2007*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 35–51. [[CrossRef](#)]
39. Zhu, B. Bacteriophage T7 DNA polymerase-sequenase. *Front. Microbiol.* **2014**, *5*, 181. [[CrossRef](#)]
40. Cox, M.P.; Peterson, D.A.; Biggs, P.J. SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinform.* **2010**, *11*, 1–6. [[CrossRef](#)]
41. Jaganathan, D.; Bohra, A.; Thudi, M.; Varshney, R.K. Fine mapping and gene cloning in the post-NGS era: Advances and prospects. *Theor. Appl. Genet.* **2020**, *133*, 1791–1810. [[CrossRef](#)]
42. Yang, W.; Yoon, A.; Lee, S.; Kim, S.; Han, J.; Chung, J. Next-generation sequencing enables the discovery of more diverse positive clones from a phage-displayed antibody library. *Exp. Mol. Med.* **2017**, *49*, e308. [[CrossRef](#)]
43. Hunt, M.; Gall, A.; Ong, S.H.; Brener, J.; Ferns, B.; Goulder, P.; Nastouli, E.; Keane, J.A.; Kellam, P.; Otto, T.D. IVA: Accurate de novo assembly of RNA virus genomes. *Bioinformatics* **2015**, *31*, 2374–2376. [[CrossRef](#)] [[PubMed](#)]
44. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
45. Quinlan, A.R.; Hall, I.M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **2010**, *26*, 841–842. [[CrossRef](#)] [[PubMed](#)]