Species and genetic variability of sweet potato viruses in China

Yongjiang Wang¹,²†, Yanhong Qin¹,²†, Shuang Wang¹,², Desheng Zhang¹,², Yuting Tian¹,², Fumei Zhao¹,², Yingzhi Wang¹,², Hui Lv¹,², Qi Qiao¹,²* and Zhenchen Zhang¹,²*

Abstract

China is the world’s largest producer of sweet potato (Ipomoea batatas (L.) Lam.). Considering that there are numerous sweet potato-producing regions in China and sweet potato is a vegetatively propagated crop, the genetic diversity of sweet potato viruses could be high in the country. However, studies on species and genetic variabilities of sweet potato viruses in China are limited, making it difficult to prevent and control viral diseases in this crop. During 2014–2019, sweet potato samples with viral disease-like symptoms were randomly collected from sweet potato fields in 25 provinces in China. Twenty-one virus species, including 12 DNA and 9 RNA viruses, were identified in the samples using next-generation sequencing, polymerase chain reaction and rolling-circle amplification methods. One novel sweepovirus species, Sweet potato leaf curl Hubei virus (SPLCHbV), was identified. Two species, Sweet potato collusive virus and Tobacco mosaic virus, were identified for the first time in sweet potato in China. Full-length or nearly full-length genomic sequences of 111 isolates belonging to 18 viral species were obtained. Genome sequence comparisons of potyvirus isolates obtained in this study indicate that the genome of sweet potato virus 2 is highly conserved, whereas the other four potyviruses, sweet potato feathery mottle virus, sweet potato virus G, sweet potato latent virus and sweet potato virus C, exhibited a high genetic variability. The similarities among the 40 sweepovirus genomic sequences obtained from eight sweepovirus species are 67.0–99.8%. The eight sweepoviruses include 14 strains, of which 4 novel strains were identified from SPLCHbV and 1 from sweet potato leaf curl Guangxi virus. Five sweet potato chlorotic stunt virus (SPCSV) isolates obtained belong to the WA strain, and the genome sequences of SPCSV are highly conserved. Together, this study for the first time comprehensively reports the variability of sweet potato viruses in China.

Keywords: Sweet potato virus, Species, Genetic diversity, China

Background

Sweet potato, an important food crop worldwide, ranks third in production and fifth in calorific contribution to human diet among all crops globally (de Albuquerque et al. 2019). China is the world’s largest producer of sweet potato, with a plantation area of 2.37 × 10⁶ ha and yield of 5.20 × 10⁷ tons in 2019 (FAOSTAT 2019). Sweet potato, a vegetatively propagated crop, can be infected by viruses during propagation, resulting in yield and quality reduction (Clark et al. 2012). Viral diseases severely affect sweet potato production, particularly sweet potato virus disease (SPVD) caused by co-infection with sweet potato feathery mottle virus (SPFMV) and sweet potato chlorotic stunt virus (SPCSV). Sweet potatoes with SPVD exhibit severe symptoms including stunting, leaf distortion, vein clearing and chlorosis, leading to an 80–90% decrease in production (Clark et al. 2012).
At least 30 viruses, including DNA and RNA viruses, have been found to infect sweet potatoes worldwide (Clark et al. 2012). Among them, at least 20 have been reported in China. The RNA viruses reported in China include SPFMV, sweet potato virus C (SPVC), sweet potato latent virus (SPLV), sweet potato virus G (SPVG) and sweet potato virus 2 (SPV2) of the genus Potyvirus (Colinet and Kummert 1993; Colinet et al. 1994, 1997, 1998; Ateka et al. 2007); SPCSV of the genus Cronivirus (Qiao et al. 2011); and sweet potato chlorotic fleck virus (SPCFV) of the genus Carlavirinae (Arutia et al. 2009). Conversely, the DNA viruses reported in China include sweet potato leaf curl virus (SPLCV), sweet potato leaf curl China virus (SPLCCNV), sweet potato leaf curl Georgia virus (SPLCGoV), sweet potato leaf curl Canary virus (SPLCCV), sweet potato leaf curl Henan virus (SPLCHnV), sweet potato leaf curl Guangxi virus (SPLCGV), sweet potato leaf curl Sichuan virus 1 (SPLCSiV1), sweet potato leaf curl Sichuan virus 2 (SPLCSiV2) and sweet potato leaf curl Shandong virus (SPLCsdV), which all belong to the genus Begomovirus (Luan et al. 2006, 2007; Liu et al. 2013, 2014, 2017; Qin et al. 2013a). Begomoviruses include the following four groups: Old and New World viruses, “sweepoviruses” and “legumoviruses” (Zerbini et al. 2017). Sweepoviruses are phylogenetically distinct from Old and New World begomoviruses, and they infect plants of the Convolvulaceae family, including sweet potato (Hassan et al. 2016; Zerbini et al. 2017). Other DNA viruses infecting sweet potato in China include sweet potato badnavirus A (SPBV-A) and sweet potato badnavirus B (SPBV-B) of the genus Badnavirus (Li et al. 2018; Qin et al. 2016), and sweet potato symptomless virus 1 (SPSMV-1) of the genus Mastrevirus (Wang et al. 2015).

In China, sweet potato is produced in the northern, Yangtze River and southern regions, spanning tropical, subtropical, warm temperate and temperate zones and covering mountainous areas, hills, plains and coastal areas, and presents complex ecotypes. Accordingly, sweet potato viruses could be highly diverse in China. Practically, sweet potato viral disease control and prevention is a big challenge owing to a lack of comprehensive research on the casual viruses. Advances in molecular detection techniques of plant viruses, particularly next-generation sequencing (NGS), have provided critical tools for studying viral diversity (Kreuze et al. 2009). More than 100 novel DNA and RNA plant virus species have been discovered via NGS in recent years (Gu et al. 2014; Hadidi et al. 2016). The aim of the present study was to investigate the sweet potato virus species and their genetic variability using molecular detection techniques. Here, 21 virus species were identified in sweet potato samples in China, and genome variability of the major viruses was analyzed. The results provide a basis for future research on the control and management of sweet potato viral diseases in China.

Results
Sweet potato virus species and their distribution in China
In this study, 21 virus species were identified in sweet potato samples using the NGS, PCR, RT-PCR and RCA methods (Table 1 and Additional file 1: Table S1). There are 12 DNA and 9 RNA viruses, belonging to nine genera and including one novel sweepovirus species, Sweet potato leaf curl Hubei virus (SPLCHbV). SPLCHbV was approved and ratified recently by the International Committee on Taxonomy of Viruses (ICTV), and has been listed in the ICTV website. In addition, to the best of our knowledge, here we identified tobacco mosaic virus (TMV) and sweet potato collusive virus (SPCV) in sweet potatoes for the first time in China. In our investigation, SPFMV was the most common virus infecting sweet potato and was detected in 23 provinces, followed by SPVG, SPLCV and SPCSV, which were detected in 22, 18 and 18 provinces, respectively (Table 1 and Additional file 1: Table S1).

Genetic diversity of potyviruses
Full-length or nearly full-length genomic sequences of 14 isolates of SPFMV were obtained from 9 provinces, 18 isolates of SPVG were obtained from 8 provinces, 10 isolates of SPLV were obtained from 4 provinces, 3 isolates of SPV2 were obtained from 1 province, and 11 isolates of SPVC were obtained from 5 provinces (Additional file 1: Table S2).

The phylogenetic analysis revealed that the 14 SPFMV isolates obtained in this study were classified into three major phytopgroups A, B and C based on their polyprotein nucleotide (nt) sequences (Fig. 1 and Additional file 2: Figure S1). Pairwise comparisons of the polyprotein gene sequences of these 14 isolates revealed their nt and deduced amino acid (aa) sequence identities in the range of 86.0–100% and 92.2–99.9%, respectively. Nine out of the 14 isolates, namely, O-gui-17-5, O-ji-17-49, O-lu-17-8, O-qian-17-1, O-shaan-17-5, O-su-17-10, O-wan-17-1, O-wan-17-3 and O-yue-17-9, and reference isolates of the O-II minor phytopgroup (strain O), including O-Ch1 (KY296450), O-10-O (AB439206) and O-Ordinary (AB465608), were grouped into one clade, indicating that these nine isolates belong to the O-II minor phytopgroup. The nt and aa sequence identities among the O-II minor phytopgroup isolates are 93.6–99.6% and 96.0–99.6%, respectively. The two isolates, RC-31-2 and RC-33-1, were clustered together with the B phytopgroup (strain RC), RC-Ch1 and RC-S, suggesting that they belong to the B phytopgroup. The
**Table 1** Sweet potato virus species and their distribution in China

| Species                           | Genus          | GenBank no. of representative sequences                                                                 | Distribution (provinces)                              | Methods            |
|-----------------------------------|----------------|--------------------------------------------------------------------------------------------------------|------------------------------------------------------|--------------------|
| Sweet potato feathery mottle virus| Potyvirus      | MK778785, MK778794, MK778796                                                                         | Anhui, Beijing, Chongqing, Fujian, Gansu, Guangdong, Guanxi, Guizhou, Hainan, Henan, Hebei, Hubei, Hunan, Jiangsu, Jiangxi, Liaoning, Ningxia, Shandong, Shaanxi, Shanxi, Sichuan, Yunnan and Zhejiang | NGS, PCR           |
| Sweet potato virus C              | Potyvirus      | MK778817                                                                                               | Chongqing, Fujian, Guangdong, Guanxi, Guizhou, Hainan, Hebei, Hubei, Jiangsu, Jiangxi, Shandong and Sichuan | NGS, PCR           |
| Sweet potato virus G              | Potyvirus      | MK778826, MK778834, MK778835, MK778837, MK778838                                                    | Anhui, Beijing, Chongqing, Fujian, Gansu, Guangdong, Guanxi, Guizhou, Hainan, Hebei, Henan, Hunan, Hubei, Jiangsu, Jiangxi, Liaoning, Ningxia, Shandong, Sichuan, Yunnan and Zhejiang | NGS, PCR           |
| Sweet potato latent virus         | Potyvirus      | MK778808                                                                                               | Anhui, Chongqing, Guangdong, Hebei, Hubei, Hunan, Jiangsu, Shaanxi, Shanxi, Shandong and Sichuan         | NGS, PCR           |
| Sweet potato virus 2              | Potyvirus      | MK778812                                                                                               | Anhui, Chongqing, Fujian, Hebei, Henan, Hubei, Hunan, Jiangsu, Jiangxi, Liaoing, Shaanxi, Shanxi, Shandong and Sichuan | NGS, PCR           |
| Sweet potato leaf curl Hubei virus*| Begomovirus    | MH577011, MK951974, MK931305, MK931304                                                               | Hebei, Hubei, Hunan, Jiangsu, Shandong and Yunnan    | NGS, PCR, RCA      |
| Sweet potato leaf curl Canary virus| Begomovirus    | MK931306                                                                                               | Guangdong, Hebei and Henan                           | PCR                |
| Sweet potato leaf curl Georgia virus| Begomovirus    | MK931310                                                                                               | Fujian, Hebei, Henan, Shandong, Sichuan, Jiangxi and Zhejiang | PCR                |
| Sweet potato leaf curl Guangxi virus| Begomovirus    | MK931313, MK951966                                                                                     | Guangxi, Guangdong, Hebei, Henan, Hunan, Jiangsu, Shaanxi, Shanxi, Shandong and Sichuan             | PCR                |
| Sweet potato leaf curl Sichuan virus 1 | Begomovirus | MN186254                                                                                               | Guangdong, Guangxi, Henan, Hunan, Jiangsu and Sichuan | PCR                |
| Sweet potato leaf curl Sichuan virus 2 | Begomovirus | MK931315                                                                                               | Guangdong, Hebei, Henan, Jiangsu and Sichuan         | PCR                |
| Sweet potato leaf curl virus      | Begomovirus    | MK951979, MK931320, MK931321                                                                          | Anhui, Beijing, Chongqing, Fujian, Guangdong, Guanxi, Hainan, Hebei, Hubei, Hunan, Jiangsu, Jiangxi, Jilin, Shaanxi, Shandong, Sichuan and Yunnan | PCR                |
| Sweet potato leaf curl China virus| Begomovirus    | MK951977                                                                                               | Guangdong, Guizhou, Henan, Jiangsu, and Shandong     | PCR                |
| Cucumber mosaic virus             | Cucumovirus    | MK778778                                                                                               | Anhui, Chongqing, Fujian, Guangdong, Guanxi, Hainan, Henan, Hubei, Jiangsu, Shaanxi, Shanxi, Sichuan and Zhejiang | NGS                |
| Sweet potato collusive virus†     | Cavemovirus    | MK802082                                                                                               | Chongqing and Hainan                                 | NGS, PCR           |
| Tobacco mosaic virus†             | Tobamovirus    | MN186255                                                                                               | Henan                                               | NGS                |
| Sweet potato badnavirus A         | Badnavirus     | KJ463420                                                                                               | Anhui, Chongqing, Guangxi, Hebei, Hubei, Hunan, Jiangsu, Shaanxi and Shandong                             | NGS, PCR           |
| Sweet potato badnavirus B         | Badnavirus     | KP407560                                                                                               | Henan, Hubei, Hunan, Jiangsu, Shaanxi, Shanxi, Shandong | NGS, PCR           |
| Sweet potato symptomless virus 1  | Mastrevirus    | MG603668                                                                                               | Chongqing, Guangdong, Guanxi, Hebei, Henan, Jiangsu, Shandong and Shaanxi                           | NGS, PCR           |
nt sequence and deduced aa sequence identities among the B phytogroup isolates are 98.3–100% and 98.9–99.9%, respectively. The rest three isolates, 19-2, 32-1 and 33-1, were grouped together and were phylogenetically distant from the A and B phytogroups, suggesting that these three isolates belong to the C phytogroup.

The phylogenetic analysis according to polyprotein nt sequences revealed that the 18 SPVG isolates obtained were divided into five distinct phytogroups (Fig. 1 and Additional file 2: Figure S2). Pairwise comparisons of the polyprotein gene sequences of these 18 isolates revealed nt and deduced aa sequence identities in the range of 75.8–100% and 85.1–87.6%, respectively. The remaining four isolates (MH577011, MK951969, MK951970, MK951971, MK951972, MK951973, MK951974, MK951975, MK931304, MK931305 and MK951976) exhibited the

| Species                              | Genus     | GenBank no. of representative sequences | Distribution (provinces)                                    | Methods          |
|--------------------------------------|-----------|----------------------------------------|------------------------------------------------------------|------------------|
| Sweet potato chlorotic fleck virus   | Carlavirus| MK778783                                | Anhui, Chongqing, Fujian, Guangdong, Guangxi, Guizhou, Hainan, Henan, Hunan, Jiangsu, Shaanxi, Shanxi, Shandong, Yunnan and Sichuan | NGS, PCR         |
| Sweet potato chlorotic stunt virus   | Crinivirus| MK802071                                | Anhui, Chongqing, Fujian, Guangdong, Guangxi, Guizhou, Hainan, Henan, Hubei, Hunan, Jiangsu, Ningxia, Shaanxi, Shanxi, Sichuan and Zhejiang | NGS, PCR         |

*New species; †First report of a sweet potato virus in China

**Genetic diversity of sweepoviruses**

Full-length genomic sequences of 40 sweepovirus isolates from 12 provinces (Additional file 1: Table S2) were obtained, with the nt sequence identity among these isolates ranging from 68.0 to 99.8%. According to the criteria set by the ICTV for begomovirus species demarcation (91% nt sequence identity) and the SDT v1.2 analysis result (Brown et al. 2015), 29 of these 40 isolates were classified into seven species: SPLCCV, SPLCGoV, SPLCGV, SPLCSiv1, SPLCSiv2, SPLCV and SPLCCNV (Additional file 1: Tables S2, S4). A comparison of the complete genomic sequences of the remaining 11 isolates (MH577011, MK951969, MK951970, MK951971, MK951972, MK951973, MK951974, MK951975, MK931304, MK931305 and MK951976) exhibited the
highest nt sequence identity, 73.4%, with known begomoviruses, indicating that the nt sequence identity was lower than 91% (Additional file 1: Table S4). Furthermore, the full-length genomic sequence identity among the 11 isolates ranged from 89.6% to 99.8%. Therefore, these 11 isolates belong to a same and novel species, with Sweet potato leaf curl Hubei virus (SPLCHbV) as the proposed name. According to the strain demarcation criteria for begomoviruses (ranges 91–94% nt sequence identity), as proposed by the ICTV (Haible et al. 2006), these 11 isolates of SPLCHbV are composed of four strains with SPLCHbV-Hb, SPLCHbV-Js, SPLCHbV-Hn and SPLCHbV-Sd as the proposed names (Fig. 2, Additional file 1: Table S5 and Additional file 2: Figure S3).

Complete genome sequence comparisons using SDT v1.2 revealed that the genomic sequence identity of the 6 SPLCGV isolates from 3 provinces ranged from 90.4 to 99.4%. Further phylogenetic analysis revealed that four (MK951965, MK951966, MK951967 and MK951968) of these isolates presented an nt sequence identity of 98.2–99.4% (Additional file 1: Table S6 and Additional file 2: Figure S4) and are phylogenetically distant from other SPLCGV isolates analyzed (Fig. 2). They share the highest identity of 93.1% with other known SPLCGV isolates, which is lower than the criteria required for begomoviruses strain classification (94% limit); therefore, these four isolates constitute a novel strain, designated as SPLCGV-Gd. In addition, the complete genomic sequences of 7 SPLCV isolates were obtained from 6 provinces, with their nt sequence identity ranging from 82.4 to 98.1%. The SDT and phylogenetic analyses revealed that these 7 isolates belong to strains SPLCV-US, SPLCV-KR and SPLCV-Fu (Fig. 2 and Additional file 1: Table S7). The genome sequences of 5 other sweepoviruses, namely, SPLCCV, SPLCGoV, SPLCSIv1, SPLCSIv2 and SPLCCNV, are conservative, with 94.8–99.7% nt sequence identity.

The 40 sweepovirus isolates showed typical genomic structural features of sweepoviruses: genome size ranging from 2770 to 2821 bp; six open read frames (ORFs); viral-sense strand with two ORFs (AV1 and AV2), and complementary-sense strand with four ORFs (AC1, AC2, AC3 and AC4). The genomes of these 40 sweepovirus isolates all encompass a highly conserved stem-loop structure containing the TAATATTAC/T sequence.
Fig. 2  Phylogenetic tree of 79 genome sequences of sweepoviruses. The tree was generated using MEGA X (the Maximum Likelihood method, General Time Reversible model and 1000 bootstrap replicates). New species and new strains are marked with dot and triangle, respectively. BCTV-PeYD (EU921828) was used as an outgroup.
**Fig. 3** Phylogenetic tree of RNA1 segments of SPCSV isolates. The tree was generated using MEGA X (the Maximum Likelihood method, General Time Reversible model and 1000 bootstrap replicates). Liyv RNA1 segment (NC_003617) was used as an outgroup.

**Fig. 4** Phylogenetic tree of RNA2 segments of SPCSV isolates. The tree was generated using MEGA X (the Maximum Likelihood method, General Time Reversible model and 1000 bootstrap replicates). Liyv RNA2 segment (NC_003618) was used as an outgroup.
Molecular variation of SPCSV

Nearly full-length genomic sequences of 5 SPCSV isolates were obtained from 5 provinces (Additional file 1: Table S2). The RNA1 and RNA2 segments of the 5 isolates have 4 and 9 ORFs, respectively, similar to those of the Jiangsu isolate of the SPCSV WA strain (Qin et al. 2013b). The nt sequence identities of RNA1 and RNA2 among the five isolates are 98.9–100% and 98.8–99.9%, respectively. The phylogenetic analysis revealed that the five isolates belong to the WA strain (Figs. 3, 4). Therefore, the WA strain could be the dominant strain in sweet potatoes in China, and this is consistent with the findings of our previous study (Qin et al. 2013b).

Molecular variations in other viruses

Full-length or nearly full-length genome sequences of 2 cucumber mosaic virus (CMV) isolates, 2 SPCFV isolates, 1 TMV isolate and 5 SPSMV-1 isolates were obtained (Additional file 1: Table S2).

For CMV, segments RNA1, RNA2 and RNA3 in the lu-17-14 isolate presented the highest nt sequence identities of 98.0%, 98.7% and 98.4% with RNA1 (KT350980) and RNA2 (KT350981) of a maize isolate from China, and RNA3 (JX993913) of a Chinese thorowax root isolate from China, respectively. RNA1, RNA2 and RNA3 of the yang-SS3-2 isolate presented the highest nt sequence identity with RNA1 (AB179764) of a *Raphanus sativus* isolate from Japan, RNA2 (KX883764) of a true fly isolate from China, and RNA3 (EF213025) of a *Brassica chinensis* isolate from China, at 96.8%, 97.5% and 97.1%, respectively. Phylogenetic analysis based on the nt sequences of coat protein (CP) genes with known subgroup CMV isolates (Roossinck 2002; Bald-Blumea et al. 2017) showed that lu-17-14 and yang-SS3-2 belong to subgroup IB (Additional file 2: Figure S5).

Sequence analysis revealed that the genome of the two SPCFV isolates (gui-17-2-1 and wan-17-1) has six complete ORFs encoding RNA-dependent RNA polymerase (RdRp), triple gene block protein 1 (TGB1), TGB2, TGB3, CP and nucleic acid binding protein (NaBP), and exhibited the typical genomic structural features of carlaviruses (Deng et al. 2015). The nt sequence identity between these two isolates is 73.8%. The phylogenetic analysis results indicate that gui-17-2-1 and wan-17-1 are distant (Additional file 2: Figure S6), suggesting a high molecular variability within SPCFV genomic sequences.

Nearly full-length genomic sequence of one TMV isolate was obtained (Additional file 1: Table S2). The isolate has the highest nt sequence identity of 99.8% with a *Nicotiana benthamiana* isolate from Spain (MK087763). To the best of our knowledge, this is the first study to report TMV infecting sweet potatoes in China.

The full-length genomic sequences of 5 SPSMV-1 isolates were obtained (Additional file 1: Table S2). Sequence comparisons indicate that the nt sequence identity of these 5 isolates with other isolates of the virus in GenBank ranged from 97.4 to 99.8% (Additional file 1: Table S8 and Additional file 2: Figures S7), suggesting that the SPSMV-1 genomic sequences are highly conserved.

Discussion

In the present study, viruses infecting sweet potato in China were identified using NGS, PCR and RCA approaches. Twenty-one virus species were identified, including 1 novel virus and 5 novel viral strains, and 2 viruses were found to infect sweet potato in China for the first time. The results highlight the high diversity of the sweet potato viromes in China. The full-length or nearly full-length genomic sequences of 111 isolates were obtained, belonging to 18 sweet potato virus species, and the molecular diversity of major sweet potato viruses was analyzed. The present study for the first time comprehensively reports the variability of sweet potato viruses in China.

Genomic sequence comparisons indicate that three SPFMV isolates (19-2, 32-1 and 33-1) obtained in this study share nt and aa sequence identities of 86.0–90.7% and 92.3–95.8% with the Δ, RC and EA strain isolates, respectively. Previous studies have reported that the polyprotein nt and aa sequence identities between different SPFMV strains are lower than 94.0% and 96.0%, respectively (Adams et al. 2005; Yamasaki et al. 2010; Wylie et al. 2017). Therefore, these three isolates could be a novel SPFMV strain, and further research on biological characteristics of these three isolates is warranted.

The detection of eight sweepoviruses in sweet potatoes in China is the highest number of virus species detected in a study to date in the country. The eight sweepoviruses include one novel species, which was identified in the present study, and four novel species recently identified in China (Luan et al. 2006; Liu et al. 2013, 2014, 2017). The findings indicate that there is a great diversity in sweepoviruses found on sweet potatoes in China. Sweepoviruses are transmitted by the whitefly (*Bemisia tabaci*), an important agricultural pest transmitting numerous economically important plant viruses, which has recently exhibited increasing trends in sweet potato cultivation regions in China (Gilbertson et al. 2015). The diversity of sweepovirus could be associated with the activities of whiteflies on sweet potatoes in China. Sweepoviruses are another whitefly-transmitted virus. In the present study, SPCSV was detected from 18 provinces. SPCSV was reported in Guangdong Province of China for the first time in 2011 (Qiao et al. 2011). The rapid spread of SPCSV could also be associated with the expansion of...
whitefly populations in China. Therefore, controlling whitefly populations in sweet potato fields is a potential strategy to manage SPCSV incidence and proliferation. SPCSV can also synergistically infect sweet potatoes with several other viruses (Karyeija et al. 2000; Mukasa et al. 2006; Cuellar et al. 2011, 2015). Co-infection by SPCSV and SPFMV causes SPVD, a destructive disease in sweet potatoes (Schaefers and Terry 1976; Clark et al. 2012). According to the NGS results, SPVD detection rate was as high as 48.4%, suggesting that SPVD is a common disease (Additional file 1: Table S9) and poses a major threat to sweet potato production in China. Therefore, it is vital to bolster SPVD monitoring and control activities.

In the present study, sweet potato samples were collected randomly from sweet potato fields in 25 provinces, where the areas under sweet potato cultivation vary considerably. The amount of sweet potato plants sampled was not proportional to sweet potato cultivation area in the provinces. Although as many as 21 virus species were identified in the present study, their distribution and incidence remain unclear in China. Consequently, more samples should be collected and more reliable sampling efforts should be implemented in future studies. Nonetheless, our results provide an important basis for research on the control and management of sweet potato viral diseases in China.

Conclusions

We identified 21 virus species, including 12 DNA and 9 RNA viruses, in the sweet potato samples from China. One novel sweepovirus, SPLCHbV, was identified. Two species, SPCV and TMV, were identified for the first time in sweet potato in China. Genome sequence comparisons of five potyvirus indicate that the genome of SPV2 is highly conserved, and the other four potyviruses, SPFMV, SPVG, SPLV and SPVC, exhibited a high genetic variability. The similarity among the sweepovirus genomic sequences obtained from eight sweepovirus species varied greatly. The genome sequences of SPCSV are highly conserved. The present study for the first time comprehensively reports the variability of sweet potato viruses in China. The results provide a basis for future research on the control and management of sweet potato viral diseases in China.

Methods

Collection of samples

Approximately 2000 samples with viral disease-like symptoms such as leaf shriveling, curling, malformation, vein clearing, chlorosis and mosaic were randomly collected from sweet potato fields in 25 provinces in China during 2014–2019. Most of the samples were leaves, and only a few were storage root and seed samples. The collected leaf and storage root samples were stored at −80 °C until further analyses, and the seed samples were stored under dry conditions at room temperature until use.

Primers

Primers for virus detection and viral genome sequence amplification were designed based on the nucleotide sequences of sweet potato viruses obtained from GenBank or NGS-assembled sequences (Additional file 1: Tables S10, S11). All primers were synthesized by Sangon Bioengineering and Technology Service Co., Ltd. (Shanghai, China).

Nucleic acid extraction

DNA and RNA were extracted from samples ground in liquid nitrogen, using the DNA extraction kit (Omega Bio-tek, Norcross, USA) and the Total Plant RNA Extraction Miniprep System (Sangon, Shanghai, China) according to the manufacturer’s instructions. The extracted nucleic acid samples were stored at −80 °C until use.

Next-generation sequencing and data processing

Sixty-four samples were used for NGS (Additional file 1: Table S9). The total RNA was extracted from sweet potato samples using the TRIzol kit (Invitrogen, Carlsbad, CA, USA), and its concentration and purity was measured using the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA) and by agarose gel electrophoresis. Subsequently, the small RNAs were enriched using the sRNAeasy kit (Qiagen, Dusseldorf, Germany) to construct small RNA libraries for small RNA sequencing (sRNA-seq); the sRNA libraries were built using the TruSeq Small RNA Sample Prep Kit (Illumina, San Diego, USA). An Illumina HiSeq2500 platform (Illumina) set with 50-bp read lengths was used for sequencing (BGI Tech Company, Shenzhen, China). The ribosome total RNA was removed using the RiboZero Magnetic Kit (Epicenter, USA) to construct ribo-depleted RNA libraries for RNA sequencing (RNA-Seq). The libraries were built using the TruSeq RNA Sample Prep Kit (Illumina). An Illumina HiSeq X-ten platform (Illumina) set with 150-bp pair-end reads was used for sequencing (BGI Tech Company). After deep sequencing, adaptor and low-quality sequences were removed from raw reads, using Trimmomatic (Bolger et al. 2014), yielding more than 8-Gb clean reads for each sample library. The remaining RNA-Seq reads were mapped to the genome sequences of sweet potato (downloaded from http://public-genomes-ngs.molgen.mpg.de/SweetPotato/) (Yang et al. 2017) using CLC Genomic Workbench 10.0 (Qiagen). The reads with sequence similarities of >60% to the sweet potato genome
sequences were eliminated to reduce interference of the host background, and the remaining unique reads were de novo assembled using CLC Genomic Workbench 10.0 (Qiagen). The clean reads of 18–26 nt from sRNA-Seq were subjected to de novo assembly using CLC Genomic Workbench 10.0 (Qiagen). The resulting contigs were subjected to BLASTx and BLASTn searches against viral (taxid:10239) and viroidal (taxid:2559587) sequences of local datasets retrieved from the National Center for Biotechnology Information (NCBI) databank (Wu et al. 2015; Zhang et al. 2021). These processes enabled the identification of contigs with viral sequence attributes.

Rolling-circle amplification
One hundred samples with leaf curl symptoms were used in rolling-circle amplification (RCA). DNA extracted from the samples was amplified using the TempliPhi™ 100 Amplification Kit (GE Healthcare, USA) according to the method of Haible et al. (2006). The RCA products (5 µL) were digested in a 30-µL container using the restriction enzyme BamHI (10 U) (TaKaRa, Japan) at 37 °C for 4 h. The digested products were separated on 1% agarose gel. The DNA fragments (2.8 kb) were purified using the Gel Extraction Kit (AXYGEN, USA), ligated into the PUC-118 vector, and then transformed into Escherichia coli strain JM109. Positive clones were identified by PCR and sequenced by TaKaRa Bioengineering Co., Ltd. (Dalian, China).

(RT)-PCR and sequencing
For RNA viruses, the extracted RNA samples were reverse transcribed into cDNA using M-MLV reverse transcriptase (TaKaRa) and reverse primers of the corresponding viral species (Additional file 1: Table S1), followed by PCR amplification with the synthesized cDNA as the template. The PCR mixture (25 µL) consisted of 2 µL of cDNA, 2.5 µL of 10× Ex Taq or LA Taq Buffer (MgCl₂), 2 µL of 2.5 mM dNTP mixture, 2 µL of each forward and reverse primers (Additional file 1: Table S10), and 0.2 µL of Ex Taq DNA polymerase or LA Taq DNA polymerase. PCR amplifications were performed under the following conditions: an initial denaturation for 5 min at 94 °C, followed by 35 cycles at 94 °C for 1 min, 55 °C for 1 min, 72 °C for 1 min/kb, and a final extension at 72 °C for 10 min. The PCR products were separated on 1% agarose gels for 30 min.

Fragments of interest were recovered using the Gel Extraction Kit (Omega Bio-tek) according to the manufacturer's instructions and cloned into E. coli strain JM109. Plasmids containing the amplified viral sequence were sequenced by TaKaRa Bioengineering Co., Ltd.. For DNA viruses, DNA was extracted and used as the template for PCR amplification. The PCR mixture, cycling conditions and sequencing procedures were similar to those for RNA viruses.

Cloning and sequencing of the genome of sweet potato viruses
Full-length genomic sequences of potyviruses were cloned by restriction enzyme digestion, and the primers used are listed in Additional file 1: Table S1. RT-PCR was performed as previously described. Overlapping fragments of 500–1000 bp were amplified and sequenced after cloning. At least three independent clones for each fragment were sequenced. 5’-RACE and 3’-RACE were conducted by TaKaRa Bioengineering Co., Ltd. Full-length genomic sequences of sweepoviruses and SPSMV-1 were cloned using the back-to-back primers (Additional file 1: Table S11) or RCA. The genomic sequences of other viruses, including TMV, CMV, SPFV and SPCSV were obtained by NGS assembly.

Sequence analysis
Sequence analysis and comparisons were performed using DNAMAN v8.1 (Lynnon Biosoft, Quebec), Sequence Demarcation Tool (SDT) v1.2 (Muhire et al. 2014) and BLAST (NCBI). A phylogenetic tree was constructed using the maximum likelihood (ML) method with MEGA X v10.1 (Kumar et al. 2018). The obtained viral nucleotide sequences and the reference sequences are listed in Additional file 1: Tables S1, S2, respectively.

Abbreviations
CMV: Cucumber mosaic virus; CP: Coat protein; GLV: Garlic latent virus; ICTV: International Committee on Taxonomy of Viruses; MSV: Maize streak virus; NaBP: Nucleic acid binding protein; NCBI: National Center for Biotechnology Information; NGS: Next-generation sequencing; ORFs: Open read frames; PSV: Peanut stunt virus; RCA: Rolling-circle amplification; RdRp: RNA-dependent RNA polymerase; SDT: Sequence demarcation tool; SPBV: Sweet potato badnavirus A, SPBV-B: Sweet potato badnavirus B; SPBV-C: Sweet potato colllusive virus; SPCSV: Sweet potato chlorotic stunt virus; SPPMV: Sweet potato mottle virus; SPLCMV: Sweet potato leaf curl China virus; SPLCCV: Sweet potato leaf curl Canary virus; SPLCV: Sweet potato leaf curl Georgia virus; SPLCVG: Sweet potato leaf curl Guangxi virus; SPLCHV: Sweet potato leaf curl Hubei virus; SPLCIV: Sweet potato leaf curl Henan virus; SPLCSSV: Sweet potato leaf curl Shandong virus; SPLCSV1: Sweet potato leaf curl Suzhou virus; SPLCSV2: Sweet potato leaf curl Suzhou virus 2; SPLCVC: Sweet potato leaf curl virus; SPLPV: Sweet potato latent virus; SPMIV: Sweet potato mild mottle virus; SPSMV-1: Sweet potato symptoms virus 1; SPV2: Sweet potato virus 2; SPVC: Sweet potato virus C; SPVD: Sweet potato virus disease; SPVG: Sweet potato virus G; TGB1: Triple gene block protein 1; TMV: Tobacco mosaic virus.
**Supplementary Information**

The online version contains supplementary material available at https://doi.org/10.1186/s42483-021-00097-8.

**Additional file 1**: Table S1. Virus species detected in sweet potato fields in different provinces. Table S2. Virus isolates obtained in the present study. Table S3. Reference sequences used in the present study. Table S4. SDT analysis of sweepoviruses. Table S5. SDT analysis of SPLCV. Table S6. SDT analysis of SPLCGV. Table S7. SDT analysis of SPCFV. Table S8. SDT analysis of SPLCHbV. Table S9. Viruses detected from 64 samples by next-generation sequencing. Table S10. Primers used for virus detection in the present study. Table S11. Primers used for complete genome amplification in the present study.

**Additional file 2**: Figure S1. Phylogenetic tree reconstructed using the polyprotein nucleotide sequences of the SPFMV isolates. The tree was generated using MEGA X (MUSCLE alignment, neighbor-joining and p-distance model, and 1000 bootstrap replicates). Sweet potato mild mottle virus (SPMMV) (Z73124) was used as an outgroup taxon. Figure S2. Phylogenetic trees reconstructed using the polyprotein nucleotide sequences of the SPVG isolates. The tree was generated using MEGA X (MUSCLE alignment, neighbor-joining and p-distance model, and 1000 bootstrap replicates). Sweet potato mild mottle virus (SPMMV) (Z73124) was used as an outgroup taxon. Figure S3. Phylogenetic tree illustrating the relationship among 9 SPLCHbV sequences. The tree was generated using MEGA X (MUSCLE alignment, neighbor-joining and p-distance model, and 1000 bootstrap replicates). Figure S4. Phylogenetic tree illustrating the relationship among 8 SPLCGV sequences. The tree was generated using MEGA X (MUSCLE alignment, neighbor-joining and p-distance model, and 1000 bootstrap replicates). Figure S5. Phylogenetic tree based on the nucleotide sequences of CP of CMV isolates. The tree was generated using MEGA X (MUSCLE alignment, neighbor-joining and p-distance model, and 1000 bootstrap replicates). Peanut stunt virus (PSV) (U15730) was used as an outgroup taxon. Figure S6. Phylogenetic tree constructed using near complete genome sequences of SPCFV isolates. The tree was generated using MEGA X (MUSCLE alignment, neighbor-joining and p-distance model, and 1000 bootstrap replicates). Garlic latent virus (GLV) (NC_003557) was used as an outgroup taxon. Figure S7. Phylogenetic tree constructed using the complete genome sequences of SPSMV-1 isolates. The tree was generated using MEGA X (MUSCLE alignment, neighbor-joining and p-distance model, and 1000 bootstrap replicates). Maize streak virus (MSV) (HQ693373) was used as an outgroup taxon.

**Acknowledgements**

We thank Dr. Cao Mengji from Southwest University, China for conducting partial NGS analysis and critically reading the manuscript.

**Authors’ contributions**

ZCZ conceived the project. YJW, YHQ and QQ performed most of the experiments with assistance from SW, DSZ, YTT, FMZ, YZW and HL. All authors read and approved the final manuscript.

**Funding**

This work was supported by China Agriculture Research System of MOF and MARA (CARS-10-B13), the project from Industry-University-Research Cooperation of Henan Province of China (172107000038) and the Science-Technology Foundation for Outstanding Young Scientists of Henan Academy of Agricultural Sciences (Grant No. 2018YQ19).

**Availability of data and materials**

Not applicable.

**Declarations**

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

1 Institute of Plant Protection, Henan Academy of Agricultural Sciences, Zhengzhou, Henan, China. 2 IPM Key Laboratory in Southern Part of North China for Ministry of Agriculture, Zhengzhou, Henan, China.

**Received**: 18 May 2021 **Accepted**: 11 August 2021

**Published online**: 30 August 2021

**References**

Adams MJ, Antonij F, Fauquet CM. Molecular criteria for genus and species discrimination within the family Potyviridae. Arch Virol. 2005;150:459–79. https://doi.org/10.1007/s00705-004-0440-6.

Aritua V, Barg E, Adipala E, Gibson RW, Lesemann DE, Vereten HJ. Host range, purification, and genetic variability in sweet potato chlorotic fleck virus. Plant Dis. 2009;93(1):87–93. https://doi.org/10.1094/PDIS-93-1-0087.

Ateka EM, Barg E, Njuru RW, Thompson G, Vereten HJ. Biological and molecular variability among geographically diverse isolates of sweet potato virus 2. Arch Virol. 2007;152:479–88. https://doi.org/10.1007/s00705-006-0879-8.

Bald-Blumea N, Bergervoetb JH, Maissa E. Development of a molecular assay for the detection of cucumber mosaic virus and the discrimination of its subgroups I and II. J Virol Methods. 2017;243:35–43. https://doi.org/10.1016/j.jviromet.2017.01.011.

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexibletrimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20. https://doi.org/10.1093/bioinformatics/btu170.

Brown JK, Zerbini FM, Navas-Castillo J, Moriones E, Ramos-Sobrinho R, Silva JCF, et al. Revision of Begomovirus taxonomy based on pairwise sequence comparisons. Arch Virol. 2015;160:1593–619. https://doi.org/10.1007/s00705-015-2398-y.

Clark CA, Davis JA, Abad JA, Cuelar WJ, Fuentes S, Kreuze JF, et al. Sweet potato viruses: 15 years of progress on understanding and managing complex diseases. Plant Dis. 2012;96(2):168–85. https://doi.org/10.1094/PDIS-07-11-0550.

Colinet D, Kummert J. Identification of a sweet potato feathery mottle virus isolate from China (SPFMV-CH) by the polymerase chain reaction with degenerate primers. J Virol Methods. 1993;45(2):149–59. https://doi.org/10.1016/0166-0934(93)90099-D.

Colinet D, Kummert J, Lepoiivre P. The complete nucleotide sequences of the coat protein cistron and the 3' non-coding region of a newly-identified potyvirus infecting sweetpotato, as compared to those of sweetpotato feathery mottle virus. Arch Virol. 1994;139:327–36. https://doi.org/10.1007/BF01310795.

Colinet D, Kummert J, Lepoiivre P. Evidence for the assignment of two strains of SPLV to the genus Potyvirus based on coat protein and 3’ non-coding region sequence data. Virus Res. 1997;49(1):91–100. https://doi.org/10.1016/S0168-1702(97)01456-1.

Colinet D, Nguyen M, Kummert J, Lepoiivre P, Xia FZ. Differentiation among potyviruses infecting sweet potato based on genus- and virus-specific reverse transcription polymerase chain reaction. Plant Dis. 1998;82(2):223–9. https://doi.org/10.1094/PDIS.1998.82.2.223.

Cuelar WJ, De Souza J, Barrantes I, Fuentes S, Kreuze JF. Distinct cavemoviruses interact synergestically with sweet potato chlorotic stunt virus (genus Crinivirus) in cultivated sweet potato. J Gen Virol. 2011;92(5):1233–43. https://doi.org/10.1099/jgv.0.029975-0.

Cuelar WJ, Galvez M, Fuentes S, Tugume J, Kreuze J. Synergistic interactions of begomoviruses with sweet potato chlorotic stunt virus (genus Crinivirus) in sweet potato (Ipomoea batatas L.). Mol Plant Pathol. 2015;16(5):459–71. https://doi.org/10.1111/mpp.12200.
Deng XG, Peng XJ, Zhu F, Chen YJ, Zhu T, Qin SB, et al. A critical domain of sweet potato chlorotic fleck virus nucleotide binding protein (NaBp) for RNA silencing suppression, nuclear localization and viral pathogenesis. Mol Plant Pathol. 2015;16(4):365–75. https://doi.org/10.1111/mpp.12186.

Gilbertson RL, Batuman O, Webster CG, Adkins S. Role of the insect supervector *Bernisia tabaci* and *Frankliniella occidentalis* in the emergence and global spread of plant viruses. Annu Rev Virol. 2015;2:67–93. https://doi.org/10.1146/annurev-virology-031413-085410.

Kreuze JF, Perez A, Untiveros M, Quispe D, Fuentes S, Barker I, et al. Combinatorial genetics analysis across computing platforms. Mol Biol Evol. 2010;27(6):1547–54. https://doi.org/10.1093/molbev/msq017.

Kreuze JF, Perez A, Untiveros M, Quispe D, Fuentes S, Barker I, et al. Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: a generic method for diagnosis, discovery and sequencing of viruses. Virology. 2009;388(1):1–7. https://doi.org/10.1016/j.viro.2009.03.024.

Karyeija RF, Kreuze JF, Gibson RW, Valkonen JPT. Synergistic interactions of a potyvirus and a phloem-limited crinivirus in sweet potato plants. Virology. 2000;269(1):26–36. https://doi.org/10.1006/viro.1999.0169.

Kreuze JF, Perez A, Untiveros M, Quispe D, Fuentes S, Barker I, et al. Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: a generic method for diagnosis, discovery and sequencing of viruses. Virology. 2009;388(1):1–7. https://doi.org/10.1016/j.viro.2009.03.024.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary analyses across computing platforms. Mol Biol Evol. 2018;35(6):1547–9. https://doi.org/10.1093/molbev/msy096.

Li XC, Zhang ZC, Qiao Q, Qin YH, Wang S, Zheng WM. Detection of the coat protein gene of sweet potato badnavirus B in Escherichia coli. J Plant Prot. 2016;43(2):345–6 (in Chinese) https://doi.org/10.13802/cnki.zwjwjbhs.2016.02.025.

Liu Q, Zhang Z, Qiao Q, Qin Y, Zhang D, Tian Y, et al. Complete genome sequence of a novel monopartite begomovirus infecting sweet potato in China. Virus Genes. 2013;47(1):591–4. https://doi.org/10.1007/s11262-013-0982-3.

Liu Q, Zhang Z, Li J, Qiao Q, Qin Y, Zhang D, et al. Complete genome sequence of a novel monopartite begomovirus infecting sweet potato in China. Arch Virol. 2014;159:1537–40. https://doi.org/10.1007/s00717-013-1958-2.

Liu Q, Wang Y, Zhang Z, Lv H, Qiao Q, Qin Y, et al. Diversity of sweepoviruses infecting sweet potato in China. Plant Dis. 2017;101(12):2098–103. https://doi.org/10.1094/PDIS-04-17-0524-RP.

Luan YS, Zhang J, An LJ. First report of sweet potato leaf curl virus in China. Plant Dis. 2006;90(9):1111. https://doi.org/10.1094/PD-90-1111C.

Luan YS, Zhang J, Liu DM, Li WL. Molecular characterization of sweet potato leaf curl virus isolate from China (SPLCV-CN) and its phylogenetic relationship with other members of the Geminiviridae. Virus Genes. 2007;35:379–85. https://doi.org/10.1007/s11262-007-0084-1.

Muhire BM, Varsani A, Martin DP. SdT: A virus classification tool based on pairwise sequence alignment and identity calculation. PLOS ONE. 2014;9(9):e108277. https://doi.org/10.1371/journal.pone.0108277.