Genotyping of *Cucumber mosaic virus* isolates in western New York State during epidemic years

(Characterization of an emergent plant virus population)

Jeremy R. Thompson^{1*}, Jamie L. Langenhan¹, Marc Fuchs², and Keith L. Perry¹

¹Plant Pathology and Plant-Microbe Biology Section, School of Integrative Plant Science, Cornell University, 334 Plant Science Building, Ithaca, NY 14853-5904, USA; ²Plant Pathology and Plant-Microbe Biology Section, School of Integrative Plant Science, Cornell University, New York State Agricultural Experiment Station, Geneva, NY 14456, USA.

*Corresponding Author: jrt36@cornell.edu

Abstract : 301

Text: 4,788

Figs: 6

Tables: 2

Suppl: 7

Keywords: *Cucumber mosaic virus*, evolution, infectious cDNA clones, *Phaseolus vulgaris*, virus population

(Bn57-CMV full-length accession numbers: RNA1 - HF572914, RNA2 - HF572915, RNA3 - HF572916)

In the early 2000s an epidemic of **cucumber mosaic virus** (CMV) spread within the Midwestern and Eastern US affecting snap and dry bean (*Phaseolus vulgaris* L.) cultivation. Fifty one CMV isolates from this period were partially characterized from varied hosts by sequencing a section from each of the three genomic RNAs. Aside from one subgroup II strain from pepper, all isolates, including those from snap bean, fell within the IA subgroup. The nucleotide sequence diversity of virus populations sampled at multiple sites and at different years was significantly higher than that of a population from single site in a single year, although in general the number of polymorphisms was low (<11%). Complementary DNA (cDNA) clones of Bn57, a representative isolate from snap bean, were engineered for the production of infectious *in vitro* RNA transcripts initiated from a T7 promoter. Infections from these cDNAs resulted in symptoms consistent with those of the original field isolate, indicating that a satellite **RNA** is not involved in symptom expression in snap bean. These infectious clones were used to assess symptom determinants and the effects of virus infection on plant growth. Inoculations with pseudorecombinants derived from Bn57 and the non-bean infecting strain Fny confirmed RNA2 as a specific determinant for snap bean infection. Bn57, along with almost all isolates identified in this study contained the Y631 locus in the 2a protein, a determinant for systemic infection in bean. The presence of this locus extended to all non-bean hosts except two pepper infecting isolates. Infection by Bn57 in snap bean had a significant effect on pod number and mass with a 55 and 41 percent reduction in greenhouse assays, respectively. To our knowledge Bn57 is the first CMV strain isolated from *P. vulgaris* to be fully sequenced and cloned, providing a useful tool for analyses of CMV-host interactions.

1. Introduction

Recent decades have seen the emergence of **cucumber mosaic virus** (CMV) as a serious pathogen of common bean (*Phaseolus vulgaris* L.) in the Great Lakes region of the United States. This sudden increase in virus incidence coincided with the introduction of the soybean aphid (*Aphis glycines*) from Asia (Losey et al., 2002; Ragsdale et al., 2011). Severe virus symptoms were observed initially in the Midwest in 2000, and the incidence of virus diseases increased to the east and south reaching New York (NY) State the following years (Larsen et al., 2002; Shah et al., 2006). Reported field incidence and associated crop loss were sometimes as high as 100% (Larsen et al., 2002; Nault et al., 2006), the impact on yield of CMV infection being both dependent on cultivar (Taylor and Shail, 2006) and plant stage at the time of inoculation (Larsen et al., 2002). The overall total economic cost in NY State alone from 2001 to 2009 was estimated to be more than \$10 million (Reiners, 2010).

CMV has perhaps the largest host **range** of any virus with at least 1,241 susceptible species in 101 plant families, including both monocots and dicots. It can be transmitted by over 80 species of aphid in a non-persistent manner (Edwardson and Christie, 1991). The genome of CMV consists of three segments RNA1, RNA2 and RNA3 of approximately 3.3, 3.0 and 2.2 kilonucleotides in length, respectively. RNA1 is monocistronic and encodes an open reading frame (ORF) (1a) containing the replication associated methyltransferase and helicase domains. RNA2 is bistrionic with a large ORF (2a) containing a RNA-dependent RNA polymerase (RdRp) domain and a small overlapping ORF (2b) encoding a suppressor of RNA silencing. RNA3 is bicistronic: the movement protein (MP) associated ORF separated by the intergenic region (IGR) from the three-prime proximal coat protein (CP) ORF. All five viral proteins have been implicated in movement, but aside from the MP their role is assumed to be indirect (Jacquemond, 2012; Palukaitis and Garcia-Arenal, 2003). Phylogenetically, CMV can be subdivided into two main subgroups I and II which share 70-75% nucleotide identity and as such they taxonomically approximate two separate species. Subgroup I can be further subdivided into IA and IB; the latter, being predominantly

Asian in distribution (Jacquemond, 2012). Natural admixtures between these different subgroups usually occur infrequently in the form of reassortants and recombinants (Bonnet et al., 2005; Fraile et al., 1997; Kim et al., 2014; Lin et al., 2004; Maoka et al., 2010; Nouri et al., 2014).

In this study we describe the collection from Upstate NY of 51 CMV isolates collected during and immediately after the epidemic years of the early 2000s. Using partial sequences from each genomic RNA segment we analyzed the evolutionary trends underlying the virus population. From these data we identified a typical virus variant (Bn57) from snap bean and produced infectious clones for which infection in snap bean was partly characterized.

2. Materials and Methods

2.1. Plant sampling

Plants were collected over a five year period (2009-2014) sampling in and around cultivated bean fields located in the NY counties of Ontario, Orleans, Steuben and Tompkins (Fig. S1). In late July, 2009, 80 bean samples were collected from a single field displaying virus symptoms located in Geneva (Ontario County) NY. A structured sampling was undertaken, irrespective of whether the plants showed symptoms. The field was divided into four sections; samples were taken every ten plants in every other row moving diagonally across the field. In total twenty bean samples per section were collected. Weed samples were also collected in or next to the same field including one mayweed (*Anthemis arvensis*) plant (Mayweed-3). The same field was again surveyed in July, 2010, and while this time no symptomatic bean infection was noted, a symptomatic horsenettle (*Solanum carolinense*) plant (Horsenettle-1) was collected from the center of the field. Also in 2009 a collection was made of bean plants displaying symptoms in various fields in Steuben County. The following year, symptomatic plants were collected from various fields in Ontario County. In 2010 and four years later, additional cultivated plants showing virus-like symptoms (and shown to be infected by CMV) were collected. Archival material in the form of frozen or lyophilized leaf tissue with CMV either from bean or from summer squash growing in NY was also included in the analyses. The only non-NY sample was an archival bean collected from Wisconsin. This

was included as a representative sample of an isolate of CMV from the early 2000s where the disease in bean was first noted (Fig. S1).

2.2. ELISA

CMV infection for all samples was tested using either a plate trapped antibody-enzyme linked immunosorbent assay (PTA-ELISA (Mowat and Dawson, 1987)) or a triple antibody sandwich- enzyme linked immunosorbent assay (TAS-ELISA; Agdia Inc., Elkhart, IN). PTA-ELISA was carried out using an in-house anti-CMV polyclonal antibody. TAS-ELISA was carried out using commercial reagents according to the manufacturer's instructions. Each assay included two healthy bean controls, one known CMV-infected bean positive control, and one purified CMV virion positive control. Assays were considered positive if OD405nm values were three or more times that of the healthy control and greater than 0.1.

2.3. RNA extraction, RT-PCR and IC-RT-PCR

RNA extraction and sample storage. RNA extraction was done with the Qiagen (Valencia, CA) RNeasy Plant Mini Kit. Buffer RTL (Qiagen, Hilden, Germany) was used as the lysis buffer and extractions were eluted with 30 µl RNase-free water. RNA extracts were stored at -70°C. Infection in all ELISA-positive samples was initially confirmed using the universal detection primers p766 (5'-GCCCTTACTTTCTCATGGATGC-3') and p767 (5'-GACTGACCATTTTAGCCGT-3') in combination with internal control primers AtropaNad2.1a and AtropaNad2.2b (Thompson et al., 2003) (Table. S1). Sequencing primer pairs (F1 and R1, F2 and R2, and F3 and R3) (Integrated DNA Technologies, Coralville, IA) were designed to flank variable regions of interest in each of the genomic RNAs (Fig. 1). Their position was chosen based on conserved positions in sequence alignments between representatives of CMV subgroups IA, IB, and II; strains Fny (Acc. no. RNA1, D00356; RNA2, D00355; and RNA3, D10538), NT9 (Acc. no. RNA1, D28778; RNA2, D28779; and RNA3, D28780), and Q (Acc. no. RNA1, X02733; RNA2, X00985; and RNA3, J02059), respectively. Reverse transcription (RT) was

carried out using random hexamers (Promega, Madison, WI) and Superscript III (Life Technologies, Grand Island, NY) following the manufacturer's directions. PCR was done using cloned *Pfu* DNA polymerase AD (Agilent Technologies, Santa Clara, CA) according to the manufacturer's instructions. An immunocapture RT-PCR (IC-RT-PCR) was used to amplify the RNA of archival lyophilized samples. Plant material was prepared in 2.0 ml tubes. 0.2 g of each sample was flash frozen in liquid nitrogen and pulverized with a pestle. 300 µl PBS + 0.1 % Tween® was added to each tube and the tubes then centrifuged at 4 °C for 3 min. CMV in-house polyclonal antibody (1:500 dilution in coating buffer (0.1 M sodium carbonate, pH 9.5)) was used to coat 0.3 ml reaction tubes for 2 h at room temperature. Tubes were then washed with PBS + 0.1% Tween® and distilled water. 100 µl of the supernatant from each plant preparation was added to each tube and incubated at 4 °C overnight. Following incubation, the tubes were washed with PBS-Tween® and distilled water. Thermocycler conditions for all PCRs were: 1 cycle of 94°C for 5 min, 34 cycles of 94°C for 1 min, 50°C for 1 minute, 72°C for 5 min, and 1 cycle of 72°C for 10 min.

2.4. Nucleic acid sequence analysis

PCR amplicons derived from primer pairs F1 and R1, F2 and R2, and F3 and R3 were sequenced at the Cornell Core Facility in both directions and twice using the same primers used in PCR. Contigs were assembled using the Vector NTI® suite (Life Technologies, Grand Island, NY). Trimmed and edited sequences were aligned with T-Coffee (Tommaso et al., 2011). Substitution model selection and neighbor-joining, maximum likelihood (PhyML) (Guindon et al., 2010) and Bayesian (MrBayes) (Ronquist et al., 2012) inferred phylogenetic trees were generated in TOPALi2 (Milne et al., 2009). The 1a, 2a and MP open reading frames were screened for recombination using RDP3 (Martin et al., 2010), GARD and SBP (Pond et al., 2006) prior to identifying positively and negatively selected codons using the MEME (Murrell et al., 2012) and FUBAR (Murrell et al., 2013) algorithms. Global dN/dS ratios for each ORF were calculated using SLAC (Kosakovsky Pond and Frost, 2005). DNA polymorphisms were

analyzed using DnaSP 5.10 (Librado and Rozas, 2009). Significance in diversity between different CMV populations was assessed by using the collective pairwise distances in a Mann-Whitney U test.

2.5. Bn57 infectious clones

Degenerate primers based on subgroup IA alignments (not shown) for each genomic RNAs were designed. From these ends, cloning primers containing a minimal RNA polymerase T7 recognition sequence and restriction sites were designed (Table S2). PCR was done using cloned *Pfu* DNA polymerase AD (Agilent Technologies, Santa Clara, CA) according to the manufacturer's instructions. Thermocycler conditions were; 1 cycle of 94°C for 5 min, 34 cycles of 94°C for 1 min, 50°C for 1 minute, 72°C for 2 min, and 1 cycle of 72°C for 10 min. PCR amplicons were cleaned and concentrated using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corp. Irvine, CA) before ligating into pGEM®-T Easy vector (Promega, Fitchburg, WI) according to the manufacturer's instructions. Subcloning was achieved by excising fragments with the corresponding restriction enzymes and re-ligating into cleaved pUC19; for RNA2 (*Bam*HI/*Pst*I) and RNA3 (*Bam*HI alone). RNA1 in pGEM-T was used as a template without subcloning. For transcription, the resulting clones pBn57-1 (RNA1), pBn57-2 (RNA2) and pBn57-3 (RNA3) were linearized/excised by *Sac*I/*Sph*I, *Pst*I, and *Bam*HI, respectively, cleaned using the Zymoclean™ Gel DNA Recovery Kit and transcribed by the mMESSAGE mMACHINE® Kit (Life Technologies) according to the supplier's protocol. One µl of the final reaction (approximately 4µg transcribed RNA) was used to inoculate one plant by rubbing a carborundum dusted leaf. Snap beans (*P. vulgaris* 'Hystyle' (Harris Moran, Modesto, CA), *Chenopodium quinoa* and tobacco (*Nicotiana tabacum*) were used to assay infectivity. Fny-CMV transcripts were generated in a similar fashion from clones 109, 209 and 309 (Rizzo and Palukaitis, 1990).

2.6. Bn57 characterization

All plants were grown in greenhouse conditions of 22 ± 3 °C and 16/8 h (light/day). Lyophilized archival leaf samples of Bn57 were revived by inoculating in 0.1 M Na₂HPO₄ onto carborundum dusted bean

leaves. Pseudorecombinants were produced by mixing each genomic RNA from Bn57 (B) and Fny (F) in every possible combination (FBB, BFB, BBF, BFF, FBF and FFB) and inoculating to both bean and tobacco. Three weeks post-inoculation, viruses from symptomatic and asymptomatic bean and tobacco were mechanically back-inoculated onto both bean and tobacco. Virus infection was tested for using CMV lateral flow sticks (Agdia, Elkhart, IN) and plants were scored for infection on a weekly basis for four weeks. To document the effect of Bn57 infection in beans, bean seedlings were inoculated at the two true-leaf stage and allowed to grow until fruiting. The same number of plants was mock inoculated and allowed to grow alongside the infected ones, every week their positions being changed to minimize the local environment effects.

3. Results

3.1. Virus population diversities

In total the sequences of 51 isolates were analyzed (Table 1, Table S3). Except for one bean isolate from Wisconsin (Le) all viruses were taken from four counties in Upstate New York (Fig. S1). The positions of sequences on each genomic RNA analyzed in this study were: RNA1, 1895-2572 nt, RNA2 1577-1995 nt and RNA3 454-1349 nt (using CMV-Fny genomic RNAs as references) (Fig. 1). Altogether there were 40 variants isolated from bean, five from spinach, two each from pepper and squash, and one each from mayweed (*A. arvensis*) and horsenettle (*S. carolinense*). Aside from Ace-Betsey (isolated from pepper in 2010) all variants were identified as belonging to CMV subgroup IA. CMV populations were grouped at four levels: Bean single-site (BSS) – 17 samples isolated from snap bean growing in a single field in a single year (as indicated in Table 1 by the upper gray box), Bean multiple site (BMS) – 23 samples isolated from snap bean at various locations from 2004 to 2014, Bean multiple and single site (BMSS) – a combination of the above two snap bean isolates (40 samples), Beanfield (BF) – all 51 virus samples isolated in this study and Subgroup I (SGI) – 17 global virus strains from subgroup I including Bn57. The overall trend across all RNAs was comparable, with a significantly higher diversity (Mann-Whitney U test $P < 0.001$) in BMS samples compared to the single field (Table 2). The only significant Tajima's D

value obtained was for the RNA1 of infected bean at a single site and year with a value of -2.12, suggesting a post-bottleneck expanding population or a recent selective sweep specific for this gene.

3.2. Beanfield virus population phylogenetics

The overall low genetic diversity is reflected in phylogenetic trees obtained for each partial sequence of genomic RNA (Fig. 2). The single-site isolates group together roughly, but there is no visible correlation between relatedness and location or host or year of isolation, except for the two pepper-infecting isolates Ace-Betsey and 1177. In this case the Ace-Betsey isolate groups outside all other isolates while for RNA2, 1177 appears on a distinct branch in between Ace-Betsey and the remaining isolates. Interestingly, the RNA1 tree is the most structured of the three with a distinct clade containing the Wisconsin 2004 isolate Le, the NY 2005 isolate Bn28 and five 2010 multiple field isolates. Although less defined this clade is also monophyletic for RNA2 and 3, and therefore is probably representative of a distinct strain.

3.3. Beanfield virus population selection pressures

Three methods (FUBAR, MEME and SLAC) were used to calculate selection pressures across coding regions of the 1a, 2a and MP open reading frames (ORFs). As a reference we also analyzed the same codon positions for 17 global CMV subgroup I strains (Table S4). Both alignments analyzed contained a single aligned outgroup of Ace-Betsey and subgroup II strain R-CMV. Both SGI and BF selection pressures were comparable, being strongly purifying (<0.0782) for the replication associated genes (1a and 2a) and less so (<0.2198) for the MP (Fig. 3). These values were reflected in the codon specific analyses where the number of individual codons undergoing pervasive purifying selection (FUBAR) were visibly less (posterior probability >0.9) for the MP gene and under half the number detected for the 1a and 2a genes. Only a handful of codons in total were detected as undergoing episodic diversifying selection (MEME $p<0.1$). Three codons (ORF1a S696, ORF2a Y631 and MP A270) were common to both global and beanfield isolate alignments. SGI MPs had seven codons undergoing positive selection compared to

one for the BF (Fig. 3). An alignment of the 2a protein of both global and beanfield isolates around RdRP motifs IV and VI clearly identifies the Y631 position as unique to the bean infecting isolates of this study and the global CMV Iizuka which is a leguminous strain (Fig. 4).

3.4. Bn57 and CMV subgroup I phylogenetics

Full-length PCR amplicons of each RNA segment were produced using 5'- and 3'-terminus specific primers (Table S2). The infectivity of transcripts from these amplicons was initially individually confirmed on tobacco prior to amplicon ligation into a plasmid vector backbone. All three genomic segments RNA1, 2 and 3 were first ligated into the pGEM®-T Easy vector (Promega, Fitchburg, WI) and then subcloned into pUC19. While transcripts from pUC19 constructs with RNA2 (pBn57-2) and RNA3 (pBn57-2) were confirmed as infectious, the cloned RNA1 fragment proved unstable in the latter vector for unknown reasons and the pGEM®-T Easy construct (pBn57-1) with a duplicated T7 promoter was used instead (Fig. S2). The full-length sequences of each Bn57 segment (RNA1 - HF572914, RNA2 - HF572915, RNA3 - HF572916) were obtained and compared with known full-length subgroup I global strains. Neighbor-joining phylogenetic trees show that each Bn57 segment forms a distinct and consistent monophyletic group with the 1A subgroup strains Fny, I17F, Rs and Y (Fig. 5). Strain Bn57 most closely associates with the leguminous strain Iizuka, while the other legume infecting strains 209 (soybean), CS (peanut) tend to associate with a more loosely defined 1B clade.

3.5. Characterization of Bn57-CMV infection

Inoculation assays of pseudorecombinants between Bn57 (B) and Fny (F), a cucurbit isolate, in all possible combinations (FBB, BFB, BBF, BFF, FBF and FFB) were undertaken. Although all pseudorecombinants were infectious with comparable symptoms in tobacco, only those containing the RNA2 of Bn57 (FBB, BBF, FBF) systemically infected snap bean (Fig. 6). Virus symptoms in snap bean cv. Hystyle and tobacco of transcript derived Bn57 (BBB) were comparable to those of the original archival Bn57 isolate (Fig. 6). Typically in bean, around two weeks post inoculation, expanding leaves

would show symptoms, ranging in severity, of mosaic, leaf distortion and blister-like bumps. Those pseudorecombinants containing the RNA2 of Fny were not able to systemically infect snap bean, producing small necrotic lesions on the inoculated leaf. The full effects of Bn57 infection was assessed in four separate greenhouse experiments in which pod number, pod mass and upper plant fresh weight were compared between infected and uninfected beans. On average, infected plants produced 4.2 (55%) fewer pods and 7.6 g (41%) lower pod mass than their uninfected counterparts (Fig. S3). Differences for both measurements were significant ($p < 0.05$) in a two-tailed t-test. Plant fresh weight was also reduced by 18% ($p = 0.051$)

4. Discussion

A goal of this present study was to gain a better understanding of the CMV bean epidemic that began in the early 2000s in the Great Lakes region. Dissemination of the virus was likely facilitated by the arrival in the US from Asia of the soybean aphid (Losey et al., 2002; Ragsdale et al., 2011). The 2004/5 bean isolates Le (from Wisconsin), Bn11, Bn28 and Bn57 are representative from when the epidemic was at its peak. These earlier isolates vary little phylogenetically from later isolates, with Le and Bn28 forming a strain distinct from the other two. All of the isolates except the pepper-infecting Ace-Betsey and 1177 contain the legume-associated tyrosine (Y) at position 631 in the 2a protein, instead of a phenylalanine (F). This position along with an alanine (A) to serine (S) change at position 641 were identified as necessary for systemic infection in cowpea (*Vigna sinensis*); the presence of either 631Y or 641S without the other resulting in systemic necrosis or local lesions, respectively (Kim and Palukaitis, 1997). The significance of 641S was disputed in a similar study where 631Y was found to be sufficient for the elicitation of symptoms in cowpea (Karasawa et al., 1999), an observation consistent with symptoms in *Nicotiana benthamiana* plants transgenic for the resistance gene *RT4-4* of common bean (Seo et al., 2006). Although sequencing of RNA2 for the beanfield isolates did not extend to position 641, both 631Y and 641S are present in the Bn57 sequence. The mechanisms underlying these shifts in host responses are unclear and plausible hypotheses involving predicted conformational changes of the 2a protein and/or

phosphorylation at the positions 631 and 641, such as those proposed by Karasawa et al. (1999) and Kim and Palukaitis (1997), respectively, have so far yet to be tested beyond confirming that a phosphorylation mimic mutant F631D was phenotypically comparable to F631Y (Seo et al., 2006). In a yeast two-hybrid study a specific interaction between the MP and the 2a protein was mapped to the region surrounding the GDD motif in the latter (Hwang et al., 2007). Although strain-specific compatibility between the MP and 2a is clearly not required, as inferred in this and previous studies using pseudorecombinants (Edwards et al., 1983; Kim and Palukaitis, 1997), there might exist possible additional host factors that facilitate this interaction in a host-compatible manner. The maintenance of at least Y631 in isolates from non-bean hosts (spinach and summer squash) suggests it may be selectively beneficial (or at least not deleterious) in other hosts. Alternatively, CMV bean populations might be so dominant in the sampled regions that spillover to other hosts becomes inevitable. This phenomenon is further highlighted by its detection in weeds (mayweed and horsenettle) found growing in the single collection site in 2009. Although not demonstrated here, such infections could easily serve as reservoirs for infection in bean in subsequent seasons.

A closer look at the global legume-infecting sequence data exposes a conflicting picture of genetic compatibilities between virus and host. The legume infecting Iizuki strain from cowpea also contains both mutations, however legume infecting 209 (wild soybean – *Glycina soja*) and CS (peanut - *Arachis hypogaea*) strains curiously do not contain either. In fact, strain 209 has uncovered a complex relationship with legumes, causing necrotic local lesions in cowpea, severe stunting in pea (*Pisum sativum*) and mosaic and stunting in soybean, while being unable to infect cultivated soybean (*G. max*) (Phan et al., 2014).

The population structure of the beanfield isolates is relatively simple with low nucleotide diversity. Recombination was not detected. By analyses using all three genome segments concatenated, one isolate, 1177, was identified as a possible reassortant with significant breakpoints detected for all methods employed (not shown). The phylogenetic trees (Fig. 2) substantiate these results showing that 1177 is less ‘bean-like’ for RNA2 than for RNAs 1 and 3. The genomic structure, although only partial

for each segment, could be interpreted as reflecting adaptive transitioning between hosts; 1177 is the only subgroup I beanfield isolate that does not contain the Y631 amino acid in the 2a protein (Fig. 5). A recent genetic study of CMV populations of pepper in South Korea identified geographically distinct subpopulations (Kim et al., 2014), a phenomenon also reported in Tunisia (Ben Tamarzizt et al., 2013). Genetic diversities in the Korean populations were highest at 0.060 in the 1a ORF and lowest at 0.029 in the MP ORF (Kim et al., 2014). At the other extreme, a Californian CMV-cucurbit population taken from the same site in the same year exhibited very low variability; averaging a mean diversity of 0.006 (Lin et al., 2004). Although these values for diversity are consistent with those found in here in bean, comparing calculated global selection pressures between CMV beanfield isolates and both CMV Korean pepper isolates (Kim et al., 2014) and multiple-location, -host and -year Californian isolates (Lin et al., 2004) show distinct differences; ORF1a (0.08 vs. 0.06 and 0.07), ORF2a (0.06 vs. 0.13 and 0.25) and MP (0.18 vs. 0.1 and 0.05). A stronger purifying selection on the 2a ORF given its role in host specificity is not unexpected, however the reasons for the differences in the MP are unclear.

The first stable infectious clones of CMV were for the cucurbit-infecting subgroup I strains Fny (Rizzo and Palukaitis, 1990) and Y (Suzuki et al., 1991). Since then a number of isolates from both subgroups have been fully sequenced and cloned (Carrere et al., 1999; Ding et al., 1995; Salanki et al., 1997; Tousignant et al., 1996) the most recent being strain 209 isolated from soybean (Phan et al., 2014). As far as we know Bn57 is the first common bean CMV isolate from which complete infectious clones have been constructed. Importantly, the symptoms induced by transcript derived Bn57 were comparable to those obtained by the original isolate. Satellite RNAs have been reported in bean-infecting CMVs (Nouri et al., 2012). Although the presence of satellite RNAs may play a role in CMV-associated symptoms, our results indicate that a satellite RNA is not required for symptom development in snap bean. The effect of transcript originating-Bn57 infection on bean yield under controlled conditions was found to be significant. Whether this translates to field conditions is beyond the scope of this work, but these results correlate well with field trials where up to 60% loss of yield was calculated, and 31% in 2004 specifically for the Hystyle variety used in these experiments (Taylor and Shail, 2006).

Unlike other viruses that infect bean, such as **bean common mosaic virus** and **bean common mosaic necrosis virus**, where the genetics of resistance are well characterized (Morales, 2003; Provvidenti, 2001), resistance to CMV is poorly understood. In a screen of 109 snap and dry bean cultivars, and 405 common bean accessions, no resistance to CMV was identified (Griffiths, 2004). A more recent study of 49 Iranian lines identified one moderately resistant line (Azizi and Shams-Bakhsh, 2014). Resistance in other *Phaseolus* species does exist (Griffiths, 2004; Provvidenti, 1976), although to-date successful introgressions have not been reported. The resistance gene *RT4-4* found in the common bean cultivar Othello induced systemic necrosis when challenged by seven non-bean infecting strains of CMV. However, infection of Othello by a bean isolate resulted in systemic infection (Seo et al., 2006). In response to the epidemic of the early 2000s and in the absence of any well-defined resistance, field trials were carried out in Upstate NY in 2008 and 2009 in order to identify tolerance to CMV infection in twenty-one varieties. Of five varieties exhibiting tolerance, one, Huntington (Syngenta), has now been adopted by most growers throughout the Great Lakes region (S. Reinert, Cornell University - personal communication).

In conclusion, this study provides an invaluable insight into an emergent plant virus population during and immediately after an epidemic, highlighting a limited genetic diversity across a decade with bean-adapted genotypes found in a number of non-bean hosts. The construction of infectious clones of a representative bean-infecting isolate was able to show RNA2 as the specific determinant for systemic infection in bean and that a satellite RNA is not required for symptom expression.

Acknowledgements

Thanks to Pawan Angara for gathering phylogenetic tree data. Thanks also to Kevin Meyers for his support in the greenhouse, and Carol McNeil and Julie Kikkert from Cornell Vegetable Program of Cornell Cooperative Extension, for their help in the field. Brian Nault provided some of the archival beanfield samples. Alan Taylor provided seeds of cv. Hystyle. **Work was in part funded by Gates Foundation Grand Challenges grant no. OPP 1068482.**

Figure 1. The relative positions (black bars) of sequences on each genomic RNA (RNAs 1, 2 and 3) (Fny strain as reference) analyzed. RNA1, nucleotide (nt) positions 1895-2572; RNA2, 1577-1995; and RNA3, 454-1349. Open reading frames – gray shaded boxes. MP – movement protein, CP – capsid protein. Length of each RNA in nt shown on right. Oval – 5' cap structure, cloverleaf – tRNA-like structure.

Figure 2. Neighbor-joining derived phylogenetic trees using the partial sequences of genomic RNAs 1, 2 and 3 characterized from CMV isolates in this study. Branch tips highlighted in gray represent non-bean-infecting isolates. Bar shows genetic distance. Numbers at nodes show bootstrap values. Nodes with values less than 70% in the neighbor-joining tree were collapsed. Nodes with black circles were also supported (>70% bootstrap) by the maximum likelihood method PhyML (Guindon et al., 2010).

Figure 3. – Selection pressures in sequences from regions of the three CMV genomic RNAs (encoding proteins 1a, 2a and MP) for global subgroup I isolates (SGI) (upper graphs) and beanfield isolates (BF) (lower graphs). Histogram bars (blue) depict posterior probabilities of pervasive purifying selection for each codon as determined by FUBAR (Murrell et al., 2013). Values are considered significant if they exceed the threshold of 0.9 (black dotted line). The proportion of codons significant for pervasive purifying selection is shown as a percentage. In red are shown amino acid positions (in isolate Bn57) undergoing episodic diversifying selection as determined by MEME (Murrell et al., 2012). In curly brackets are the global dN/dS ratios for each alignment.

Figure 4. Amino acid alignment of a conserved stretch of the 2a protein in beanfield (BF) and global subgroup I (SGI) CMV sequences. Isolate/strain names are shown on the left. Representative isolate sequences from this study (in gray) are included as most of them were identical. White or lighter blue boxes highlight variable positions. Letters at bottom indicate conserved polymerase motifs (roman numerals) as described by Koonin (1991). The legume associated F631Y mutation is shown in red.

Circles adjacent to names indicate legume-infecting isolates. The dot adjacent to the name indicates a non-subgroup I isolate.

Figure 5. Neighbor-joining derived phylogenetic trees of CMV genomic RNAs 1, 2, and 3 of Bn57 and other full-length global CMV subgroup I isolates. Subgroup IA sequences are highlighted by the shaded blue rectangle. Bar shows genetic distance. Numbers at nodes show bootstrap values. Nodes with values less than 70% in the neighbor-joining tree were collapsed. Nodes with black circles were also supported (>70% bootstrap) by both maximum likelihood (Guindon et al., 2010) and Bayesian inference (Ronquist and Huelsenbeck, 2003). Nodes with white circles were only supported by Bayesian inference. Node with gray circle was only supported by maximum likelihood. Names in bold are viruses isolated from legumes.

Figure 6. Photographs of symptoms in expanding leaves of bean following infection with Bn57 and pseudorecombinants at 18 days post inoculation. Each capital letter in a triplet refers to the origin of the genomic RNAs (1, 2, and 3); Bn57 (B) or Fny (F).

References

- Azizi, A., Shams-Bakhsh, M., 2014. Impact of *Cucumber mosaic virus* infection on the varietal traits of common bean cultivars in Iran. *Virus Dis.* 25(4), 447-454.
- Ben Tamarzizt, H., Montarry, J., Girardot, G., Fakhfakh, H., Tepfer, M., Jacquemond, M., 2013. *Cucumber mosaic virus* populations in Tunisian pepper crops are mainly composed of virus reassortants with resistance-breaking properties. *Plant Pathol.* 62(6), 1415-1428.
- Bonnet, J., Fraile, A., Sacristan, S., Malpica, J.M., Garcia-Arenal, F., 2005. Role of recombination in the evolution of natural populations of *Cucumber mosaic virus*, a tripartite RNA plant virus. *Virology* 332(1), 359-368.

417 Carrere, I., Tepfer, M., Jacquemond, M., 1999. Recombinants of *Cucumber mosaic virus* (CMV):
418 determinants of host range and symptomatology. Arch. Virol. 144(2), 365-379.

419 Ding, S.W., Rathjen, J.P., Li, W.X., Swanson, R., Healy, H., Symons, R.H., 1995. Efficient infection
420 from cDNA clones of cucumber mosaic cucumovirus RNAs in a new plasmid vector. J Gen.
421 Virol. 76 (Pt 2), 459-464.

422 Edwards, M.C., Gonsalves, D., Provvidenti, R., 1983. Genetic-analysis of *Cucumber mosaic virus* in
423 relation to host-resistance - location of determinants for pathogenicity to certain legumes and
424 lactuca-saligna. Phytopathology 73(2), 269-273.

425 Edwardson, J.R., Christie, R.G., 1991. Cucumoviruses, CRC Handbook of Viruses Infecting Legumes.,
426 CRC Press, Boca Raton, FL, pp. 293-319.

427 Fraile, A., AlonsoPrados, J.L., Aranda, M.A., Bernal, J.J., Malpica, J.M., GarciaArenal, F., 1997. Genetic
428 exchange by recombination or reassortment is infrequent in natural populations of a tripartite
429 RNA plant virus. J Virol. 71(2), 934-940.

430 Griffiths, P., 2004. Breeding snap beans for *Cucumber mosaic virus* (CMV) resistance. Hortscience 39(4),
431 869-869.

432 Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms
433 and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML
434 3.0. Syst. Biol. 59(3), 307-321.

435 Hwang, M.S., Kim, K.N., Lee, J.H., Park, Y.I., 2007. Identification of amino acid sequences determining
436 interaction between the cucumber mosaic virus-encoded 2a polymerase and 3a movement
437 proteins. J Gen. Virol. 88, 3445-3451.

438 Jacquemond, M., 2012. Cucumber mosaic virus. Adv. Virus Res. 84, 439-504.

439 Karasawa, A., Okada, I., Akashi, K., Chida, Y., Hase, S., Nakazawa-Nasu, Y., Ito, A., Ehara, Y., 1999.
440 One amino acid change in cucumber mosaic virus RNA polymerase determines virulent/avirulent
441 phenotypes on cowpea. Phytopathology 89(12), 1186-1192.

- Kim, C.H., Palukaitis, P., 1997. The plant defense response to cucumber mosaic virus in cowpea is elicited by the viral polymerase gene and affects virus accumulation in single cells. *EMBO J* 16(13), 4060-4068.
- Kim, M.K., Seo, J.K., Kwak, H.R., Kim, J.S., Kim, K.H., Cha, B.J., Choi, H.S., 2014. Molecular genetic analysis of *Cucumber mosaic virus* populations infecting pepper suggests unique patterns of evolution in Korea. *Phytopathology* 104(9), 993-1000.
- Koonin, E.V., 1991. The phylogeny of RNA-dependent RNA-polymerases of positive-strand RNA viruses. *J Gen. Virol.* 72, 2197-2206.
- Kosakovsky Pond, S.L., Frost, S.D., 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol. Biol. Evol.* 22(5), 1208-1222.
- Larsen, R.C., Miklas, P.N., Eastwell, K.C., Grau, C.R., Mondjana, A., 2002. A virus disease complex devastating late season snap bean production in the Midwest. *Annu. Rep. Bean Improv. Coop.* 45, 36-37.
- Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25(11), 1451-1452.
- Lin, H.X., Rubio, L., Smythe, A.B., Falk, B.W., 2004. Molecular population genetics of *Cucumber mosaic virus* in California: Evidence for founder effects and reassortment. *J Virol.* 78(12), 6666-6675.
- Losey, J.E., Waldron, J.K., Hoebeke, E.R., Macomber, L.E., Scott, B.N., 2002. First record of the soybean aphid, *Aphis glycines* Matsumura (Hemiptera : Sternorrhyncha : Aphididae), in New York. *Great Lakes Entomol* 35(2), 101-105.
- Maoka, T., Hayano, Y.S., Iwasaki, M., Yoshida, K., Masuta, C., 2010. Mixed infection in tomato to ensure frequent generation of a natural reassortant between two subgroups of *Cucumber mosaic virus*. *Virus Genes* 40(1), 148-150.
- Martin, D.P., Lemey, P., Lott, M., Moulton, V., Posada, D., Lefevre, P., 2010. RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* 26(19), 2462-2463.

468 Milne, I., Lindner, D., Bayer, M., Husmeier, D., McGuire, G., Marshall, D.F., Wright, F., 2009. TOPALi
469 v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters
470 and multi-core desktops. *Bioinformatics* 25(1), 126-127.

471 Morales, F.J., 2003. Virus and Virus-like Diseases of Major Crops in Developing Countries. In:
472 Loebenstein, G., Thottappilly, G. (Eds.), *Virus and virus-like diseases of major crops in*
473 *developing countries*. Kluwer Academic Publishers, Dordrecht, pp. 425-445.

474 Mowat, W.P., Dawson, S., 1987. Detection and identification of plant-viruses by ELISA using crude sap
475 extracts and unfractionated antisera. *J Virol. Methods* 15(3), 233-247.

476 Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Pond, S.L.K., Scheffler, K., 2013.
477 FUBAR: A fast, unconstrained Bayesian approximation for inferring selection. *Mol. Biol. Evol.*
478 30(5), 1196-1205.

479 Murrell, B., Wertheim, J.O., Moola, S., Weighill, T., Scheffler, K., Pond, S.L.K., 2012. Detecting
480 individual sites subject to episodic diversifying selection. *PLOS Genet.* 8(7).

481 Nault, B.A., Shah, D.A., Taylor, A.G., 2006. Viruses and aphids everywhere in New York snap bean
482 fields in 2005,. *Proceedings of the 2006 Empire State Fruit and Vegetable EXPO*, 74-76.

483 Nouri, S., Arevalo, R., Falk, B.W., Groves, R.L., 2014. Genetic structure and molecular variability of
484 *Cucumber mosaic virus* isolates in the United States. *PLOS One* 9(5).

485 Nouri, S., Falk, B.W., Groves, R.L., 2012. A new satellite RNA is associated with natural infections of
486 *Cucumber mosaic virus* in succulent snap bean. *Arch. Virol.* 157(2), 375-377.

487 Palukaitis, P., Garcia-Arenal, F., 2003. Cucumoviruses. *Adv. Virus Res.* 62, 241-323.

488 Phan, M.S., Seo, J.K., Choi, H.S., Lee, S.H., Kim, K.H., 2014. Molecular and biological characterization
489 of an isolate of *Cucumber mosaic virus* from *Glycine soja* by generating its infectious full-
490 genome cDNA clones. *Plant Pathol. J* 30(2), 159-167.

491 Pond, S.L.K., Posada, D., Gravenor, M.B., Woelk, C.H., Frost, S.D.W., 2006. Automated phylogenetic
492 detection of recombination using a genetic algorithm. *Mol. Biol. Evol.* 23(10), 1891-1901.

493 Provvidenti, R., 1976. Reaction of Phaseolus and Macroptilium species to a strain of *Cucumber mosaic*
 494 *virus*. Plant Dis. Rep. 60(4), 289-293.

495 Provvidenti, R., 2001. Inheritance of resistance to the NL-8 strain of *Bean common mosaic necrosis virus*
 496 in bean. Hortscience 36(5), 958-960.

497 Ragsdale, D.W., Landis, D.A., Brodeur, J., Heimpel, G.E., Desneux, N., 2011. Ecology and management
 498 of the soybean aphid in North America. Annu. Rev. Entomol. 56, 375-399.

499 Reiners, S., J. Ballerstein and J. Shail. 2010. Evaluating NYS snap bean varieties for CMV susceptibility.
 500 Proceedings of the Empire State Fruit & Vegetable Expo, January 25-27, 2010, Syracuse, NY,
 501 pp.51-53.

502 Rizzo, T.M., Palukaitis, P., 1990. Construction of full-length cDNA clones of *Cucumber mosaic virus*
 503 RNAs 1, 2 and 3: generation of infectious RNA transcripts. Mol. Gen. Genet. 222(2-3), 249-256.

504 Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models.
 505 Bioinformatics 19(12), 1572-1574.

506 Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L.,
 507 Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference
 508 and model choice across a large model space. Syst. Biol. 61(3), 539-542.

509 Salanki, K., Carrere, I., Jacquemond, M., Balazs, E., Tepfer, M., 1997. Biological properties of
 510 pseudorecombinant and recombinant strains created with *Cucumber mosaic virus* and *Tomato*
 511 *aspermy virus*. J Virol. 71(5), 3597-3602.

512 Seo, Y.S., Rojas, M.R., Lee, J.Y., Lee, S.W., Jeon, J.S., Ronald, P., Lucas, W.J., Gilbertson, R.L., 2006.
 513 A viral resistance gene from common bean functions across plant families and is up-regulated in
 514 a non-virus-specific manner. Proc. Natl. Acad. Sci. U S A 103(32), 11856-11861.

515 Shah, D.A., Dillard, H.R., Mazumdar-Leighton, S., Gonsalves, D., Nault, B.A., 2006. Incidence, spatial
 516 patterns, and associations among viruses in snap bean and alfalfa in New York. Plant Dis. 90(2),
 517 203-210.

- Suzuki, M., Kuwata, S., Kataoka, J., Masuta, C., Nitta, N., Takanami, Y., 1991. Functional-analysis of deletion mutants of *Cucumber mosaic virus* RNA3 using an in vitro transcription system. *Virology* 183(1), 106-113.
- Taylor, A.G., Shail, J.W., 2006. The impact of CMV in reducing yield of selected snap bean cultivars. *Annu. Rep. Bean Improv. Coop.* 49, 199-200.
- Thompson, J.R., Wetzel, S., Klerks, M.M., Vaskova, D., Schoen, C.D., Spak, J., Jelkmann, W., 2003. Multiplex RT-PCR detection of four aphid-borne strawberry viruses in *Fragaria* spp. in combination with a plant mRNA specific internal control. *J Virol. Methods* 111(2), 85-93.
- Tommaso, P., Moretti, S., Xenarios, I., Orobittg, M., Montanyola, A., Chang, J.M., Taly, J.F., Notredame, C., 2011. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res.* 39, W13-W17.
- Tousignant, M.E., McGarvey, P.M., Piazzolla, P., Kaper, J.M., 1996. *Cucumber mosaic virus* from *Ixora*: infectious RNA transcripts confirm deficiency in satellite support and unusual symptomatology. *Res.Virology* 147(6), 363-373.

545

546

547

548

549

550

551

Name	Year	Host Species	Location#
G5	2009	<i>Phaseolus vulgaris</i>	Ontario
G8	2009	<i>P. vulgaris</i>	Ontario
G10	2009	<i>P. vulgaris</i>	Ontario
G15	2009	<i>P. vulgaris</i>	Ontario
G19	2009	<i>P. vulgaris</i>	Ontario
G26	2009	<i>P. vulgaris</i>	Ontario
G33	2009	<i>P. vulgaris</i>	Ontario
G35	2009	<i>P. vulgaris</i>	Ontario
G38	2009	<i>P. vulgaris</i>	Ontario
G42	2009	<i>P. vulgaris</i>	Ontario
G46	2009	<i>P. vulgaris</i>	Ontario
G49	2009	<i>P. vulgaris</i>	Ontario
G50	2009	<i>P. vulgaris</i>	Ontario
G55	2009	<i>P. vulgaris</i>	Ontario
G63	2009	<i>P. vulgaris</i>	Ontario
G76	2009	<i>P. vulgaris</i>	Ontario
G80	2009	<i>P. vulgaris</i>	Ontario
Mayweed-3	2009	<i>Anthemis arvensis</i>	Ontario
Horsenettle-1	2010	<i>Solanum carolinense</i>	Ontario
CMV111	1979	<i>Cucurbita pepo</i>	Ontario
Bn11	2004	<i>P. vulgaris</i>	Ontario
Le	2004	<i>P. vulgaris</i>	(WI)
CMV3	2005	<i>C. pepo</i>	Ontario
Bn57	2005	<i>P. vulgaris</i>	Orleans
Bn28	2005	<i>P. vulgaris</i>	Orleans
F-A	2009	<i>P. vulgaris</i>	Steuben
F-B	2009	<i>P. vulgaris</i>	Steuben
F-D	2009	<i>P. vulgaris</i>	Steuben
F-E	2009	<i>P. vulgaris</i>	Steuben
F-G	2009	<i>P. vulgaris</i>	Steuben
B1	2010	<i>P. vulgaris</i>	Ontario
B2	2010	<i>P. vulgaris</i>	Ontario
B3	2010	<i>P. vulgaris</i>	Ontario
B4	2010	<i>P. vulgaris</i>	Ontario
B5	2010	<i>P. vulgaris</i>	Ontario
B6	2010	<i>P. vulgaris</i>	Ontario
B7	2010	<i>P. vulgaris</i>	Ontario
B8	2010	<i>P. vulgaris</i>	Ontario
B9	2010	<i>P. vulgaris</i>	Ontario
B12	2010	<i>P. vulgaris</i>	Ontario
B14	2010	<i>P. vulgaris</i>	Ontario
B16	2010	<i>P. vulgaris</i>	Ontario
B18	2010	<i>P. vulgaris</i>	Ontario
1177	2010	<i>Capsicum annuum</i>	Tompkins
Ace-Betsey	2010	<i>C. annuum</i>	Tompkins
Bn2014A	2014	<i>P. vulgaris</i>	Orleans
Sp-MB1	2014	<i>Spinacia oleracea</i>	Orleans
Sp-MB2	2014	<i>S. oleracea</i>	Orleans
Sp-J1	2014	<i>S. oleracea</i>	Orleans
Sp-TP1	2014	<i>S. oleracea</i>	Orleans
Sp-TP2	2014	<i>S. oleracea</i>	Orleans

552 Table 1. List of CMV infected plant samples analyzed in this study. Groups of isolates highlighted in gray
553 were taken together from one of two single fields, one planted to bean and the other to spinach. The
554 upper gray box from G5 to Horsenettle-1 includes all bean single-site (BSS) samples. # - NY State
555 counties listed in full. WI – Wisconsin State.

556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578

RNA1

Population	# Isolates	Diversity (π)	SD (π)	Tajima's D	# haplotypes	Haplotype diversity (Hd)	SD (Hd)	Total # nt	Polymorphic sites	Singletons
BSS	17	0.00672	0.00254	-2.1194*	9	0.787	0.101	678	32	28
BMS	23	0.02945	0.00241	0.4612	14	0.929	0.035	678	73	14
BMSS	40	0.02276	0.00283	-0.4941	22	0.935	0.024	678	73	14
BF	50	0.02270	0.00247	-0.5257	28	0.946	0.019	678	77	17
SGI	17	0.10124	0.00875	-0.6190	17	1.000	0.02	678	231	84

Diversity significance between single (BSS) and multiple sites (BMS) only (Mann Whitney pairwise distances) $U_A=5098$ $z=11.45$ $P<0.001$.

RNA2

Population	# Isolates	Diversity (π)	SD (π)	Tajima's D	# haplotypes	Haplotype diversity (Hd)	SD (Hd)	Total # nt	Polymorphic sites	Singleton
BSS	17	0.00639	0.00134	-0.6583	7	0.794	0.078	419	11	6
BMS	23	0.00858	0.00119	0.3683	7	0.763	0.061	419	12	3
BMSS	40	0.00861	0.00111	-0.7697	13	0.806	0.052	419	20	9
BF	50	0.01179	0.00185	-1.4160	18	0.859	0.039	419	37	23
SGI	17	0.07711	0.00532	-0.4652	17	1.000	0.020	419	108	38

Diversity significance between single (BSS) and multiple sites (BMS) only (Mann Whitney pairwise distances) $U_A=12526$ $z=4.33$ $P<0.001$

RNA3

Population	# Isolates	Diversity (π)	SD (π)	Tajima's D	# haplotypes	Haplotype diversity (Hd)	SD (Hd)	Total # nt	Polymorphic sites	Singletons
BSS	17	0.00719	0.00044	0.6926	10	0.904	0.050	871(896)*	18	5
BMS	23	0.01506	0.00123	0.2957	16	0.945	0.034	871(896)*	43	9
BMSS	40	0.01259	0.00122	-0.3059	24	0.940	0.022	871(896)*	49	13
BF	50	0.01241	0.00112	-0.8548	31	0.955	0.016	871(896)*	60	23
SGI	17	0.06775	0.00715	-0.8830	17	1.000	0.020	878(945)*	225	117

Diversity significance between single (BSS) and multiple sites (BMS) only (Mann Whitney pairwise distances) $U_A=8221$ $z=8.42$ $P<0.001$.

Table 2. Population genetics parameters of different CMV populations for different genomic RNAs.

FIG1

FIG 2

FIG3

FIG 4

Fig.5

FIG 6