

# SCIENTIFIC REPORTS



OPEN

## Engineering of soybean mosaic virus as a versatile tool for studying protein–protein interactions in soybean

Jang-Kyun Seo<sup>1</sup>, Hong-Soo Choi<sup>1</sup> & Kook-Hyung Kim<sup>2</sup>

Received: 27 November 2015

Accepted: 15 February 2016

Published: 29 February 2016

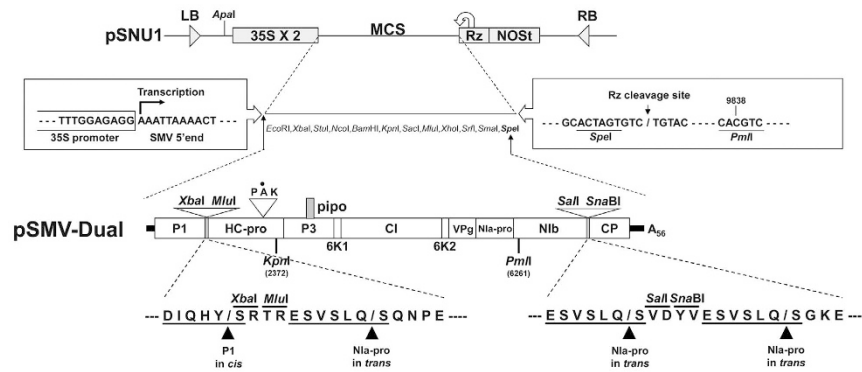
Transient gene expression approaches are valuable tools for rapid introduction of genes of interest and characterization of their functions in plants. Although agroinfiltration is the most effectively and routinely used method for transient expression of multiple genes in various plant species, this approach has been largely unsuccessful in soybean. In this study, we engineered soybean mosaic virus (SMV) as a dual-gene delivery vector to simultaneously deliver and express two genes in soybean cells. We further show the application of the SMV-based dual vector for a bimolecular fluorescence complementation assay to visualize *in vivo* protein–protein interactions in soybean and for a co-immunoprecipitation assay to identify cellular proteins interacting with SMV helper component protease. This approach provides a rapid and cost-effective tool for transient introduction of multiple traits into soybean and for *in vivo* characterization of the soybean cellular protein interaction network.

Transient gene expression systems, including *Agrobacterium*-mediated gene delivery (agroinfiltration), particle bombardment, and virus-based gene expression and silencing vectors, have proven to be powerful and versatile tools for gain-of-function and loss-of-function approaches in plant molecular, cellular, genetic, and proteomic studies. In particular, agroinfiltration has been used widely and effectively for ectopic expression of genes of interest and examination of their functions<sup>1,2</sup>. Most importantly, agroinfiltration facilitates synchronized delivery of multiple transgenes into the same cell, offering the advantage of expressing multiple recombinant proteins in the same cell and examining their interactions and functions<sup>3,4</sup>. However, despite efficient application of agroinfiltration in various plant systems, this method has been largely unsuccessful in soybean.

Alternatively, virus-mediated gene delivery systems have been developed for systemic expression of recombinant proteins and gene silencing in soybean plants<sup>5–7</sup>. Virus-mediated expression systems are superior to other transient gene expression systems as well as the transgenic approach because viruses infect their host plants systemically and replicate to high titers, so that large amounts of recombinant proteins accumulate within a short period of time<sup>8</sup>. To date, a few viruses, including bean pod mottle virus (BPMV), cucumber mosaic virus (CMV), and soybean mosaic virus (SMV), have been engineered as gene delivery vectors for systemic expression of recombinant proteins in soybean<sup>5,7,9</sup>. However, the viral vectors derived from multipartite viruses (i.e., BPMV and CMV) have some restrictions in introducing and expressing desired genes in plants because the estimated maximum DNA fragment sizes that can be inserted into BPMV and CMV genomes are approximately 1.8 kb and 0.9 kb, respectively<sup>5</sup>. In addition, expensive and complicated inoculation procedures, such as particle bombardment and *in vitro* transcription, are required to achieve high infection rates with these multipartite virus-derived vectors.

SMV, which has a monopartite single-stranded RNA genome of approximately 9.6 kb, is a member of the genus *Potyvirus*. Various potyviruses have been developed as viral gene delivery vectors<sup>8</sup>. Advantageous features of potyvirus-mediated expression systems include simultaneous equimolecular expression of multiple desired genes and relatively flexible length of the foreign genes (up to 4 kb) that can be expressed<sup>10,11</sup>. In our previous study, we developed an SMV-based gene delivery vector by engineering the cloning sites and the additional NIa-Pro

<sup>1</sup>Crop Protection Division, National Academy of Agricultural Science, Rural Development Administration, Wanju 565-851, Republic of Korea. <sup>2</sup>Department of Agricultural Biotechnology and Plant Genomics and Breeding Institute, Seoul National University, Seoul 151-921, Republic of Korea. Correspondence and requests for materials should be addressed to J.-K.S. (email: jakyseo@korea.kr)



**Figure 1. Schematic representation of the construction of pSMV-Dual vector.** The binary vector, pSNU1, contains, in sequential order, a left border of T-DNA (LB), a double 35S promoter (35S), multiple cloning site (MCS), a cis-cleaving ribozyme sequence (Rz), a NOS terminator (NOSI), and a right border of T-DNA (RB). pSMV-Dual contains two gene insertion cassettes between the P1 and HC-Pro cistron and between the Nib and CP cistrons. Each gene insertion cassette contains two unique cloning sites as indicated. Amino acid sequences of the peptide cleavage sites recognized by either the P1 or Nla-Pro viral proteases are underlined and arrowheads indicate the location of the cleaved peptide bond. A non-aphid-transmissible mutation (substitution of Thr to Ala in the PTK motif of HC-Pro) was introduced into pSMV-Dual.

cleavage site between SMV P1 and helper component protease (HC-Pro) cistrons and successfully expressed single recombinant proteins in soybean<sup>7</sup>. As in other potyvirus-mediated expression systems<sup>10</sup>, recombinant proteins expressed by the SMV-based vector are synthesized as part of the viral polyprotein. We also showed that simple rub-inoculation of plasmid DNAs of the SMV-based viral vectors was successful to cause infection and systemically express recombinant proteins in soybean plants<sup>7</sup>.

In the present study, we further engineered SMV as a dual-gene delivery vector to simultaneously deliver and express two genes in soybean cells. We successfully visualized subcellular localization of two different fluorescent proteins in soybean cells using the SMV-based dual vector. In addition, we applied the SMV-based dual vector system to a bimolecular fluorescence complementation (BiFC) assay to visualize *in vivo* protein–protein interactions in soybean. We described the detailed procedure for a co-immunoprecipitation (co-IP) assay in combination with the SMV-based dual vector to identify cellular proteins interacting with SMV HC-Pro. We expect that our procedures will provide useful tools to the soybean research community.

## Results and Discussion

**Engineering of SMV as a dual-gene delivery vector.** We previously developed a promising gene delivery system by engineering the full-length infectious cDNA clone of SMV strain G7H (pSMV-G7H)<sup>7</sup>. The viral vector, named pSMV-MCS, contains a single gene insertion cassette between P1 and HC-Pro<sup>7</sup>. The desired genes can be stably and systemically delivered into soybean by simple rub-inoculation with intact plasmid DNA of this recombinant SMV-based vector.

In the current study, we modified pSMV-MCS by engineering an additional gene insertion cassette between nuclear inclusion b (Nib) and coat protein (CP) cistrons. The second gene insertion cassette is composed of two restriction enzyme sites (*SalI* and *SnaBI*) and an additional Nla-Pro cleavage site (ESVSLQ) (Fig. 1). To minimize the potential for homologous recombination during plasmid DNA replication and subsequent instability, each residue of the inserted Nla-Pro cleavage site was selected based on the codon usage frequency for SMV. Abolishment of plant-to-plant transmission capacity of a plant virus-based vector is an important environmental issue that may facilitate its field application. Therefore, a non-aphid-transmissible mutation (substitution of Thr to Ala at amino acid position 310 of SMV HC-Pro; T310A) was introduced into the P<sub>309</sub>T<sub>310</sub>K<sub>311</sub> motif of HC-Pro, which is the critical motif for aphid transmission of potyviruses<sup>12</sup>. The resulting dual-gene delivery vector was named pSMV-Dual (Fig. 1). Two different genes can be delivered simultaneously by the pSMV-Dual vector upon utilizing two gene insertion cassettes to create an in-frame translational fusion. Therefore, the recombinant protein expressed from the first gene insertion cassette will have two additional amino acids (SR) at the N-terminus and 10 amino acids (SRTRESVSLQ) at the C-terminus when cloned by utilizing the *XbaI* site. Proteolysis will produce a recombinant protein having three (SVD) and 10 (VDYVESVSLQ) additional amino acids at the N-terminus and C-terminus, respectively, when cloned by utilizing the *SalI* site (Fig. 1).

In the previous study, we showed that DNA-mediated rub-inoculation of the SMV infectious cDNA construct yielded highly efficient infection on soybean plants<sup>7</sup>. Thus, in the present study, we sought to examine whether the additional insertion of the gene cassette between the Nib and CP cistrons affects the infectivity of the pSMV-Dual plasmid. To this end, soybean seedlings (cv. Lee74, used throughout this study) were rub-inoculated with different quantities (10 µg, 5 µg, or 1 µg) of plasmid DNA of pSMV-Dual. Experiments were carried out three times independently, using 45 plants in total (Table 1). Infection of the inoculated plants with SMV was investigated by observing the appearance of symptoms on systemic leaves and by RT-PCR analysis using SMV-specific primers as described previously<sup>7</sup>. At 15 days post inoculation (dpi), all of the soybean plants inoculated with 10 µg, 5 µg, or 1 µg of plasmid DNA showed typical mild mosaic symptoms in upper uninoculated leaves. RT-PCR analysis further confirmed that all the inoculated plants were systemically infected (Table 1).

Amount of inoculated DNA ( $\mu\text{g/plant}$ )	Infectivity <sup>a</sup>			
	1	2	3	Total
10	5/5 <sup>b</sup>	5/5	5/5	15/15
5	5/5	5/5	5/5	15/15
1	5/5	5/5	5/5	15/15
0	0/5	0/5	0/5	0/15

**Table 1. Dose response of soybean plants to pSMV-Dual plasmid DNA by rub-inoculation.** <sup>a</sup>Soybean seedlings (cv. Lee74) were rub-inoculated with the corresponding plasmids. <sup>b</sup>Number of systemically infected plants/number of plants inoculated. Virus infection of the upper non-inoculated leaves was confirmed by RT-PCR using SMV-specific primers at 15 dpi.

SMV construct	Aphid transmission		
	Exp. 1	Exp. 2	Exp. 3
Mock	0/5	0/5	0/5
pSMV-G7H	5/5	5/5	5/5
pSMV-MCS	5/5	5/5	5/5
pSMV-Dual	0/5	0/5	0/5

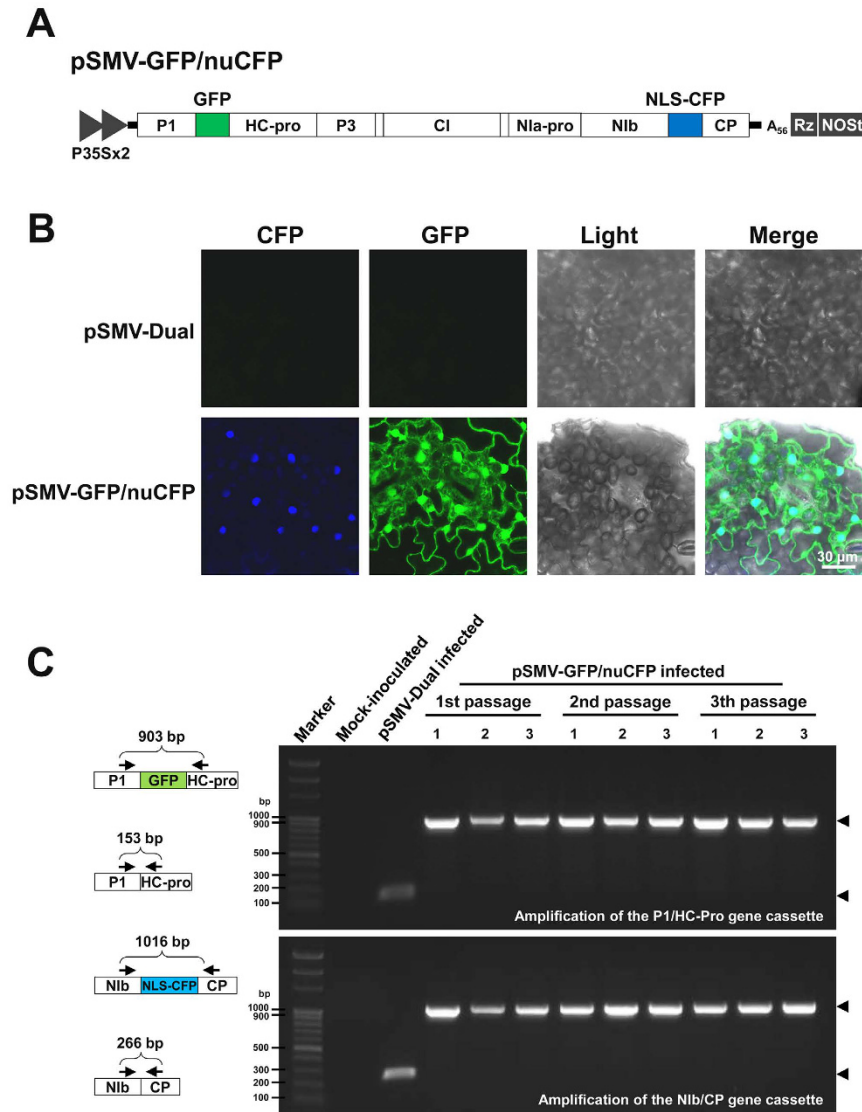
**Table 2. Abolishment of aphid transmissibility of pSMV-Dual.** <sup>a</sup>Number of plants infected over number of plants inoculated by aphid-transmission. Virus infection was confirmed by RT-PCR using SMV-specific primers.

To confirm the effect of the introduced non-aphid-transmissible mutation into pSMV-Dual, we conducted a plant-to-plant transmission assay using aphids (*Aphis glycines*). The transmission assays were performed three times independently, comprising 15 plants for each construct (Table 2). Although the progeny viruses of pSMV-G7H and -MCS were successfully transmitted, no aphid transmission of the progeny viruses of pSMV-Dual was observed, demonstrating that the introduced mutation abolished aphid transmissibility of pSMV-Dual.

**Simultaneous expression of two recombinant proteins and visualization of their subcellular accumulation.** To evaluate simultaneous expression of two recombinant proteins from the SMV dual vector in soybean, two fluorescence reporter genes, *gfp* and *cfp*, were cloned. In addition, to specifically visualize the expression of the resulting fluorescence proteins (i.e., GFP and CFP), CFP was expressed as a fusion protein tagged with the nuclear localization signal of SV40 T antigen (NLS). The *gfp* and NLS-tagged *cfp* genes were inserted into pSMV-Dual utilizing the cloning sites *Xba*I and *Sal*I, respectively, resulting in a construct designated as pSMV-GFP/nuCFP (Fig. 2B). Five micrograms of pSMV-GFP/nuCFP plasmid DNA were inoculated onto the primary leaves of soybean seedlings by direct rub-inoculation. At 15 dpi, typical mild mosaic symptoms appeared on the systemic leaves of the soybean plants inoculated with pSMV-GFP/nuCFP, similar to infection by the pSMV-Dual (data not shown). To verify whether the inoculated soybeans were infected systemically with the recombinant SMV, we extracted total RNA from upper uninoculated leaves at 15 dpi and subjected the extracts to RT-PCR using SMV-specific primers. The RT-PCR results confirmed that all of the inoculated soybean plants were infected systemically (data not shown). Next, systemic leaves of soybean plants inoculated with pSMV-GFP/nuCFP were subjected to confocal microscopy at 15 dpi to monitor fluorescent signals. As expected, the GFP signals were observed throughout the cytoplasm as well as in the nucleoplasm and the plasma membrane of the soybean cells, while the CFP signals were detected specifically in the nucleus (Fig. 2B), indicating that GFP and NLS-tagged CFP proteins were successfully expressed in the systemic leaves via the SMV-based dual vector. As a negative control, plants were inoculated with pSMV-Dual (empty vector) and no fluorescent signal was evident in the systemic leaves of the inoculated plants (Fig. 2B).

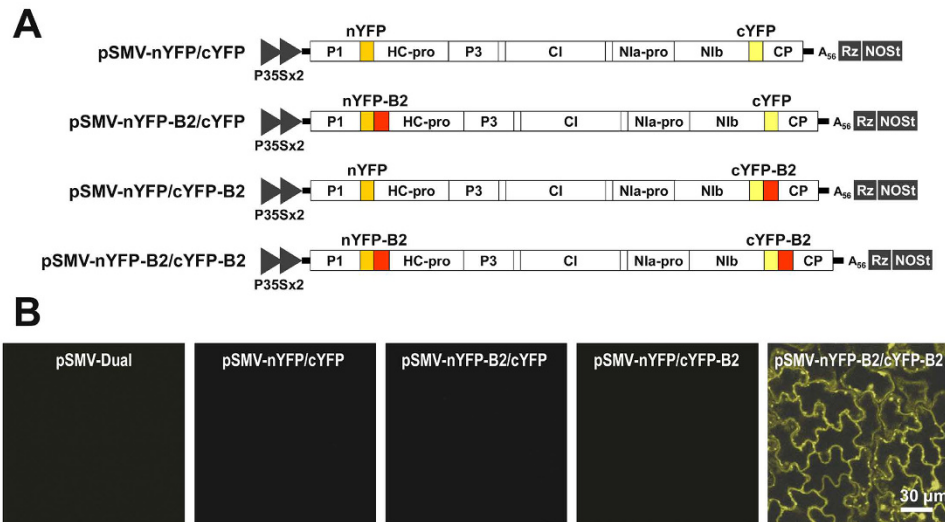
Few studies have been performed to the visualize subcellular distribution of cellular proteins in soybean cells mainly because of the unavailability of agroinfiltration in soybean leaves. In our confocal microscopy with high magnification, we clearly observed free GFP in the cytoplasm and the nuclear localization of NLS-tagged CFP in soybean cells. Thus, our approach will be useful to examine the subcellular localizations of host cellular proteins by expressing target proteins tagged with a fluorescence protein.

To test the stability of heterologous gene insertion in the SMV genome, the recombinant virus was transferred three times from plant to plant by mechanical sap-inoculation. Total RNA was isolated from each inoculated plant and analyzed for stable insertion of the *gfp* and *cfp* genes in the viral genome by RT-PCR using primer pairs spanning the gene insertion cassettes (Fig. 2C). Only amplicons with the predicted sizes were detected for viral genomes carrying both *gfp* and *cfp* genes (Fig. 2C). In addition, the fluorescence signals of GFP and CFP were readily detected in all soybean plants infected with pSMV-GFP/nuCFP or its progeny viruses through serial passages (data not shown). These results indicate that the dual-gene insertion in the viral genome was stably maintained during virus replication even after three serial passages.



**Figure 2.** Simultaneous expression and visualization of two recombinant proteins in soybean cells using the SMV-based dual vector (A) SMV genome organization and simultaneous insertion of the *gfp* gene between P1 and HC-Pro and the NLS-tagged *cfp* gene between Nib and CP. (B) Confocal images showing subcellular distribution of GFP and NLS-tagged CFP expressed in soybean cells by the SMV-based dual vector. No fluorescence was detected in the soybean cells infected by pSMV-Dual. (C) Analysis of stability of heterologous gene insertions in the SMV genome. Schematic maps of SMV P1/HC-Pro, P1/GFP/HC-Pro, Nib/CP, and Nib/NLS-CFP/CP regions are shown. Arrows indicate the regions that have been amplified. Total RNAs, isolated from mock inoculated, pSMV-Dual-infected, or pSMV-GFP/nuCFP-infected leaves of each soybean plant, were analyzed by RT-PCR. The progeny viruses were transferred three times from plant to plant by mechanical sap-inoculation. The numbered lanes indicate three individual plants tested in each passage. The arrowheads point at the RT-PCR products of the expected sizes.

**Application of the SMV-based dual vector for *in vivo* visualization of protein–protein interactions in soybean.** Protein–protein interactions are basic cellular events in the control of many cellular processes. Thus, characterizing protein–protein interactions is essential for unraveling the biological functions of proteins and is becoming increasingly important in understanding various aspects of cell biology. Various approaches have been developed to examine protein–protein interactions *in vitro* and *in vivo*<sup>13–16</sup>. Among them, BiFC has been used as a convenient and powerful tool for identifying and visualizing protein–protein interactions in living cells<sup>13,17</sup>. Currently, BiFC, in association with agroinfiltration, has been used widely to examine protein–protein interactions and the subcellular localization of the interacting protein partner in various plant systems<sup>13</sup>. BiFC requires co-expression of two target proteins fused with the N- and C-terminal nonfunctional halves of a fluorescent protein. Only when the two target proteins interact can the N- and C-terminal YFP fragments be brought into close proximity to reconstitute functional YFP as a result of specific protein interactions. However,



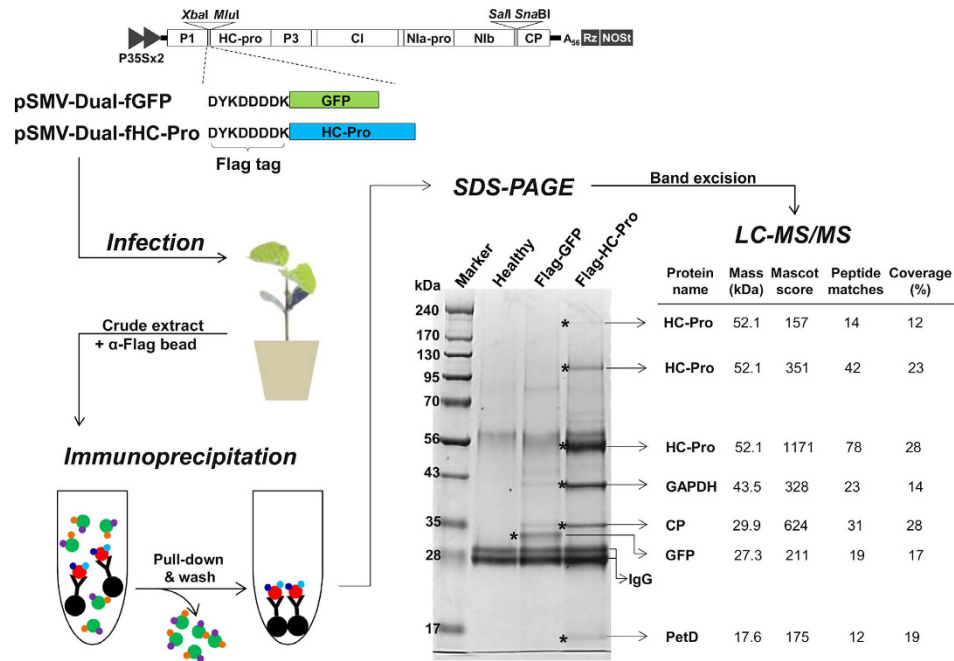
**Figure 3.** *In vivo* characterization of FHV B2 self-interaction in soybean cells by BiFC in combination with SMV-based dual-gene delivery. (A) Schematic representation of SMV recombinant constructs applied for BiFC assay to examine FHV B2 self-interaction. The open reading frames (ORFs) of nYFP, cYFP, nYFP-B2, and cYFP-B2 were in-frame inserted into the gene insertion cassettes of pSMV-Dual vectors for simultaneous expression. (B) *In vivo* visualization of FHV B2 self-interaction in soybean cells. Each SMV recombinant construct was mechanically inoculated into soybeans (cv. Lee74) as indicated at the top of each image. The reconstructed YFP signals were observed in the upper uninoculated leaves using confocal microscopy at 15 days post infiltration (dpi).

*in planta* application of BiFC analysis has not been performed in soybean because of a lack of efficient methods for synchronized expression of two recombinant proteins in soybean cells.

Because the SMV-based dual vector was used successfully to simultaneously express two recombinant proteins in a single cell and to visualize subcellular compartments emitting fluorescence signals using confocal microscopy, we sought to examine whether dual-gene delivery by the SMV-based dual vector could be applied for BiFC analysis of protein–protein interactions in soybean cells. A previous study showed *in vivo* self-interaction of protein B2 of flock house virus (FHV) by employing a BiFC assay in association with agroinfiltration in *Nicotiana benthamiana*<sup>18</sup>. Thus, we decided to test whether the B2 self-interaction could be detected in soybean cells when the same fusion recombinant B2 proteins are expressed by the SMV-based dual vector. To this end, we constructed four recombinant SMV constructs that express one of the following: (i) the N-terminal region (amino acids 1 to 156; nYFP)+the C-terminal region (amino acids 157 to 239; cYFP) of yellow fluorescent protein (YFP) (pSMV-nYFP/cYFP); (ii) nYFP-fused B2 + cYFP (pSMV-nYFP-B2/cYFP); (iii) nYFP + cYFP-fused B2; or (iv) nYFP-fused B2 + cYFP-fused B2 (Fig. 3A). Each construct was rub-inoculated on the leaves of soybean seedlings and the reconstructed YFP signals were monitored at 15 dpi using a confocal microscope. Strong fluorescence signals were observed in soybean cells when nYFP-B2 and cYFP-B2 were co-expressed by the SMV-based dual vector (Fig. 3A). However, the expression of other combinations employed as negative controls did not induce any detectable fluorescence signals, indicating that the fluorescence signals observed in the co-expression of nYFP-B2 and cYFP-B2 resulted from specific self-interaction of B2. The results suggest that our approach utilizing a viral dual vector can be a useful and efficient tool for *in vivo* characterization of protein–protein interactions in soybean and other plant systems.

**Application of the SMV-based dual vector for co-immunoprecipitation-based identification of cellular interacting protein partners.** The identification of interacting protein partners in a given pathway often provides decisive clues to establish a hierarchical mechanism of a system biology. Among the various strategies, co-IP coupled with mass spectrometric analysis is one of the most popular techniques for identification of interacting protein partners<sup>19,20</sup>. Co-IP uses an antibody that specifically binds to a target protein to isolate this protein and its interacting partners from cellular lysates. Co-IP can be performed by ectopically expressing a recombinant protein tagged with a small epitope such as HA and Flag<sup>21</sup>. In this case, the recombinant protein can be immunoprecipitated by epitope-specific antibodies. The immunoprecipitated protein complexes then can be identified directly by mass spectrometric analyses, such as liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) and matrix-assisted laser desorption/ionization time-of-flight MS (MALDI-TOF MS)<sup>20,21</sup>.

Despite the recent abundance of soybean genomic data, only limited information is available on soybean protein–protein interaction networks when compared with other model plant systems. This is mainly due to difficulty of transient expression of recombinant proteins in soybean as mentioned above. The SMV-based dual vector is capable of expressing recombinant proteins at a high level because the recombinant proteins are synthesized as part of the viral polyprotein<sup>7,8</sup>. Thus, we sought to examine whether transient expression of a recombinant protein



**Figure 4. Workflow of the SMV-based gene delivery for identification of cellular interacting protein partners.** Schematic representation of SMV recombinant constructs shows in-frame insertion of the Flag-tagged GFP and Flag-tagged HC-Pro into the P1/HC-Pro gene insertion cassette of pSMV-Dual vectors. Each SMV recombinant construct was rub-inoculated on the leaves of soybean seedlings. At 15 dpi, crude plant extracts prepared by homogenizing the upper symptomatic leaves were subjected to immunoprecipitation using anti-Flag antibody-conjugated agarose beads. The resulting co-immunoprecipitated products were analyzed by SDS-PAGE and the bands of interest (indicated by asterisks) were excised from the gel and subjected to LC-MS/MS analysis. The identified proteins and MS/MS spectral information are shown.

tagged with an epitope using the SMV-based vector is applicable for identification of cellular interacting protein partners by co-IP followed by mass spectrometric analysis.

The potyvirus HC-Pro is a multifunctional protein involved in crucial steps of virus infection<sup>22</sup>. HC-Pro has proteolytic activity to cleave at its carboxyl-terminus and is required not only for aphid transmission but also for long-distance systemic movement in plants, symptom expression, and suppression of RNA silencing. The multifunctional activities of HC-Pro may be regulated by interactions with other viral and cellular proteins. Indeed, direct interaction between potyvirus HC-Pro and CP mediates aphid transmission<sup>23,24</sup>. In addition, it has been shown that HC-Pro self-interacts to form oligomers, including dimers, tetramers, and hexamers<sup>25</sup>.

To confirm these known interactions by co-IP and to identify additional cellular interacting partner proteins, we decided to transiently express the SMV HC-Pro tagged with the Flag epitope (Asp-Tyr-Lys-Asp-Asp-Asp-Lys) at the N-terminus using the SMV-based dual vector. To this end, we generated a recombinant SMV construct, pSMV-Dual-fHC-Pro, that expresses Flag-tagged HC-Pro (Fig. 4). In parallel, an additional SMV construct, pSMV-Dual-fGFP, that expresses Flag-tagged GFP, was generated and used as a negative control in co-IP experiments. Each construct was rub-inoculated on the leaves of soybean seedlings. At 15 dpi, crude plant extracts were prepared by homogenizing the upper symptomatic leaves in the extraction buffer. After removing cell debris by centrifugation, the resulting extracts were incubated with anti-Flag antibody-conjugated agarose beads followed by precipitation by centrifugation. The resulting co-immunoprecipitated products were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and stained with Coomassie Blue (Fig. 4). After staining, the specific dominant bands in each lane were excised from the gel and subjected to LC-MS/MS analysis. The identified proteins and MS/MS spectral information are shown in Fig. 4. As expected, one dominant band in the Flag-GFP co-IP sample was identified as GFP. A total of six bands were analyzed in the Flag-HC-Pro co-IP sample. The bands at position 220, 110, and 55 kDa were identified as HC-Pro whereas the band at position 35 kDa was identified as SMV CP. Therefore, our co-IP approach confirmed the previous findings of the HC-Pro oligomerization and the HC-Pro-CP interaction. The identification of HC-Pro from the bands at positions 220, 110, and 55 kDa indicates tetrameric, dimeric, and monomeric forms of HC-Pro, respectively. The HC-Pro self-interaction is likely quite strong because the oligomeric forms were detected under the denaturing conditions of SDS-PAGE. In addition, we identified two novel cellular interacting partners of HC-Pro, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) at position 43 kDa and cytochrome b6/f complex subunit IV (PetD) at position 17 kDa. GAPDH is an important enzyme that plays a pivotal role in energy metabolism<sup>26</sup>. Recent studies have shown that GAPDH has multiple functions in DNA replication/repair, RNA transport, apoptosis, oxidative stress, membrane fusion, and cytoskeleton assembly<sup>27-30</sup>. In addition, a few studies have suggested that GAPDH may play a role in RNA virus replication<sup>31-33</sup>. On the other hand, PetD is required

for photosynthetic electron transport<sup>34</sup>. The involvement of PetD in virus infection cycle has not been studied yet. We are currently performing a separate study to further characterize the interactions between HC-Pro and either GAPDH or PetD and to examine the roles of GAPDH and PetD in HC-Pro functions and SMV infection.

The accumulation of a large amount of genetic information through high-throughput technologies has increased the need for rapid and simple analyses of gene function. Agroinfiltration as a transient gene expression approach has been used extensively for *in vivo* gene function analysis in *N. benthamiana* and several other plant species<sup>1,2</sup>. However, agroinfiltration is not equally successful in different plant species because the compatibility between the plant and bacterium varies<sup>35</sup>. In addition, the leaf architecture of some plant species including soybean is likely unsuitable for the application of agroinfiltration<sup>36</sup>.

Soybean is one of the most important crop plants because it serves as a major ingredient for a wide range of foods and beverage products. Considering the importance of soybean as a major food crop, it is highly desirable to identify and evaluate valuable traits to improve productivity, environmental and disease resistance, and the commercial value. The whole genome of the soybean has been sequenced<sup>37,38</sup> and high-throughput transcriptome analyses have identified numerous genes specifically involved in various biological processes<sup>39–43</sup>. However, most of the identified soybean genes still remain uncharacterized because of the lack of a versatile transient expression method such as agroinfiltration in soybean.

The interest in using plant systems as biofactories for production of valuable proteins has led to the development of various plant virus-based vectors<sup>8,44</sup>. In addition, plant virus-based vectors provide attractive and powerful tools for rapid introduction of genes of interest and characterization of their functions in plants. In the present study, we developed the SMV-based dual-gene delivery vector to simultaneously express two genes in soybean. We also showed the possible applications of the SMV-based dual vector in visualizing and characterizing protein subcellular localization and protein–protein interaction at the cellular level (Figs 2 and 3). In addition, we demonstrated that our approach in combination with co-IP and mass spectrometric analysis is useful for identification of cellular interacting protein partners in soybean (Fig. 4). Recently, we have identified a number of genes that are expressed differentially in susceptible and resistant responses against SMV in soybean by high-throughput transcriptome analysis<sup>43</sup>. We are currently characterizing the gain-of-function effects of some of these genes by simultaneously expressing two different genes using the SMV-based dual vector to evaluate synergistic effects of the genes.

Potyvirus constitute the largest genus of plant viruses and adopt the same gene expression strategy of proteolytic processing of polyprotein precursors, making heterologous technologies broadly applicable for manipulation of potyvirus genomes<sup>10,45,46</sup>. So far, various potyviruses including tobacco etch virus (TEV), turnip mosaic virus (TuMV), and zucchini yellow mosaic virus (ZYMV) that have a broad host range have been engineered to express genes of interest<sup>46–48</sup>. Therefore, our approaches shown in this study with the SMV-dual vector can be applied to other potyvirus–host plant systems for simultaneous expression of multiple genes, visualization of protein subcellular localization and identification and characterization of protein–protein interactions in plants in which agroinfiltration is unsuccessful. We expect that our method will become a versatile and powerful tool for studying protein–protein interactions and for rapid analysis of gene function in various plant species.

## Materials and Methods

**Construction of pSMV-Dual vector.** The SMV-based dual-gene delivery vector was constructed by engineering an additional gene insertion cassette between NIB and CP cistrons of pSMV-MCS<sup>7</sup>. The NIB region spanning from the *PmlI* site to the 3' end of NIB was amplified using a primer pair (5'-GTCAGATGTTCACGTCGCCAAA-3' and 5'-TAAAGATACGGACTCTACGTAGTCGACTGATGTAAAGGACACTGATTCACAACA-3', *PmlI*, *Sall*, and *SnaBI* sites are shown in bold). The CP and 3' untranslated region (from the 5' end of CP to the second *PmlI* site) was amplified using a primer pair (5'-GTCGACTACGTAGAGTCCGATCTTTACAGTCAGGTAAGGAGAAGGAAGGA-3' and 5'-GTCACCTGTAATTCACACGTGG-3', *PmlI*, *Sall*, and *SnaBI* sites are shown in bold, the nucleotide sequence for the NIa-Pro cleavage site is underlined). The two PCR fragments were joined by joint PCR as described previously<sup>49</sup>. The resulting joined fragment was digested with *PmlI* and inserted into pSMV-MCS, which was opened with *PmlI*. The resulting construct was named pSMVG7H-Dual. Next, introduction of the non-aphid-transmissible mutation, T310A, into HC-Pro was performed as follows. The HC-Pro region spanning from the 5' terminus to the *KpnI* site was amplified from pSMVdHC-HC<sub>T310A</sub>, which contains the non-aphid-transmissible mutation (T301A) in HC-Pro<sup>12</sup>, using a primer pair (5'-TAGAACGCGTGAGTCTGTCTCGTTGCAGTCCAAAATCCTGAAGCTCAGTT-3' and 5'-CCAGCTTTAAGAACATGGTACC-3', *MluI* and *KpnI* sites are shown in bold, the nucleotide sequence for NIa-Pro cleavage site is underlined). The resulting PCR fragment was digested with *MluI* and *KpnI* and inserted into pSMVG7H-Dual, which was opened with *MluI* and *KpnI*. This final construct, which was named pSMV-Dual, contains two gene insertion cassettes and a non-aphid-transmissible mutation.

**Insertion of heterologous genes into pSMV-Dual vector.** The *gfp* gene was amplified using a primer pair harboring *XbaI* sites (5'-GCTCTAGAATGGTGAGCAAGGGCGA-3' and 5'-GCTCTAGAGAGGATCCCCTTGACAG-3', *XbaI* sites are shown in bold). The NLS-tagged *cfp* gene was amplified using a primer pair harboring *Sall* sites (5'-ACGCGTCGACATGGTGAGCAAGGGCGAGGA-3' and 5'-ACGCGTCGACACTTCTCTTCTTTGGAG-3', *Sall* sites are shown in bold). The resulting amplicons were digested with *XbaI* and *Sall* and cloned into the first and second gene insertion cassettes of pSMV-Dual, respectively, for simultaneous expression of the two genes. A similar cloning strategy was applied for cloning of the nYFP, cYFP, nYFP-fused B2, and cYFP-fused B2 genes into pSMV-Dual. The nYFP, cYFP, nYFP-fused B2, and cYFP-fused B2 genes were amplified from the PZPn-nYFP-B2 and PZPn-cYFP-B2 clones<sup>18</sup> using appropriate primer pairs (the list of the primers is available on request) and cloned into pSMV-Dual utilizing the *XbaI* and *Sall* sites in the gene insertion cassettes, accordingly. The

Flag-tagged *gfp* and HC-Pro genes were amplified using the primer pairs (for Flag-tagged *gfp*, 5'-GCTCTAGA GACTACAAGGACGACGATGACAAGATGGTGGCAAGGGCGAGG-3' and 5'-GCTCTAGAGAGGATC CCCTTGTTACAGCT-3'; for Flag-tagged HC-Pro, 5'-GCTCTAGAGACTACAAGGACGACGATG ACAAGTCCCAAAATCCTGAAGCTCAGT-3' and 5'-GCTCTAGAACCACTCTGTAGAATTTTCATCTC-3'; *Xba*I sites are shown in bold, the nucleotide sequence for the Flag epitope is underlined). The resulting amplicons were digested with *Xba*I and cloned into pSMV-Dual, which was opened with *Xba*I.

**Plant growth and inoculation.** Soybean plants were grown in a growth chamber at 25 °C under a 16/8-h photoperiod. Seedlings were selected for inoculation when the cotyledons were fully expanded. Plasmid DNAs of the SMV constructs were prepared using the Plasmid Maxi Kit (QIAGEN, Valencia, CA, USA). Each cDNA plasmid was rub-inoculated as described previously<sup>7</sup>. To detect virus accumulation in the inoculated and upper uninoculated leaves, RT-PCR was performed using an SMV-specific primer pair (5'-GATTGGAAGCATGGCGATTT-3' and 5'-TTCACATACYTTCATGCCGTCAA-3'). We evaluated the stability of the heterologous gene insertion in the recombinant progeny viruses as described here. Total RNA was isolated from each inoculated plant and analyzed for stable insertion of heterologous genes in the viral genome by RT-PCR using the primer pairs spanning the gene insertion cassettes (for the P1/HC-Pro gene cassette, 5'-GAATGGGAAGCTCGTTAACGC-3' and 5'-GGAGGCATTTTATCAAACACCTT-3'; for the N1b/CP gene cassette, 5'-ATATTGCAGAGACAGCTTTGAGAA-3' and 5'-TCCAACATTTACATCTTTGCTGCT-3').

**Aphid transmission assay.** The aphid transmission assay was performed as described previously<sup>7</sup>. In brief, aphids (*A. glycines*) were reared in controlled environment chambers on soybean. Aphids, previously starved for 2 h, were placed on soybean leaves showing SMV symptoms and allowed to probe for 10 min. Then, 15 aphids were transferred to new healthy soybean seedlings and allowed to feed for 24 h before being killed with an insecticide. Inoculated plants were maintained for 3 weeks and SMV infection was verified by RT-PCR.

**Confocal microscopy.** Fluorescence signals emitted by GFP, CFP, and YFP in soybean leaves were visualized by confocal microscopy. At 15 dpi, upper uninoculated leaves were observed for emission of fluorescence using a Leica SP8 laser-scanning confocal microscope (Leica, Wetzlar, Germany) equipped with a specific laser/filter combination to detect CFP (excitation at 458 nm), GFP (excitation at 488 nm), and YFP (excitation at 514 nm).

**Immunoprecipitation assay.** Total protein extracts were prepared from the systemic leaves of the soybean plants inoculated with pSMV-Dual-fGFP and pSMV-Dual-fHC-Pro. At 15 dpi, the leaves were homogenized in three volumes of protein extraction buffer (20 mM Tris-HCl at pH 7.5, 300 mM NaCl, 5 mM MgCl<sub>2</sub>, 5 mM dithiothreitol, 0.5% Triton X-100, proteinase inhibitor cocktail [Sigma, St Louis, MO, USA]). Cell debris was removed by centrifugation at 18,000 g for 20 min at 4 °C. The resulting supernatants were incubated with anti-Flag antibody conjugated agarose beads (Takara, Japan) for 16 h at 4 °C. The immunocomplexes were then precipitated by centrifugation for 1 min at 8,200 g and washed five times in 1 mL of the protein extraction buffer. The resulting samples were analyzed by 10% SDS-PAGE and stained with Coomassie blue. The Xpert prestained protein marker (GenDEPOT, Barker, TX, USA) was used as the molecular mass. After staining, bands of interest were excised from the gel and subjected to in-gel digestion followed by LC-MS/MS analysis as described previously<sup>50</sup>.

**Peptide Sequence Analysis by LC-MS/MS and Database Search.** The entire LC-MS/MS procedure was performed at Yonsei Proteome Research Center (Seoul, South Korea). Briefly, LC was performed with an Easy n-LC 1000 system (Thermo Fisher Scientific, Rockford, IL, USA). A C18-nanobore column (150 mm × 0.1 mm, 3-μm pore size, Agilent) was used for peptide separation. LTQ-Orbitrap mass spectrometry (Thermo Fisher, San Jose, CA, USA) was used to identify and quantify peptides. Xcalibur (version 2.1, Thermo Fisher Scientific, Waltham, MA, USA) was used to generate peak lists. The peak lists were examined by searching the National Center for Biotechnology Information database using the MASCOT search engine (<http://www.matrixscience.com>, Matrix Science, Boston, MA, USA). The acquired data were compared to the whole database with search parameters set as follows: enzyme, trypsin; allowance of up to one missed cleavage peptide; mass tolerance ±0.5 Da and MS/MS tolerance ±0.5 Da; modifications of methionine oxidation and cysteine carbamidomethylation when appropriate, with auto hits allowed and only significant hits to be reported. The proteins were identified on the basis of two or more peptides whose ion scores exceeded the threshold,  $P < 0.05$ , which indicated the 95% confidence level for these matched peptides.

## References

- Lee, M. W. & Yang, Y. Transient expression assay by agroinfiltration of leaves. *Methods Mol. Biol.* **323**, 225–229 (2006).
- Schob, H., Kunc, C. & Meins, F. Silencing of transgenes introduced into leaves by agroinfiltration: a simple, rapid method for investigating sequence requirements for gene silencing. *Mol. Gen. Genet.* **256**, 581–585 (1997).
- Vaghchhipawala, Z., Rojas, C. M., Senthil-Kumar, M. & Mysore, K. S. Agroinoculation and agroinfiltration: simple tools for complex gene function analyses. *Methods Mol. Biol.* **678**, 65–76 (2011).
- Annamalai, P. & Rao, A. L. Delivery and expression of functional viral RNA genomes in planta by agroinfiltration. *Curr. Protoc. Microbiol.* Chapter 16, Unit 16B 12 (2006).
- Zhang, C. & Ghabrial, S. A. Development of Bean pod mottle virus-based vectors for stable protein expression and sequence-specific virus-induced gene silencing in soybean. *Virology* **344**, 401–411 (2006).
- Zhang, C., Bradshaw, J. D., Whitham, S. A. & Hill, J. H. The development of an efficient multipurpose bean pod mottle virus viral vector set for foreign gene expression and RNA silencing. *Plant Physiol.* **153**, 52–65 (2010).
- Seo, J. K., Lee, H. G. & Kim, K. H. Systemic gene delivery into soybean by simple rub-inoculation with plasmid DNA of a Soybean mosaic virus-based vector (vol 154, pg 87, 2009). *Arch. Virol.* **154**, 1575–1575 (2009).
- Gleba, Y., Klimyuk, V. & Marillonnet, S. Viral vectors for the expression of proteins in plants. *Curr. Opin. Biotechnol.* **18**, 134–141 (2007).



9. Matsuo, K. *et al.* Development of Cucumber mosaic virus as a vector modifiable for different host species to produce therapeutic proteins. *Planta* **225**, 277–286 (2007).
10. Bedoya, L., Martinez, F., Rubio, L. & Daros, J. A. Simultaneous equimolar expression of multiple proteins in plants from a disarmed potyvirus vector. *J. Biotechnol.* **150**, 268–275 (2010).
11. Kelloniemi, J., Makinen, K. & Valkonen, J. P. Three heterologous proteins simultaneously expressed from a chimeric potyvirus: infectivity, stability and the correlation of genome and virion lengths. *Virus Res.* **135**, 282–291 (2008).
12. Seo, J. K., Kang, S. H., Seo, B. Y., Jung, J. K. & Kim, K. H. Mutational analysis of interaction between coat protein and helper component-proteinase of Soybean mosaic virus involved in aphid transmission. *Mol. Plant Pathol.* **11**, 265–276 (2010).
13. Walter, M. *et al.* Visualization of protein interactions in living plant cells using bimolecular fluorescence complementation. *Plant J.* **40**, 428–438 (2004).
14. Chen, Y., Mills, J. D. & Periasamy, A. Protein localization in living cells and tissues using FRET and FLIM. *Differentiation* **71**, 528–541 (2003).
15. Fields, S. & Song, O. K. A Novel Genetic System to Detect Protein Protein Interactions. *Nature* **340**, 245–246 (1989).
16. Tsuda, K. *et al.* An efficient Agrobacterium-mediated transient transformation of Arabidopsis. *Plant J.* **69**, 713–719 (2012).
17. Kerppola, T. K. Design and implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells. *Nat. Protoc.* **1**, 1278–1286 (2006).
18. Seo, J. K., Kwon, S. J. & Rao, A. L. Molecular dissection of Flock house virus protein B2 reveals that electrostatic interactions between N-terminal domains of B2 monomers are critical for dimerization. *Virology* **432**, 296–305 (2012).
19. Ho, Y. *et al.* Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* **415**, 180–183 (2002).
20. Mann, M., Hendrickson, R. C. & Pandey, A. Analysis of proteins and proteomes by mass spectrometry. *Annu. Rev. Biochem.* **70**, 437–473 (2001).
21. Vasilescu, J. & Figeys, D. Mapping protein-protein interactions by mass spectrometry. *Curr. Opin. Biotech.* **17**, 394–399 (2006).
22. Maia, I. G., Haenni, A. L. & Bernardi, F. Potyviral HC-Pro: A multifunctional protein. *J. Gen. Virol.* **77**, 1335–1341 (1996).
23. Roudet-Tavert, G. *et al.* Interaction between potyvirus helper component-proteinase and capsid protein in infected plants. *J. Gen. Virol.* **83**, 1765–1770 (2002).
24. Blanc, S. *et al.* A specific interaction between coat protein and helper component correlates with aphid transmission of a potyvirus. *Virology* **231**, 141–147 (1997).
25. Ruiz-Ferrer, V. *et al.* Structural analysis of tobacco etch potyvirus HC-Pro oligomers involved in aphid transmission. *J. Virol.* **79**, 3758–3765 (2005).
26. Sirover, M. A. Role of the glycolytic protein, glyceraldehyde-3-phosphate dehydrogenase, in normal cell function and in cell pathology. *J. Cell. Biochem.* **66**, 133–140 (1997).
27. Glaser, P. E., Han, X. L. & Gross, R. W. Tubulin is the endogenous inhibitor of the glyceraldehyde 3-phosphate dehydrogenase isoform that catalyzes membrane fusion: Implications for the coordinated regulation of glycolysis and membrane fusion. *Proc. Natl. Acad. Sci. USA* **99**, 14104–14109 (2002).
28. Ishitani, R. & Chuang, D. M. Glyceraldehyde-3-phosphate dehydrogenase antisense oligodeoxynucleotides protect against cytosine arabinonucleoside-induced apoptosis in cultured cerebellar neurons. *Proc. Natl. Acad. Sci. USA* **93**, 9937–9941 (1996).
29. Berry, M. D. & Boulton, A. A. Glyceraldehyde-3-phosphate dehydrogenase and apoptosis. *J. Neurosci. Res.* **60**, 150–154 (2000).
30. Sirover, M. A. On the functional diversity of glyceraldehyde-3-phosphate dehydrogenase: biochemical mechanisms and regulatory control. *Biochim. Biophys. Acta* **1810**, 741–751 (2011).
31. Yang, S. H., Liu, M. L., Tien, C. F., Chou, S. J. & Chang, R. Y. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) interaction with 3' ends of Japanese encephalitis virus RNA and colocalization with the viral NS5 protein. *J. Biomed. Sci.* **16** (2009).
32. Prasanth, K. R. *et al.* Glyceraldehyde 3-phosphate dehydrogenase negatively regulates the replication of Bamboo mosaic virus and its associated satellite RNA. *J. Virol.* **85**, 8829–8840 (2011).
33. Kishimoto, N. *et al.* Glyceraldehyde 3-phosphate dehydrogenase negatively regulates human immunodeficiency virus type 1 infection. *Retrovirology* **9**, 107 (2012).
34. Sakamoto, W., Kindle, K. L. & Stern, D. B. *In vivo* analysis of *Chlamydomonas* chloroplast petD gene expression using stable transformation of beta-glucuronidase translational fusions. *Proc. Natl. Acad. Sci. USA* **90**, 497–501 (1993).
35. Wroblewski, T., Tomczak, A. & Michelmore, R. Optimization of Agrobacterium-mediated transient assays of gene expression in lettuce, tomato and Arabidopsis. *Plant Biotechnol. J.* **3**, 259–273 (2005).
36. King, J. L., Finer, J. J. & McHale, L. K. Development and optimization of agroinfiltration for soybean. *Plant Cell Rep.* **34**, 133–140 (2015).
37. Schmutz, J. *et al.* Genome sequence of the palaeopolyploid soybean. *Nature* **463**, 178–183 (2010).
38. Kim, M. Y. *et al.* Whole-genome sequencing and intensive analysis of the undomesticated soybean (*Glycine soja* Sieb. and Zucc.) genome. *Proc. Natl. Acad. Sci. USA* **107**, 22032–22037 (2010).
39. Severin, A. J. *et al.* RNA-Seq Atlas of *Glycine max*: a guide to the soybean transcriptome. *BMC Plant Biol.* **10**, 160 (2010).
40. Kim, K. H. *et al.* RNA-Seq analysis of a soybean near-isogenic line carrying bacterial leaf pustule-resistant and -susceptible alleles. *DNA Res.* **18**, 483–497 (2011).
41. Jones, S. I. & Vodkin, L. O. Using RNA-Seq to profile soybean seed development from fertilization to maturity. *PLoS One* **8**, e59270 (2013).
42. Vidal, R. O., do Nascimento, L. C., Mondego, J. M., Pereira, G. A. & Carazzolle, M. F. Identification of SNPs in RNA-seq data of two cultivars of *Glycine max* (soybean) differing in drought resistance. *Genet. Mol. Biol.* **35**, 331–334 (2012).
43. Seo, J. K., Kwon, S. J., Cho, W. K., Choi, H. S. & Kim, K. H. Type 2C Protein Phosphatase Is a Key Regulator of Antiviral Extreme Resistance Limiting Virus Spread. *Sci. Rep.* **4** (2014).
44. Pogue, G. P., Lindbo, J. A., Garger, S. J. & Fitzmaurice, W. P. Making an ally from an enemy: plant virology and the new agriculture. *Annu. Rev. Phytopathol.* **40**, 45–74 (2002).
45. Urququi-Inchima, S., Haenni, A. L. & Bernardi, F. Potyvirus proteins: a wealth of functions. *Virus Res.* **74**, 157–175 (2001).
46. Arazi, T. *et al.* Engineering zucchini yellow mosaic potyvirus as a non-pathogenic vector for expression of heterologous proteins in cucurbits. *J. Biotechnol.* **87**, 67–82 (2001).
47. Dolja, V. V., McBride, H. J. & Carrington, J. C. Tagging of plant potyvirus replication and movement by insertion of beta-glucuronidase into the viral polyprotein. *Proc. Natl. Acad. Sci. USA* **89**, 10208–10212 (1992).
48. Beauchemin, C., Bougie, V. & Laliberte, J. F. Simultaneous production of two foreign proteins from a polyvirus-based vector. *Virus Res.* **112**, 1–8 (2005).
49. Yu, J. H. *et al.* Double-joint PCR: a PCR-based molecular tool for gene manipulations in filamentous fungi. *Fungal Genet. Biol.* **41**, 973–981 (2004).
50. Park, M. *et al.* Distinct Protein Expression Profiles of Solid-Pseudopapillary Neoplasms of the Pancreas. *J. Proteome Res.* **14**, 3007–3014 (2015).

## Acknowledgements

This research was supported by a grant from the Agenda Program (PJ011306) funded by the Rural Development Administration of Korea.

## Author Contributions

J.-K.S., H.-S.C. and K.-H.K. designed the experiments. J.-K.S. performed the experiments. J.-K.S., H.-S.C. and K.-H.K. analyzed the data. J.-K.S. and K.-H.K. wrote the manuscript. All authors have read and approved the manuscript.

## Additional Information

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Seo, J.-K. *et al.* Engineering of soybean mosaic virus as a versatile tool for studying protein-protein interactions in soybean. *Sci. Rep.* **6**, 22436; doi: 10.1038/srep22436 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>