Increased accumulation of recombinant proteins in soybean seeds via the combination strategy of polypeptide fusion and suppression of endogenous storage proteins

Jing Yang
Jilin Academy of Agricultural Sciences
Yuanyu Zhang
Jilin Academy of Agricultural Sciences
Lu Niu
Jilin Academy of Agricultural Sciences
Guojie Xing
Jilin Academy of Agricultural Sciences
Qianqian Zhao
Jilin Academy of Agricultural Sciences
Xiaofang Zhong (xfzhong649@163.com)
Jilin Academy of Agricultural Sciences
Xiangdong Yang
Jilin Academy of Agricultural Sciences

Research Article

Keywords: Soybean, Seed storage protein, RNAi, Polypeptide fusion, GFP fusion

Posted Date: April 28th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1437102/v1

Abstract

Background: Soybean seeds show great potential as a safe and cost-effective host for the large-scale production of biopharmaceuticals and industrially important macromolecules. However, the yields of desired recombinant proteins in soybean seeds are usually lower than the economic threshold (1% of total soluble protein, TSP) for potential commercialization. Here, a recombination strategy of the polypeptide (maize *γ*-zein, or elastin-like polypeptide, ELP) fusions and suppression of endogenous storage proteins (glycinin or conglycinin) *via* RNA interference (RNAi) was exploited to improve yield of the target protein in soybean seeds.

Results: Transgenic soybean plants harboring both the polypeptide-fused GFP and glycinin/conglycinin RNAi expression cassettes were generated and confirmed by molecular analysis. On the glycinin suppression background, average accumulation levels of recombinant zein-GFP and GFP-ELP proteins were significantly increased as compared to that of the unfused GFP, reaching to 2.69% and 2.30% TSP in the immature seeds, and 3.49% and 4.31 % TSP in the mature seeds, respectively. However, no significant differences were detected in the glycinin or conglycinin suppression backgrounds for the same polypeptide fusion constructs, though suppression of one of the storage proteins in soybean seeds led to a significant increase of the other. Moreover, the increases in the recombinant protein yield did not affect the total protein content or the protein/oil ratio in soybean seeds

Conclusions: These results demonstrate that the fusion of the foreign protein with polypeptide tags together with the depletion of endogenous storage proteins can significantly increase the expression of recombinant proteins without affecting the total protein content or the protein/oil ratio in soybean seeds.

Background

Over the past few decades, many plant systems have been explored as expression hosts for the cost-effective and scalable production of recombinant pharmaceutical proteins and industrial macromolecules [1, 2]. Seed-based systems are particularly attractive because of their relatively high protein/mass ratios favorable for downstream processing, stable biochemical environment optimized for long-term storage, and abundant seed-specific organelles such as protein storage vacuoles (PSVs), or endoplasmic reticulum (ER)-derived protein bodies (PBs) specialized for protein accumulation [3, 4]. Compared with the existing expression systems, soybean seeds provide a practical and promising production platform, as they naturally accumulate large amounts of protein (approximately 40% of dry mass) and represent one of the most protein-rich plant seeds. Such a high protein/mass ratio even makes it possible to produce over 1 mg of recombinant proteins in a single soybean seed, which is a much higher yield than could be reached in the other plant or non-plant host systems [4]. Furthermore, high stability at ambient temperature, a simple proteome, and relative homogeneity of soybean seeds can greatly facilitate the downstream processing and purification of recombinant proteins [5–7]. Additional advantages such as low production costs, easy scalability, well-established processing procedures, and

the possibility of oral delivery of the expressed proteins and vaccines make soybean seeds a very attractive system for the industrial-scale production of recombinant proteins [3, 4].

To date, a number of recombinant pharmaceutical proteins have been successfully expressed in the seeds of transgenic soybeans [6, 8–12]. However, the accumulation levels of some recombinant proteins in soybean seeds are usually less than 1% total soluble protein (TSP) of economic threshold [13, 14]. Significant efforts have been made to improve the yield of target proteins in soybean seeds up to the level acceptable for practical application, including the use of constitutive or seed-specific promoters, codon optimization, and subcellular targeting of the synthesized proteins by signal peptides, ER retention signals (H/KDEL), or polypeptide partners [8, 9, 11, 15]. The selection of an appropriate promoter is essential to achieve the optimal transcription level of heterologous proteins, and strong soybean seed-specific promoters such as *β*-conglycinin and glycinin are usually used to increase the yield of foreign proteins expressed in soybean seeds [8, 9, 11].

The targeting of the expressed proteins to the subcellular compartments such as the ER or PSVs can not only influences their high-level accumulation in cells, but also protein stability and post-translational modifications [3, 4]. Increased accumulation of recombinant proteins in soybean and other crop seeds has been achieved through manipulations of the secretory pathway and protein targeting to different subcellular locations [8, 9, 11, 15–17]. Among the strategies used to promote protein accumulation, fusion with small polypeptides such as 27 kDa maize γ-zein, 13.56 kDa human elastin-like polypeptide (ELP), and 12 kDa fungal hydrophobin I have shown great potential [15, 18–22]. Our previous studies in soybean seeds indicated that fusion with *γ*-zein or ELP could significantly increase the accumulation of recombinant proteins in ER-derived vesicles [15], which not only provided protection from proteolysis but also facilitated target protein purification [11, 15].

Another factor to consider in using seed-based expression systems is endogenous storage proteins such as prolamins or glutelins, which hamper the accumulation of foreign proteins by competing for the biosynthesis and transport pathways and deposition space [23]. Yang et al. [26] have reported that a decrease in the expression of rice seed prolamins or glutelins increases the accumulation of recombinant human IL-10 by approximately 3-fold. Complete inhibition or reduction of endogenous storage protein expression could even lead to changes in the localization of foreign proteins in rice seeds [28] or the formation of novel ER-derived PBs in soybean seeds [15, 29, 30].

The main storage proteins accumulated in specialized PSVs of soybean seeds are 11S- and 7S-globulins (glycinin and conglycinin, respectively), which account for approximately 70% of soybean storage proteins [29]. It is reasonable to redirect the intrinsically high transcriptional and translational activity in soybean seeds towards heterologous proteins by suppressing the synthesis of these storage proteins. In the present study, a recombination strategy of the polypeptide (γ-zein or ELP) fusions and specific suppression of endogenous storage proteins (glycinin or conglycinin) by RNA interference (RNAi) was exploited to explore its potential in increasing the accumulation of foreign proteins in soybean seeds. Our results indicated that the suppression of glycinin or conglycinin synthesis led to a significant increase in

the expression of zein- or ELP-fused green fluorescent protein (GFP) in soybean seeds without affecting their total protein content and protein/oil ratio.

Results

Generation and molecular analysis of transgenic soybeans

To suppress the expression of endogenous glycinin or conglycinin in soybean seeds, two tandem RNAi fragments based on conservative sequences of five glycinin subunit or three conglycinin subunit-encoding genes were synthesized and used for the construction of the expression vectors (Additional file 1: Fig. S1, S2; Fig. 1). To optimize the transcription of the foreign genes, two soybean seed-specific promoters, *BCSP* and *Pgy1*, were used to drive the expression of the *gfp* and *Gy/ Cgy* RNAi fragments, respectively. A series of transformation experiments produced 84 transgenic soybean plants, including 12 Gyless/zein-GFP, 19 Gyless/GFP-ELP, 21 Gyless/GFP plants, 13 Cgyless/zein-GFP, and 19 Cgyless /GFP-GFP, as shown by LibertyLink® strip analysis (Additional file 1: Fig. S3). The presence of foreign *gfp* and RNAi fragments was further confirmed by PCR, with expected sizes of amplified fragments observed in these transgenic plants, but absent in the non-transformed (Nt) controls (Additional file 1: Fig. S4). Transgenic soybean lines from each construct were further advanced to the subsequent generations until homozygous lines were obtained using glufosinate spraying and PCR detection.

The relative mRNA expression of *gfp* in T_3 transgenic soybean seeds was determined by qRT-PCR. Similar expression levels of the foreign genes were also observed at the same time point of seed development in some transgenic plants (Fig. 2a). To avoid the influence of differential mRNA expression on the recombinant protein accumulation in seeds, only transgenic lines with similar expression levels were selected for further expression analysis, including three Gyless/zein-GFP (G45, G54, and G59), three Gyless/GFP-ELP (A44, A58, and A59), three Gyless/GFP (B14, B43, and B66), two Cgyless/zein-GFP (F24 and F32) plants, and three Cgyless/GFP-ELP (J41, J47, and J53).

Western blotting was further conducted to detect the expression of the recombinant proteins in soybean seeds. The results showed that the bands representing zein-GFP, GFP-ELP, and GFP with the expected sizes of 26.8, 51.8, and 40.5 kDa, were observed in the corresponding transgenic seeds, confirming the accurate translation of these recombinant proteins (Fig. 2b-2f). However, the cleavage of the fusion proteins was also detected, as evidenced by the 26.8-kDa band corresponding to a GFP monomer detected in both zein-GFP and GFP-ELP seed samples (Fig. 2c-2f), which was consistent with previous reports [15].

Suppression of glycinin and conglycinin expression in the transgenic soybean seeds

The expressions of glycinin and conglycinin family proteins were analyzed in immature seeds of selected transgenic plants at 45 DAF using primers specific for five glycinin (*Gy1, Gy2, Gy3, Gy4*, and *Gy5*) and

three conglycinin (*Cgy1*, *Cgy2*, and *Cgy3*) subunit-encoding genes. The results showed that in the Gyless transgenic plants (B66, G45, and A58), the relative mRNA expressions of glycinin subunits (*Gy1–Gy5*) were markedly downregulated with 3.16–78.46 fold decreases compared with the Nt plants (Fig. 3a). Similar to glycinin, transcriptional suppression was also observed for three conglycinin subunits (*Cgy1–Cgy3*) in the Cgyless transgenic plants (F24 and J41) with 3.75–8.42 fold decreases compared with the Nt control (Fig. 3b).

Next, we analyzed the protein expression in immature seeds of transgenic plants at 45 DAF. The results showed that in Gyless/zein-GFP and Gyless/GFP-ELP plants carrying the *Gy* RNAi cassette, the glycinin proteins had a significantly decreased accumulation compared with Nt control or DG44 and DE9 plants, which harbored the same *zein-gfp* and *gfp-elp* cassettes but not the *Gy* RNAi cassette in our previous study [15] (Fig. 3c). Meanwhile, conglycinin proteins were visibly increased as compared with the Nt plants in these soybean seeds with the suppressed glycinin expression. A similar effect was observed in Cgyless/zein-GFP and Cgyless/GFP-ELP plants with the downregulated conglycinin expression. In this case, the accumulation of conglycinin proteins was markedly decreased compared with Nt control or DG44 and DE9 plants, whereas that of glycinin proteins was markedly increased (Fig. 3c). Taken together, these results indicated that the RNAi-mediated suppression of one of the storage proteins (glycinin or conglycinin) significantly inhibited its transcription and translation, and also led to a significant increase of the other endogenous storage proteins in soybean seeds.

Influence of RNAi-mediated suppression of glycinin or conglycinin synthesis on the accumulation of recombinant polypeptides

The increased accumulation of conglycinin to compensate for RNAi-mediated glycinin suppression suggested that a concomitant increase may be expected for the expression of foreign proteins. As expected, the results revealed a marked increase in the accumulation of the recombinant fusion proteins in the Gyless immature soybean seeds (Table 1). The average accumulation levels of zein-GFP and GFP-ELP were 2.69% and 2.30% TSP, which indicated 2.2- and 7.7-fold increases compared to those in control DG44 and DE9 plants, respectively (Table 1). The accumulation of unfused GFP (1.02% TSP) on the Gyless background also showed a marked increase compared to its counterpart DF8 (0.08% TSP) (Table 1). Similar observations were found in mature soybean seeds, which the accumulation of zein-GFP, GFP-ELP, and unfused GFP on the Gyless background reached the average levels of 3.49%, 4.31%, and 1.53% TSP, representing 1.68, 4.89, and 2.68 times higher than those in their counterparts DG44, DE9, and DF8, respectively (Table 1).

Table 1
Recombinant protein content in soybean seeds with the suppressed expression of glycinin and conglycinin

Transgenes	Lines	Immature seeds		Mature seeds	
		% TSP	Mean ± SD	% TSP	Mean ± SD
Gyless/zein-GFP	G45	2.83	2.69 ± 0.20	3.15	3.49 ± 0.33
	G54	2.78		3.80	
	G59	2.46		3.51	
Cgyless/zein-GFP	F24	2.72	2.80 ± 0.12	3.22	4.49 ± 1.80
	F32	2.89		5.76	
Zein-GFP	DG44	/	1.20 ± 0.22	/	2.08 ± 0.21
Gyless/GFP-ELP	A44	1.84	2.30 ± 0.55	2.78	4.31 ± 1.50
	A58	2.90		4.37	
	A59	2.14		5.78	
Cgyless/GFP-ELP	J41	2.96	3.04 ± 0.08	3.66	5.03 ± 1.9
	J47	3.04		5.74	
	J53	3.12		5.69	
GFP-ELP	DE9	/	0.30 ± 0.12	/	0.88 ± 0.19
Gyless/GFP	B14	1.03	1.02 ± 0.08	1.94	1.53 ± 0.37
	B43	0.94		1.21	
	B66	1.09		1.45	
GFP	DF8	/	0.08 ± 0.03	/	0.57 ± 0.18

Total protein was extracted from immature seeds (45 DAF) and mature dry seeds, and recombinant GFP expression was quantified by ELISA using purified GFP as the standard.

Furthermore, compared with unfused GFP on the Gyless background, the average accumulation levels of zein-GFP and GFP-ELP were also increased by 2.28- and 3.29-fold, respectively (Table 1). These results indicated that both the suppression of endogenous glycinin and the fusion with zein or ELP increases the accumulation of the recombinant protein in soybean seeds, and the depletion of glycinin in soybean seeds could be compensated for by the increased abundance of both conglycinin and foreign proteins. Moreover, there appeared to be no significant difference between zein and ELP in terms of influence on the accumulation of fused GFP in this study. In Cgyless transgenic plants with suppressed conglycinin expression, the average accumulation levels of zein-GFP and GFP-ELP in immature seeds (2.80% and 3.04%TSP, respectively) showed no significant difference with those in their Gyless counterparts

(Table 1). Similar trends were observed in mature soybean seeds, in which the accumulation levels of zein-GFP and GFP-ELP were 4.49% and 5.03% TSP. Taken together, these results indicated that the suppression of both endogenous glycinin or conglycinin could lead to a marked increase in the accumulation of recombinant proteins in soybean seeds.

Increased accumulation of recombinant proteins did not affect total protein and oil contents in soybean seeds

Total protein and oil accumulation in transgenic soybean seeds showed that despite protein profile changes by introduction of foreign protein and exchanges of glycinin/conglycinin, no significant changes were observed in the total protein and protein/oil ratio in seeds between the transgenic and non-transformed plants (Table 2). The results were consistent with the previous reports that limited plasticity of total protein content resulted from the loss of one major storage proteins compensated by the increase in accumulation of foreign and other heterologous proteins, possibly through the rebalancing of protein synthesis in the remodelled seeds to maintain stable levels of protein accumulation [15, 25, 29].

Table 2
Protein and oil content (%) in mature soybean seeds

Transgenes	Lines	Total protein (%)		Protein/oil ratio	
		%	Average(%)		Average (%)
Gyless/zein-GFP	G45	40.76 ± 0.15	41.29	2.19	2.19
	G54	41.55 ± 0.25		2.22	
	G59	41.56 ± 0.17		2.17	
Cgyless/zein-GFP	F24	40.74 ± 0.73	40.67	2.08	2.07
	F32	40.60 ± 0.06		2.05	
Gyless/GFP-ELP	A44	39.93 ± 0.65	40.95	2.00	2.18
	A58	41.19 ± 1.18		2.22	
	A59	41.72 ± 0.44		2.30	
Cgyless/GFP-ELP	J41	41.61 ± 0.65	41.39	2.08	2.07
	J47	41.16 ± 0.77		2.07	
	J53	41.08 ± 0.35		2.06	
Gyless/GFP	B14	40.25 ± 0.04	40.55	2.13	2.17
	B43	40.24 ± 1.07		2.08	
	B66	41.15 ± 0.39		2.30	
Nt control	P03	40.60 ± 0.02	/	2.06	/
Total protein and oil contents were determined based on dry weight.					

Discussion

The applicability of plant systems to large-scale production of the pharmaceutical and industrial products of significance critically depends on the high protein yield and the ease of downstream recovery and purification [31]. Our previous studies have shown that fusion with ELP or γ-zein significantly increases recombinant protein accumulation in soybean seeds [15], which might result from increased transportation and deposition in cell compartments such as PBs or PSVs, and protection from proteolytic digestion. Furthermore, ELP and γ-zein can be used as tags to facilitate the recovery and purification of heterologous proteins from plant materials. In this study, we further explored the capability of the combination strategy of specific suppression of endogenous storage proteins (glycinin or conglycinin) and polypeptide fusion in increasing the yield of recombinant protein production in soybean seeds. Our results showed that the accumulation of the foreign polypeptide (zein or ELP) fused GFP proteins (zein-

GFP and GFP-ELP) was significantly increased in soybean seeds with suppressed glycinin or conglycinin biosynthesis compared to their counterparts. These results indicate that both depletion of endogenous storage proteins (glycinin or conglycinin) and polypeptide fusion contribute to the increase in the yield of foreign protein accumulation, suggesting a potential of the developed combination strategy for cost-effective large-scale production of industrially important proteins in soybean seeds.

An interesting question is the influence of endogenous proteins on the accumulation of recombinant proteins in plant seeds, especially in soybean seeds which naturally accumulate higher amounts of proteins than most of other crops. It has been reported that the mutation or suppression of endogenous storage proteins in seeds increase the expression of other seed proteins at the transcriptional and translational levels [32–34]. Kawakatsu et al. [34] have shown that in transgenic rice seeds, the downregulation of one or several storage proteins such as glutelins, prolamins, and globulins significantly affects the expression of other seed proteins and PB formation, which could be due to a compensatory mechanism to maintain seed protein homoeostasis. In this study, we also observed that the suppression of one major storage protein (glycinin) resulted in the upregulation of another endogenous protein (conglycinin) in soybean seeds. The observed compensatory effect and/or the extra deposition space created by the loss of an endogenous protein may also promote the accumulation of foreign proteins in seeds [25].

However, a question arises whether the increased accumulation of recombinant proteins could affect the total protein and oil content in soybean seeds. It has been shown that the suppression of the α and α' subunits of β-conglycinin does not influence the total oil and protein content and ratio of soybean seeds, and the decrease in β-conglycinin protein can be compensated by increased accumulation of glycinin [15, 29]. Consistent with previous reports, our results indicated that no significant differences were observed in the total protein content and the protein/oil ratio among transgenic and non-transformed plants, suggesting that the loss of one endogenous protein was compensated by increased synthesis of foreign and other endogenous storage proteins, so that the total protein content in soybean seeds was stably maintained despite the changes in protein composition.

Conclusions

We developed an effective strategy for the expression of foreign proteins in soybean seeds based on the polypeptide fusions of the protein of interest and RNAi-mediated downregulation of endogenous storage proteins, which resulted in a significant increase in the accumulation of the recombinant protein without affecting the total protein content or the protein/oil ratio.

Materials And Methods

Vector constructs for soybean genetic transformation

Expression vectors were constructed based on the binary vectors pTF101-zein-GFP and pTF101-GFP-ELP carrying the GFP-encoding gene fused with zein- or ELP-coding sequences under the control of soybean seed-specific promoter of *β*-conglycinin alpha subunit (*BCSP*) as described previously [15]. Another soybean seed-specific promoter *Pgy1* of the glycinin subunit Gy1 was amplified by polymerase chain reaction (PCR) from the pCB2004B-Gy1 plasmid and inserted into the pHANNIBALG RNAi vector at *Sac*I/*Xho*I sites to replace the native *CaMV 35S* promoter. A 723-bp tandem RNAi fragment *Gy* was synthesized based on the conserved sequences of five genes encoding soybean glycinin subunits *A1aB2* (*Gy1*, X15121), *A2B1a* (*Gy2*, NM_001248881), *A1bB1b* (*Gy3*, X15123), *A5A4B3* (*Gy4*, AB195712), and *A3B4* (*Gy5*, FJ599666) (Additional file 1: Fig. S1), and sequentially inserted at the *Xho*I/*Kpn*I and *Hind*III/*Xba*I sites of the pHANNIBALG-Pgy1 vector in the sense and antisense directions, respectively. The whole *Gy* RNAi cassette, including the *Pgy1* promoter, two inverted *Gy* fragments separated by the *pdk* intron, and the octopine synthase terminator *ocs*, was then ligated into the corresponding binary vectors to generate constructs pGyless/zein-GFP and pGyless/GFP-ELP, respectively (Fig. 1). For construction of the control vector pGyless/GFP, the Gy RNAi cassette was ligated into the pTF101-GFP vector, which contained the unfused *gfp* under the *BCSP* promoter as described previously [15].

All three vectors contained the phosphinothricin acetyltransferase marker gene (*bar*) for the selection of transgenic plants based on glufosinate resistance. Similarly, a 721-bp tandem RNAi fragment *Cgy* was synthesized based on the conserved sequences of three genes encoding soybean conglycinin subunits *Cgy1* (GU723691), *Cgy2* (AB237643), and *Cgy3* (NM_001249943) (Additional file 1: Fig. S2), and sequentially inserted at the *Xho*I/*Kpn*I and *Hind*III/*Xba*I sites of the pHANNIBALG-Pgy1 vector as described above. The whole *Cgy* RNAi cassette was then subcloned into the corresponding binary vectors to obtain pCgyless/zein-GFP and pCgyless/GFP-ELP, respectively (Fig. 1).

Generation of transgenic soybean plants

In this study, the acceptor material was soybean Cultivar P03, belonging to Jilin Academy of Agricultural Sciences, and agreed to carry out scientific experiments.For genetic transformation, mature seeds of soybean cultivar P03 were sterilized with chlorine gas in a sealed chamber for 12–16 h and imbibed with three layers of filter paper soaked in sterile water for 12 h. Then, the seed coats and primary shoots were removed for explant preparation. Transformation with *Agrobacterium tumefaciens* EHA101, induction and selection of resistant shoots, elongation, and rooting of shoots were performed as described previously [35]. The regenerated plantlets were transferred into a greenhouse with controlled temperature (25°C) and light/dark cycle (16-h/8-h) for growth and seed setting. The primary transformants were detected using the LibertyLink® strip (EnviroLogix Inc., Portland, ME, USA) and PCR with gene-specific primers (Table S1). The progenies of transgenic plants were further screened by glufosinate (1000 mg/L) spraying to obtain homozygous transgenic lines.

Transcription analysis of transgenic soybean seeds

To determine the expression of foreign genes in seeds at the transcription level, quantitative reverse transcription PCR (qRT-PCR) was carried out with *GFP*-specific primers (Table S1). Briefly, total RNA was

extracted from immature seeds (~ 10 mm in length) at 45 days after flowering (DAF) using the EasyPure Plant RNA Kit (TransGen Biotech, Beijing, China). First-strand cDNA was synthesized using TransScript® One-Step gDNA Removal and cDNA Synthesis SuperMix (TansGen Biotech). qRT-PCR was performed using TransStart® Top Green qPCR SuperMix (TansGen Biotech) according to the manufacturer's protocol under the following reaction conditions: uracil-*N*-glycosylase treatment at 50°C for 2 min, denaturation at 95°C for 10 min, and 35 cycles of denaturation at 95°C for 30 s, annealing at 60°C for 30 s, and extension at 72°C for 30 s. We also analyzed the transcription of soybean glycinin (*Gy1*, *Gy2*, *Gy3*, *Gy4*, and *Gy5*) and conglycinin (*Cgy1*, *Cgy2*, and *Cgy3*) genes using specific primers (Table S1). The native soybean gene *Gmactin6* (GenBank No. NM_001289231) was amplified as the internal control. Relative gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method as previously described [15]. All qRT-PCR experiments were conducted with three biological replicates for each sample.

Protein expression and quantification analysis

The expression of recombinant GFP proteins was analyzed by western blotting and enzyme-linked immunosorbent assay (ELISA). The accumulation of the fusion proteins in transgenic soybean plants with the inhibited glycinin synthesis was analyzed using ELISA and transgenic soybean plants without such inhibition (DF8, DG44, and DE9) were used as controls as described in our previous study [15]. Immature seeds at 45 DAF were collected from the transgenic plants, grounded to fine powder in liquid nitrogen with a mortar and pestle. Each sample (100 mg) was used for total soluble protein extraction performed as previously described [15]. Total protein was quantified using the BCA Protein Quantification Kit (Yeasen Biotechnology, Wuhan, China) according to the manufacturer's protocol. Proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in 12% gels, transferred to PVDF membranes (GE Healthcare, USA), and subjected to western blotting using primary anti-GFP antibodies (1:500 dilution; Abcam, Cambridge, UK) and secondary horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG (1:5,000 dilution; Abcam) as previously described [15].

For quantification of recombinant protein expression, the total protein extracts of immature (45 DAF) and dry mature soybean seeds were subjected to ELISA. Blocking, incubation with primary anti-GFP antibodies (1:500 dilution) and secondary HRP-conjugated secondary antibodies (1:500 dilution), washing, and reaction with the substrate 3,3,5,5-tetramethyl benzidine (TMB) were performed as previously described [15]. The reaction was terminated with 2 N H_2SO_4, and the optical density (OD) was determined at 450 nm using a plate reader (ELx800; BioTek, Winooski, VT, USA).

Analysis of total protein and oil content in soybean seeds

The protein content of soybean seeds was measured as described previously [36]. Briefly, mature soybean seeds were ground in a mortar, and 200 mg of the powder was used for analysis. Total nitrogen was determined in a LECO CHN 2000 analyzer (LECO, St. Joseph, MI, USA) as described by Hwang et al. [37]. Protein content was calculated as the amount of nitrogen multiplied by 6.25 based on the dry weight basis. Seed oil was extracted with methanol:chloroform (2:1 v/v), and total fatty acids were quantified by

gas chromatography-mass spectrometry (GC-MS) after methyl esterification using 5% sulfuric acid in methanol. The oil content was calculated as the percentage of dry weight [36].

Statistical analysis

All statistical analyses were performed using SPSS software v.17.0 (SPSS Inc., Chicago, IL, USA). The data were presented as the mean ± standard deviation (SD) of three biological replicates.

Declarations

Acknowledgments

This study was financially supported by the Jilin Provincial Agricultural Science & Technology Innovation Project (KYJF2021ZR005, CXGC2021TD015) and the National Natural Science Foundation of China (31671764). These funds played an important role in experimental reagents, consumables, and also provided strong support for preparing the manuscript. We thank the Editage (www.editage.cn) for English language editing of the manuscript.

Author Contribution

Yang X. and Zhong X. designed the experiments. Yang J., Zhang Y., and Niu L. performed most of the experiments, analyzed the data and drafted the manuscript. Xing G. and Zhao Q. conducted soybean transformation experiments. All authors have read and approved the final version of the manuscript.

Availability of data and materials

All the datasets used and/or analyzed throughout this study can be available from the corresponding authors on reasonable request.

Ethics approval and consent to participate

The contents of this study are not contrary to human ethics and are only used for scientific research, all methods used in the study were conducted in accordance with relevant institutional/national/international standards

Consent for publication

Not applicable

Competing interests

The authors claim no conflict of interest regarding publication of this paper.

Statement

The data we use are publicly available, and the sources have been noted in the article. No data from the store is required. All authors agree to publish without conflict of interest.

Funding

This research was carried out by the Innovation Engineering Basic Scientific Research Fund of Jilin Academy of Agricultural Sciences (Project Code: KYJF2021ZR005)

References

1. Ma JKC, Christou P, Chikwamba R, Haydon H, Paul M, Ferrer MP, Ramalingam S, Rech E, Rybicki E, Wigdorowitz A, Yang DC, Thangaraj H (2013) Realising the value of plant molecular pharming to benefit the poor in developing countries and emerging economies. Plant Biotechnol J 11:1029–1033
2. Stoger E, Fischer R, Moloney M, Ma JK (2014) Plant molecular pharming for the treatment of chronic and infectious diseases. Annu Rev Plant Biol 65:743–768
3. Vianna GR, Cunha NB, Murad AM, Rech EL (2011) Soybeans as bioreactors for biopharmaceuticals and industrial proteins. Genet Mol Res 10:1733-1752
4. Hudson LC, Garg R, Bost KL, Piller KJ (2014) Soybean seeds: a practical host for the production of functional subunit vaccines. Biomed Res Int 2014:340804. doi:10. 1155/2014/340804
5. Müntz K (1998) Deposition of storage proteins. Plant Mol Biol 38:77-99
6. Cunha NB, Murad AM, Ramos GL, Maranhão AQ, Brígido MM, Araújo AC, Lacorte C, Aragão FJ, Covas DT, Fontes AM (2011a) Accumulation of functional recombinant human coagulation factor IX in transgenic soybean seeds. Transgenic Res 20:841-855
7. Morandini F, Avesani L, Bortesi L, Van DB, De WK, Arcalis E, Bazzoni F, Santi L, Brozzetti A, Falorni A (2011) Non-food/feed seeds as biofactories for the high-yield production of recombinant pharmaceuticals. Plant Biotechnol J 9:911-921
8. Ding SH, Huang LY, Wang YD, Sun HC, Xiang ZH (2006) High-level expression of basic fibroblast growth factor in transgenic soybean seeds and characterization of its biological activity. Biotechnology Letters 28:869-875
9. Moravec T, Schmidt MA, Herman EM, Woodfordthomas T (2007) Production of *Escherichia coli* heat labile toxin (LT) B subunit in soybean seed and analysis of its immunogenicity as an oral vaccine. Vaccine 25:1647-1657
10. Yamada Y, Nishizawa K, Yokoo M, Zhao H, Onishi K, Teraishi M, Utsumi S, Ishimoto M, Yoshikawa M (2008) Anti-hypertensive activity of genetically modified soybean seeds accumulating novokinin. Peptides 29:331-337
11. Cunha NB, Murad AM, Cipriano TM, Araújo ACG, Aragão FJL, Leite A, Vianna GR, Mcphee TR, Souza GHMF, Waters MJ (2011b) Expression of functional recombinant human growth hormone in transgenic soybean seeds. Transgenic Res 20:811-826

12. O'Keefe BR, Murad AM, Vianna GR, Ramessar K, Saucedo CJ, Wilson J, Buckheit KW, Cunha NB, Araújo ACG, Lacorte CC (2015) Engineering soya bean seeds as a scalable platform to produce cyanovirin-N, a non-ARV microbicide against HIV. Plant Biotechnol J 13:884-892
13. Sirko A, Vanek T, Gora-Sochacka A, Redkiewicz P (2011) Recombinant cytokines from plants. Int J Mol Sci 12:3536–3552
14. Kusnadi AR, Nikolov ZL, Howard JA (2015) Production of recombinant proteins in transgenic plants: practical considerations. Biotechnol Bioeng 56:473-484
15. Yang J, Xun HW, Niu L, He HL, Cheng YQ, Zhong XF, Zhao QQ, Xing GJ, Liu JF, Yang XD (2021) Elastin-like polypeptide and γ-zein fusions significantly increase recombinant protein accumulation in soybean seeds. Transgenic Res 30:675-686
16. He Y, Ning T, Xie TT, Qiu QC, Zhang LP, Sun YF, Jiang DM, Fu K, Yin F, Zhang WJ, Shen L, Wang H, Li JJ, Lin QS, Sun YX, Li HZ, Zhu YG, Yang DC (2011) Large-scale production of functional human serum albumin from transgenic rice seeds. Proc Natl Acad Sci USA 108:19078-19083
17. Zhang DS, Lee HF, Pettit SC, Zaro JL, Huang N, Shen WC (2012) Characterization of transferrin receptor-mediated endocytosis and cellular iron delivery of recombinant human serum transferrin from rice (*Oryza sativa* L.). BMC Biotechnol 12:92. doi: 10.1186/1472-6750-12-92
18. Scheller J, Leps M, Conrad U (2006) Forcing single-chain variable fragment production in tobacco seeds by fusion to elastin-like polypeptides. Plant Biotechnol J 4:243-249
19. Torrent M, Llompart B, Lasserre-Ramassamy S, Llop-Tous I, Bastida M, Marzabal P, Westerholm-Parvinen A, Saloheimo M, Heifetz PB, Ludevid MD (2009) Eukaryotic protein production in designed storage organelles. BMC Biol 7:5. doi: 10.1186/1741-7007-7-5
20. Hofbauer A, Peters J, Arcalis E, Rademacher T, Lampel J, Eudes F, Vitale A, Stoger E (2014) The induction of recombinant protein bodies in different subcellular compartments reveals a cryptic plastid-targeting signal in the 27-kda γ-zein sequence. Front Bioeng Biotechnol 2:67. doi: 10.3389/fbioe.2014.00067
21. Phan HT, Hause B, Hause G, Arcalis E, Stoger E, Maresch D, Altmann F, Joensuu J, Conrad U (2014) Influence of elastin-like polypeptide and hydrophobin on recombinant hemagglutinin accumulations in transgenic tobacco plants. PLoS One 9: e99347. doi: 10.1371/journal.pone.0099347
22. Davide M, Marrano C A, Bhakti P, Dario M, Marc T, Luca E, Stoger E, Faoro F, Pedrnzzini E, Vitale A (2018) Maize 16kd γ-zein forms very unusual disulfide-bonded polymers in the endoplasmic reticulum: implications for prolamin evolution. J Exp Bot 69:5013-5027
23. Hegedus DD, Baron M, Labbe N, Coutu C, Lydiate D, Lui H, Rozwadowski K (2014) A strategy for targeting recombinant proteins to protein storage vacuoles by fusion to *Brassica napus* napin in napin-depleted seeds. Protein Expression & Purification 95:162-168
24. Goossens A, Montagu MV, Angenon G (1999) Co-introduction of an antisense gene for an endogenous seed storage protein can increase expression of a transgene in Arabidopsis thaliana seeds. FEBS Letters 456:160–164

25. Schmidt MA, Herman EM (2008) Proteome rebalancing in soybean seeds can be exploited to enhance foreign protein accumulation. Plant Biotechnol J 6:832–842
26. Yang L, Hirose S, Takahashi H, Kawakatsu T, Takaiwa F (2012) Recombinant protein yield in rice seed is enhanced by specific suppression of endogenous seed proteins at the same deposit site. Plant Biotechnology J 10:1035-1045
27. Kim YM, Lee JY, Lee T, Lee YH Lim SH (2012) The suppression of the glutelin storage protein gene in transgenic rice seeds results in a higher yield of recombinant protein. Plant Biotechnology Rep 6:347-353
28. Kurokawa S, Kuroda M, Mejima M, Nakamura R, Takahashi Y, Sagara H, Takeyama N, Satoh S, Kiyono H, Teshima R, Masumura T, Yuki Y (2014) RNAi-mediated suppression of endogenous storage proteins leads to a change in localization of overexpressed cholera toxin B-subunit and the allergen protein RAG2 in rice seeds. Plant Cell Rep 33:75-87
29. Kinney AJ, Herman JEM (2001) Cosuppression of the α subunits of β-conglycinin in transgenic soybean seeds induces the formation of endoplasmic reticulum- derived protein bodies. Plant Cell 13:1165-1178
30. Tada Y, Utsumi S, Takaiwa F (2003) Foreign gene products can be enhanced by introduction into low storage protein mutants. Plant Biotechnol J 1:411–422
31. Gutiérrez SP, Saberianfar R, Kohalmi SE, Menassa R (2013) Protein body formation in stable transgenic tobacco expressing elastin-like polypeptide and hydrophobin fusion proteins. BMC Biotechnol 13:40. doi: 10.1186/1472-6750- 13- 40
32. Iida S, Aman E, Nishio T (1993) A rice (*Oryza sativa* L.) mutant having a low content of glutelin and a high content of prolamine. Theor Appl Genet 87, 374–378
33. Iida S, Kusaba M, Nishio T (1997) Mutants lacking glutelin subunits in rice: mapping and combination of mutated glutelin gene. Theor Appl Genet 94, 199–183
34. Kawakatsu T, Hirose S, Yasuda H, Takaiwa F (2010) Reducing rice seed storage protein accumulation leads to changes in nutrient quality and storage organelle formation. Plant Physiol 154, 1842–1854
35. Yang X, Niu L, Zhang W, He H, Yang J, Xing G, Guo D, Du Q, Qian X, Yao Y, Li Q, Dong Y (2017) Robust RNAi-mediated resistance to infection of seven potyvirids in soybean expressing an intron hairpin *Nlb* RNA. Transgenic Res 26:665-676
36. Yang J, Xing G, Niu L, He H, Guo D, Du Q, Qian X, Yao Y, Li H, Zhong X, Yang X (2018) Improved oil quality in transgenic soybean seeds by RNAi-mediated knockdown of *GmFAD2-1B*. Transgenic Res 27:155-166
37. Hwang EY, Song QJ, Jia GF, Specht JE, Hyten DL, Costa J, Cregan PB (2014) A genome-wide association study of seed protein and oil content in soybean. BMC Genomics 15:1. doi: 10.1186/1471-2164-15-1

Figures

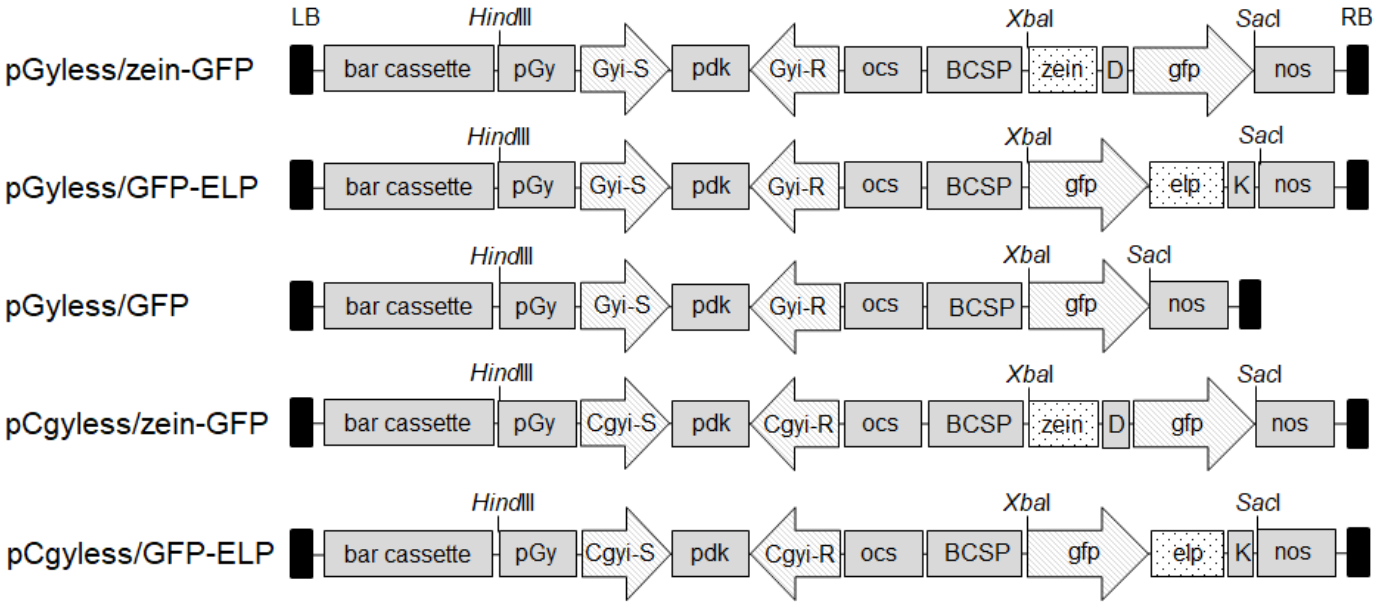

Figure 1

Schematic maps of vector constructs used for soybean transformation. The constructs contained the *Gy* or *Cgy* RNAi cassette and *gfp* (*zein*- or *elp*-fused, and unfused) expression cassettes. *BCSP* and *Pgy1*, soybean seed-specific promoters of β-conglycinin alpha subunit and glycinin Gy1 subunit genes, respectively; GFP, green fluorescent protein; zein, maize 27 kDa γ-zein; ELP, a 13.56 kDa elastin-like polypeptide; D, enterokinase cleavage site (DDDDK); K, endoplasmic reticulum (ER) localization silGyi-S and Gyi-R, based on glycinin subunit-encoding genes (*Gy1*– *Gy5*); Cgyi-S and Cgyi-R, based on conglycinin subunit-encoding genes (*Cgy1*– *Cgy3*); ocs, octopine synthase terminator; pdk, pdk intron spacer; LB, left border; RB, right border.

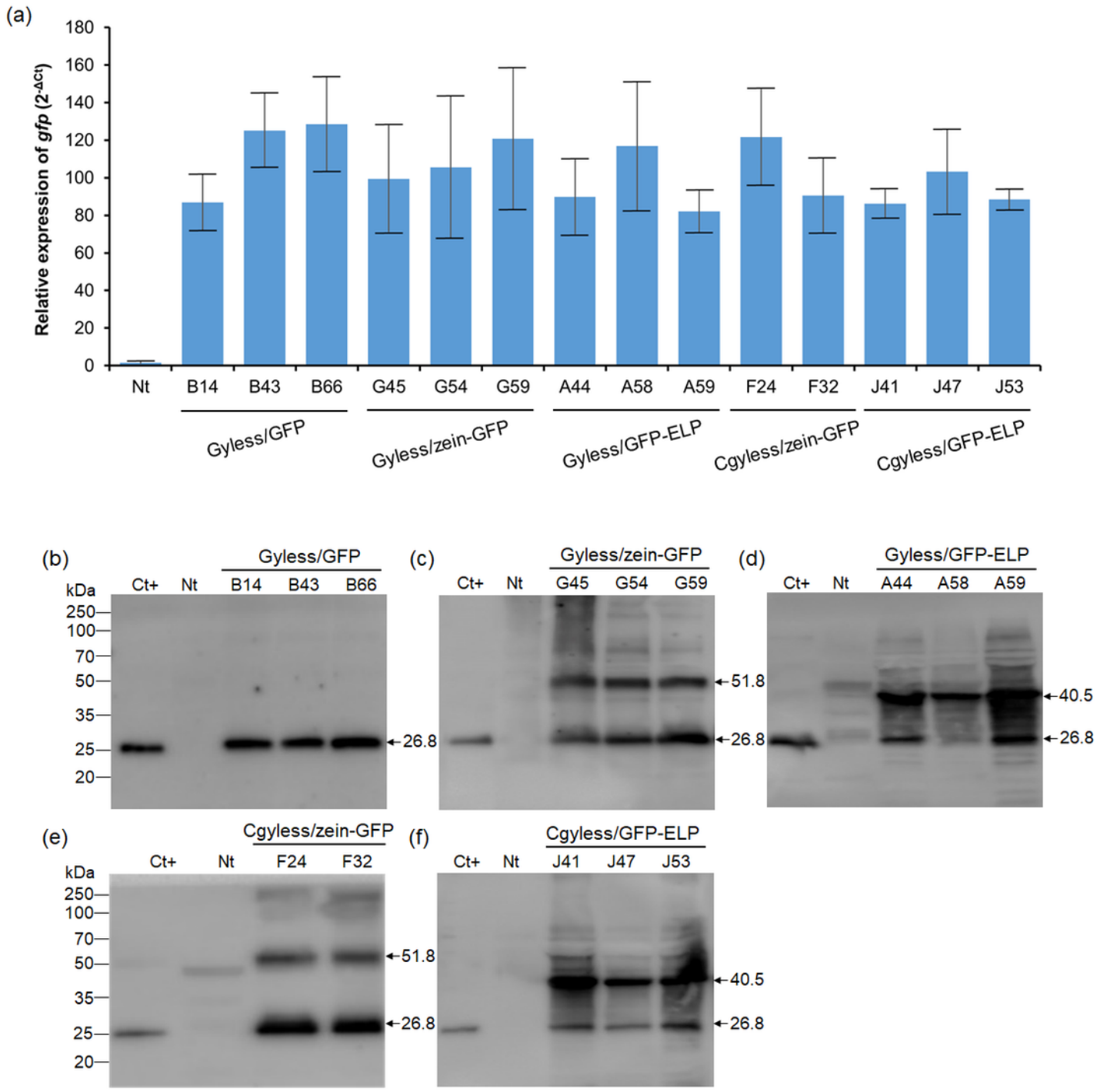

Figure 2

Expression analysis of the foreign genes in transgenic soybean plants. (**a**) Quantification of *gfp* expression in immature T3 transgenic seeds at 45 days after flowering (DAF) using qRT-PCR (n = 3). (**b–f**) Western blotting analysis of GFP expression in T3 soybean seeds at 45 DAF. The bands represented GFP (26.8 kDa), zein-GFP (51.8 kDa), and GFP-ELP (40.5 kDa). Ct+, GFP standard; Nt, non-transformed plant.

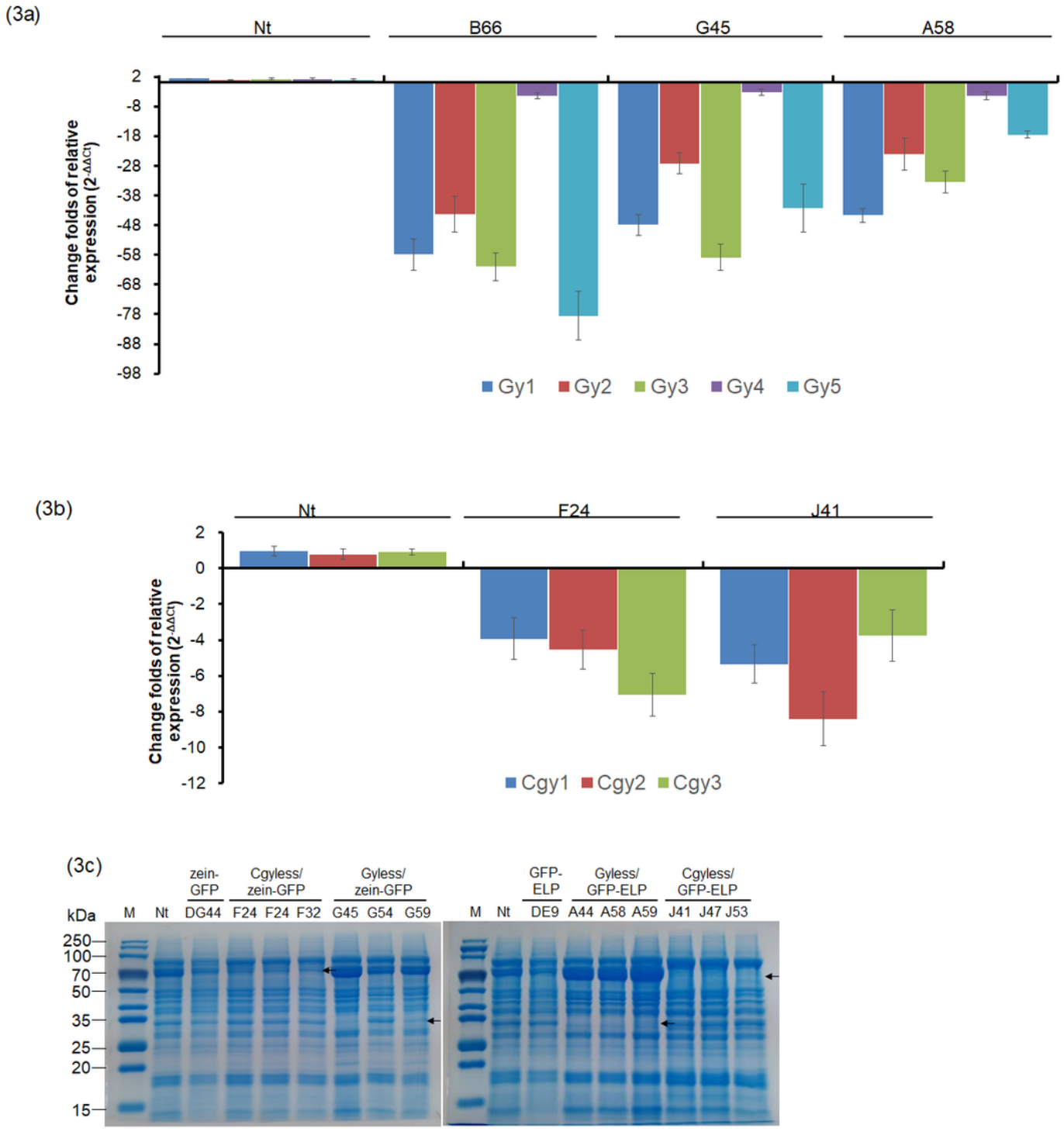

Figure 3

Expression analysis of endogenous storage proteins glycinin and conglycinin in transgenic soybean seeds carrying *Gy* and *Cgy* RNAi cassettes. qRT-PCR analysis of glycinin (**a**) and conglycinin (**b**) gene expression. Nt, non-transformed plants; B66, Gyless/GFP plants; G45, Gyless/zein-GFP plants; A58,

Gyless/GFP-ELP plants; F24, Cgyless/zein-GFP plants; J41, Cgyless/GFP-ELP plants. (**c**) Protein expression profiles analyzed using SDS-PAGE. M, protein marker; Nt, non-transformed plants; DG44, zein-GFP plants without the RNAi cassette; DE9, GFP-ELP plant without the RNAi cassette; G45, G54, and G59, Gyless/zein-GFP plants; F24 and F32, Cgyless/zein-GFP plants; A44, A58, and A59, Gyless/GFP-ELP plants; J41, J47, and J53, Cgyless/GFP-ELP plants.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformation.pptx
- qpcr.xlsx