


SHORT REPORT

Open Access



# Full genome sequence analysis and putative host-shifting of *Milk vetch dwarf virus* infecting tobacco (*Nicotiana tabacum*) in China

Ali Kamran<sup>1,2†</sup> , Han Hou<sup>1,2†</sup>, Yi Xie<sup>1</sup>, Cunxiao Zhao<sup>3</sup>, Xiaomin Wei<sup>3</sup>, Chaoqun Zhang<sup>4</sup>, Xiangwen Yu<sup>5</sup>, Fenglong Wang<sup>1\*</sup> and Jinguang Yang<sup>1\*</sup>

## Abstract

**Background:** Tobacco production in China has been affected by plant viruses with Milk vetch dwarf virus (MDV) as a recent invader posing serious concern. According to most of the studies, MDV mainly infects hosts from Fabaceae family but in our previous study we reported its infection in tobacco plant (*Nicotiana tabacum* L.) in Shandong province.

**Findings:** In current study (2016–2017), tobacco plants (*Nicotiana tabacum*) with severe stunting, yellowing and axillary bunch of new leaves were observed in Zhengning, Gansu province. Isolate GSZN yielded into eight genomic circular single-stranded DNA components while no alphasatellite DNA was obtained. High percent identity of this isolate was recorded in overall nucleotide and amino acid assembly with reported MDV isolates worldwide. Phylogenetic analysis fetched into a separate sub-clade comprising of new isolate along with other tobacco infecting isolates of MDV. While recombination was predicted in DNA-C encoding Clink protein and DNA-U1, which may attribute towards the potential host-shifting phenomenon and ability of this virus to expand its host range.

**Conclusion:** To our knowledge this is the first full genome annotation of a Nanovirus, infecting tobacco in natural field conditions, also this is the first extended analysis on host-shifting behavior of MDV.

**Keywords:** Nanovirus, Rolling circle amplification, Full genome sequence, *Milk vetch dwarf virus*, MDV, *N. tabacum*, Recombination, Host-shifting

## Main text

Tobacco is an important economic crop worldwide, with half of its farming in China; world's largest producer with 1,259,549 ha of cultivated area, producing 2,806,770 t annually [1]. Viral diseases have been reported as a potential threat to tobacco production in China, where 25 plant viruses are considered to be involved [2–4]. Among them, *Milk vetch dwarf virus* (MDV) has caused a serious damage to legume crops throughout the Japan as well as in central and south China, since its first detection in 1968 [5–8]. Despite the preference of legumes, recently a report

was published regarding MDV infection causing severe stunting and crinkling in tobacco (*N. tabacum*) from an eastern province of China [4].

MDV is a member of genus Nanovirus in family Nanoviridae and transmit by aphid vector in a persistent manner. The genome is circular single-stranded DNA (cssDNA), comprising of at least eight genomic (~ 1 kb each) and occasionally three or four alphasatellite DNAs that encapsidate separately as isometric particles of 18 nm giving multipartite genomic assembly to MDV. Each genomic DNA encode a distinct protein i.e., DNA-C, DNA-N, DNA-M, DNA-R, and DNA-S encodes; cell-cycle link protein (Clink), nuclear-shuttle protein (NSP), movement protein (MP) master replication protein (M-Rep) and capsid protein (CP), respectively [7, 9, 10], while the functions for DNA-U1, DNA-U2 and DNA-U4 are unknown [11].

\* Correspondence: wangfenglong@caas.cn; yangjinguang@caas.cn

†Ali Kamran and Han Hou contributed equally to this work.

<sup>1</sup>Key Laboratory of Tobacco Pest Monitoring Controlling & Integrated Management, Tobacco Research Institute of Chinese Academy of Agricultural Sciences, Qingdao 266101, China

Full list of author information is available at the end of the article

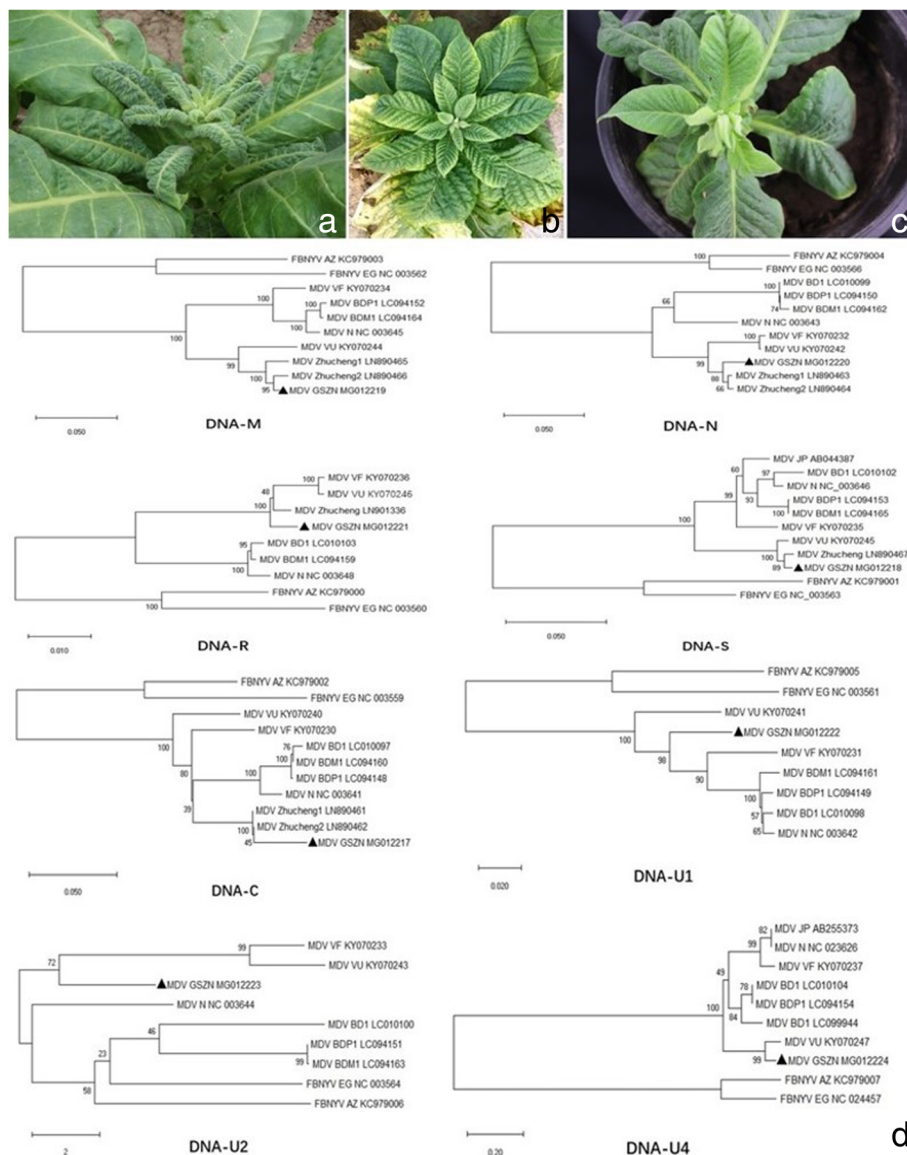


Yang et al. [4], presented preliminary research, including analysis of only two genomic DNAs i.e. DNA-R and DNA-S sequences that provoke further study for this potential and emerging constraint to tobacco production as well as host-shifting of MDV in natural field conditions. In current study we have identified, cloned and sequenced all eight genomic cssDNAs sequences of GSZN isolate of MDV, infecting tobacco plants in Zhengning city, Gansu province in northwest of China. Due to the economic importance of the crop and persistence of the disease in two consecutive years, we further investigated the etiological agent to adopt appropriate measures accordingly, that will

surely help farmers to avoid the particular viral attack that might be turn into epidemic if not addressed at this stage.

**Materials and methods**

During the spring, 2016-2017, tobacco plants (*N. tabacum*) were observed showing severe stunting, interveinal yellowing, leaf crinkling and axillary bunch of new leaves, in Zhengning city with disease incidence of 4.1% (Fig. 1a and b). On the bases of the characteristic symptoms of Nanovirus infection observed in the field from eight different locations, total DNA was extracted from 49 samples, using



**Fig. 1** Symptoms observed on tobacco plants during field survey: **a** axillary bunch of new leaves **(b)** stunting and interveinal yellowing **(c)** symptoms expression by aphid transmission of MDV **(d)** Phylogenetic analysis based on the nucleotide sequences of all genomic components of reported MDV isolates and GSZN isolate (reported herein) along with another Nanovirus; FBNNV. The tree was constructed by the neighbor-joining method using the MEGAX program. Bootstrap analysis was applied using 1000 replicates

Easy Pure Plant Genomic DNA Kit (TransGen Biotech, Beijing, China) for downstream experiments to confirm the potential etiological agent of the disease. Owing to the single-stranded circular genome of Nanoviruses, rolling circle amplification (RCA) [12] was carried out using TempliPhi™ Kit (GE Healthcare, Fisher Scientific, USA) according to manufacturer's protocol. Viral DNA and pUC19 (Fisher Scientific, USA) digested with the same restriction enzymes i.e. *Xba*I, *Sma*I, *Sall*, *Pst*I, *Bam*HI, *Eco*RI, *Hind*III, and *Sac*I (Additional file 1: Figure S1) and resulted fragments were recovered by Easy pure gel DNA extraction kit (TransGen Biotech, Beijing, China). The extracted fragments were then dephosphorylated and ligated using T4-DNA ligase (Fisher Scientific, USA) followed by transformation into *Escherichia coli*. Universal primer pair: M13F:5'-TGTA AAACGACGGCCAGT-3', M13R:5'-CAGGAAACAGCT ATGACC-3 was used for colony identification and the fragments of interest were sequenced from all randomly selected clones at Sangon Biotech Co. Ltd. (Shanghai, China), through Sanger sequencing. However, no alpha-satellite DNA was amplified using three sets of specific primers i.e. C2F:5'-ACAGATGTAGAGAGAGAAA-CAT-3', C2R: 5'-AAGAAGGTTTCATTTAATTGTGT-3', C3F: 5'-TAATGACCGCGTTCAGTACG-3' C3R:5'- TC CCAAGAGTAGCGTCTGAG-3' C10F:5'- CATGAGA GAGTGAATCACGA-3', and C10R: 5'-TTAATTACGC AGTAATTGAG-3' for DNAC2, DNAC3, and DNA C10, respectively. All the obtained sequences were analyzed and submitted to GenBank under the accession

numbers: MG012217- MG012224. Phylogenetic analysis was performed using MegaX software using neighbour joining method [13], whereas recombination events were predicted using RDP4.95 software [14]. Insect transmission experiment was performed using *Myzus persicae* in greenhouse followed by PCR assay using primer pair: MDVCP-F: 5'-CTGGTGC GGGGCTT AGTATTACCCCGCA-3', and MDVCP-R:5'- CGGGA TCAAATGACGTCACAAGGTCATAT-3' (amplicon size: 1003 bp) to confirm the MDV infection, 30 days post-inoculation.

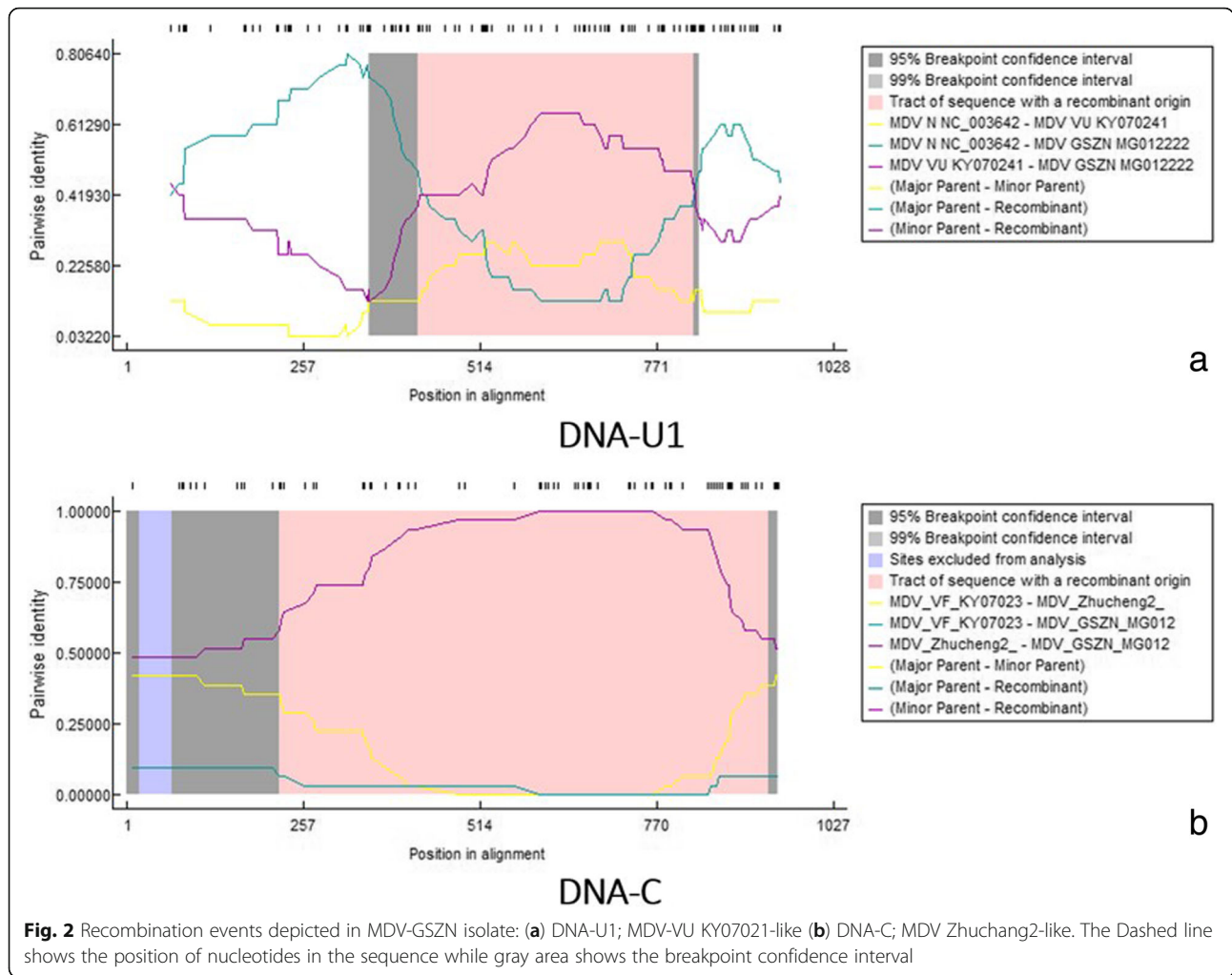
**Results and discussion**

Two Samples showed presence of Nanovirus infection through RCA. All eight genomic components were retrieved from selected isolate GSZN by direct sequencing of RCA product. BLAST showed similarity with the corresponding sequences of MDV. Whereas, using DNAMAN software, high level of percent nucleotide (nt) (99.20–79.33%) as well as amino acid (aa) identity (100–82.67%) except DNA-U2 (73.60%) was estimated between GSZN isolate and most of the corresponding sequences available in GenBank, while an overall low percent identity was depicted in DNA-U4 (Table 1). The trend of low percent similarity in the above mentioned genomic components, makes the isolate GSZN more vulnerable of possible recombination. Moreover, the phylogenetic analysis performed using MegaX software, depicted that isolate GSZN is placed among the tobacco infecting isolates i.e. Zhucheng 1 and 2 in all

**Table 1** Percent identity of genomic DANs of isolate GSZN and corresponding sequences in GenBank

	Zhucheng1	Zhucheng2	VU <sup>b</sup>	VF <sup>b</sup>	BDP1 <sup>b</sup>	BDM1 <sup>b</sup>	BD1 <sup>b</sup>	N <sup>b</sup>	JP <sup>b</sup>
<b>DNA-U1</b>	–	–	<b>88.89%</b> <sup>a</sup>	<b>87.90%</b>	<b>92.82%</b>	<b>92.72%</b>	<b>92.92%</b>	<b>88.64%</b>	–
	–	–	90.67% <sup>a</sup>	–	82.67%	84.00%	84.87%	85.33%	–
<b>DNA-U2</b>	–	–	<b>93.89%</b>	<b>82.95%</b>	<b>83.92%</b>	<b>84.12%</b>	<b>81.68%</b>	<b>86.17%</b>	–
	–	–	92.00%	73.60%	90.40%	90.40%	88.00%	89.60%	–
<b>DNA-U4</b>	–	–	<b>91.12%</b>	<b>80.10%</b>	<b>75.70%</b>	–	<b>73.75%</b>	<b>80.18%</b>	<b>80.18%</b>
	–	–	96.30%	83.33%	90.74%	–	89.81%	75.61%	75.61%
<b>DNA-C</b>	<b>96.38%</b>	<b>96.38%</b>	<b>87.99%</b>	<b>88.90%</b>	<b>84.76%</b>	<b>84.76%</b>	<b>84.27%</b>	<b>89.44%</b>	–
	100%	100%	90.53%	92.31%	92.31%	92.31%	91.12%	92.31%	–
<b>DNA-S</b>	<b>98.11%</b>	–	<b>98.50%</b>	<b>91.48%</b>	<b>88.02%</b>	<b>88.10%</b>	<b>87.52%</b>	<b>92.28%</b>	<b>91.98%</b>
	98.84%	–	98.84%	94.77%	93.60%	93.68%	93.02%	94.19%	93.60%
<b>DNA-R</b>	<b>99.20%</b>	–	<b>98.70%</b>	<b>98.70%</b>	–	<b>91.88%</b>	<b>91.78%</b>	<b>95.31%</b>	–
	98.60%	–	98.95%	98.95%	–	96.50%	96.15%	95.80%	–
<b>DNA-N</b>	<b>97.60%</b>	<b>97.81%</b>	<b>92.62%</b>	<b>92.52%</b>	<b>82.85%</b>	<b>82.45%</b>	<b>82.85%</b>	<b>88.25%</b>	–
	97.39%	97.39%	95.42%	94.77%	96.08%	94.77%	96.08%	95.42%	–
<b>DNA-M</b>	<b>96.66%</b>	<b>97.98%</b>	<b>92.02%</b>	<b>84.26%</b>	<b>79.33%</b>	<b>79.53%</b>	<b>79.92%</b>	<b>83.15%</b>	–
	97.32%	100%	99.11%	94.64%	92.86%	92.86%	93.75%	93.75%	–

<sup>a</sup> = amino acid identity; italic, nucleotide identity; bold; <sup>b</sup> = name of the isolate of MDV



**Fig. 2** Recombination events depicted in MDV-GSZN isolate: (a) DNA-U1; MDV-VU KY07021-like (b) DNA-C; MDV Zhuchang2-like. The Dashed line shows the position of nucleotides in the sequence while gray area shows the breakpoint confidence interval

the available sequences of these two isolates that strongly justify the grouping of MDV on the basis of host plant infectivity (Fig. 1d) Furthermore, leaf crinkling, stunting and axillary bunch of new leaves were observed after 30 days of insect transmission experiment (Fig. 1c) followed by confirmation of MDV in 12 out of 14 tested plants through PCR assay depositing 1003 bp product on 1% agarose gel electrophoresis (Additional file 2: Figure S2). Same results were also confirmed by loop-mediated isothermal amplification (LAMP) [15].

Moreover, both Faba bean necrotic yellows virus (FBNV) isolates grouped together in same clade expect in case of DNA-N where Azerbaijani isolate is placed with Bangladeshi isolates of MDV and Egyptian isolate with two Chinese and one Japanese isolates of MDV that might be the result of possible reassortment reported to be occur in DNA-N of MDV [8]. Recombination was predicted using RDP4.95 [14] to further validate the evolutionary changes. The analysis detected

potential recombination events in DNA-C and DNA-U1 segment of MDV-GSZN isolate (Fig. 2) that supported by six out of the seven algorithms: RDP, GENECONV, RecScan, SisScan, MaxChi, Chimera and 3Seq [16–20]. In case of DNA-C, the potential minor parent in Zhucheng 1 (LN890461) with 80% similarity, while the potential major parent is unknown. Whereas, FBNYV (KC979005) is minor parent and BDM1 (LC094161) detected as potential major parent for DNA-U1. The putative multiple parental origin may contribute towards the unusual infectivity behavior of GSZN isolate.

The results of recombination analysis along with phylogenetic study, support the phenomenon of host-shifting and formation of two different groups i.e. legume-infecting and tobacco-infecting in MDV that confirms the ability of this virus to infect plant species out of Fabaceae family. The possible recombination in DNA-C segment of the genome, might play a key role in host specificity i.e. narrow host range of MDV and successful



infection in tobacco plant in particular, by initiating viral DNA replication through interacting key regulatory pathways of the host plant cell cycle regulator protein families i.e. pRB and SKP1 [21]. To our knowledge, this is the first full genome analysis of MDV infecting *N. tabacum* in natural field conditions. However, still a comprehensive study is required to further clarify the host shifting dilemma of this virus.

## Additional files

**Additional file 1: Figure S1.** Agarose gel (1.5%) showing the results of restriction enzymes used for digesting RCA products. (left to right) Lane M: 5 kb DNA marker, 1–4: no band, 5: XbaI, 6: 5 kb DNA marker, 7: SmaI, 8: Sall, 9: no band, 10–11: HindIII, 12:PstI, 13–14: SacI (no band), 15:BamHI, 16: EcoRI. (PDF 244 kb)

**Additional file 2: Figure S2.** Agarose gel (1%) showing the results of the PCR detection assay of MDV-CP in selected samples of insect transmission experiment. Lane M: DL2,000 bp DNA Marker (Takara); lane C represents the control. (PDF 199 kb)

## Abbreviations

aa: Amino acid; Clink: Cell-cycle link protein; CP: Capsid protein; cssDNA: Circular single-stranded DNA; GSZN: Name of virus isolate used in current study; LAMP: Loop-mediated isothermal amplification; MDV: Milk vetch dwarf virus; MP: Movement protein; M-Rep: Master replication protein; NSP: Nuclear-shuttle protein; nt: Nucleotide; RCA: Rolling circle amplification

## Acknowledgements

Authors are thankful to Mr. Jun Zhang and Mr. Shengping Wang for their assistance in the cloning of GSZN isolate during this work.

## Funding

This study was subsidized by tobacco industry: Screening of new high quality flue-cured tobacco varieties in Qingyang (201862100020017), Research and application of tobacco and water fertilizer integration in Qingyang (201862100020016), Development and application of rapid detection method for Tobacco virus (SCYC201804) and Research and Application of Green Prevention and Control Technology in Tobacco virus in Jiangxi (201701002).

## Availability of data and materials

Full length sequences of all eight genomic segments of MDV isolate GSZN were submitted in GenBank with the accession numbers: MG012217- MG012224.

## Author's contribution

CZ, XW, XY collected the samples, FW and JY conceived and designed the experiments. AK, HH and YX performed the experiments. AK, HH and JY analyzed the data. AK and HH wrote the paper. All authors read and approved the final manuscript.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>Key Laboratory of Tobacco Pest Monitoring Controlling & Integrated Management, Tobacco Research Institute of Chinese Academy of Agricultural Sciences, Qingdao 266101, China. <sup>2</sup>Graduate School of Chinese Academy of Agricultural Sciences, Beijing 100081, China. <sup>3</sup>Qingyang Tobacco Company, Gansu Tobacco Cooperation, Xifeng 745000, China. <sup>4</sup>Jiangxi Tobacco Science Institute, Nanchang 330025, China. <sup>5</sup>Sichuan Tobacco Science Institute, Chengdu 610000, China.

Received: 18 November 2018 Accepted: 13 February 2019

Published online: 27 March 2019

## References

1. FAOSTAT. The International Food and Agricultural Statistics Database of the Food and Agriculture Organization (FAO) of the United Nations. 2016. <http://www.fao.org/faostat/en/#data/QC>. Accessed 1st August 2018.
2. Akinyemi IA, Wang F, Zhou BG, Qi SS, Wu Q. Ecogenomic survey of plant viruses infecting tobacco by next generation sequencing. *Virology*. 2016;13:181.
3. Sun HJ, Liu W, Yang JG, Wang FG, Shen LL, Li Y, Sun DM, Qian YM. Complete genome sequence of a novel wild tomato mosaic virus isolate infecting *Nicotiana tabacum* in China. *J Phytopathol*. 2016;164:686.
4. Yang JG, Wang SP, Liu W, Li Y, Shen LL, Qian YM, Wang FL, Du ZG. First report of Milk vetch dwarf virus associated with a disease of *Nicotiana tabacum* in China. *Plant Dis*. 2016;100:1255.
5. Inouye T, Inouye N, Mitsuhata K. Yellow dwarf of pea and broad bean caused by Milk-vetch dwarf virus. *Jpn J Phytopathol*. 1968;34:28.
6. Kumari SG, Rodoni B, Vetten HJ, Loh MH, Freeman A, Van Leur J, Bao S, Wang X. Detection and partial characterization of Milk vetch dwarf virus isolates from faba bean (*Vicia faba* L.) in Yunnan Province, China. *J Phytopathol*. 2010;158:35.
7. Sano Y, Wada M, Hashimoto Y, Matsumoto T, Kojima M. Sequences of ten circular ssDNA components associated with the milk vetch dwarf virus genome. *J Gen Virol*. 1998;79:3111.
8. Zhang C, Zheng H, Yan D, Han K, Song X, Liu Y, Zhang D, Chen J, Yan F. Complete genomic characterization of milk vetch dwarf virus isolates from cowpea and broad bean in Anhui province. *China Arch Virol*. 2017;162:2437.
9. Timchenko T, Katul L, Sano Y, De KF, Vetten HJ, Gronenborn B. The master rep concept in nanovirus replication: identification of missing genome components and potential for natural genetic reassortment. *Virology*. 2000;274:189.
10. Vetten HJ, Chu PWG, Dale JL, Harding R, Hu J, Katul L, Kojima M, Randles JW, Sano Y, Thomas JE. Nanoviridae. In: Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA, editors. *Virus taxonomy*. 7th rep ICTV. London: Elsevier; 2004.
11. Shirasawa-Seo N, Sano Y, Nakamura S, Murakami T, Seo S, Ohashi Y, Hashimoto Y, Matsumoto T. Characteristics of the promoters derived from the single-stranded DNA components of Milk vetch dwarf virus in transgenic tobacco. *J Gen Virol*. 2005;86:1851.
12. Daniela Haible, Sigrid Kober, Holger Jeske. Rolling circle amplification revolutionizes diagnosis and genomics of geminiviruses. *Journal of Virological Methods*. 2006;135(1):9–16
13. Kumar S, Stecher G, Li M, Knyaz C, Tamura KMEGAX. Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 2018;35:1547.
14. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol*. 2015;1:vev003.
15. Jun Zhang, Xiaowei Liu, Wei Li, Jing Zhang, Zhixin Xiao, Zhicheng Zhou, Tianbo Liu, Ying Li, Fenglong Wang, Songbai Zhang, Jinguang Yang. Rapid detection of milk vetch dwarf virus by loop-mediated isothermal amplification. *Journal of Virological Methods*. 2018;261:147–152
16. Boni MF, Posada D, Feldman MW. An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics*. 2007;176:1035.
17. Martin DP, Posada D, Crandall KA, Williamson C. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS Res Hum Retroviruses*. 2005;21:98.
18. Padidam M, Sawyer S, Fauquet CM. Possible emergence of new geminiviruses by frequent recombination. *Virology*. 1999;265:218.
19. Posada D, Crandall KA. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proc Natl Acad Sci*. 2001;98:13757.
20. Smith JM. Analyzing the mosaic structure of genes. *J Mol Evol*. 1992;34:126.
21. Aronson MN, Meyer AD, Györgyey J, Katul L, Vetten HJ, Gronenborn B, Timchenko T. Clink, a nanovirus-encoded protein, binds both pRB and SKP1. *Virology*. 2000;74:2967.