A newly-identified inactive starch synthase simultaneously regulates starch synthesis and carbon allocation in storage roots of cassava (*Manihot esculenta* Crantz)

Running title: Function of a newly-identified starch synthase in cassava

Shutao He[1,2,3], Xiaomeng Hao[1,3], Shanshan Wang[1,3], Wenzhi Zhou[1], Qiuxiang Ma[1], Xinlu Lu[1], Luonan Chen[2,*], Peng Zhang[1,3,*]

[1] National Key Laboratory of Plant Molecular Genetics, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China

[2] Key Laboratory of Systems Biology, CAS Center for Excellence in Molecular Cell Science, Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, Shanghai, China

[3] University of Chinese Academy of Sciences, Beijing, China

*** Corresponding authors:**

Luonan Chen (lnchen@sibcb.ac.cn) or Peng Zhang (zhangpeng@cemps.ac.cn)

Abstract: Starch is a glucose polymer synthesized by green plants for energy storage, and is crucial for plant growth and reproduction. The biosynthesis of starch polysaccharides is mediated by members of the large starch synthase (SS) protein superfamily. Here, we report the functionality of a cassava SS that clusters with a previously unreported lineage, SSVI. A phylogenetic analysis based on SS sequences from 44 plant species showed that SSVI is distributed in eudicots, but not in monocots. Functional analyses showed that in cassava storage roots, SSVI is inactive but plays an important role in both starch biosynthesis and carbon allocation via sugar-induced feedback regulation and also by forming protein complexes with starch biosynthetic enzymes. Contrary to current dogma, our results show a close coordination between granule-bound starch synthase and amylopectin biosynthetic enzymes, which implies that the processes of amylose synthesis and amylopectin synthesis are directly interrelated. These findings shed light on the key components of the starch biosynthesis machinery in root crops.

Keywords: cassava, carbon allocation, protein–protein interaction, starch biosynthesis, SSVI, storage root

Introduction

Cassava (*Manihot esculenta* Crantz) accumulates large amounts of starch (up to 32% by fresh weight) in its storage roots even under unfavorable growth conditions in tropical and subtropical regions (Cock, 1982). Cassava starch is pure white and contains low levels of fat (0.08–1.54%) and proteins (0.03–0.6%), providing an irreplaceable raw material for bioindustrial applications (Li et al., 2017). Cassava starch granules are generally round and truncated, ranging in size from 5-40 μm, and the amylose content varies from 15% to 30%, depending on the cultivar and growth conditions (Moorthy, 2002). Various modified cassava starch derivatives are widely used in the food, textile, paper manufacturing, pharmaceutical, and bioethanol production industries (Zhu, 2015). Despite this, the regulation of starch biosynthesis in cassava storage roots is largely uncharacterized.

Similar to starches from other plant sources, cassava starch in its native form consists of both amylose and amylopectin. Basically, amylose is a largely linear polymer in which glucosyl residues are joined by α(1→4) bonds, and is an important and interspersed component of starch, but is not crucial for the spatial structure in the formation of starch granules. Amylopectin is composed of a clustered arrangement of crystalline, densely packed glucosyl residues organized as parallel chains, in which linear α-helices are joined by α(1→4) bonds, and interspersed by more disorganized α(1→6) branches at regular intervals of ~20-30 glucose units. Such organization eventually forms a classic semicrystalline model that determines the stable and insoluble granular structure of starch (Hizukuri, 1986). Typically, branched amylopectin chains are classified as A, with a degree of polymerization (DP) of 6-12, B_1 (DP 13-24), B_2 (DP 25-36), and B_{3+} (DP ≥37) (Hizukuri, 1985). The chain length of cassava starch averages 23.2 DP, and the majority consists

of up to 43.7% B_1 chains; the molecular weight of native cassava starch varies from 1.885×10^7 to 3.323×10^7 g/mol, depending on cultivars and growth conditions (Liu et al., 2018).

Amylose is synthesized by a single enzyme, granule-bound starch synthase (GBSS), and mutants of cassava lacking GBSS produce amylose-free waxy starches(Bull et al., 2018; Ceballos et al., 2007; Zhao et al., 2011). Recent studies have shown that the activity of GBSS can be regulated by post-translational modifications. It was found that in maize, GBSS is phosphorylated in the starch granule (Grimaud et al., 2008), which is supported by a recent report of rice endosperm GBSS (Liu et al., 2013). Liu *et al.* (2013) also noted that GBSS can form oligomers, and ATP and protein kinases increase the degree of oligomerization. In addition, with increasing concentrations of ADP-glucose, GBSS oligomerization increases[12]. However, it is still unknown whether GBSS associates with amylopectin synthases directly in the plastids.

The synthesis of amylopectin, which mainly described from grain seeds, requires the coordinated participation of at least 10 different enzymes, including soluble starch synthases (SSs), starch branching enzymes (BEs or SBEs), and debranching enzymes (DBEs). SSs (EC 2.4.1.21) elongate α-1,4-linked linear glucan polymers using ADP-glucose as a substrate. BEs (EC 2.4.1.18) produce α-1,6-linked branches and are crucial for amylopectin formation because they are the only enzymes that can form branches in amylopectin. DBEs such as isoamylase (ISA; EC 3.2.1.68) and pullulanase (PUL; EC 3.2.1.41) hydrolyze and remove α-1,6-linked branches, which is essential to organize crystalline structure formation in amylopectin (Jeon et al., 2010; Nakamura, 2002). In addition, starch phosphorylase (Pho; EC 2.4.1.1) elongates α-1,4-linked linear glucan polymers using glucose 1-phosphate (G1P) as a substrate, and is thought to be involved in the initiation process of starch biosynthesis (Jeon et al., 2010; Satoh et al., 2008). Although GBSSI has an effect on the extra-long amylopectin chains in addition to participating in amylose synthesis (Hanashiro et al., 2008), no direct evidence of its interaction with other amylopectin biosynthetic enzymes has been reported. Recently, starches with remarkably altered physico-chemical properties and amylopectin structure have been observed in transgenic cassava in which the expressions of genes for GBSS1 and BEs have been silenced (Zhao et al., 2011; Zhou et al., 2019). Nevertheless, the function and regulation of these SSs in cassava remain unclear.

Previous studies have shown that SS can interact with BE isozymes in wheat, maize, barley, and rice seeds (Ahmed et al., 2015; Crofts et al., 2015; Hennen-Bierwagen et al., 2008; Liu et al., 2012a; Liu et al., 2012b; Tetlow et al., 2008; Tetlow et al., 2004). In maize, the trimeric complex composed of SSI, SSIIa, and BEIIb is one of the best studied, and the formation of the trimer depends on protein phosphorylation (Liu et al., 2012a; Liu et al., 2012b; Makhmoudova et al., 2014). Protein-protein interactions have also been found between BEI and BEIIb, as well as in protein complexes consisting of SSIII, pyruvate orthophosphate dikinase (PPDK), BEIIb, and AGPase (Hennen-Bierwagen et al., 2009; Liu et al., 2009). In wheat and maize, Pho1 can interact with SBEI and SBEII (Subasinghe et al., 2014; Tetlow et al., 2004). Moreover, functional and

mutually synergistic relationships between BEs and Pho1, and BEs and SSI, have been observed using recombinant rice starch biosynthetic enzymes *in vitro* (Abe et al., 2014; Nakamura et al., 2014; Nakamura et al., 2012). In addition, there is also the protein-protein interaction between ISA and a carbohydrate-binding module (CBM)-containing protein, FLO6 (Peng et al., 2014). Recently, GBSSI has been reported to bind to the surface of starch granules by interacting with the PTST protein (Seung et al., 2015), and PTST2 also can interact with SSIV to control starch granule initiation in Arabidopsis leaves (Seung et al., 2017).

To date, the majority of studies on multi-enzyme complexes have focused on amylopectin synthesis and are limited to the seeds of major gramineous crop species including wheat, barley, maize, and rice. It remains unknown whether there are similar starch synthase complexes in dicotyledonous species and those with underground storage organs. In this study, we identified a novel subfamily of starch synthases, SSVI, that is present in some eudicot species and also the earliest plants, but not in monocots. We found that in cassava, in spite of being in an inactive form, MeSSVI not only affects starch metabolism by forming protein complexes with other active starch biosynthetic enzymes, but also plays a crucial role in carbon allocation via sugar-induced feedback regulation. In particular, in contrast to the current dogma, our results showed a close coordination between granule-bound starch synthase (GBSSI) and amylopectin biosynthesis enzymes, which implies that the processes of amylose synthesis and amylopectin synthesis are directly interrelated.

Results

The SSVI family is a newly identified cluster of starch synthases found in dicot species but not in monocots

Phylogenetic analysis of the deduced SS protein sequences from cassava and 43 other plant species showed that the starch synthase protein family can be classified into seven categories: the GBSS, SSI, SSII, SSIII, SSIV, and SSV families as well as a new group named the SSVI family (Fig. 1A). Among these seven groups, six had >80% bootstrap support, and one (SSII) had 49.5% bootstrap support. Furthermore, SS proteins from monocot species were not found in the SSVI family because of gene loss events, whereas the other six groups contained proteins from both monocot and dicot species.

In order to estimate the selective pressure on starch synthase (*SS*) genes, the ML method and the codon substitution model were used to evaluate the K_a/K_s ratios (Nielsen and Yang, 1998; Yang, 2000, 2007). In this analysis, we used the site-specific model to identify whether positive selection had acted on the evolution of each *SS* gene family (Table S1). Firstly, a comparison of the M0 and M3 models was used to evaluate whether there were K_a/K_s ratio variations among the codon positions in each subfamily. Overall, the ML estimates of the K_a/K_s ratios for all subfamilies under the model M0 were <1, suggesting that purifying selection was the principal constraint acting on the evolution of *SS* gene subfamilies. However, the LRTs for the comparisons of the M3 and M0

models were statistically significant for all subfamilies, revealing the existence of extreme variation in selective pressure among amino acid sites for all SS subfamilies. Secondly, the LRTs used to compare the data fit to models M2a versus M1a and M7 versus M8 were used to address whether positive selection promoted the divergence of each subfamily. Unexpectedly, none of the seven SS subfamilies was found to be influenced by positive selection. These results indicate that purifying selection contributed to the functional stability of each plant *SS* gene subfamily.

To obtain further insights into the possible structural evolution of the *SSVI* genes in the plant kingdom, the exon-intron structures and conserved domains in these SSVI genes were visualized using the Gene Structure Display Server (GSDS; http://gsds.cbi.pku.edu.cn/) as shown in Figure 1B. The numbers of exons varied between 13 and 17, except for one gene in *Vitis vinifera* that had six exons. The exon positions also varied between the genes. Although there were various exon-intron patterns present in related species, the conserved domains and conserved motifs were found to be highly similar (Fig. 1B, Supplementary Fig. S1). All of the deduced SSVI proteins were predicted to have a single Glyco_transf_5 domain, which is the catalytic domain of starch synthases that use ADP-glucose as the glucose donor. In addition, the majority of these proteins, except the SSVIs from *Citrus sinensis* and *Vitis vinifera*, contained a Glycos_transf_1 domain that can transfer activated sugars to a variety of substrates, including glycogen, fructose-6-phosphate, and lipopolysaccharides. Furthermore, motif 4 was continually detected in positions close to the Glyco_transf_5 domain, indicating that this motif may be essential for the function of the SSVI proteins (Supplementary Fig. S1).

Down-regulated *MeSSVI* gene expression leads to reduced plant growth and altered carbon allocation in cassava

After carefully surveying the cassava genome, 10 genes were defined as cassava *SS* genes, including one SSVI family gene (Table S2). MeSSVI had both the Glyco_transf_5 and Glyco_transf_1 domains, which is the typical pattern of conserved domains and motifs found in SSVI family proteins (Fig. 1B, Supplementary Fig. S1). To further investigate the function of MeSSVI, we first characterized the tissue-specific gene expression patterns of cassava starch synthase genes in the seven SS families (Fig. 1C). Among the 10 cassava starch synthase genes, *MeSSIII-2* and *MeSSV* were almost undetectable in all tissues. In storage roots, *MeSSVI* had the highest transcript level among soluble starch synthase genes, which indicates that *MeSSVI* might play an important role in starch metabolism.

A 200 bp conserved DNA sequence of cassava *MeSSVI* gene (base pairs 978-1177), which encodes the starch synthase catalytic domain, was used as the target sequence for RNAi. After primary growth and phenotypic screening, three *MeSSVI*-RNAi transgenic (SSC) plant lines that showed normal growth and development, together with one empty vector (EV) line, were chosen for further study. Single transgene integrations were detected in all transgenic lines by Southern blot analysis using an *HPT* DIG-labeled probe following digestion of genomic DNA with either

*Hin*dIII (Supplementary Fig. S2, upper panel) or *Eco*RI (Supplementary Fig. S2, lower panel). The results of qRT-PCR (Fig. 2A) and Western blotting (Fig. 2B) analyses suggested that the expression of *MeSSVI* in the storage roots of the SSC lines was significantly decreased compared to the WT and EV controls.

After six months of growth in the field, phenotypic measurements that included plant height, fresh-weight biomass, root number, root length, and root diameter were performed on the WT and transgenic lines. Plant heights ranged from 160.8 cm to 166.6 cm for the SSC lines, which were shorter than both WT and EV line plants (Supplementary Fig. S3A, B). The biomass per plant ranged from 1.39 kg to 2.15 kg for the SSC lines (Supplementary Fig. S3A, C), which was considerably less than the average fresh weight of the WT (4.33 kg) and EV (4.01 kg) plants. The average storage root length was 23.42 cm for WT and 23.33 cm for EV, and it ranged from 19.41 cm to 21.26 cm for the SSC lines (Supplementary Fig. S3A, D). Average root diameter was 2.67 cm for WT, 2.64 cm for EV, and 2.52 cm to 2.88 cm for the SSC lines (Supplementary Fig. S3A, E). The average root number per plant varied from 11.67 to 14 for the SSC lines (Supplementary Fig. S3A, F), which was dramatically less than the numbers for WT (25.1) and EV (25) plants. Thus, for the SSC transgenic lines, total biomass was reduced due to retarded growth of both the aerial and subterranean organs.

More importantly, slices of dried storage roots from the SSC lines were more yellow in color than were slices of WT and EV roots (Supplementary Fig. 2C). To investigate which metabolic pathways were changed significantly in the SSC lines, we conducted RNA-seq analysis of the storage roots of WT and SSC lines grown in the field. KEGG analysis showed that the differentially expressed genes (DEGs) were significantly enriched in genes associated with carbohydrate metabolism and lipid metabolism in the SSC lines (Supplementary Fig. S4). Furthermore, the total lipid content increased five-fold in the storage roots of the SSC lines compared with WT and EV, whereas the starch content was decreased significantly in the SSC lines, although the resistant starch content was increased (Fig. 2D, E and Supplementary Fig. S5B). Protein contents were not altered in roots of the SSC lines (Supplementary Fig. S5A). These results show that MeSSVI is involved in regulating carbon allocation in cassava storage roots.

MeSSVI influences starch properties in cassava storage roots

Microscopic analysis of storage root starch granule morphology by SEM and TEM did not show obvious differences between the WT and SSC transgenic lines. The starch in both the WT and SSC lines consisted of a mixture of round, truncated, and dome-shaped granules (Fig. 3A and Supplementary Fig. S5C). The starch granule size distributions in the SSC transgenic lines differed from those of the WT and EV. The SSC lines exhibited a broader granule size distribution compared to the WT and EV, and contained a greater number of larger starch granules which were largely responsible for the increased average granule size in the SSC lines (Fig. 3B).

The relative proportion of amylose was dramatically increased in the SSC lines compared with starch from the WT and EV (Supplementary Fig. S5D). In the SSC lines, the number of chains with DP 6–9 and DP 24–40 increased, while the number of chains with DP 10–23 decreased (Fig. 3C). The CLD differential patterns in the SSC lines indicated that MeSSVI could play a key role in the synthesis of short to intermediate length amylopectin chains (DP 10-23) in cassava storage roots.

RVA analysis showed that root starches from the SSC lines exhibited different pasting patterns compared to WT and EV starches (Fig. 3D). The pasting temperature (PT) of WT starch was 71.05°C, while all three SSC starches showed increased PTs (Table S3). Overall, the values for peak viscosity (PV), hot paste viscosity (HV), breakdown (BD), final viscosity (FV), and setback (SB) were significantly lower in the SSC lines than in starches from the WT and EV lines. The DSC (differential scanning calorimeter) parameters of the starch samples were determined in order to evaluate the gelatinization properties (Table S4). Starch isolated from WT roots gelatinized with a temperature range of 56.26°C (T_o) to 78.21°C (T_c) and an enthalpy (ΔH) of 11 J/g. The gelatinization temperatures (T_o, T_p, and T_c) and enthalpies (ΔH) were decreased in starch isolated from the SSC transgenic lines, and showed clear differences when compared with starches from the WT and EV lines.

X-ray diffraction (XRD) analysis of normal cassava storage starch showed typical A-type crystals with one doublet around 17° (2θ) and two singlets predominantly around 15° and 23° (2θ) (Fig. 3E)[10]. Compared to storage root starch from WT and EV plants, SSC storage starches underwent pronounced changes in their diffraction patterns, including replacement of a doublet with a singlet emerging around 17° (2θ). The peak intensities around 15°, 17°, and 23° (2θ) of SSC starches were weaker than that of WT and EV starches.

MeSSVI is an inactive SS, but it affects the enzyme activities of AGPase, GBSS, and ISA

In vitro enzyme assays showed that MeSSVI is inactive (Fig. 4A), indicating that the function of MeSSVI could be due to its interaction with other starch biosynthetic enzymes to effect the phenotypic differences between the SSC lines and WT (Supplementary Fig. 3A).

To investigate the mechanisms that cause decreased starch contents in storage roots of the SSC lines, the *in vivo* activities of key enzymes involved in starch synthesis (AGPase) and starch degradation (α-amylase and β-amylase) were determined. The activity of AGPase was found to be slightly increased in the SSC lines compared with WT and EV (Fig. 4B). However, supplementation of the supernatants extracted from the SSC lines with recombinant MeSSVI protein caused a reduction in the soluble AGPase activity in the corresponding lines. No significant differences in the activities of α-amylase and β-amylase were detected between the SSC lines and WT (Fig. 4C, D). These results indicate that increased AGPase activity contributes to the decreased starch content. Additionally, transcription of *MeAGPL* and *MeAGPS* was

up-regulated in the SSC lines, although the protein levels did not differ compared with WT and EV, which implies that the increased AGPase activity is not due to the increase in AGPase gene expression (Fig. 4E-G).

Previous studies of WRINKLED1 in Arabidopsis have reported that increased sugar content can trigger the expression of lipid biosynthesis genes (Zhai et al., 2017). Therefore, we analyzed the mRNA levels of 12 lipid biosynthesis genes and the sugar contents between the transgenic lines and WT (Supplementary Fig. S6A, B). The SSC lines had higher mRNA levels of lipid biosynthesis genes and related transcription factors (Supplementary Fig. S6A) compared to WT, and increases in the amounts of sucrose, fructose, and glucose were also detected in the SSC lines as expected (Supplementary Fig. S6B).

To further uncover how MeSSVI affects starch properties, the relative mRNA levels of six key starch biosynthesis genes were determined by qRT-PCR (Fig. 5A). Although only *MeSSI* and *MeSBEII* transcription was found to be up-regulated in the SSC lines, all six of the enzymes encoded by these genes showed no significant differences in protein levels in the SSC lines compared to the WT and EV (Fig. 5B). These results indicate that the phenotypic differences between the transgenic SSC lines and WT are not due to the expression of these six starch biosynthesis genes. Therefore, the relative activities levels of these starch biosynthesis enzymes (GBSS, SS, SBE, and ISA) were further analyzed. The activity of GBSS was significantly higher in the SSC lines compared with WT and EV (Fig. 5C). Addition of MeSSVI recombinant protein to the supernatants extracted from storage roots of the SSC lines caused a reduction in the soluble GBSS activity of the corresponding lines. The activities of both SSI (Fig. 5D) and SBE (Fig. 5E) showed no significant changes compared with WT and EV, while the activity of ISA was lower in the SSC lines (Fig. 5F). Nevertheless, increased ISA activity was observed in the SSC supernatants to which recombinant MeSSVI protein had been added (Supplementary Fig. S7).

MeSSVI forms protein complexes with other starch biosynthetic enzymes

Peptide-specific antibodies were used in reciprocal co-immunoprecipitation experiments to analyze protein-protein interactions among MeSSVI, MeAGPL and MeAGPS. As shown in Figure 6, antibodies raised to each of these proteins co-precipitated the others (Fig. 6A-C), and these three proteins were detected in the higher molecular weight fraction (Fig. 6D) by gel permeation chromatography (GPC), indicating that MeSSVI may regulate the enzyme activity of AGPase (ADP-glucose pyrophosphorylase) through the formation of protein complexes.

To illustrate the molecular weight distribution patterns of seven key starch biosynthesis enzymes (MeSSI, MeSSVI, MeSBEI, MeSBEII, MeISAI, MeISAII and MeGBSSI), soluble proteins extracted from mature cassava storage roots were fractionated using GPC. All of the seven enzymes were detected in the higher molecular weight fraction (Fig. 7A), which is consistent with the notion that they can form protein complexes.

To further investigate the pattern of starch biosynthesis enzyme complexes, we performed blue-native polyacrylamide gel electrophoresis (BN-PAGE), because this technique maintains protein-protein interactions and accurately reflects the molecular weight of the protein complexes (Eubel et al., 2005). In order to make the molecular weight of the proteins comparable, Western blotting of BN-PAGE was performed by using slices of lanes from a single gel. The result showed that these enzymes can also be present at higher molecular weights on BN-PAGE (Fig. 7B). These results are consistent with the molecular weight distribution patterns detected by GPC (Fig. 7A).

To investigate possible interacting partners among the starch biosynthesis isozymes, immunoprecipitation was performed by using soluble cassava storage root protein extracts and isozyme-specific antibodies (Fig. 7C). The results demonstrated that the pairwise protein interactions obtained by reciprocal co-immunoprecipitation includes MeSSI-MeSSVI, MeSSI-MeSBEI, MeSSI-MeSBEII, MeSSI-MeISAI, MeSSI-MeISAII, MeGBSSI-MeSSI, MeSSVI-MeSBEII, MeSSVI-MeGBSSI, MeSBEI-MeISAI, MeSBEI-MeISAII, MeSBEI-MeGBAII, MeSBEII-MeISAII, MeISAI-MeISAII, MeISAI-MeISAII, MeISAI-MeISAII, and MeISAI-MeISAII. Western blot signals obtained from only one side of the co-immunoprecipitation included AbMeSSVI-MeSBEI, AbMeSSVI-MeISAI, AbMeSSVI-MeISAII, AbMeSBEI-MeSBEII, and AbMeISAII-MeGBSSI (first acronym, antibody used for immunoprecipitation; second acronym, isozyme detected by Western blotting).

In order to further explore the components of the starch biosynthesis isozyme complexes in mature cassava storage roots, protein complexes immunoprecipitated by each antibody were analyzed separately by LC-MS/MS, and the identified proteins are summarized in Table S6. The results showed that MeSSI, MeSSVI, MeSBEI, MeSBEII, MeISAI, MeISAII, and MeGBSSI can be co-immunoprecipitated by each antibody, indicating that these enzymes form protein complexes through specific protein-protein interactions, further confirming the Co-IP experiment results (Fig. 7C). In addition, many other proteins were also detected. Among these proteins, the 14-3-3 protein was co-immunoprecipitated by each antibody. Starch phosphorylase (MePho) can be co-immunoprecipitated with MeSBEI and MeSBEII antibodies, which is consistent with previous reports showing that Pho can interact with SBEI and SBEII (Subasinghe et al., 2014; Tetlow et al., 2004). MePho is also co-immunoprecipitated by antibodies against MeSSI, MeSSVI, MeISAI, MeISAII, and MeGBSSI. Furthermore, each antibody can co-immunoprecipitate with MeAGPase and MeSuS.

To investigate direct protein-protein interaction patterns among these seven enzymes, we performed yeast-two-hybrid (Y2H) assays. We found that MeSSI, MeSSVI, and MeSBEII can interact with one another (Fig. 7D and Supplementary Fig. S8A). MeSSVI can interact with itself and MeISAII directly, and the interaction between MeSSVI and MeISAII was a very interesting finding, indicating that SSs, SBEs and ISAs may assemble into multi-enzymes complexes to perform their functions. Furthermore, we found that MeGBSSI could interact with itself and was

also capable of interacting with MeSSI and MeISAI (Supplementary Fig. S8A). In addition, we detected strong interactions between MeISAI and MeISAII, and MeISAII with itself.

The Y2H results were validated by testing these interactions in bimolecular fluorescence complementation (BiFC) experiments. The cDNAs encoding the above seven enzymes were cloned into BiFC vectors to generate protein fusions with either the N-terminal half of YFP (nYFP) or the C-terminal half of YFP (cYFP). Vector combinations of nYFP-MeSSI+cYFP-MeSSVI, nYFP-MeSSI+cYFP-MeSBEII, nYFP-MeSSI+cYFP-MeGBSSI, nYFP-MeSSVI+cYFP-MeSBEII, nYFP-MeSSVI+cYFP-MeISAII, nYFP-MeISAI+cYFP-MeISAII, and nYFP-MeISAI+cYFP-MeGBSSI were transiently co-expressed in *N. benthamiana* leaves, and strong YFP fluorescence signals were detected in the chloroplasts (Supplementary Fig. S8B), which were consistent with the Y2H results (Supplementary Fig. S8A).

MeSSVI-associated protein complexes are affected by protein phosphorylation and redox potential

To investigate whether the cassava starch biosynthesis enzymes can be regulated by protein phosphorylation, crude protein extracts isolated from cassava storage roots were treated with phosphatase or ATP, and co-immunoprecipitation experiments were performed with these protein extracts to determine whether the protein-protein interactions between these enzymes could be phosphorylation dependent (Fig. 8A). The data showed that phosphorylation (preincubation with ATP) or dephosphorylation (preincubation with PPases) did not affect the ability of any of the antibodies to bind to the respective proteins. Phosphorylation and dephosphorylation had no significant effect on the MeISAI-MeISAII and MeSSI-MeSSVI protein-protein interactions. However, the interactions between MeSSI-MeSBEII, MeSSVI-MeSBEII, and MeSSVI-MeISAII were significantly weaker after pre-incubation with PPase, suggesting that these interactions are phosphorylation dependent.

To investigate the effects of redox potential (DTT or NADP treatment) on the protein-protein interactions among these enzymes, crude proteins extracted from mature cassava roots were incubated with either 10 mM DTT or 10 mM NADP, and co-immunoprecipitation experiments were then performed using these extracts (Fig. 8B). The results showed that DTT and NADP did not affect the ability of any of the antibodies to bind to the respective proteins. DTT and NADP had no significant effects on the MeSSI-MeSBEII, MeSSI-MeSSVI, and MeSSVI-MeSBEII protein-protein interactions. However, the MeISAI-MeISAII and MeSSVI-MeISAII interactions were significantly enhanced after pre-incubation with NADP, suggesting that redox potential affects the formation and/or stability of these interactions.

MeSSVI influences GPC mobility and granule-bound ability of starch biosynthetic enzymes

To investigate whether the absence of MeSSVI would alter the GPC mobility of other members of the complex, soluble proteins extracted from mature roots of the WT and SSC lines were

fractionated using GPC (Fig. 9A). MeSSI and MeSBEII mobility in the GPC column was significantly affected by the absence of MeSSVI. In the WT protein extract, MeSBEII was not detected in the 232 kDa fraction, and the amount of MeSSI present in the 232 kDa fraction was reduced compared with the SSC lines. No obvious changes in the GPC mobility of other proteins (MeSBEI, MeISAI, MeISAII, and MeGBSSI) were detected in the SSC lines compared with WT. The fact that MeSSI and MeSBEII exhibited altered GPC elution volumes when MeSSVI was depleted, compared with WT, is consistent with the proposal that the three proteins associate with one another in the same quaternary structure.

Furthermore, to investigate the effect of MeSSVI on starch granule-bound protein, granule-bound proteins isolated from starch samples extracted from storage roots of WT and the SSC lines were detected by Western blotting (Fig. 9B). MeSSI, MeSBEII, MeISAI, and MeISAII were not detected in the SSC lines, although they were present as granule-bound proteins in WT, whereas no differences were observed in the granule-bound protein profiles for MeSBEI and MeGBSSI between WT and the SSC lines.

Discussion

At present, all members of the starch synthase family in plants are classified into six subfamilies that include GBSS, SSI, SSII, SSIII, SSIV, and SSV based on the earliest available genomic sequences (Yang et al., 2013), and mainly limited to economically important crop species. The functions of five starch synthases, including GBSSI, SSI, SSII, SSIII, and SSIV, have been well-documented in previous studies (Craig et al., 1998; Delvalle et al., 2005; Edwards et al., 1999; Fujita et al., 2006; Fujita et al., 2007; Fulton et al., 2002; Kossmann et al., 1999; McPherson and Jane, 1999; Morell et al., 2003; Nakamura et al., 2005; Roldan et al., 2007; Ryoo et al., 2007; Song and Jane, 2000; Takahata et al., 2010; Wang et al., 2017; Yamamori et al., 2000; Yoo and Jane, 2002; Zhang et al., 2008; Zhao et al., 2011; Zhou et al., 2016). It has been demonstrated that amylopectin biosynthesis isozymes can interact with each other in the endosperms of wheat, maize, barley, and rice seeds (Ahmed et al., 2015; Crofts et al., 2015; Hennen-Bierwagen et al., 2009; Hennen-Bierwagen et al., 2008; Liu et al., 2009; Tetlow et al., 2008; Tetlow et al., 2004). Also, amylose synthase GBSSI is also able to form oligomers in rice, and GBSSI activity and oligomerization are affected by phosphorylation, redox regulation, and ADPG (Liu et al., 2013). Nevertheless, these mechanisms are limited to the seeds of the major gramineous crop species. It is still unclear whether similar starch synthase complexes are present in dicotyledonous species and in underground storage organs. In this study, we identified a novel subfamily of starch synthase, SSVI, which appears to be a lineage that was lost from monocots but was conserved in many eudicot species and the earliest plants based on their genomic/transcriptomic sequences (Figs. 1A, B and Supplementary Fig. S1 and Table S1). By focusing on cassava SSVI, we have elucidated its function and the molecular mechanism by which it chaperones the starch synthase complex to simultaneously regulate carbon allocation and starch biosynthesis in cassava storage

roots (Figs. 2, 3 and Supplementary Fig. S5). Based on our findings and those of previously published studies (Crofts et al., 2015; Hennen-Bierwagen et al., 2009; Zhou et al., 2016), we propose a functional model of MeSSVI that uncovers new features of starch biosynthesis in plants, especially in eudicot species (Fig. 10).

As a newly classified member of the starch synthase family, MeSSVI was shown to be enzymatically inactive in cassava (Fig. 4A, Fig. 5D), which implies that, unlike the published starch synthases, the function of MeSSVI might be not achieved through its catalytic activity. Further investigation showed that MeSSVI is important for growth and development in cassava (Supplementary Fig. S3), and the altered carbon allocation in the SSC lines is due to the increased activity of AGPase, not the starch degradation process (Fig. 4 B-D). A protein complex consisting of MeSSVI and MeAGPase was detected in cassava storage roots (Fig. 6A-D), showing that MeSSVI can inhibit AGPase activity by complexing with other proteins. Previous studies have shown that SSIII may regulate the activity of AGPase by forming protein complexes with AGPase in order to control carbon allocation in rice and maize seeds (Hennen-Bierwagen et al., 2009; Zhou et al., 2016). Indeed, the cassava SSC lines accumulated more sugar in storage roots by reducing starch synthesis. As a result, these plants showed a significant increase in the expression levels of lipid biosynthesis genes in their storage roots, which ultimately resulted in an increase in total lipid content (Fig. 2C,D; Supplementary Fig. S6A, B). In Arabidopsis, when intracellular sugar levels are elevated, KIN 10-dependent degradation of WRI1 provides a homeostatic mechanism that favors lipid biosynthesis due to KIN10 inhibition (Zhai et al., 2017). As such, we propose that loss of MeSSVI disrupts the protein complex consisting of MeSSVI and MeAGPase, and this disruption will reduce the influence on AGPase activity so that more ADPGlc is directed toward glycolytic intermediates to support lipid synthesis. In addition, elevated sugar concentrations would trigger the transcription of lipid biosynthesis genes, which is more conducive to lipid synthesis (Fig. 10).

The starch isolated from storage roots of the SSC lines showed altered physico-chemical properties, such as amylose content, pasting, gelatinization properties, and crystallinity (Fig. 3 and Supplementary Fig. S5), indicating that MeSSVI plays a vital role in starch biosynthesis in cassava. Among these properties, the increased amylose content in roots of the SSC lines is an interesting phenomenon (Supplementary Fig. S5D), and implies that as an amylopectin synthase, MeSSVI has an important influence on amylose synthesis. Analyses of the mRNA and protein expression levels of key enzymes involved in starch synthesis demonstrated that the effect of MeSSVI on starch biosynthesis was not achieved by affecting the expression levels of the genes that encode these enzymes (Fig. 5A, B). Our results show that MeSSVI can regulate the enzymatic activities of GBSSs and ISAs, but not SSs and SBEs, in cassava storage roots (Fig. 5C-F and Supplementary Fig. S7), and SBE can enhance the activity of SS (Boyer and Preiss, 1979; Pollock and Preiss, 1980). Moreover, MeSSVI influenced the ability of MeSSI, MeSBEII, MeISAI ,and MeISAII to bind to starch granules (Fig. 9B), which is similar to the effect of SSIIa on the ability

of SSI and SBEII to bind to starch (Morell et al., 2003; Umemoto and Aoki, 2005; Yamamori et al., 2000), although the presence of a carbohydrate binding module was not predicted in MeSSVI (Fig. 1B). This indicates that MeSSVI may bind to starch granules by interacting with other proteins that contain a CBM domain. Furthermore, obvious changes in the GPC mobility of MeSSI and MeSBEII in the SSC lines also implies the significance of MeSSVI in these protein complexes that consist of starch biosynthesis enzymes (Fig. 9A). Therefore, we propose that owing to the disruption of these protein complexes, a deficiency in MeSSVI decreases ISA activity and increases GBSS activity, leading to a shift in carbon allocation toward amylose biosynthesis. Consequently, the increased levels of amylose and lipids together give rise to an increase in the resistant starch content (Supplementary Fig. S5B), which is consistent with previous results obtained in rice (Zhou et al., 2016). These results also show that in the SSC lines, in order to avoid cell toxicity caused by high sugar concentrations, the storage cells not only fix excess sugars by accelerating lipid biosynthesis, but also inhibit starch degradation through the formation of resistant starch, indicating that MeSSVI provides a homeostatic mechanism to stabilize intracellular sugar levels.

The results of our study also revealed previously-unknown protein interactions among starch biosynthesis enzymes in cassava, a dicot species. Although protein-protein interactions among SSs including GBSSI have been reported (Seung et al., 2015; Zhou et al., 2016), it is still unclear whether GBSSI participates in amylopectin synthesis. Here, we found that MeGBSSI can interact with MeSSI and MeISAI directly (Supplementary Fig. S8), strongly implying that GBSSI is directly involved in the process of amylopectin synthesis. With regard to the function of the interactions between MeGBSSI and MeSSI or MeISAI, GBSSI may interact with amylopectin enzymatic complexes through SSI, and contribute to the synthesis of extra-long amylopectin chains directly, as shown in previous studies in potato and rice (Fulton et al., 2002; Hanashiro et al., 2008). Furthermore, GBSSI may interact with SBEII through SSI to produce more branches in amylose and the extra-long amylopectin chains, leading to the synthesis of more amylopectin clusters. It should also be noted that pea amylose was reported to be debranched completely by isoamylase (ISA) in aqueous 40% dimethyl sulfoxide (Colonna and Mercier, 1984). The interaction between GBSSI and ISAI indicates that ISAI may excise the extra erroneous branched chains in amylopectin in favor of producing long chains or the synthesis of amylose by interacting with GBSSI. In *Chlamydomonas reinhardtii*, amylose can be formed by extending the branches on amylopectin, which are then removed from the amylopectin chains (van de Wal et al., 1998). The interaction between GBSSI and the amylopectin enzymatic complexes through SSI may play a role in regulating the activity of one another, which is consistent with the enhanced enzyme activity of GBSSs in the cassava SSC lines (Fig. 5C). In wheat, SSI, SSII, SBEIIa, or SBEIIb can form heterotrimers after phosphorylation of SBEIIa or SBEIIb, which increases the enzymatic activity of SBEIIa and SBEIIb significantly (Tetlow et al., 2008). In addition, PHOI can improve the activity of SBE by interacting with SBE in rice (Crofts et al., 2015).

To date, a direct interaction between ISAs and the protein complexes that consist of SS and SBE has not been reported. It is clear that the removal of erroneous branches by ISAs plays an important role in the extension and branching of amylopectin chains, and, therefore, protein interactions with ISAs are essential (Delatte et al., 2005; Hussain et al., 2003; Kubo et al., 2010; Peng et al., 2014; Utsumi et al., 2011). Lin *et al.* (2012) hypothesized that ISAI may interact with SSIII, either directly or indirectly, so that ISAI can bind to the growing amylopectin molecules. The interaction between ISA and a CBM-containing FLO6 protein shows that ISA could potentially bind to starch granules via FLO6 (Peng et al., 2014). In cassava, we found that MeSSVI can interact with MeISAII directly (Supplementary Fig. S8), suggesting that MeISAI and MeISAII are part of the protein complexes that consist of SS and SBE via their interactions with MeSSVI, and hence, contributing directly to amylopectin synthesis by removing the erroneous branches. Our results show that MeSSVI may increase the activities of the ISA enzymes through direct protein-protein interactions in cassava (Fig. 5F and Supplementary Fig. S7).

In this study, we have confirmed the presence of multienzyme complexes for starch biosynthesis in cassava storage roots. Our results showed that except for the protein-protein interactions MeISAI-MeISAII and MeSSI-MeSSVI, the interactions MeSSI-MeSBEII, MeSSVI-MeSBEII, and MeSSVI-MeISAII were phosphorylation dependent (Fig. 8A). In cereal endosperm development, it has been found that the formation of multienzyme complexes is dependent on protein phosphorylation (Hennen-Bierwagen et al., 2008; Lin et al., 2012; Tetlow et al., 2008; Tetlow et al., 2004). ATP can promote the formation of high molecular weight multienzyme complexes, and phosphatase treatment leads to the depolymerization of protein complexes (Lin et al., 2012; Liu et al., 2009). AGPase, GBSS, SSIIa, SBEIIb, BEIIa, SBEI, and PHO1 have been shown to be phosphorylated by direct biochemical analysis (Grimaud et al., 2008; Subasinghe et al., 2014; Tetlow et al., 2004; Walley et al., 2013). Moreover, disulfide linkages and hydrogen bonds also affect the formation of protein complexes (Chothia, 1975). In cassava, DTT and NADP affected the formation of the MeISAI-MeISAII and MeSSVI-MeISAII interactions, and had no obvious effect on the MeSSI-MeSBEII, MeSSI-MeSSVI, and MeSSVI-MeSBEII interactions (Fig. 8B), implying that redox potential plays a significant role in regulating protein-protein interactions between the amylopectin biosynthesis enzymes, which has been shown previously. In Arabidopsis, the enzymatic activity of AGPase is regulated by redox modification, and AGPase activity is activated by opening intermolecular disulfide bonds (Tetlow et al., 2004). In rice, the activity of GBSSI decreased after DTT treatment and increased after NADP treatment, and the percentage of GBSSI monomers increased after DTT treatment and decreased after NADP treatment (Liu et al., 2013). All these findings indicate that in cassava, protein phosphorylation and redox modulation are important post-translational modifications that affect the formation of multienzyme complexes, and play a key role in starch synthesis by regulating the activities of these enzymes.

In conclusion, we have identified a novel subfamily of starch synthases, SSVI, and elucidated the function of MeSSVI in starch biosynthesis and carbon allocation in cassava storage roots. Our

work describes a new mechanism of starch accumulation in underground storage organs of a dicot species. Contrary to current dogma, our results suggest that starch synthesis requires a close coordination between GBSSI and amylopectin biosynthetic enzymes, and that amylose synthesis and amylopectin synthesis are directly interrelated. Further molecular investigations into the regulatory mechanisms of starch biosynthesis in cassava, such as the transcription factors that regulate the expression of the genes that encode these enzymes, will improve our understanding of the mechanisms governing starch accumulation in cassava storage roots. Additionally, this work provides a novel source of cassava starch with a high resistant starch content for various industrial applications.

Materials and methods

Phylogenetic tree analysis

The genomes of 44 plant species were sampled for starch synthase superfamily sequences with assignments to the PF08323 Pfam hidden Markov models (HMM), including the algal species *Chlamydomonas reinhardtii*, *Coccomyxa subellipsoidea*, and *Volvox carteri*; the mosses *Marchantia polymorpha*, *Physcomitrella patens*, *Sphagnum fallax*, *Selaginella moellendorffii*; the basal angiosperm *Amborella trichopoda*; the eudicots *Arabidopsis thaliana*, *Aquilegia coerulea* (Colorado blue columbine), *Capsella rubella* (pink shepherd's purse), *Citrus sinensis* (orange), *Cucumis sativus* (cucumber), *Eucalyptus grandis* (rose gum), *Glycine max* (soybean), *Gossypium raimondii* (New World cotton), *Manihot esculenta* (cassava), *Mimulus guttatus* (monkeyflower), *Populus trichocarpa* (poplar), *Prunus persica* (peach), *Ricinus communis* (castor bean), *Solanum lycopersicum* (tomato), *Vitis vinifera* (grape); the monocots *Ananas comosus* (pineapple), *Brachypodium distachyon*, *Hordeum vulgare* (barley), *Musa acuminata* (banana), *Oryza sativa* (rice), *Panicum virgatum* (switchgrass), *Setaria italica* (foxtail millet), *Sorghum bicolor* (sorghum or great millet), *Spirodela polyrhiza* (duckweed), *Zea mays* (maize), and *Zostera marina* (eelgrass). Additionally, 10 monocot transcriptomes were retrieved from the National Center for Biotechnology Information Sequence Read Archive (Leinonen et al., 2011), including *Lachnocaulon anceps* (SRX1639018), *Flagellaria indica* (SRX1639020), *Stegolepis ferruginea* (SRX1639022), *Elegia fenestrata* (SRX1639024), *Cyperus alternifolius* (umbrella papyrus; SRX1639031), *Juncus effusus* (common rush; SRX1639021), *Mayaca fluviatilis* (SRX1639017), *Neoregelia carolinae* (SRX1639025), S*treptochaeta angustifolia* (SRX1639030), and *Typha latifolia* (bulrush; SRX1639026).

Monocot transcript sequences were assembled using the Trinity RNA assembly program (Haas et al., 2013). Translated predicted transcript isoforms were searched against the PF08323 Pfam HMMs using hmmsearch (Finn et al., 2011). For the GBSS, SSI, SSII, SSIII, SSIV, SSV, and SSVI families from the fully sequenced genomes, the Pfam (Finn et al., 2016) HMMs for Glyco_transf_5 PF08323 were used with hmmsearch. We aligned sequences from fully sequenced plants and translated predicted transcript isoforms with MUSCLE (Edgar, 2004) using the default

parameters that are tuned for accuracy. BMGE (Criscuolo and Gribaldo, 2010) was then used with reduced stringency parameters to strip the alignment to biologically meaningful positions with a permitted gap rate of 0.7, BLOSUM30, and a block size of 2. Best-fit models of sequence evolution for the final alignments were selected using ProtTest (Darriba et al., 2011; Lanfear et al., 2017; Whelan et al., 2015). ML phylogenies of the PF08323 sequences from the fully sequenced genomes and the translated predicted transcript isoforms were reconstructed from amino acid sequences, using RAxML version 8.2 (Stamatakis, 2014). Bootstrap values were obtained with 1,000 replications. Notung (version 2.9) reconciliation of species and gene trees for the starch synthase family was used to analyze histories of gene duplications and losses in these clades (Stolzer et al., 2012). The codeml program in PAML 4.5 (Yang, 2007) was then used to calculate the K_a/K_s ratios using the nucleotide alignment and corresponding phylogenetic tree as inputs. We detected variation in K_a/K_s among sites using likelihood ratio tests (LRTs), and the models tested were M3, M2a, and M8 versus the null models M0, M1a, and M7, respectively. To identify the distributions of conserved amino acid motifs within the SSVI proteins, the conserved motifs were detected with MEME Suite (Bailey et al., 2009) based on the SSVI protein sequences that were used in the phylogenetic analysis with the default settings, except that the maximum number of motifs to find was defined as 5 and the maximum width was set to 300.

Cloning of the *MeSSVI* genes and production of *MeSSVI* RNAi transgenic cassava

Sequence information for *MeSSVI* was derived from the cassava genome database (http://www.phytozome.net). The full length cDNA of *MeSSVI* was cloned from a cassava cDNA library. The binary RNAi gene silencing expression vector p35S::*MeSSVI*-RNAi was constructed from the plasmid pRNAi-dsAC1 (Vanderschuren et al., 2009). The AC1 sequence was excised and replaced with a partial *MeSSVI* cDNA sequence from 978 bp to 1177 bp. The control construct was pRNAi-dsAC1 without the *MeSSVI* cDNA. The two constructs were introduced into *Agrobacterium tumefaciens* LBA4404 and the cassava cultivar TMS60444 was used to generate transgenic plants as described previously by Zhang *et al.* (2000). The integration pattern of T-DNA in transgenic cassava was determined by Southern blot analysis as described by Zhao *et al.* (2011). The DIG-labeled hygromycin phosphotransferase (HPT) gene was used as the hybridization probe.

The resulting transgenic lines and wild type (WT) plants were propagated *in vitro* and then transferred to pots in the greenhouse for macro-propagation (16 h/8 h of light/dark, 30°C/22°C day/night). The *MeSSVI*-RNAi transgenic lines are here given the designation SSC, which is an abbreviation for the "starch synthase catalytic domain" used for constructing the p35S::*MeSSVI*-RNAi vector, and the control lines are designated EV, which is the abbreviation for "empty vector". For field evaluation, 10 stems per transgenic line and WT were planted in early May in the Wushe Plantation for Transgenic Crops, Shanghai, China (31°13948.0099 N, 121°28912.0099E), and harvested in early November. The phenotypic performance of the plants

was recorded.

Transcriptional and translational expression analysis

Leaves, petioles, stems, and storage roots from at least three plants per line were harvested and then ground to a fine powder in liquid N_2 for mRNA extraction. To quantify the expression of genes, qRT-PCR was performed as described (Xu et al., 2013). The *β-Actin* gene was used as a reference to normalize gene expression. The sequences of the PCR primers used in these experiments are shown in Table S6. For RNA-seq, extraction of total RNA from WT and SSC transgenic lines and library preparation from each sample were performed as described previously (Garg et al., 2016). All four libraries (two samples with two biological replicates per sample) were sequenced on an Illumina HiSeq 2000 instrument to generate sequence reads. The raw sequence data were assessed using various quality parameters, and high-quality reads were filtered using the NGS QC Toolkit (v2.3) (Patel and Jain, 2012). The filtered high-quality reads were mapped onto the cassava genome (v7.0) using Tophat (v2.0.0) with the default parameters. The mapped output was processed via Cufflinks (v2.0.2) to obtain FPKM values for all cassava genes in each sample. The differential expression of genes between the different samples was determined by Cuffdiff. The genes that showed a difference of at least two-fold in expression with corrected *P*-values after adjusting for the false discovery rate (*q*-value) <0.05 were considered to be significantly differentially expressed. Furthermore, pathway enrichment analysis of differentially expression genes was performed using clusterProfiler in the R package. The RNA-seq data are available in the NCBI Sequence Read Archive under accession number PRJNA590285.

Protein extraction was performed as described by Crofts *et al.* (2015) with some modifications. Mature cassava storage roots (diameter >2.5 cm; 1 g samples) were used for each extraction. Total protein was extracted with 1 vol (w/v) of a denaturing buffer containing 0.125 M Tris-HCl (pH 6.8), 8 M urea, 4% SDS, and 5% β-mercaptoethanol. Samples were extracted overnight at room temperature, centrifuged at 12,000 rpm to remove gelatinized starch and other particulate matter, and the supernatants were used for denaturing polyacrylamide gel electrophoresis (SDS–PAGE) and immunoblotting. Soluble proteins were extracted on ice with 6 vols (w/v) (three repeats with 2 vols each) of an extraction buffer containing 10 mM HEPES-KOH (pH 7.5) and 100 mM NaCl. After extraction, samples were centrifuged at 12,000 rpm at 4°C for 15 min. The residual pellet was extracted with 1 vol (w/v) of denaturing buffer as described above, centrifuged as before, and the supernatant was then used to represent the insoluble, starch granule-associated proteins.

Proteins were transferred to polyvinylidene fluoride (PVDF) membranes after SDS–PAGE, Native-PAGE, or blue native (BN)-PAGE. Membranes were treated as described by Crofts *et al.* (2015). Polyclonal antibodies against MeSSI, MeSSVI, MeSBEI, MeSBEII, MeISAI, MeISAII, MeGBSSI, MeAGPL (ADP-glucose pyrophosphorylase large subunit), and MeAGPS (ADP-glucose pyrophosphorylase small subunit) were prepared by ABclonal Biotech Co., Ltd using purified recombinant proteins (purity >95%) as antigens. Primary antibodies were used at a

dilution of 1:1,000. HRP-conjugated goat anti-rabbit IgG (Sigma) was used as a secondary antibody at a dilution of 1:10,000. The colorimetric signal was detected with the Pierce ECL Western Blotting Substrate Reagent Kit (catalog no. 32106; Thermo Fisher) and was visualized using the Tanon-5500 Chemiluminescent Imaging System (Tanon Science and Technology, Shanghai, China).

Physico-chemical property assays of cassava storage starches

The granule size distribution of starch extracted from cassava storage roots was determined as described by Zhou *et al.* (2015). A Mastersizer 2000 laser diffraction instrument (Malvern Instruments Ltd., Worcestershire, UK) was used in wet-well mode. The starch was added to the reservoir and sonicated for 30 s at 6 W until an obscuration value of 12–17% was reached. The refractive indices used for the water and starch were 1.330 and 1.50, respectively. The particle size distributions are presented as the diameter versus volume[94]. To examine whether starch granule morphology was altered in the different transgenic lines, the starch granules were coated with gold after they were spread on a metal stub and then observed under a scanning electron microscope (SEM, JSM6360lv, JEOL, Tokyo, Japan). Pieces of fresh mature leaves (2 mm^3) were coated with 1,3-diformal propane and subjected to ultra-thin sectioning. The sections were then photographed under a transmission electron microscope (TEM; Hitachi H7650, Tokyo, Japan).

The chain length distribution (CLD) of the starches was measured as described by Zhou *et al.* (2015). Starch from the transgenic plants and WT was digested with *Pseudomonas amyloderamosa* isoamylase (Sigma), and the CLD of the amylopectin was analyzed using high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD; Dionex-ICS 3000; Dionex Corporation, USA). A D8 Advance Bruker X-ray diffractometer (Bruker AXS, Germany) was used to study the crystal type of the starch granules. Starch powders were scanned with the 2θ diffraction angle at 4–50°. The crystal patterns and degree of crystallinity were analyzed using Jade 5.0 software.

The thermal properties of the starch samples were determined using a differential scanning calorimeter (DSC Q2000; TA Instrument Ltd., UK). A mixture of starch and distilled water (w/w, 1:3) was sealed in an aluminum pan and scanned at 10°C/min from 30°C to 95°C with an empty aluminum pan as a reference. The gelatinization temperature and enthalpy were calculated using Universal Analysis 2000 software. The pasting properties of the starch were analyzed using a rapid viscosity analyzer (model RVA-Super 3; Newport Scientific Pty. Ltd., Australia). The starch was suspended in distilled water (5% w/v, dry weight basis) and tested using a dedicated program for cassava. The temperature of the starch slurry was increased from 30°C to 95°C at a rate of 5°C/min and held at 95°C for 6 min, followed by cooling to 50°C at the same rate, and the temperature was maintained at 50°C for 10 min. The rotating speed of the paddle was 960 rpm during the first 10 s, after which it was kept constant (160 rpm) throughout the analysis.

The amylose content of the starch was measured using colorimetric amylose content determination as previously described (Knutson and Grove, 1994). Amylose (Type III, Sigma A0512, St. Louis, MO) and amylopectin (Sigma 10118) from potato were used to establish standard curves.

Analysis of sugar, total lipid, starch, and resistant starch contents

Storage roots of six-month-old transgenic and WT cassava were harvested from three independent plants and then baked at 80°C for two days to obtain a constant dry weight (DW). The same types of samples from each plant line were mixed and ground together. The dried samples (30 mg) were dissolved in 0.7 ml of 80% ethanol, thoroughly mixed by vortexing, and incubated at 70°C for 2 h. Aliquots of 0.7 ml HPLC-grade water and 0.7 ml chloroform were then added to the samples. After shaking several times, the mixtures were centrifuged at 12,000 × *g* for 10 min. The pellets were collected for starch content analysis while the supernatants contained the soluble sugars. The starch pellets were washed three times with 80% ethanol, and total starch contents were analyzed using the Total Starch Assay Kit (Megazyme, Wicklow, Ireland). A 0.7 ml aliquot of each aqueous supernatant was transferred to a 1.5 ml micro-centrifuge tube and mixed with 0.7 ml chloroform. After centrifugation at 12,000 × *g* for 10 min, 0.5 ml of the supernatant was transferred to a glass tube for HPLC analysis of each sugar component. The sugar-separation method used was slightly modified from the original protocol described by the manufacturer; the Agilent technologies HPLC column (ZORBAX Carbohydrate column; 4.6 × 150 mm, 5 μm) was then used with a differential refraction detector. The mobile phase consisted of 75% acetonitrile with a flow rate of 0.8 ml/min and the temperature of the column was maintained at 35°C. The sugar types were identified based on the retention times of the standards, and the sample concentration was calculated from the external standard curve. Data was calculated from three biological replicates. Resistant starch contents were analyzed using the Resistant Starch Assay Kit (Megazyme, Wicklow, Ireland). Total lipid contents were measured according to the National Standards of the China, GB/T 5512-2008. Protein contents were determined by the Bradford method according to the manufacturer's protocol (Zoman, Beijing).

Enzyme activity and regulation assays

The activities of SS, GBSS, and AGPase in the storage roots of the transgenic and WT plants were assayed as described by Nakamura *et al.* (1989). SS-native-PAGE/activity staining was performed as described (Fujita et al., 2006; Nishi et al., 2001). DBE native-PAGE/activity staining and BE native-PAGE/activity staining were performed as described (Fujita et al., 1999). The activities of α-amylase and β-amylase were determined using α-amylase and β-amylase activity kits (Megazyme, Wicklow, Ireland).

The treatments of ATP or PPase were performed as described by Tetlow *et al.* (2004). Root lysates (1.2 to 1.5 mg protein cm^{-3}) were either preincubated with 1 mM ATP or with 10 units of PPase for 20 min at 25°C, and immunoprecitated. The root lysates without treatment were used as controls.

After immunoprecipitation (IP) with protein-specific antibodies, immunoprecipitated proteins were separated by SDS-PAGE, electroblotted onto polyvinylidene fluoride (PVDF) membranes, and developed with various antisera. PPase is composed of APase (catalog number P-0762; Sigma-Aldrich), alkaline phosphatase from bovine intestinal mucosa (type VII-L, catalog number P-6772; Sigma-Aldrich), and λ protein phosphatase (λ-Pase from *E. coli*; New England Biolabs, Beverly, MA). DDT or NADP treatments were performed as described by Liu *et al.* (2013). Root lysates (1.2 to 1.5 mg protein cm^{-3}) were either preincubated with 10 mM DTT or with 10 mM NADP for 30 min at 25°C, and immunoprecitated.

Analysis of protein interactions

Samples of cassava storage roots (1.5 g) were extracted in 1.5 ml of gel filtration buffer containing 50 mM Tris-acetate (pH 7.5), 100 mM NaCl, 10 µl/ml plant protease inhibitor cocktail (Sigma-Aldrich LLC) and centrifuged at 12,000 rpm for 15 min. The supernatants were injected into a 500 µl sample loop, and fractionated by gel permeation chromatography (GPC) using a Superdex 200 10/300 GL column (GE Healthcare, catalog no. 17-5175-01) connected to an AKTA FPLC system (GE Healthcare) at 4°C. The column was equilibrated with 50 mM Tris-acetate buffer (pH 7.5) containing 100 mM NaCl; the flow rate was 1 ml/min and fraction sizes of 0.5 ml were collected. Molecular mass standards run under identical conditions were from GE Healthcare (catalog no. 17-0445-01). Samples from each fraction (40 µl) were analyzed by SDS-PAGE and immunoblotting.

The BN-PAGE procedures used in this study were performed as previously described (Crofts et al., 2015) with some modifications. Cassava storage roots were extracted with 1 vol (w/v) of 50 mM Bis-Tris (pH 7.5), 50 mM NaCl, 10% glycerol, 0.001% Ponceau S, and centrifuged at 12,000 × *g* for 10 min. The supernatants were subjected to 3–12% acrylamide Bis-Tris native-PAGE (Invitrogen, catalog no. BN1001BOX) and electrophoresed with an anode buffer containing 50 mM Bis-Tris and 50 mM tricine, and a cathode buffer containing 50 mM Bis-Tris, 50 mM tricine and 0.004% CBB G-250 stain at 80 V for an initial 1 h and at 120 V for the remaining time.

Co-immunoprecipitation experiments were performed using the methods described by Liu *et al.* (2012a). Purified peptide-specific antibodies (each approximately 2 µg) were individually used for the co-immunoprecipitation (Co-IP) experiments with the soluble protein extracts described above (1 ml, between 0.8 and 1.2 mg/ml protein). The protein extract/antibody mixtures were incubated at room temperature on a rotator for 50 min, and immunoprecipitation of the antibody/protein complexes was performed by adding 50 µl of Protein A-Sepharose (Sigma-Aldrich) made up as a 50% (w/v) slurry with phosphate buffered saline (PBS; 137 mM NaCl, 10 mM Na_2HPO_4, 2.7 mM KCl, 1.8 M KH_2PO_4, pH 7.4) at room temperature for 40 min. The Protein A-Sepharose/antibody/protein complexes were recovered by centrifugation at 2,000 × *g* for 5 min at 4°C in a refrigerated microcentrifuge, and the supernatants were discarded. The pellets were washed five times with PBS (1.3 ml each), followed by washing five times with a buffer

containing 10 mM HEPES-NaOH (pH 7.5) and 150 mM NaCl. The washed pellets were boiled in SDS-containing gel loading buffer and separated by SDS-PAGE, followed by immunoblot analysis. In order to exclude the possibility that co-immunoprecipitation of the proteins observed in the immunoprecipitation pellet was a result of AGPL, AGPS, SS, SBE, ISA or GBSS binding to the same glucan chain, the soluble protein extracts used for immunoprecipitation were pre-incubated with 5 U each of the glucan degrading enzymes amyloglucosidase (EC 3.2.1.3, Sigma A7255) and α-amylase (EC 3.2.1.1, Sigma A2643) for 20 min at 25°C.

Yeast two-hybridization (Y2H) assays were performed according to the manufacture's procedure (Clontech Laboratories Inc., Palo Alto, CA, USA). For each isozyme, a DNA fragment that was PCR-amplified using the specific primer pair (Table S1) was cloned in-frame into pGBKT7 to form the bait construct. The same fragment was inserted into pGADT7 to form the prey construct. Both plasmids were transformed into yeast cells (strain AH109). The resulting yeast transformants were grown on SD medium containing 3-AT and X-α-gal, but lacking Trp, Leu, His, and Ade, for 2-3 days at 30°C.

Bimolecular fluorescence complementation was conducted as previously described (Ma et al., 2015). The primers used in these experiments are shown in Table S6.

Recombinant protein expression and purification

Recombinant MeSSI and MeSSVI proteins were expressed in *E. coli* and purified as follows: pET28a-MeSSI and pET28a-MeSSVI were constructed by inserting the *MeSSI* and *MeSSVI* cDNA fragments individually into pET28a. The inserted cDNA fragments were fused in-frame with a His-tag sequence. The fused proteins were expressed in *E.coli* BL21 (DE3). The induction conditions were as follows: 0.2 mM IPTG, 16°C, for 16 h. The cells were collected by centrifugation, washed in water, suspended in 20 ml of binding/wash buffer (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% Triton X-100, 5 mM DTT) supplemented just prior to use with 1 mM PMSF and 0.1 mg/ml lysozyme (Sigma), and incubated at 30°C for 15 min. The cell suspensions were sonicated on ice with six bursts of 20 s each, then centrifuged at 12,000 × *g* at 4°C for 20 min. The supernatants were filtered through a 0.45 μm nitrocellulose filter, and additional DTT was added to bring the final concentration to 10 mM. The His-tag fusion proteins were purified from the lysates using Nickel Magnetic Agarose Beads (GE Healthcare) according to the manufacturer's protocol. The primers used in these experiments are shown in Table S6.

Mass spectrometry

The proteins were digested in the gel with bovine trypsin as previously described (Wang et al., 2007). After digestion, the peptides were collected and detected by LC-MS/MS according to the facility procedures used at the Public Mass Spectrometry Technical Service Center, Institute of Plant Physiology & Ecology, Chinese Academy of Sciences. Data were analyzed using Proteome Discoverer software.

Statistical analysis

Root samples were collected from three independent plants per line and then mixed together for further analyses. Data from at least three biological replicates are presented as mean ± SD. Analysis of independent samples with Student's *t*-test was performed using SPSS software, version 17 (SPSS Inc., Chicago, IL, USA). An alpha value of $P < 0.05$ was considered to be statistically significant.

Acknowledgements

We thank Xinyan Liu and Chuanzhong Li for assistance in field experiments; Xiaoyan Gao, Zhiping Zhang, and Jiqin Li for assistance with the SEM and TEM experiments; and Yuanhong Shan for assistance with LC-MS/MS analysis. This work was supported by the National Science Foundation of China (No. 31871682), the National Key Research and Development Program of China (Nos. 2017YFA0505500, 2019YFD1001100), the Earmarked Fund for China Agriculture Research System (CARS-11-shzp), and Shanghai Municipal Science and Technology Major Project (No. 2017SHZDZX01).

Author contributions

S.H. performed most of the experiments, analyzed the data, and drafted the manuscript; X.H., S.W., and W.Z. conducted some of the experiments; S.W. and Q.M. produced the transgenic cassava plants; X.H., Q.M. and X.L. maintained the transgenic cassava plants; P.Z., L.C., S.H. designed and conceived this project; P.Z. and L.C. analyzed the data and revised the manuscript with input from the other authors. All authors contributed to and approved the final manuscript.

References

Abe, N., Asai, H., Yago, H., Oitome, N.F., Itoh, R., Crofts, N., Nakamura, Y., and Fujita, N. (2014). Relationships between starch synthase I and branching enzyme isozymes determined using double mutant rice lines. BMC Plant Biol. 14:80.

Ahmed, Z., Tetlow, I.J., Ahmed, R., Morell, M.K., and Emes, M.J. (2015). Protein-protein interactions among enzymes of starch biosynthesis in high-amylose barley genotypes reveal differential roles of heteromeric enzyme complexes in the synthesis of A and B granules. Plant Sci. 233:95-106.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 37:W202-208.

Boyer, C.D., and Preiss, J. (1979). Properties of citrate-stimulated starch synthesis catalyzed by starch synthase I of developing maize kernels. Plant Physiol. 64:1039-1042.

Bull, S.E., Seung, D., Chanez, C., Mehta, D., Kuon, J.E., Truernit, E., Hochmuth, A., Zurkirchen, I., Zeeman, S.C., Gruissem, W., et al. (2018). Accelerated ex situ breeding of GBSS- and PTST1-edited cassava for modified starch. Sci. Adv. 4:eaat6086.

Ceballos, H., Sanchez, T., Morante, N., Fregene, M., Dufour, D., Smith, A.M., Denyer, K.,

Perez, J.C., Calle, F., and Mestres, C. (2007). Discovery of an amylose-free starch mutant in cassava (*Manihot esculenta* Crantz). J. Agric. Food Chem. 55:7469-7476.

Chothia, C. (1975). Structural Invariants in Protein Folding. Nature 254:304-308.

Cock, J.H. (1982). Cassava: a basic energy source in the tropics. Science 218:755-762.

Colonna, P., and Mercier, C. (1984). Macromolecular structure of wrinkled-pea and smooth-pea starch components. Carbohydr. Res. 126:233-247.

Craig, J., Lloyd, J.R., Tomlinson, K., Barber, L., Edwards, A., Wang, T.L., Martin, C., Hedley, C.L., and Smith, A.M. (1998). Mutations in the gene encoding starch synthase II profoundly alter amylopectin structure in pea embryos. Plant Cell 10:413-426.

Criscuolo, A., and Gribaldo, S. (2010). BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol. Biol. 10:210.

Crofts, N., Abe, N., Oitome, N.F., Matsushima, R., Hayashi, M., Tetlow, I.J., Emes, M.J., Nakamura, Y., and Fujita, N. (2015). Amylopectin biosynthetic enzymes from developing rice seed form enzymatically active protein complexes. J. Exp. Bot. 66:4469-4482.

Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2011). ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 27:1164-1165.

Delatte, T., Trevisan, M., Parker, M.L., and Zeeman, S.C. (2005). Arabidopsis mutants *Atisa1* and *Atisa2* have identical phenotypes and lack the same multimeric isoamylase, which influences the branch point distribution of amylopectin during starch synthesis. Plant J. 41:815-830.

Delvalle, D., Dumez, S., Wattebled, F., Roldan, I., Planchot, V., Berbezy, P., Colonna, P., Vyas, D., Chatterjee, M., Ball, S., et al. (2005). Soluble starch synthase I: a major determinant for the synthesis of amylopectin in *Arabidopsis thaliana* leaves. Plant J. 43:398-412.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792-1797.

Edwards, A., Fulton, D.C., Hylton, C.M., Jobling, S.A., Gidley, M., Rossner, U., Martin, C., and Smith, A.M. (1999). A combined reduction in activity of starch synthases II and III of potato has novel effects on the starch of tubers. Plant J.17:251-261.

Eubel, H., Braun, H.P., and Millar, A.H. (2005). Blue-native PAGE in plants: a tool in analysis of protein-protein interactions. Plant Methods 1:11.

Finn, R.D., Clements, J., and Eddy, S.R. (2011). HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 39: W29-37.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., et al. (2016). The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 44: D279-285.

Fujita, N., Kubo, A., Francisco, P.B., Jr., Nakakita, M., Harada, K., Minaka, N., and Nakamura, Y. (1999). Purification, characterization, and cDNA structure of isoamylase from developing endosperm of rice. Planta 208:283-293.

Fujita, N., Yoshida, M., Asakura, N., Ohdan, T., Miyao, A., Hirochika, H., and Nakamura, Y. (2006). Function and characterization of starch synthase I using mutants in rice. Plant Physiol. 140:1070-1084.

Fujita, N., Yoshida, M., Kondo, T., Saito, K., Utsumi, Y., Tokunaga, T., Nishi, A., Satoh, H., Park, J.H., Jane, J.L., et al. (2007). Characterization of SSIIIa-deficient mutants of rice: the function of SSIIIa and pleiotropic effects by SSIIIa deficiency in the rice endosperm. Plant Physiol. 144:2009-2023.

Fulton, D.C., Edwards, A., Pilling, E., Robinson, H.L., Fahy, B., Seale, R., Kato, L., Donald, A.M., Geigenberger, P., Martin, C., et al. (2002). Role of granule-bound starch synthase in determination of amylopectin structure and starch granule morphology in potato. J. Biol. Chem. 277:10834-10841.

Garg, R., Shankar, R., Thakkar, B., Kudapa, H., Krishnamurthy, L., Mantri, N., Varshney, R.K., Bhatia, S., and Jain, M. (2016). Transcriptome analyses reveal genotype- and developmental stage-specific molecular responses to drought and salinity stresses in chickpea. Sci. Rep. 6:19228.

Grimaud, F., Rogniaux, H., James, M.G., Myers, A.M., and Planchot, V. (2008). Proteome and phosphoproteome analysis of starch granule-associated proteins from normal maize and mutants affected in starch biosynthesis. J. Exp. Bot. 59:3395-3406.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., et al. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8:1494-1512.

Hanashiro, I., Itoh, K., Kuratomi, Y., Yamazaki, M., Igarashi, T., Matsugasako, J.I., and Takeda, Y. (2008). Granule-bound starch synthase I is responsible for biosynthesis of extra-long unit chains of amylopectin in rice. Plant Cell Physiol. 49:925-933.

Hennen-Bierwagen, T.A., Lin, Q., Grimaud, F., Planchot, V., Keeling, P.L., James, M.G., and Myers, A.M. (2009). Proteins from multiple metabolic pathways associate with starch biosynthetic enzymes in high molecular weight complexes: a model for regulation of carbon allocation in maize amyloplasts. Plant Physiol. 149:1541-1559.

Hennen-Bierwagen, T.A., Liu, F., Marsh, R.S., Kim, S., Gan, Q., Tetlow, I.J., Emes, M.J., James, M.G., and Myers, A.M. (2008). Starch biosynthetic enzymes from developing maize endosperm associate in multisubunit complexes. Plant Physiol. 146:1892-1908.

Hizukuri, S. (1985). Relationship between the distribution of the chain-length of amylopectin and the crystalline-structure of starch granules. Carbohydr. Res. 141:295-306.

Hizukuri, S. (1986). Polymodal distribution of the chain lengths of amylopectins, and its significance. Carbohydr. Res. 147:342-347.

Hussain, H., Mant, A., Seale, R., Zeeman, S., Hinchliffe, E., Edwards, A., Hylton, C., Bornemann, S., Smith, A.M., Martin, C., et al. (2003). Three isoforms of isoamylase contribute different catalytic properties for the debranching of potato glucans. Plant Cell 15:133-149.

Jeon, J.S., Ryoo, N., Hahn, T.R., Walia, H., and Nakamura, Y. (2010). Starch biosynthesis in cereal endosperm. Plant Physiol. Biochem. 48:383-392.

Knutson, C.A., and Grove, M.J. (1994). Rapid method for estimation of amylose in maize starches. Cereal Chem. 71:469-471.

Kossmann, J., Abel, G.J.W., Springer, F., Lloyd, J.R., and Willmitzer, L. (1999). Cloning and functional analysis of a cDNA encoding a starch synthase from potato (*Solanum tuberosum* L,) that is predominantly expressed in leaf tissue. Planta 208:503-511.

Kubo, A., Colleoni, C., Dinges, J.R., Lin, Q., Lappe, R.R., Rivenbark, J.G., Meyer, A.J., Ball, S.G., James, M.G., Hennen-Bierwagen, T.A., et al. (2010). Functions of heteromeric and homomeric isoamylase-type starch-debranching enzymes in developing maize endosperm. Plant Physiol. 153:956-969.

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., and Calcott, B. (2017). PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34:772-773.

Leinonen, R., Sugawara, H., Shumway, M., and International Nucleotide Sequence Database, C. (2011). The sequence read archive. Nucleic Acids Res. 39:D19-21.

Li, S., Cui, Y., Zhou, Y., Luo, Z., Liu, J., and Zhao, M. (2017). The industrial applications of cassava: current status, opportunities and prospects. J. Sci. Food Agric. 97:2282-2290.

Lin, Q.H., Huang, B.Q., Zhang, M.X., Zhang, X.L., Rivenbark, J., Lappe, R.L., James, M.G., Myers, A.M., and Hennen-Bierwagen, T.A. (2012). Functional interactions between starch synthase III and isoamylase-type starch-debranching enzyme in maize endosperm. Plant Physiol. 158:679-692.

Liu, D.R., Huang, W.X., and Cai, X.L. (2013). Oligomerization of rice granule-bound starch synthase 1 modulates its activity regulation. Plant Sci. 210:141-150.

Liu, F., Ahmed, Z., Lee, E.A., Donner, E., Liu, Q., Ahmed, R., Morell, M.K., Emes, M.J., and Tetlow, I.J. (2012a). Allelic variants of the amylose extender mutation of maize demonstrate phenotypic variation in starch structure resulting from modified protein-protein interactions. J. Exp. Bot. 63:1167-1183.

Liu, F., Makhmoudova, A., Lee, E.A., Wait, R., Emes, M.J., and Tetlow, I.J. (2009). The amylose extender mutant of maize conditions novel protein-protein interactions between starch biosynthetic enzymes in amyloplasts. J. Exp. Bot. 60:4423-4440.

Liu, F., Romanova, N., Lee, E.A., Ahmed, R., Evans, M., Gilbert, E.P., Morell, M.K., Emes, M.J., and Tetlow, I.J. (2012b). Glucan affinity of starch synthase IIa determines binding of starch synthase I and starch-branching enzyme IIb to starch granules. Biochem. J. 448:373-387.

Liu, K., Zu, Y., Chi, C., Gu, B., Chen, L., and Li, X. (2018). Modulation of the digestibility and multi-scale structure of cassava starch by controlling the cassava growth period. Int. J. Biol. Macromol. 120:346-353.

Ma, W., Kong, Q., Grix, M., Mantyla, J.J., Yang, Y., Benning, C., and Ohlrogge, J.B. (2015). Deletion of a C-terminal intrinsically disordered region of WRINKLED1 affects its stability and enhances oil accumulation in Arabidopsis. Plant J. 83:864-874.

Makhmoudova, A., Williams, D., Brewer, D., Massey, S., Patterson, J., Silva, A., Vassall, K.A., Liu, F., Subedi, S., Harauz, G., et al. (2014). Identification of multiple phosphorylation sites on maize endosperm starch branching enzyme IIb, a key enzyme in amylopectin biosynthesis. J. Biol. Chem. 289:9233-9246.

McPherson, A.E., and Jane, J. (1999). Comparison of waxy potato with other root and tuber starches. Carbohydr. Polym. 40:57-70.

Moorthy, S.N. (2002). Physicochemical and functional properties of tropical tuber starches: A review. Starch-Starke 54:559-592.

Morell, M.K., Kosar-Hashemi, B., Cmiel, M., Samuel, M.S., Chandler, P., Rahman, S.,

Buleon, A., Batey, I.L., and Li, Z.Y. (2003). Barley sex6 mutants lack starch synthase IIa activity and contain a starch with novel properties. Plant J. 34:172-184.

Nakamura, Y. (2002). Towards a better understanding of the metabolic system for amylopectin biosynthesis in plants: Rice endosperm as a model tissue. Plant Cell Physiol. 43:718-725.

Nakamura, Y., Aihara, S., Crofts, N., Sawada, T., and Fujita, N. (2014). *In vitro* studies of enzymatic properties of starch synthases and interactions between starch synthase I and starch branching enzymes from rice. Plant Sci. 224:1-8.

Nakamura, Y., Francisco, P.B., Hosaka, Y., Sato, A., Sawada, T., Kubo, A., and Fujita, N. (2005). Essential amino acids of starch synthase IIa differentiate amylopectin structure and starch quality between japonica and indica rice varieties. Plant Mol. Biol. 58:213-227.

Nakamura, Y., Ono, M., Utsumi, C., and Steup, M. (2012). Functional interaction between plastidial starch phosphorylase and starch branching enzymes from rice during the synthesis of branched maltodextrins. Plant Cell Physiol. 53:869-878.

Nielsen, R., and Yang, Z. (1998). Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148:929-936.

Nishi, A., Nakamura, Y., Tanaka, N., and Satoh, H. (2001). Biochemical and genetic analysis of the effects of amylose-extender mutation in rice endosperm. Plant Physiol. 127:459-472.

Patel, R.K., and Jain, M. (2012). NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLoS one 7:e30619.

Peng, C., Wang, Y., Liu, F., Ren, Y., Zhou, K., Lv, J., Zheng, M., Zhao, S., Zhang, L., Wang, C., et al. (2014). FLOURY ENDOSPERM6 encodes a CBM48 domain-containing protein involved in compound granule formation and starch synthesis in rice endosperm. Plant J. 77:917-930.

Pollock, C., and Preiss, J. (1980). The citrate-stimulated starch synthase of starchy maize kernels: purification and properties. Arch. Biochem. Biophys. 204:578-588.

Roldan, I., Wattebled, F., Mercedes Lucas, M., Delvalle, D., Planchot, V., Jimenez, S., Perez, R., Ball, S., D'Hulst, C., and Merida, A. (2007). The phenotype of soluble starch synthase IV defective mutants of *Arabidopsis thaliana* suggests a novel function of elongation enzymes in the control of starch granule formation. Plant J. 49:492-504.

Ryoo, N., Yu, C., Park, C.S., Baik, M.Y., Park, I.M., Cho, M.H., Bhoo, S.H., An, G., Hahn, T.R., and Jeon, J.S. (2007). Knockout of a starch synthase gene OsSSIIIa/Flo5 causes white-core floury endosperm in rice (*Oryza sativa* L.). Plant Cell Rep. 26:1083-1095.

Satoh, H., Shibahara, K., Tokunaga, T., Nishi, A., Tasaki, M., Hwang, S.K., Okita, T.W., Kaneko, N., Fujita, N., Yoshida, M., et al. (2008). Mutation of the plastidial alpha-glucan phosphorylase gene in rice affects the synthesis and structure of starch in the endosperm. Plant Cell 20:1833-1849.

Seung, D., Boudet, J., Monroe, J., Schreier, T.B., David, L.C., Abt, M., Lu, K.J., Zanella, M., and Zeemana, S.C. (2017). Homologs of PROTEIN TARGETING TO STARCH control starch granule initiation in Arabidopsis leaves. Plant Cell 29:1657-1677.

Seung, D., Soyk, S., Coiro, M., Maier, B.A., Eicke, S., and Zeeman, S.C. (2015). PROTEIN TARGETING TO STARCH is required for localising GRANULE-BOUND STARCH SYNTHASE to starch granules and for normal amylose synthesis in Arabidopsis. PLoS Biol.

13:e1002080.

Song, Y., and Jane, J. (2000). Characterization of barley starches of waxy, normal, and high amylose varieties. Carbohydr. Polym. 41:365-377.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312-1313.

Stolzer, M., Lai, H., Xu, M., Sathaye, D., Vernot, B., and Durand, D. (2012). Inferring duplications, losses, transfers and incomplete lineage sorting with nonbinary species trees. Bioinformatics 28:i409-i415.

Subasinghe, R.M., Liu, F., Polack, U.C., Lee, E.A., Emes, M.J., and Tetlow, I.J. (2014). Multimeric states of starch phosphorylase determine protein-protein interactions with starch biosynthetic enzymes in amyloplasts. Plant Physiol. Biochem. 83:168-179.

Takahata, Y., Tanaka, M., Otani, M., Katayama, K., Kitahara, K., Nakayachi, O., Nakayama, H., and Yoshinaga, M. (2010). Inhibition of the expression of the starch synthase II gene leads to lower pasting temperature in sweetpotato starch. Plant Cell Rep. 29:535-543.

Tetlow, I.J., Beisel, K.G., Cameron, S., Makhmoudova, A., Liu, F., Bresolin, N.S., Wait, R., Morell, M.K., and Emes, M.J. (2008). Analysis of protein complexes in wheat amyloplasts reveals functional interactions among starch biosynthetic enzymes. Plant Physiol. 146:1878-1891.

Tetlow, I.J., Wait, R., Lu, Z., Akkasaeng, R., Bowsher, C.G., Esposito, S., Kosar-Hashemi, B., Morell, M.K., and Emes, M.J. (2004). Protein phosphorylation in amyloplasts regulates starch branching enzyme activity and protein-protein interactions. Plant Cell 16:694-708.

Umemoto, T., and Aoki, N. (2005). Single-nucleotide polymorphisms in rice starch synthase IIa that alter starch gelatinisation and starch association of the enzyme. Funct. Plant Biol. 32:763-768.

Utsumi, Y., Utsumi, C., Sawada, T., Fujita, N., and Nakamura, Y. (2011). Functional diversity of isoamylase oligomers: the ISA1 homo-oligomer is essential for amylopectin biosynthesis in rice endosperm. Plant Physiol. 156:61-77.

Vanderschuren, H., Alder, A., Zhang, P., and Gruissem, W. (2009). Dose-dependent RNAi-mediated geminivirus resistance in the tropical root crop cassava. Plant Mol. Biol. 70:265-272.

van de Wal, M., D'Hulst, C., Vincken, J.P., Buleon, A., Visser, R., and Ball, S. (1998). Amylose is synthesized in vitro by extension of and cleavage from amylopectin. J. Biol. Chem. 273:22232-22240.

Walley, J.W., Shen, Z.X., Sartor, R., Wu, K.J., Osborn, J., Smith, L.G., and Briggs, S.P. (2013). Reconstruction of protein networks from an atlas of maize seed proteotypes. Proc. Natl. Acad. Sci. USA 110:E4808-E4817.

Wang, X., Li, X., Deng, X., Han, H., Shi, W., and Li, Y. (2007). A protein extraction method compatible with proteomic analysis for the euhalophyte Salicornia europaea. Electrophoresis 28:3976-3987.

Wang, Y., Li, Y., Zhang, H., Zhai, H., Liu, Q., and He, S. (2017). A soluble starch synthase I gene, IbSSI, alters the content, composition, granule size and structure of starch in transgenic sweet potato. Sci. Rep. 7:2315.

Whelan, S., Allen, J.E., Blackburne, B.P., and Talavera, D. (2015). ModelOMatic: fast and

automated model selection between RY, nucleotide, amino acid, and codon substitution models. System. Biol. 64:42-55.

Xu, J., Duan, X., Yang, J., Beeching, J.R., and Zhang, P. (2013). Enhanced reactive oxygen species scavenging by overproduction of superoxide dismutase and catalase delays postharvest physiological deterioration of cassava storage roots. Plant Physiol. 161:1517-1528.

Yamamori, M., Fujita, S., Hayakawa, K., Matsuki, J., and Yasui, T. (2000). Genetic elimination of a starch granule protein, SGP-1, of wheat generates an altered starch with apparent high amylose. Theor. Appl. Genet. 101:21-29.

Yang, Z. (2000). Maximum likelihood estimation on large phylogenies and analysis of adaptive evolution in human influenza virus A. J. Mol. Evol. 51:423-432.

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24:1586-1591.

Yang, Z., Wang, Y., Xu, S., Xu, C., and Yan, C. (2013). Molecular evolution and functional divergence of soluble starch synthase genes in cassava (*Manihot esculenta* Crantz). Evol. Bioinform. 9:239-249.

Yoo, S.H., and Jane, J.L. (2002). Structural and physical characteristics of waxy and other wheat starches. Carbohydr. Polym. 49:297-305.

Zhai, Z.Y., Liu, H., and Shanklin, J. (2017). Phosphorylation of WRINKLED1 by KIN10 results in its proteasomal degradation, providing a link between energy homeostasis and lipid biosynthesis. Plant Cell 29:871-889.

Zhang, P., Potrykus, I., and Puonti-Kaerlas, J. (2000). Efficient production of transgenic cassava using negative and positive selection. Transgenic Res. 9:405-415.

Zhang, X., Szydlowski, N., Delvalle, D., D'Hulst, C., James, M.G., and Myers, A.M. (2008). Overlapping functions of the starch synthases SSII and SSIII in amylopectin biosynthesis in Arabidopsis. BMC Plant Biol. 8:96.

Zhao, S.S., Dufour, D., Sanchez, T., Ceballos, H., and Zhang, P. (2011). Development of waxy cassava with different Biological and physico-chemical characteristics of starches for industrial applications. Biotechnol. Bioeng. 108:1925-1935.

Zhou, H., Wang, L., Liu, G., Meng, X., Jing, Y., Shu, X., Kong, X., Sun, J., Yu, H., Smith, S.M., et al. (2016). Critical roles of soluble starch synthase SSIIIa and granule-bound starch synthase Waxy in synthesizing resistant starch in rice. Proc. Natl. Acad. Sci. USA 113:12844-12849.

Zhou, W., Yang, J., Hong, Y., Liu, G., Zheng, J., Gu, Z., and Zhang, P. (2015). Impact of amylose content on starch physicochemical properties in transgenic sweet potato. Carbohydr. Polym. 122:417-427.

Zhou, W., Zhao, S., He, S., Ma, Q., Lu, X., Hao, X., Wang, H., Yang, J., and Zhang, P. (2019). Production of very-high-amylose cassava by post-transcriptional silencing of branching enzyme genes. J. Integr. Plant Biol. https://doi.org/10.1111/jipb.12848

Zhu, F. (2015). Composition, structure, physicochemical properties, and modifications of cassava starch. Carbohydr. Polym. 122:456-480.

Figure 1

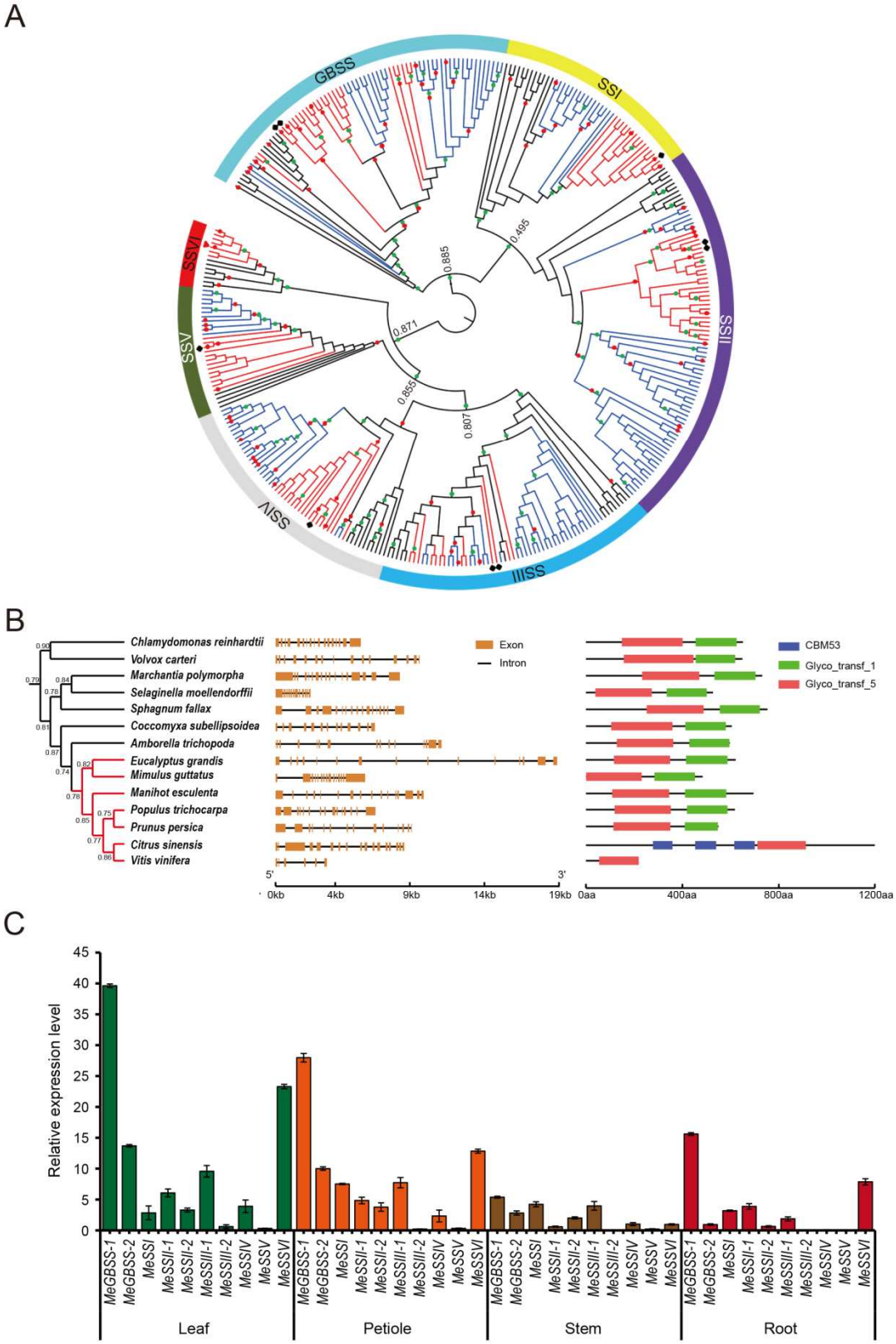


Figure 1. Molecular phylogenies of the plant soluble starch synthase families and transcription of soluble starch synthase genes in cassava.

(A) Molecular phylogenies of the plant soluble starch synthase families calculated using the maximum likelihood method as implemented in RAxML. Cassava SSs are marked with a black rhombus (MeGBSSI and MeSSI-V) and a red triangle (MeSSVI). Blue colored branches indicate monocot species and red colored branches indicate eudicot species. Gene duplications and gene losses are annotated as green circles and red circles on nodes, respectively. The numbers on the nodes show the bootstrap values.

(B) Phylogenetic tree, gene structure, and domain analyses of soluble starch synthase VI (SSVI) proteins across the plant kingdom. The phylogenetic tree of *SSVIs* constructed with RAxML using the maximum likelihood method is shown in the left panel. Red colored branches indicate eudicot species. Bootstrap values from 1,000 replicates are indicated at each branch. Exon-intron structures of *SSVI* genes are shown in the middle panel. Conserved domains annotated using the Pfam database are shown on the right panel. CBM53, carbohydrate binding module 53; Glyco_transf_1, Glycosyl transferases group 1; Glyco_transf_5, starch synthase catalytic domain.

(C) mRNA levels of transcribed soluble starch synthase genes in different cassava organs quantified by qRT-PCR assays. Data are presented as the mean ± SD (n = 3).

Figure 2

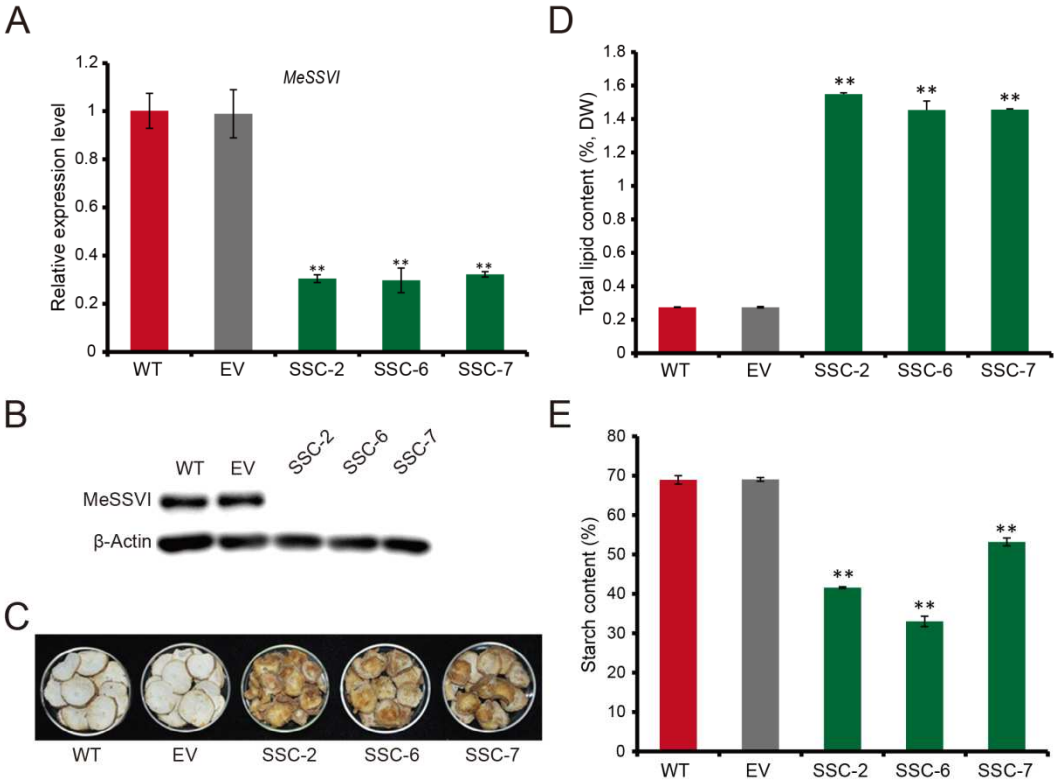


Figure 2. Repression of MeSSVI expression affects carbon allocation in cassava storage roots.

(A-B) mRNA (A) and protein (B) levels of *MeSSVI* in storage roots of wild-type (WT), empty vector (EV) and *MeSSVI*-RNAi (SSC) transgenic plants determined by qRT-PCR and Western blotting using a specific antibody against MeSSVI, respectively. The actin protein was used as a loading control.

(C) Dried slices of storage roots from WT, EV, and SSC transgenic cassava plants.

(D) Total lipid content in storage roots of WT, EV, and SSC transgenic cassava plants.

(E) Starch content in storage roots of WT, EV, and SSC transgenic plants. Data in parts A, D, and E are presented as the means ± SD (n = 3). * and ** indicate a significant difference compared to WT at $P < 0.05$ and <0.01, respectively, determined by Student's *t*-test.

Figure 3

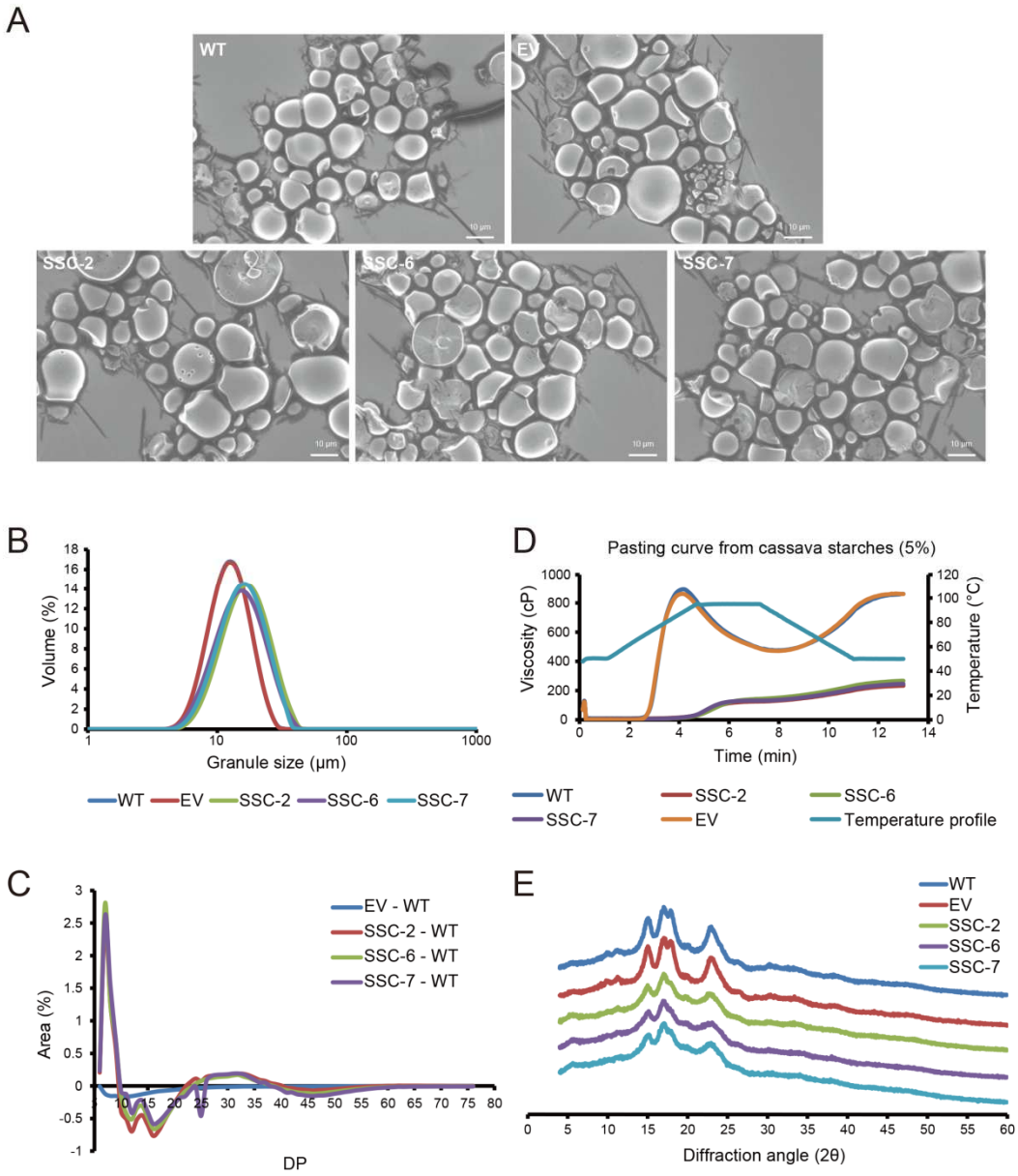


Figure 3. Physico-chemical properties of storage root starches in cassava.

(A, B) Scanning electron microscopy (A) and starch granule size distribution (B) of starch granules from wild-type (WT), empty vector (EV), and SSC transgenic line plants.

(C-E) Chain length distribution (C), Rapid Visco Analyser (RVA) pasting profiles (D), and X-ray diffractograms of storage starches isolated from plants of WT, EV, and SSC transgenic plant lines.

Figure 4

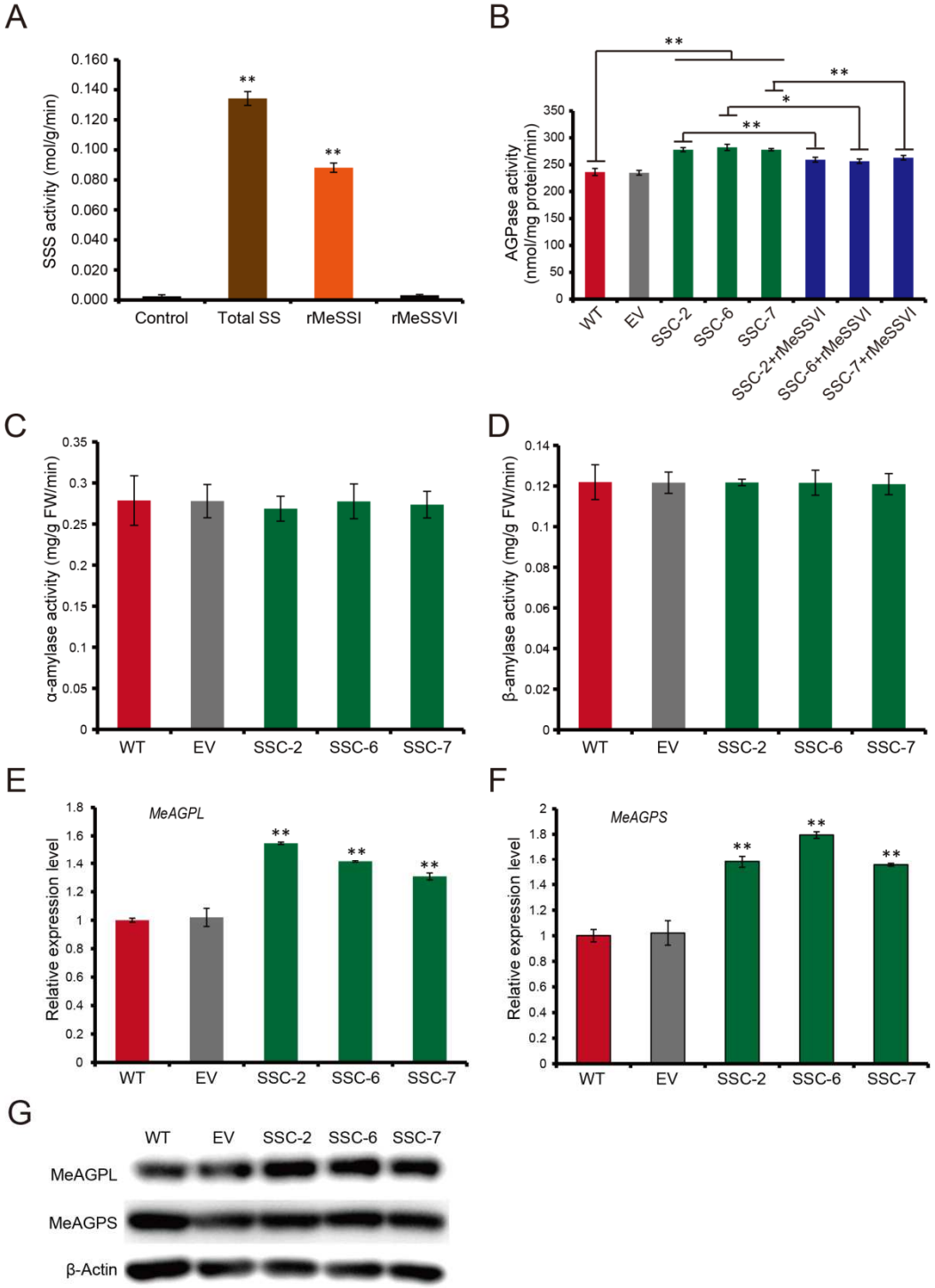


Figure 4. Enzyme activity and gene expression of MeSSVI, AGPase, and amylases in cassava storage roots.

(A) Enzyme activity of MeSSVI. Control: grade I water; Total SS: soluble protein extract from cassava storage root; rMeSSI: 1.2 μg recombinant MeSSI expressed in *E.coli* BL21 (DE3); rMeSSVI: 1.2 μg recombinant MeSSVI expressed in *E.coli* BL21 (DE3). Data are presented as the

mean ± SD (n = 3). * and ** indicate a significant difference compared to WT at $P < 0.05$ and <0.01, respectively, determined by Student's *t*-test.

(B-D) Enzyme activities of AGPase (B), α-amylase (C) and β-amylase (D) in storage roots of wild-type (WT), empty vector (EV), and SSC transgenic cassava plants. The enzyme extract (1.5 mg/ml) was a soluble protein extract from cassava storage roots; rMeSSVI: 1.2 μg recombinant MeSSVI expressed in *E.coli* BL21 (DE3); Data are presented as the means ± SD (n = 3). * and ** indicate a significant difference at P <0.05 and <0.01, respectively, determined by Student's *t*-test.

(E-F) Transcription of *MeAGPL* (D) and *MeAGPS* (E) in storage roots of WT, EV, and SSC transgenic cassava determined by qRT-PCR. Data are presented as the means ± SD (n = 3). * and ** indicate a significant difference compared WT at P <0.05 and <0.01, respectively, determined by Student's *t*-test.

(G) Protein levels of MeAGPL and MeAGPS in storage roots of WT, EV, and SSC transgenic cassava detected by Western blotting. The actin protein was used as a loading control.

Figure 5

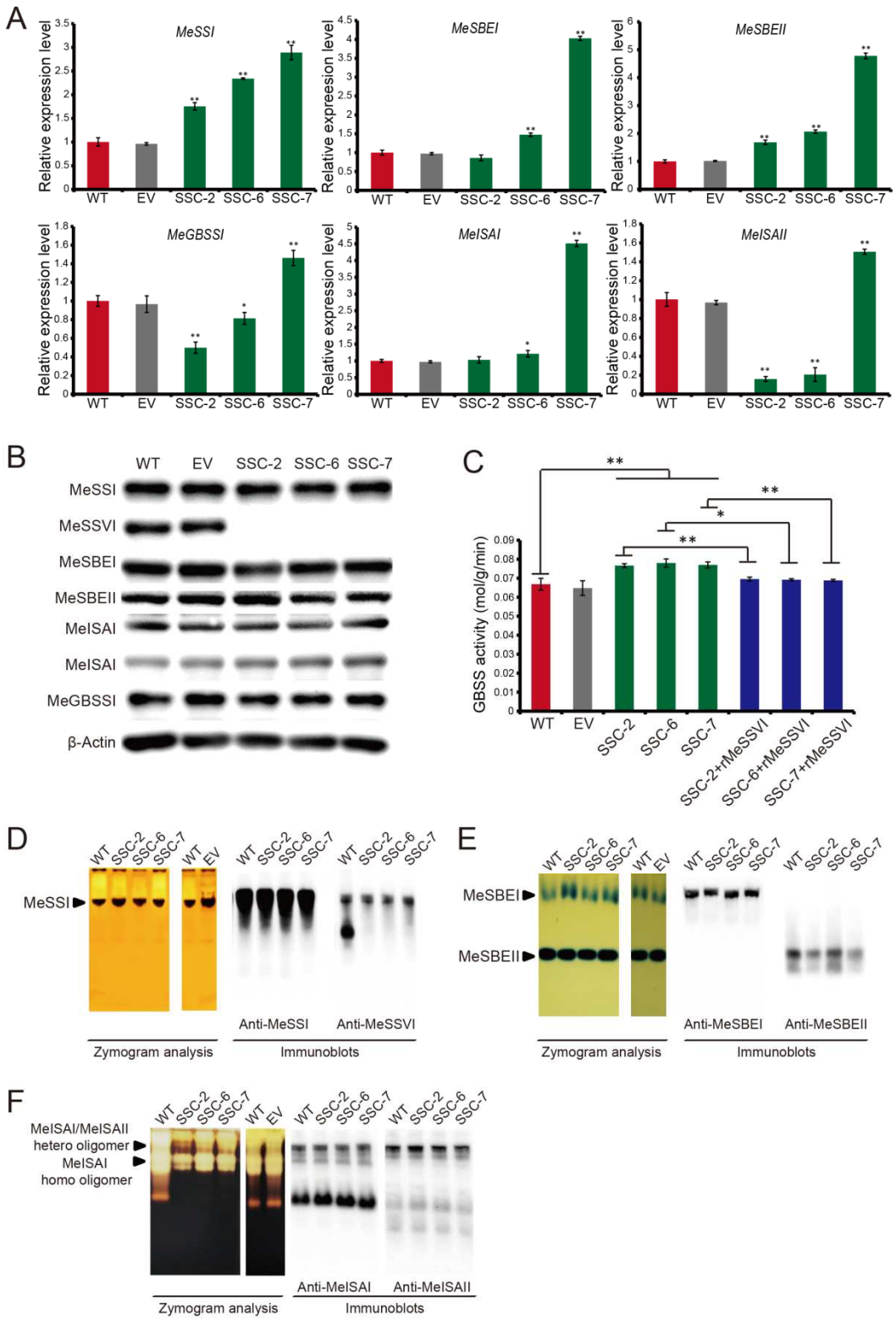


Figure 5. Expression and activity analyses of starch biosynthesis enzymes in cassava storage roots.

(A) Transcription of *MeSSI*, *MeSBEI*, *MeSBEII*, *MeGBSSI*, *MeISAI*, and *MeISAII* in storage roots

of wild-type (WT), empty vector (EV), and SSC transgenic plants determined by qRT-PCR. Data are presented as the means ± SD (n = 3). * and ** indicate a significant difference compared to WT at $P < 0.05$ and < 0.01, respectively, determined by Student's *t*-test.

(B) Protein levels of MeSSI, MeSSVI, MeSBEI, MeSBEII, MeISAI, MeISAII, and MeGBSSI in storage roots of WT, EV, and SSC transgenic cassava plants detected by Western blotting with the corresponding antibodies.

(C) Enzyme activity of GBSS in storage roots of WT, EV, and SSC transgenic cassava plants. The enzyme extract (1.5 mg/ml) was a soluble protein extract from cassava storage root; rMeSSVI: 1.2 μg recombinant MeSSVI expressed in *E.coli* BL21 (DE3). Data are presented as the means ± SD (n = 3). * and ** indicate a significant difference at $P < 0.05$ and < 0.01, respectively, as determined by Student's *t*-test.

(D-F) Enzyme activity of amylopectin biosynthesis enzymes in WT, EV, and SSC transgenic cassava plants. Enzyme names are shown on the left; sample names are shown at the top; antibodies used for Western blot detection are shown below the figure.

Figure 6

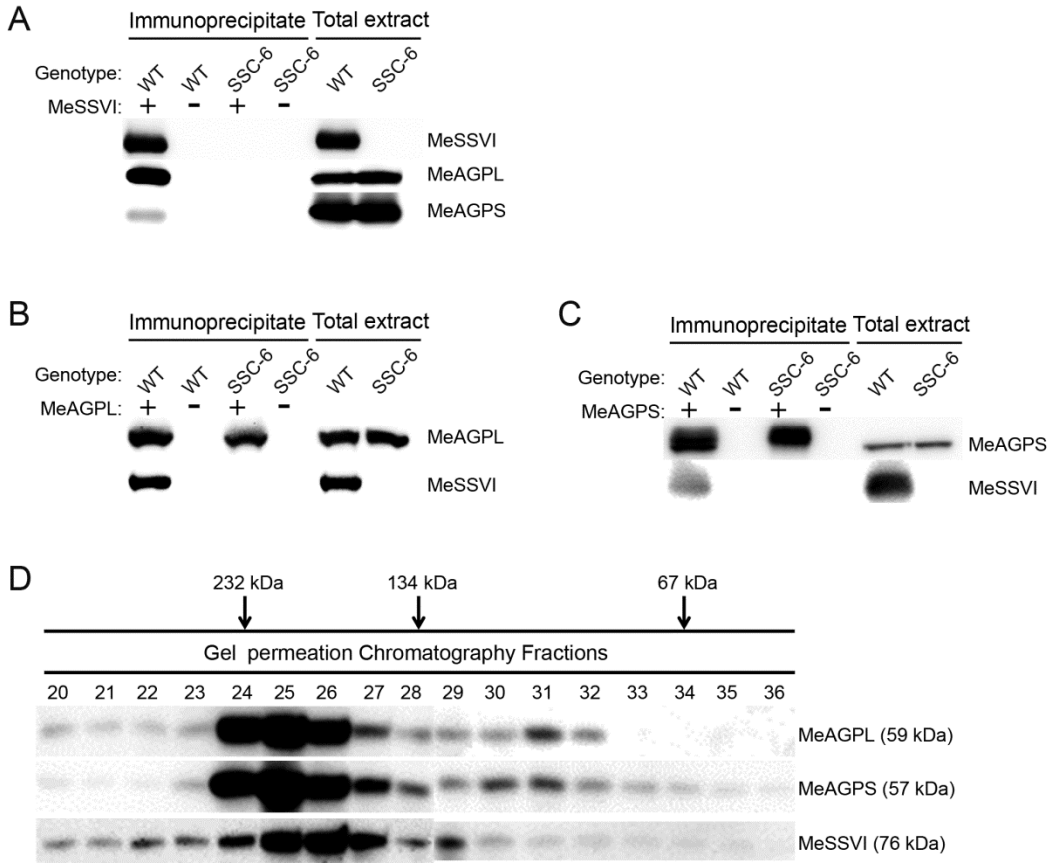
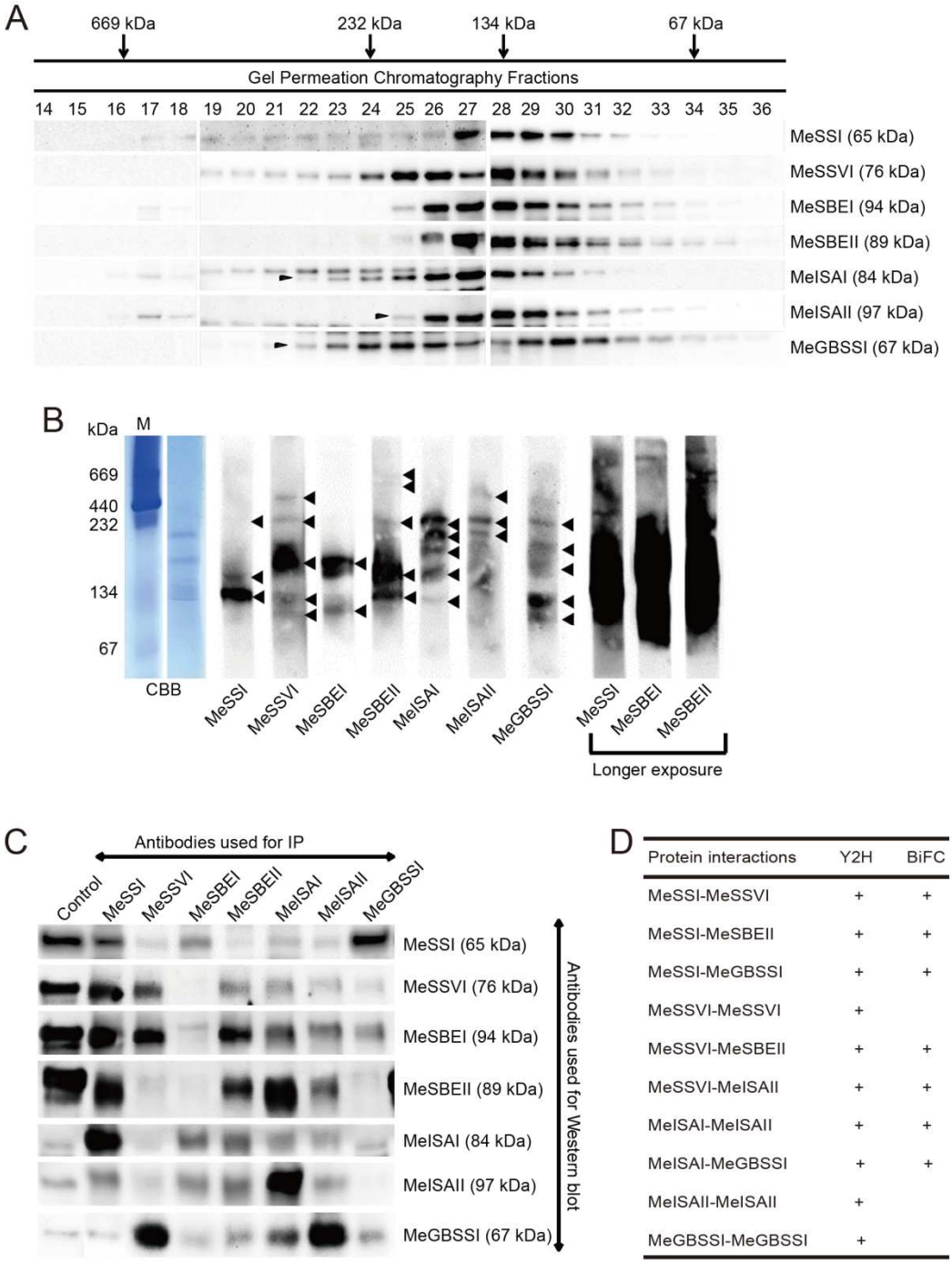


Figure 6. Protein-protein interactions among MeAGPL, MeAGPS, and MeSSVI in cassava storage roots.

(A-C) Co-immunoprecipitation (Co-IP) analyses. Soluble proteins extracted from mature cassava storage roots were immunoprecipitated using specific antibodies as indicated. A soluble protein extract from WT cassava storage roots was used as a positive control. Soluble protein extract from SSC transgenic cassava storage roots was used as a negative control. The antibodies used for Western blotting are indicated on the right.

(D) Gel permeation chromatography (GPC) analyses of proteins extracted from mature storage roots of WT cassava plants. Proteins present in the indicated fractions from the GPC column were separated by SDS-PAGE and then probed with each specific antibody in immunoblot analyses. The peak elution volumes of molecular mass standards analyzed using the same GPC protocol are indicated at the top.

Figure 7



Protein interactions	Y2H	BiFC
MeSSI-MeSSVI	+	+
MeSSI-MeSBEII	+	+
MeSSI-MeGBSSI	+	+
MeSSVI-MeSSVI	+	
MeSSVI-MeSBEII	+	+
MeSSVI-MeISAII	+	+
MeISAI-MeISAII	+	+
MeISAI-MeGBSSI	+	+
MeISAII-MeISAII	+	
MeGBSSI-MeGBSSI	+	

Figure 7. Protein-protein interactions among starch biosynthesis enzymes of cassava.

(A) GPC analyses of proteins extracted from mature storage roots of wild-type (WT) cassava. Proteins present in the indicated fractions from the GPC column were separated by SDS-PAGE and then probed with each specific antibody in immunoblot analyses. The peak elution volumes of molecular mass standards analyzed using the same GPC protocol are indicated at the top. Monomeric molecular weights of each enzyme are indicated on the right. Arrowheads indicate the band that matches the predicted molecular weight of the proteins.

(B) Blue native (BN)-PAGE analyses of starch biosynthetic enzyme complexes from mature storage roots of WT cassava. Soluble protein extracts were separated by native-PAGE and then probed with specific antibodies in immunoblot analyses. Arrowheads indicate the presence of polypeptides recognized by the antibodies.

(C) Co-immunoprecipitation (Co-IP) analyses of cassava starch biosynthesis enzymes. Soluble proteins from mature storage roots were immunoprecipitated using the corresponding specific antibodies. Soluble proteins from WT cassava storage roots were used as a positive control. The antibodies used for Western blotting are indicated on the right.

(D) Summary of protein-protein interactions among starch biosynthesis enzymes of cassava analyzed by yeast two-hybrid (Y2H) and bimolecular fluorescence complementation (BiFC) assays. +, Positive; -, Negative.

Figure 8

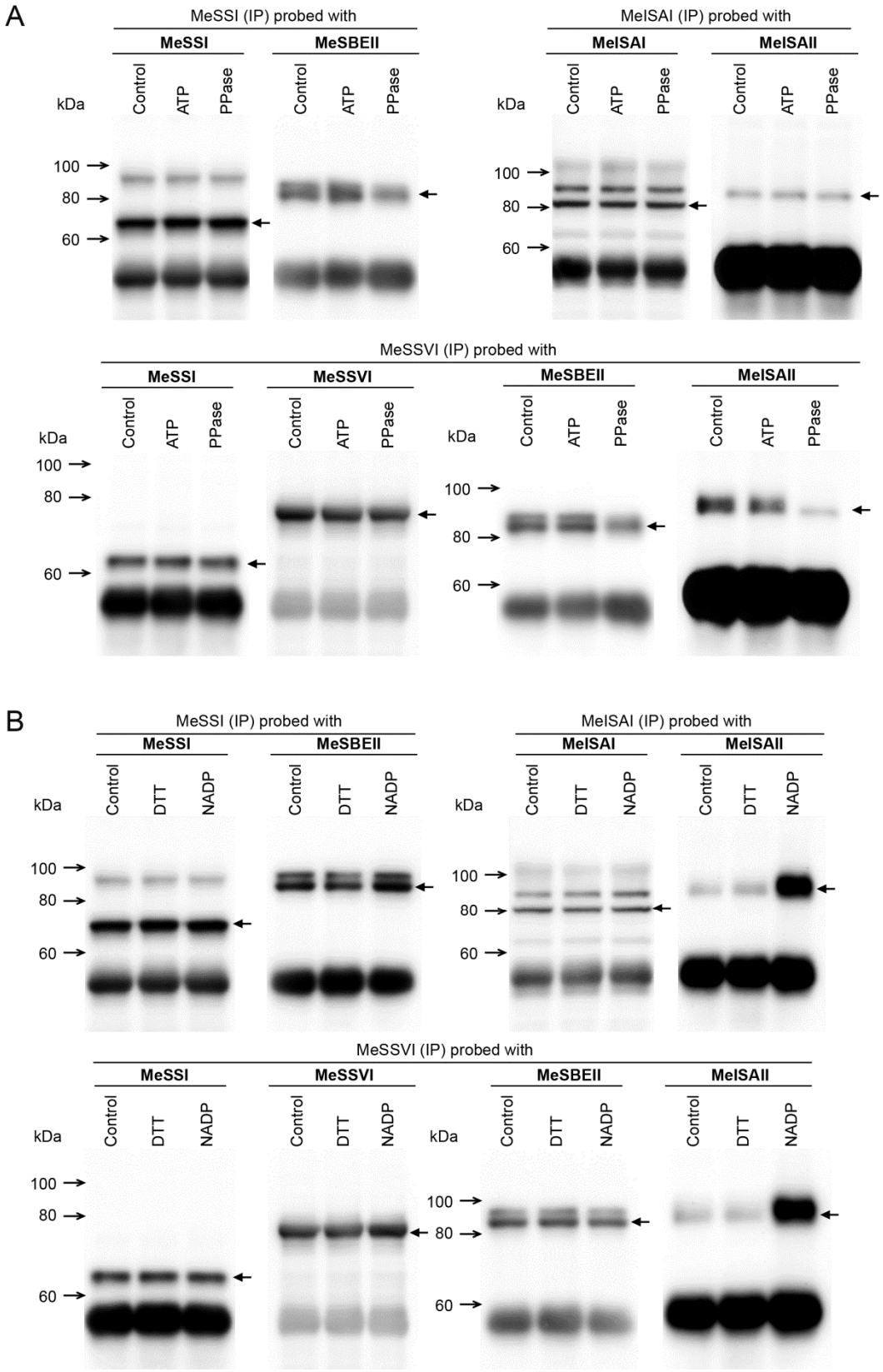


Figure 8. Effects of phosphorylation treatments and redox potential on starch biosynthesis enzymes from cassava storage roots.

(A) ATP and PPase treatments. Root lysates (1.2 to 1.5 mg protein cm^{-3}) were preincubated with either 1 mM ATP or with 10 units of PPase for 20 min at 25°C and immunoprecitated. Untreated root lysates were used as controls. After immunoprecipitation (IP) with protein-specific antibodies, immunoprecipitated proteins were separated by SDS-PAGE, electroblotted onto polyvinylidene fluoride (PVDF) membranes, and developed with various antisera as shown. Arrows indicate the positions of each protein that cross-reacted with protein-specific antisera.

(B) Dithiothreitol (DTT) or NADP treatments. Root lysates (1.2 to 1.5 mg protein cm^{-3}) were preincubated with either 10 mM DTT or with 10 mM NADP for 30 min at 25°C and immunoprecitated. Details are as in A.

Figure 9

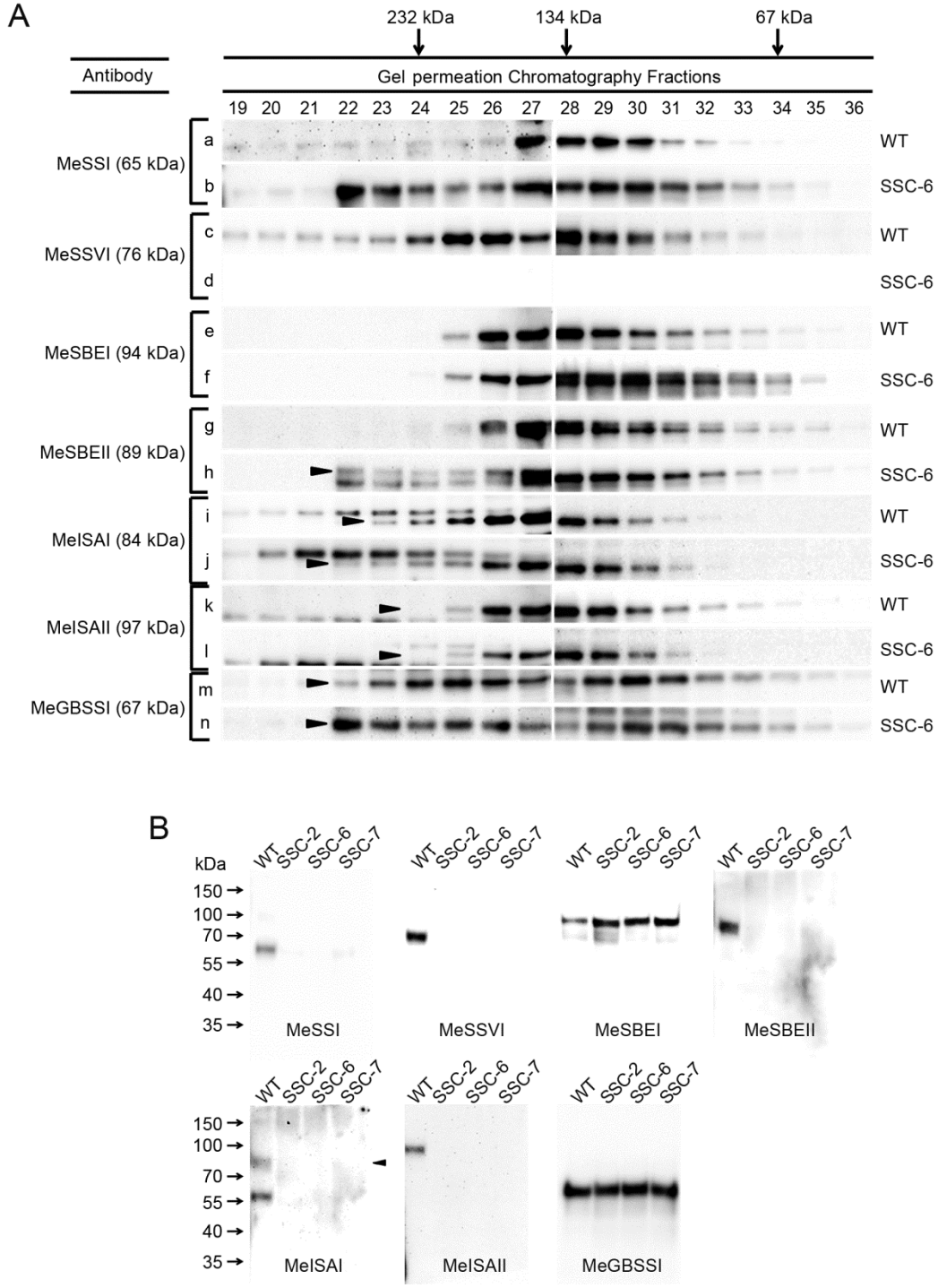


Figure 9. Characterization of protein interaction patterns and starch granule-bound proteins in storage roots between WT and SSC transgenic cassava plants.

(A) GPC analyses of proteins extracted from mature storage roots of wild-type (WT) and SSC transgenic cassava plants. Proteins present in the indicated fractions from the GPC column were separated by SDS-PAGE and then probed with specific antibodies in immunoblot analyses. The peak elution volumes of molecular mass standards analyzed using the same GPC protocol are

indicated at the top. The monomeric molecular weights of each enzyme are indicated on the left. Material names are indicated on the right. Arrowheads indicate the bands that match the predicted molecular weights.

(B) Analysis of starch granule-bound proteins in storage roots of WT and SSC transgenic cassava plants. Molecular mass standards are indicated on the left. The antibodies used for Western blotting are indicated on the bottom. Sample names are indicated on the top. Arrowheads indicate the bands that match the predicted molecular weights.

Figure 10

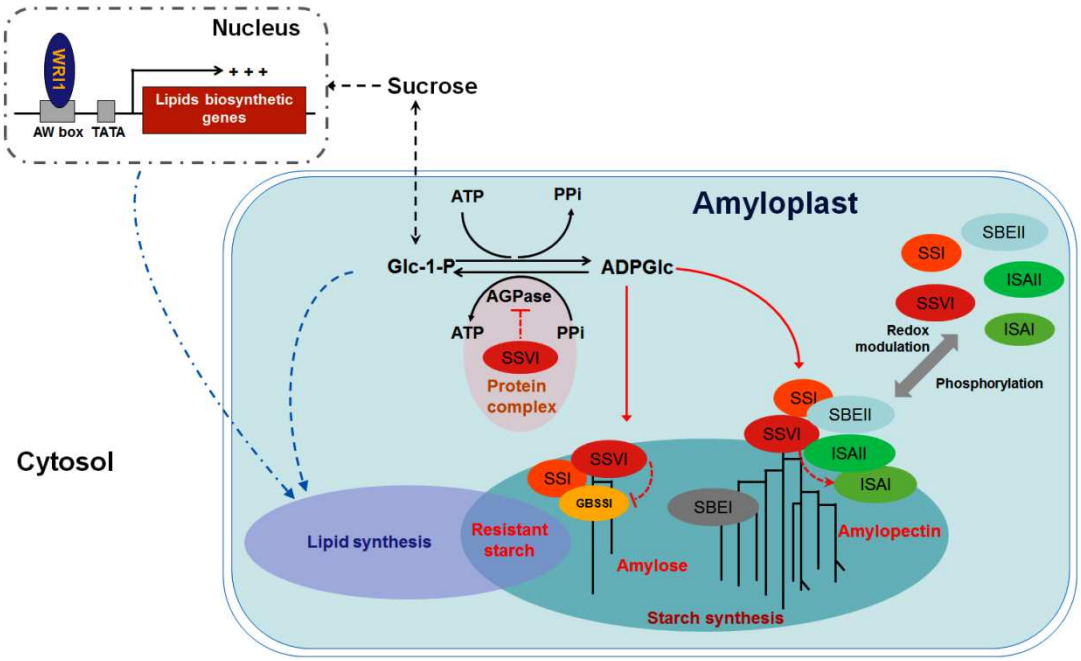


Figure 10. A molecular model of SSVI function in carbon allocation and starch biosynthesis in cassava storage roots. Normally, SSVI inhibits AGPase activity through the formation of a protein complex so that Glc-1-P is directed toward ADPGlc intermediates to support starch synthesis. At the same time, SSVI enhances ISA activity and inhibits GBSS activity through protein-protein interactions, which is more conducive to amylopectin synthesis. Down-regulated *SSVI* expression increases lipid accumulation and amylose deposition, which results in increased levels of resistant starch in storage roots. The formation of MeSSVI-associated protein complexes depends on protein phosphorylation and redox state.

Supplementary files:

Table S1. Prediction of positive selection for each starch synthase subfamily under the site specific model

Groups	K_a/K_s	LRTs (M3 versus M0)	LRTs (M2a versus M1a)	LRTs (M7 versus M8)
GBSS	0.10769	*P*<0.01	*P*>0.05	*P*>0.05
SSI	0.08615	*P*<0.01	*P*>0.05	*P*>0.05
SSII	0.09668	*P*<0.01	*P*>0.05	*P*>0.05
SSIII	0.09642	*P*<0.01	*P*>0.05	*P*>0.05
SSIV	0.16138	*P*<0.01	*P*>0.05	*P*>0.05
SSV	0.25505	*P*<0.01	*P*>0.05	*P*>0.05
SSVI	0.08595	*P*<0.01	*P*>0.05	*P*>0.05

Table S2. The list of starch synthase genes in cassava

Gene	Gene_ID	Exon	Intron	Chromosome	Location
MeGBSS-1	Manes.02G001000	13	13	Chromosome02	179291-183561
MeGBSS-2	Manes.01G055700	13	13	Chromosome01	19379772-19383339
MeSSI	Manes.01G184000	15	14	Chromosome01	32290569-32298102
MeSSII-1	Manes.02G046900	8	7	Chromosome02	4475431-4482297
MeSSII-2	Manes.01G091700	7	6	Chromosome01	25188203-25193935
MeSSIII-1	Manes.16G006900	16	15	Chromosome16	691031-702618
MeSSIII-2	Manes.S044400	15	14	Chromosome17	19664439-19676045
MeSSIV	Manes.15G118600	17	16	Chromosome15	9760010-9768511
MeSSV	Manes.13G105300	8	7	Chromosome13	27038861-27042811
MeSSVI	Manes.02G082800	14	14	Chromosome02	7377914-7387931

Table S3. Rapid visco analysis of cassava starches

Line	Peak viscosity (cP)	Hot paste viscosity (cP)	Breakdown (cP)	Final viscosity (cP)	Setback (cP)	Peak time (min)	Pasting temperature (°C)
WT	893.33±17.62	475.33±4.04	418±15.87	861.33±15.53	386±17.52	4.15±0.04	71.05±0.05
EV	861.67±2.89	471±3.61	390.67±2.52	860.67±15.18	389.67±13.43	4.11±0.03	70.25±0.05
SSC-2	124.67±1.15**	115±1**	9.67±0.58**	233±1.73**	118±1**	6.84±0.08**	84.36±0.03**
SSC-6	140.67±2.31**	121.33±1.53**	19.33±1.53**	267±4.36**	145.67±3.21**	6.98±0.04**	84.02±0.07**
SSC-7	131.33±1.53**	124.67±3.06**	6.67±1.53**	245.33±3.06**	120.67±5.77**	6.75±0.17**	83.65±0.06**

Notes: WT, wild type; EV, empty vector; SSC-x, SSC transgenic lines. Data are presented as the means ± SD (n = 3). * and ** indicate a significant difference versus WT at $P < 0.05$ and <0.01, respectively, as determined by Student's *t*-test.

Table S4. Thermal properties of cassava starches as determined by differential scanning calorimetry

Line	T_o (°C)	T_p (°C)	T_c (°C)	ΔH (J/g)
WT	56.26±0.26	63.11±0.12	78.21±1.25	11.00±0.87
EV	56.37±0.11	63.16±0.08	78.19±0.86	11.09±0.18
SSC-2	51.12±0.16**	57.86±0.20**	65.83±2.77**	2.85±0.82**
SSC-6	51.38±1.52**	54.85±1.65**	62.16±1.66**	5.38±1.02**
SSC-7	51.05±0.15**	56.39±0.46**	63.03±0.39**	5.40±0.20**

Notes: WT, wild type; EV, empty vector; SSC-x, SSC transgenic lines. T_o, Onset temperature; T_p, Peak temperature; T_c, End temperature; ΔH, Endothermic enthalpy. Data are presented as the means ± SD (n = 3). * and ** indicate a significant difference versus WT at $P < 0.05$ and <0.01, respectively, as determined by Student's *t*-test.

Table S5. Identification of co-immunoprecipitated proteins by mass spectrometry

Accession No.	Identified protein	Antibodies used for IP						
		MeSSI	MeSSVI	MeSBEI	MeSBEII	MeISAI	MeISAII	MeGBSSI
B3SRP2	MeSSI	+	+	+	+	+	+	+
B3SRP3	MeSSVI	+	+	+	+	+	+	+
Q08131	MeSBEI	+	+	+	+	+	+	+
A0A067L2P8	MeSBEII	+	+	+	+	+	+	+
B9SV81	MeISAI	+	+	+	+	+	+	+
B9T4B0	MeISAII	+	+	+	+	+	+	+
Q43784	MeGBSSI	+	+	+	+	+	+	+
A0A172R266	MeAGPL	+	+	+	+	n.d.	+	+
A0A172R259	MeAGPS	+	+	+	+	+	+	+
B9SJB6	MePho	+	+	+	+	+	+	+
D5IBV3	14-3-3	+	+	+	+	+	+	+
A0A076N642	MeSuS1	+	+	+	+	+	+	+
A0A076N3L4	MeSuS4	n.d.	n.d.	n.d.	+	+	+	n.d.

Notes: Protein family assignment is made based on identification between the peptide sequences and known protein sequences from cassava or Arabidopsis, together with BLAST analysis in other species. Database accession numbers are for the predicted cassava protein. +, Protein was detected. n.d., Protein cannot be detected.

Table S6. Primer pairs used in this study

Gene ID (gene name)	Vector	Primer sequences (5' to 3')
For construction of vector in yeast two-hybridization assay		
Manes.01G184000 (*MeSSI*)	pGADT7	CCCGGGTGCTTCCTTGACT
		TGGCTCGAGAGAAATGTAGGGTGAGTC
	pGBKT7	TTTCCATGGAGGCTTCCTTGACT
		TGGGTCGACAGAAATGTAGGGTGAGTC
Manes.02G082800 (*MeSSVI*)	pGADT7	TCGCATATGGCTACAGGTAAGGAAGG
		GTTGGATCCCCACTGGTACTTGGCT
	pGBKT7	TCGCATATGGCTACAGGTAAGGAAGG
		GTTGGATCCCCACTGGTACTTGGCT
Manes.05G133800 (*MeSBEI*)	pGADT7	CCCGGGTGCTGTGGAAAGAAAC
		CCCTCGAGTTTGTCATCCGATGTC
	pGBKT7	CCCGGGTGCTGTGGAAAGAAAC
		GCGTCGACTTTGTCATCCGATGTC
Manes.09G059400 (*MeSBEII*)	pGADT7	CCGGAATTCGCGTCTAAAAGAGTCC
		TTTGGATCCACCGGCGACAGG
	pGBKT7	CCGGAATTCGCGTCTAAAAGAGTCC
		TTTGGATCCACCGGCGACAGG
Manes.06G054000 (*MeISAI*)	pGADT7	ACCCGGGTGCCACCAGAAGAGA
		CCTCGAGAGCAATTTCATCAGGAG
	pGBKT7	ACCCGGGTGCCACCAGAAGAGA
		AAAGTCGACAGCAATTTCATCAGGAG
Manes.05G073400 (*MeISAII*)	pGADT7	ACCCGGGTGCCACCAGAAGAGA
		AAACTCGAGTCAACCACCAAGGCTAG
	pGBKT7	ACCCGGGTGCCACCAGAAGAGA
		AAAGTCGACTCAACCACCAAGGCTAG
Manes.02G001000 (*MeGBSSI*)	pGADT7	CATATGGGTCATGGAATGAATTTAA
		GAATTCCATTTTTTTTTTTTTTGTCAGA
	pGBKT7	CATATGGGTCATGGAATGAATTTAA
		GAATTCCATTTTTTTTTTTTTTGTCAGA
For construction of vector in BiFC assay		
Manes.01G184000 (*MeSSI*)	nYFP	TCTAGAATGGAATCTCTACGGATCAC
		GTCGACTCAAGAAATGTAGGGTGAGTC
Manes.02G082800 (*MeSSVI*)	nYFP	TCTAGAATGGCATTTATAGGATCACTTC
		GTCGACTCACCACTGGTACTTGGCTGC
	cYFP	TCTAGAATGGCATTTATAGGATCAC
		GTCGACTCACCACTGGTACTTGGCTG
Manes.09G059400 (*MeSBEII*)	cYFP	TCTAGAATGGGACACTACACCATATC
		GTCGACTTAACCGGCGACAGGTTCCAC
Manes.06G054000 (*MeISAI*)	nYFP	TCTAGAATGGACTTAATTCAATCTTC
		CTGCAGCTAAGCAATTTCATCAGGAG
Manes.05G073400 (*MeISAII*)	cYFP	CCCGGGATGGCAACACTTCTTCCTTCC
		GTCGACTCAACCACCAAGGCTAGTGGC
Manes.02G001000 (*MeGBSSI*)	cYFP	TCTAGAATGGCAACTGTAATAGCTGC
		GTCGACCATTTTTTTTTTTTTTGTCAG
For construction of vector in recombinant protein expression assay		

Manes.01G184000 (*MeSSI*)	pET-28a	CCCGGGTGCTTCCTTGACT
		CTCGAGAGAAATGTAGGGTGAGTC
Manes.02G082800 (*MeSSVI*)	pET-28a	GGATCCGCTACAGGTAAGGAAGGC
		CTCGAGTCACCACTGGTACTTGGC
For qRT-PCR		
Manes.13G084300 (*Meβ-Actin*)		TGATGAGTCTGGTCCATCCA
		CCTCCTACGACCCAATCTCA
Manes.01G184000 (*MeSSI*)		GCTCTGGACAAAGGTGAAGC
		CAACAGAGTAATGGGCAGCA
Manes.02G082800 (*MeSSVI*)		GTTGAGGTTCCTTGGCATGT
		CGACCCTGGTGAGCTATGTT
Manes.05G133800 (*MeSBEI*)		GCTCGCACTTGTGTGGTTTA
		CATCGGCAATCAAAGAAGGT
Manes.09G059400 (*MeSBEII*)		CAGTTCAAGCACCAGGTGAA
		AAGCTTTTTGATGCGAGGAA
Manes.06G054000 (*MeISAI*)		GCGACCACAGGATGAACTTT
		GGGCCTTTCTCATTTCCTTC
Manes.05G073400 (*MeISAII*)		GCGTGCACATTTTCAAGCTA
		AAAGCAGGCTCCTTGCATAA
Manes.02G001000 (*MeGBSSI*)		CGTGAAGGGAAGGAAAATCA
		TGTGGCATCGTAGTGGATGT
Manes.11G085500 (*MeAGPL*)		GTCTATGACCTTACAACCACC
		GCTTCCCTTAATCGGCACC T
Manes.12G067900 (*MeAGPS*)		CTCATTCTCCTCGTCTAATCTC
		GAAAATCGTGCAACTGCCTTT
Manes.10G063200 (*MeWRI1*)		ACGCCACCGCATCAATTAAG
		AACCTTCCCGTCCATCTATGC
Manes.04G145400 (*MeTAF9*)		TGAGCATGCTGGCAAATCTG
		ACCAGGTCCAGCTATTGTCTTG
Manes.02G053800 (*MeBCCP2*)		ATGCGTTGGATCCAATTCCC
		ATTGGCGCGGAATTTGAAGG
Manes.08G044800 (*MeFAD2*)		ACCTGCAGCATACTCATCCTTC
		TTGTTGCCTCCATCGCATTG
Manes.15G120000 (*MeKAS1*)		AGCAATTGCCTGTGTGAAGG
		AAGCAGAAAAGGCCACAACG
Manes.16G091400 (*MeKAS2*)		AGCCCTCATGCGTTGTTTTG
		ACTTTTGTGTCCACGCCTTG
Manes.17G101800 (*MeFAE1*)		TCGTTATCGTCGTTGTCTGC
		TGCGCTGAAATTCCAACGAG
Manes.09G160200 (*MeACP1*)		AAGCCAGAGAAGTTTCCTGTCC
		AGGGCTGCAAATTTGGACTC
Manes.11G061900 (*MeMED15*)		ACAACAGCTGGGTTTGCAAC
		TTGCGCCGTTGAATCAAGAG
Manes.14G039100 (*MeOLE1*)		CGCAGTTGGTCTCATAATCACAG
		TTGCTTGATGCCCAAACTGC
Manes.05G120900 (*MePLO2*)		TGGAAAAGTACCAAGATGTA
		TTGTCTAAGAACATTGAAGG
Manes.02G036200 (*MeKAS3*)		TGCTCTTGAAAAGGCTGGTC
		TGAAGCAGCACTTGTGTTGC

Figure S1

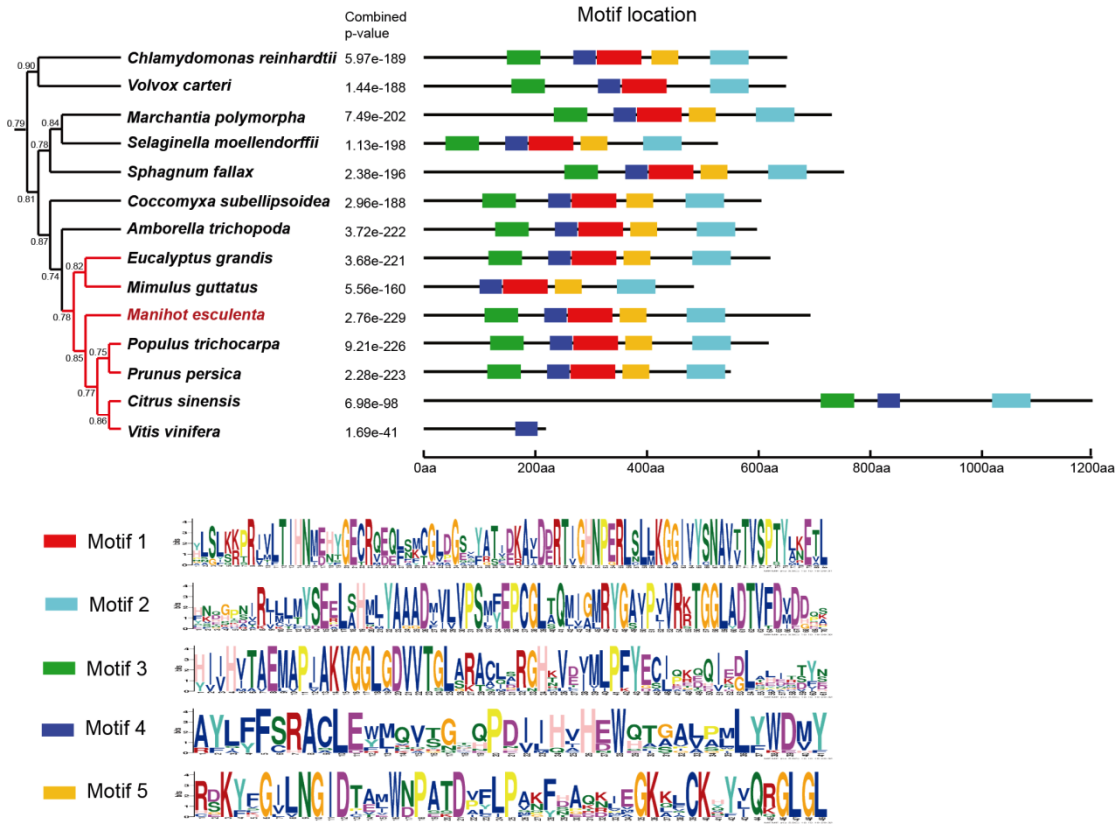


Figure S1. Phylogenetic relationships and conserved motifs of SSVI proteins. The left panel shows the phylogenetic tree that was constructed with RAxML using the maximum-likelihood method. Red colored branches indicate eudicot species. Bootstrap values from 1,000 replicates are indicated at each branch. The right panel is a schematic representation of the conserved motifs in the SSVI proteins as detected by MEME analysis. Black lines represent the non-conserved sequences, and each motif is represented by a color box numbered at the bottom. The lengths of the proteins can be estimated using the scale at the bottom.

Figure S2

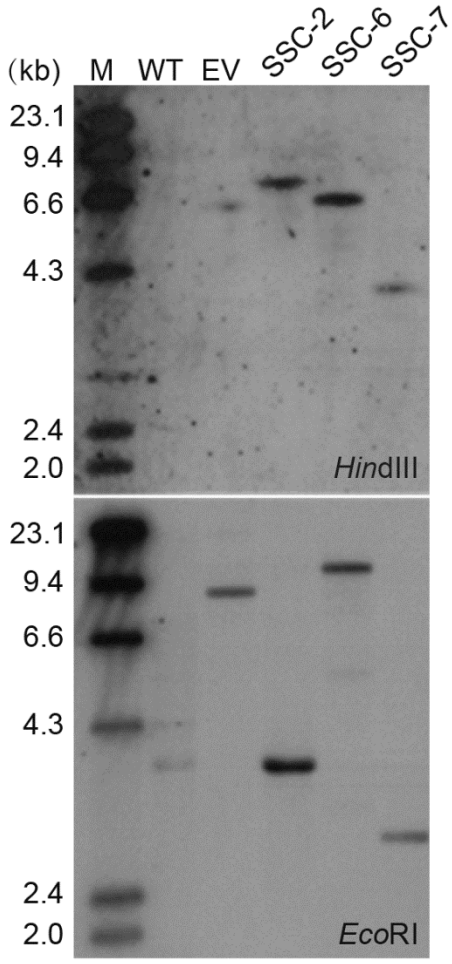


Figure S2. Integration patterns of transgenes in SSC transgenic cassava plant lines determined by Southern blotting analysis. M, DIG-labeled molecular marker; WT, wild-type control; EV, empty vector plasmid 35::RNAi; SSC-x, plant lines transformed with 35S::*MeSSVI*-RNAi.

Figure S3

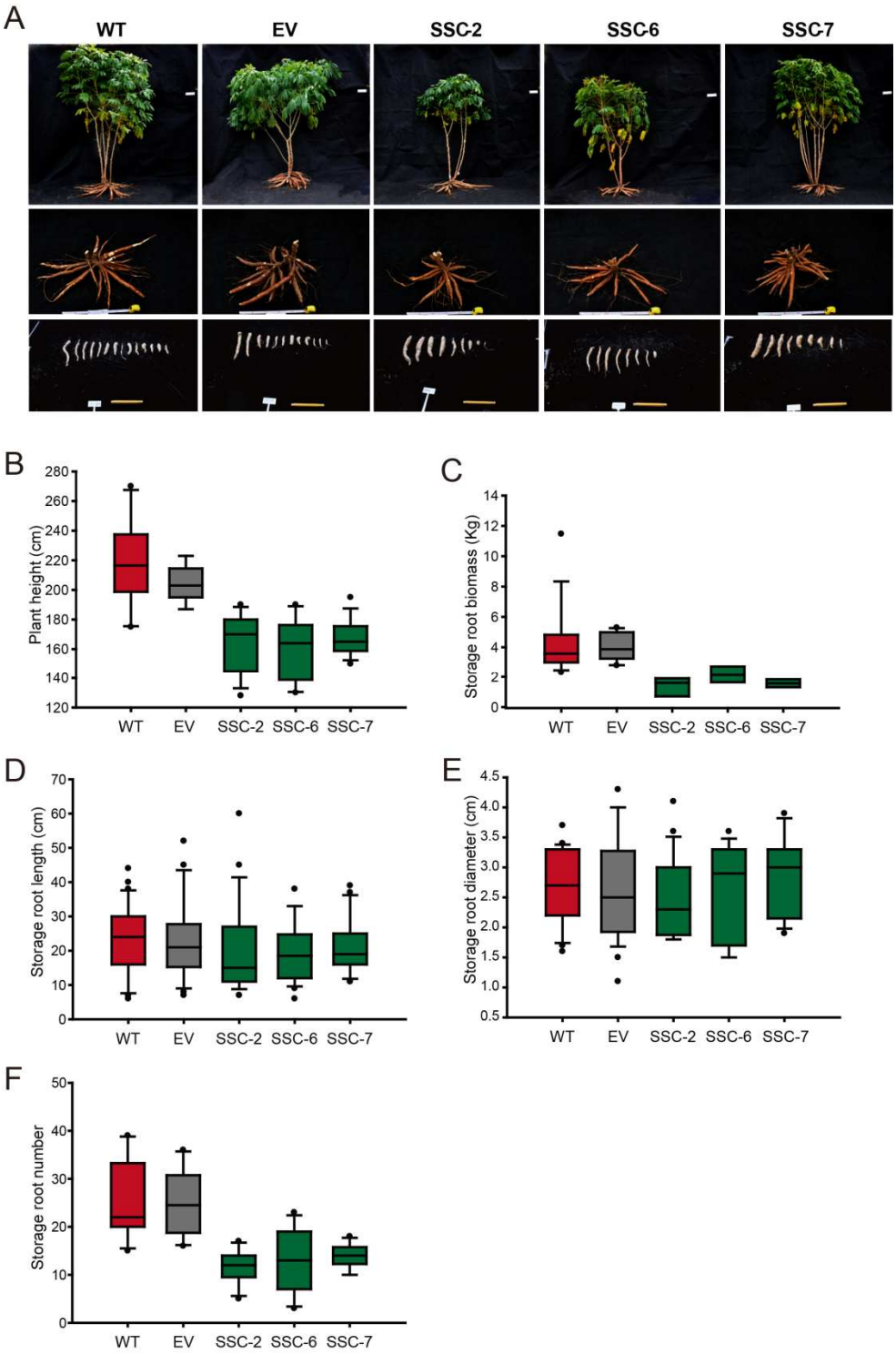


Figure S3. Phenotypes of field-grown WT and transgenic cassava plants. (A) Canopy architecture (upper panel), attached storage roots (middle panel), and harvested storage roots (lower panel) of the SSC transgenic plants in comparison with the wild-type (WT) and empty vector (EV) plants. (B-F) Comparisons of plant height (B), storage root biomass (C), storage root length (D), storage root diameter (E), and storage root number (F).

Figure S4

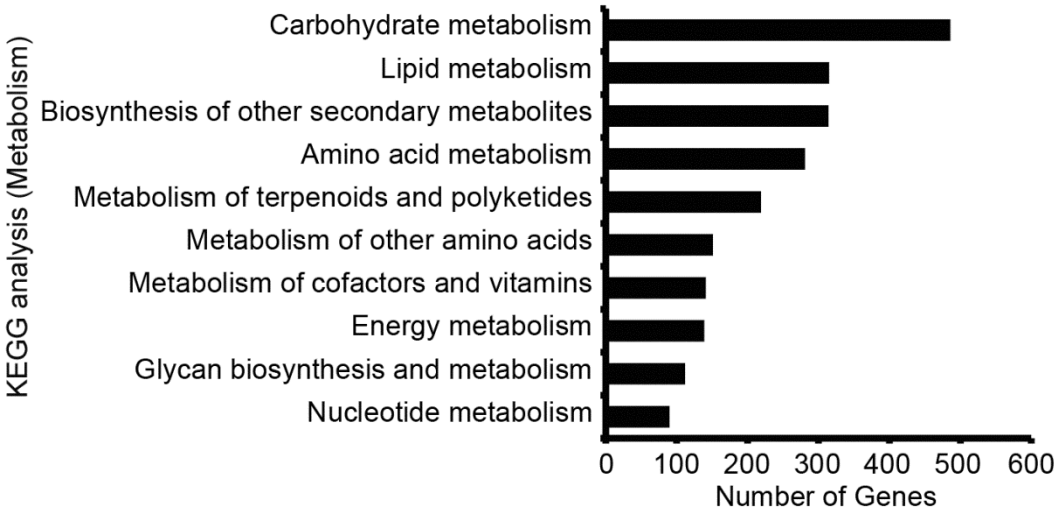


Figure S4. KEGG enrichment analysis of the differentially expressed genes (DEGs) in storage roots of the SSC line plants relative to the wild-type (WT). The KEGG metabolism terms with statistical significance (P <0.01) are shown. Two biological replicates were performed for each genotype.

Figure S5

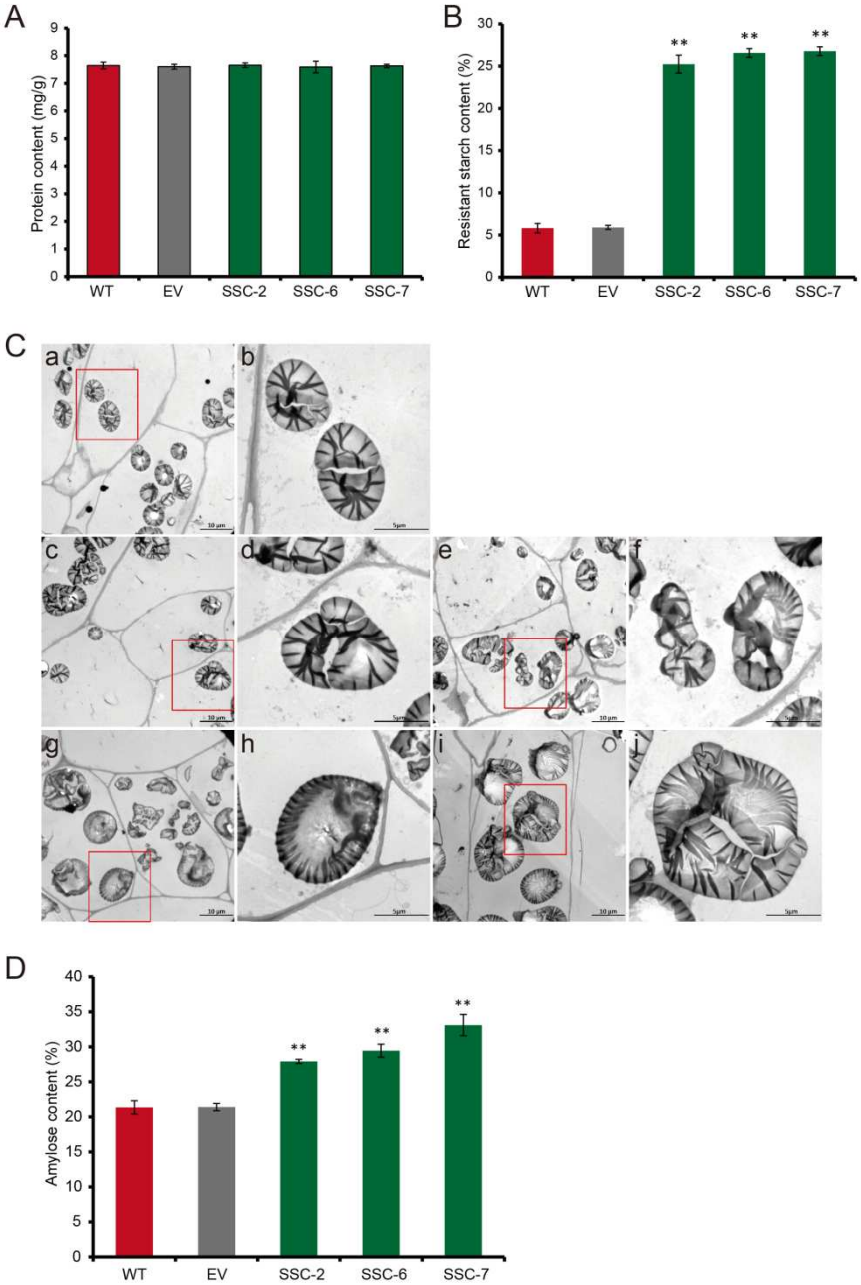


Figure S5. Protein content, resistant starch content, transmission electron microscopic observation, and amylose content of starch granules in storage roots of the wild-type (WT), empty vector (EV), and SSC transgenic cassava lines.

(A) Protein content. **(B)** Resistant starch content. **(C)** Transmission electron microscopy of starch granules in storage roots of the WT (a, b), EV (c, d), SSC-2 (e, f), SSC-6 (g, h) and SSC-7 (i, j) cassava lines. Panels b, d, f, h, and j are higher magnification views of the boxed areas in a, c, e, g, and i. **(D)** Amylose content of starches extracted from mature cassava storage roots. Data in A, B, and D are presented as the means ± SD (n = 3). * and ** indicate a significant difference compared to WT at $P < 0.05$ and <0.01, respectively, determined by Student's *t*-test.

Figure S6

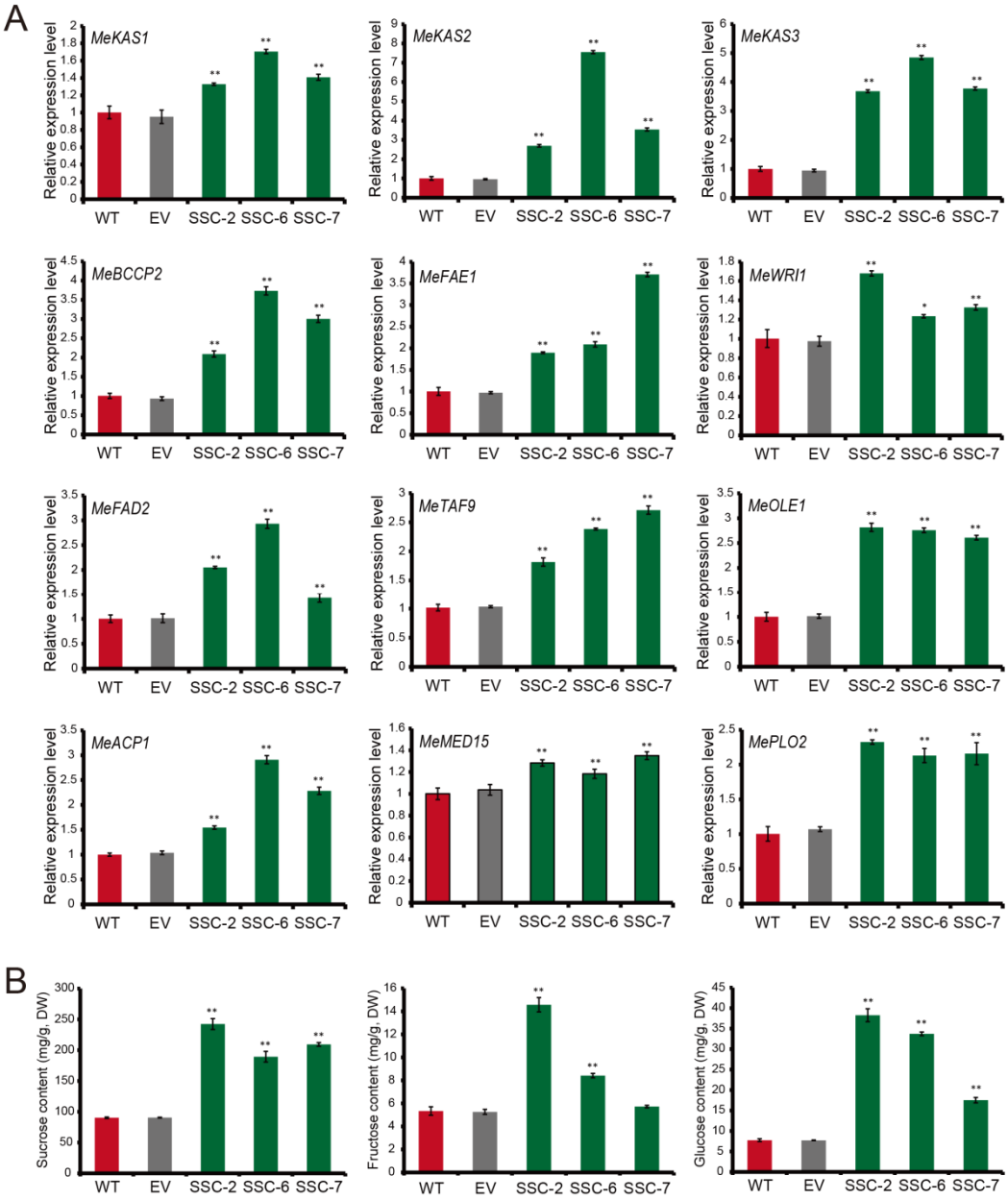


Figure S6. Transcription of lipid biosynthesis genes and sugar contents in storage roots of the wild-type (WT), empty vector (EV), and SSC transgenic cassava lines.

(A) mRNA levels of lipid biosynthesis genes determined by qRT-PCR. Data are presented as the means ± SD (n = 3). * and ** indicate a significant difference compared WT at P <0.05 and <0.01, respectively, determined by Student's *t*-test.

(B) Sucrose, fructose, and glucose contents. Data are presented as the means ± SD (n = 3). * and ** indicate a significant difference versus WT at P < 0.05 and <0.01, respectively, determined by Student's *t*-test.

Figure S7

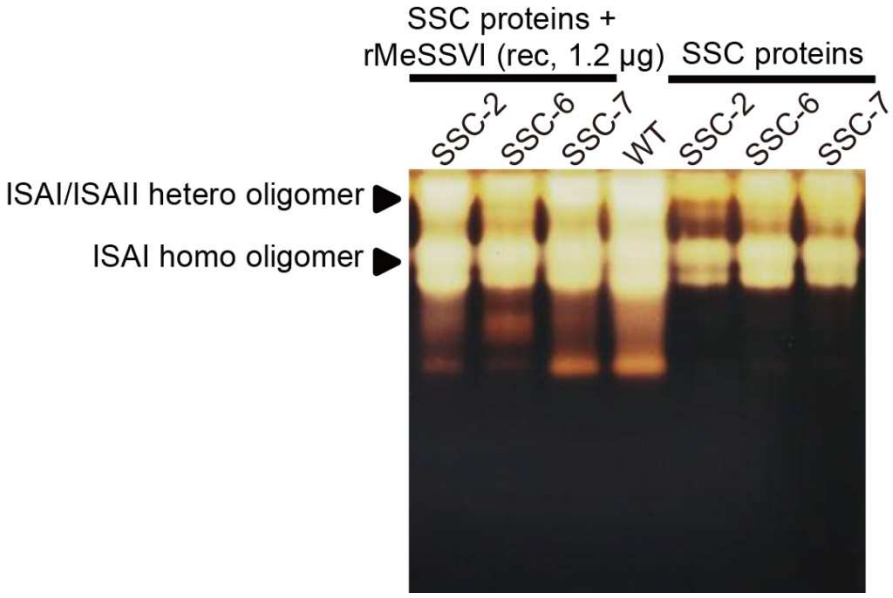


Figure S7. The effect of MeSSVI on the enzyme activities of MeISAI and MeISAII. Enzyme names are shown on the left; material names are shown at the top. Th enzyme extract (1.5 mg/ml) was a soluble protein extract from cassava storage root; rMeSSVI: 1.2 μg recombinant MeSSVI expressed in *E.coli* BL21 (DE3).

Figure S8

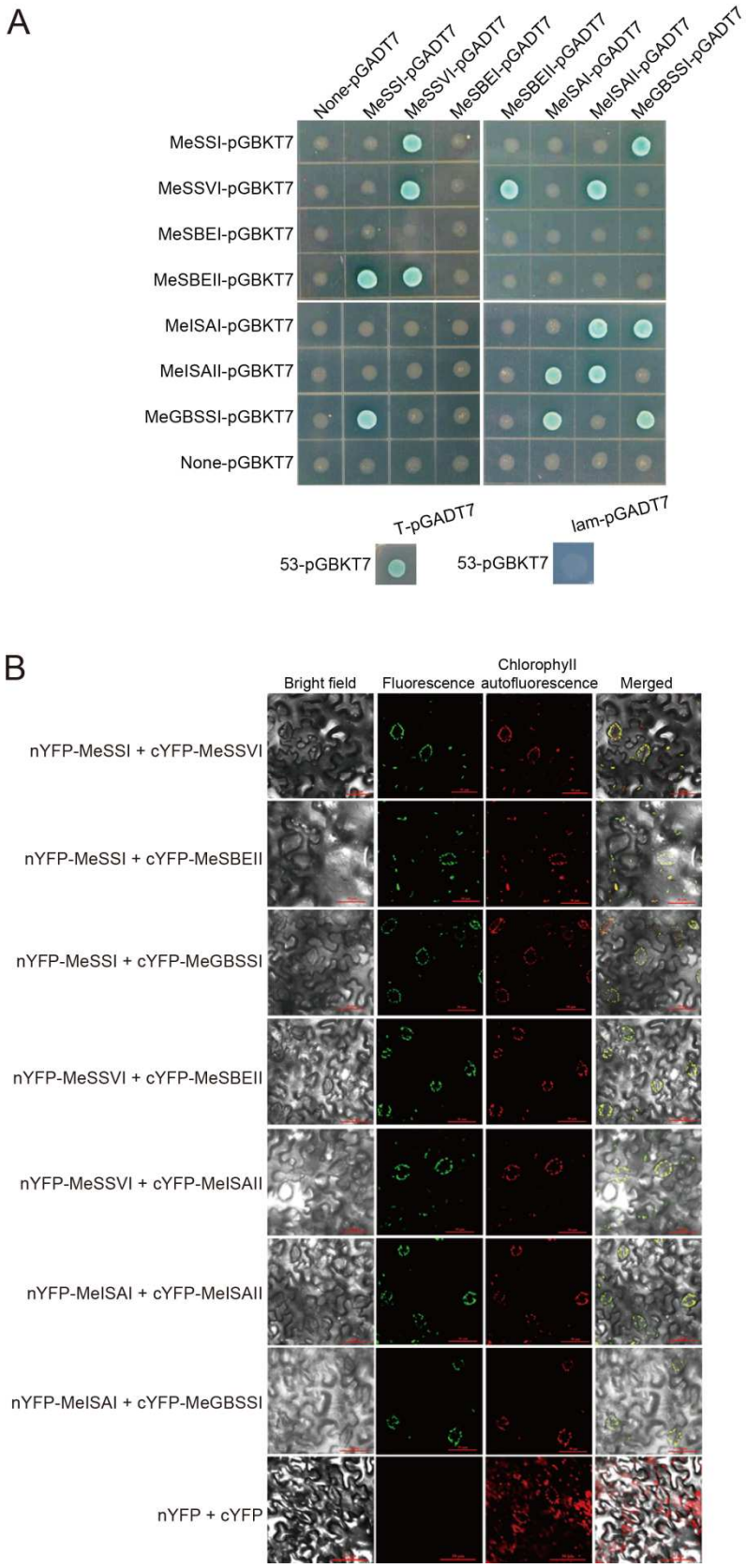


Figure S8. ***In vivo*** **protein-protein interactions between starch biosynthesis enzymes of**

cassava.

(A) Yeast two-hybrid analyses. Empty vectors (None-pGBKT7 and None-pGADT7) were used as the negative control. 53-pGBKT7/T-pGADT7 was used as the positive control, and 53-pGBKT7/lam-pGADT7 was used as the negative control.

(B) Bimolecular fluorescence complementation (BiFC) assays. The empty vectors (nYFP+cYFP) were used as a negative control.