Elastin-like polypeptide and $\gamma$-zein fusions significantly increase recombinant protein accumulation in soybean seeds

Jing Yang · HongWei Xun · Lu Niu · Hongli He · Yunqing Cheng · Xiaofang Zhong · Qianqian Zhao · Guojie Xing · Jianfeng Liu · Xiangdong Yang

Received: 9 October 2020 / Accepted: 3 May 2021 / Published online: 8 May 2021 © The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

Abstract Soybean seeds are an ideal host for the production of recombinant proteins because of their high content of proteins, long-term stability of seed proteins under ambient conditions, and easy establishment of efficient purification protocols. In this study, a polypeptide fusion strategy was applied to explore the capacity of elastin-like polypeptide (ELP) and $\gamma$-zein fusions in increasing the accumulation of the recombinant protein in soybean seeds. Transgenic soybean plants were generated to express the $\gamma$-zein- or ELP-fused green fluorescent protein (GFP) under the control of the soybean seed-specific promoter of $\beta$-conglycinin alpha subunit ($BCSP$). Significant differences were observed in the accumulation of zein-GFP and GFP-ELP from that of the unfused GFP in transgenic soybean seeds based on the total soluble protein (TSP), despite the low-copy of T-DNA insertions and similar expression at the mRNA levels in selected transgenic lines. The average levels of zein-GFP and GFP-ELP accumulated in immature seeds of these transgenic lines were 0.99% and 0.29% TSP, respectively, compared with 0.07% TSP of the unfused GFP. In mature soybean seeds, the accumulation of zein-GFP and GFP-ELP proteins was 1.8% and 0.84% TSP, an increase of 3.91- and 1.82-fold, respectively, in comparison with that of the unfused GFP (0.46% TSP). Confocal laser scanning analysis showed that both zein-GFP and GFP-ELP were abundantly deposited in many small spherical particles of transgenic seeds, while there were fewer such fluorescence signals in the same cellular compartments of the unfused GFP-expressing seeds. Despite increased recombinant protein accumulation, there were no significant changes in the total protein and oil content in seeds between the transgenic and non-transformed plants, suggesting the possible presence of threshold limits of total protein accumulation in transgenic soybean seeds. Overall, our results indicate that $\gamma$-zein and ELP fusions significantly increased the accumulation of the recombinant protein, but exhibited no significant influence on the total protein and oil content in soybean seeds.

Keywords Soybean seed · Fusion partner · Elastin-like polypeptide · $\Gamma$-Zein · Recombinant protein

Jing Yang and HongWei Xun equally contributed to this study.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11248-021-00258-7.

J. Yang · H. Xun · L. Niu · X. Zhong · Q. Zhao · G. Xing · X. Yang (✉) Jilin Provincial Key Laboratory of Agricultural Biotechnology, Jilin Academy of Agricultural Sciences, Changchun 130033, China e-mail: xdyang020918@126.com

H. He · Y. Cheng · J. Liu Jilin Normal University, Siping 136000, China
Abbreviations

ELP  Elastin-like polypeptide
GFP  Green fluorescent protein
TSP  Total soluble protein
hFIX  Human coagulation factor IX
LTB  Labile toxin B
HIV  Human Immunodeficiency Virus
HFBI  Fungal hydrophobin I
PB  Protein body
ER  Endoplasmic reticulum
PSV  Sprotein storage vacuoles
BCSP  β-Conglycinin alpha subunit
NOS  Nopaline synthase
PCR  Polymerase chain reaction
DIG  Digoxigenin
qRT  Quantitative reverse transcription
EDTA  Ethylene Diamine Tetraacetic Acid
HRP  Horseradish peroxidase
OD  Optical density
TMB  Tetramethylbenzidine
PVDF  Polyvinylidene Fluoride

Background

Soybean seeds are a practical expression system for the biotechnological production of recombinant proteins. Besides the general advantages of plant production systems, such as unlimited scale-up potentials, remarkably reduced production costs, eukaryotic post-translational modifications, and low contamination risks (Lau and Sun 2009; Kawakatsu and Takaiwa 2010), soybean seeds also possess several inherent benefits for recombinant protein expression (Hudson et al. 2014). For example, soybean seeds contain approximately 40% protein by dry mass (compared with 7%–12% in cereals such as maize, rice, and wheat), which provide ample and stable space for the deposition of recombinant proteins in seeds (Hudson et al. 2014). Despite of less yields per acre than other crops such as rice and maize, the total protein yields per acre that can be obtained for soybean is double the amount for maize or rice (Cunha et al. 2011a). Moreover, soybean seed yields can be considerably maximized through simple photoperiod manipulation under greenhouse containment which also minimizes the risk into the food chain. The general properties such as low protease activities and water content in soybean seeds provide stable biochemical environments for long-term stable post-harvest storage of recombinant proteins without any detectable degradation (Münz 1998; Cunha et al. 2011a; Morandini et al. 2011). Cunha et al. (2011a) reported that correct processing and stable accumulation of functional recombinant human coagulation factor IX (hFIX) could be maintained in soybean seeds for 6 years at room temperature. Moreover, simple proteome, relative homogeneity, and less metabolites in soybean seeds also facilitate downstream processing and purification of recombinant proteins (Lau and Sun 2009; Morandini et al. 2011). To date, several recombinant pharmaceuticals and antigen vaccines, including Escherichia coli heat labile toxin B subunit (LTB) (Moravec et al. 2007), anti-hypertensive peptide (Yamada et al. 2008), human growth hormone (Cunha et al. 2011b), human coagulation factor IX (Cunha et al. 2011a), and microbicide against HIV cyanovirin-N (O’Keefe et al. 2015), have been successfully expressed in soybean seeds. However, compared with that in other host plants such as rice seeds (He et al. 2011; Zhang et al. 2012), the accumulation of some recombinant proteins in soybean seeds is usually below the 1% total soluble protein (TSP) of economic threshold proposed by Kusnadi et al. (2015), except one report in which the accumulation level of the recombinant protein reached 2.9% TSP in soybean seeds (Cunha et al. 2011b).

Multiple strategies have been developed and tested to increase the yield and stability of recombinant proteins in plant production systems. Among those, fusion protein technology showed great potential in enhancing recombinant protein accumulation and facilitating recovery and downstream purification (Conley et al. 2011; Gutiérrez et al. 2013; Hofbauer et al. 2014; Phan et al. 2014; Heidari-Japelaghi et al. 2019). Several fusion partners such as oil body-targeted oleosin, elastin-like polypeptide (ELP), 27-kDa maize γ-zein, and fungal hydrophobin I (HFBI) have been utilized to improve both the accumulation and stability of recombinant proteins (Torrent et al. 2009; Floss et al. 2010; Joensuu et al. 2010; Conley et al. 2011; Gutiérrez et al. 2013; Phan et al. 2014; Heidari-Japelaghi et al. 2019). γ-Zein, a member of the major prolamin storage family in maize seeds, contains a signal peptide with a “CGC” motif, proline-rich amphipathic region with eight “PPPVHL” repeats, and third domain with four additional cysteine residues at its N-terminal region.
\(\gamma\)-Zein possesses an intrinsic protein body (PB)-forming property in the absence of tissue-specific factors (Hofbauer et al. 2014). When fused to a recombinant protein of interest, \(\gamma\)-zein can successfully trigger PB formation and induce high accumulation of recombinant proteins (Mainieri et al. 2004; Torrent et al. 2009; Joseph et al. 2012; Hofbauer et al. 2014; Davide et al. 2018). For example, fusion of the human immunodeficiency virus negative factor (Nef) sequence to the chimeric protein zeolin which is composed of phaseolin and 89 amino acid of \(\gamma\)-zein, led to the accumulation of 1.5\% TSP in transgenic tobacco leaves. Elastin-like polypeptide is repeats of the amino acid “VPGXG” that can undergo a temperature-dependent reversible phase transition. The use of ELP as a fusion partner has also been reported to be capable of inducing PB formation in the leaves when targeted to the endoplasmic reticulum (ER) in transient expression assays (Ceresoli et al. 2016; Reza et al. 2016; Saberianfar and Menassa 2017). The addition of ELP to the C-terminus of Influenza hemagglutinin resulted in its increased accumulation in both tobacco leaves and seeds compared with that of unfused hemagglutinin (Phan et al. 2014). The expression of single-chain variable fragments (scFvs) in transgenic tobacco seeds even increased to 40-fold, with levels approaching 25\% TSP, when fused with ELP at the C-terminus (Scheller et al. 2006).

Seed vacuoles are intracellular endpoints of the plant secretory pathways for proteolysis or storage of proteins. Different from the lytic vacuoles, as seed-specific organelles, the protein storage vacuoles (PSVs) are specialized for the deposition of large amounts of endogenous proteins (Jolliffe et al. 2005; Boothe et al. 2010; Vianna et al. 2011). The abundance of PSVs in the seeds makes it the preferred site for recombinant protein accumulation. PSV signal peptides have been used to direct recombinant proteins into PSVs of tobacco or soybean seeds in previous studies (Leite et al. 2000; Kim and Krishnan 2004; Yamada et al. 2008; Hofbauer et al. 2014). Trafficking and storage of proteins to PSVs may depend on plant species, tissues, and developmental stages (Ibl and Stoger 2012). As for PBs, although they may remain in the cytosols after formation and budding from ER, they can also be sequestered in the PSVs probably via an autophagy-like process for long-term storage, as showed in wheat, oats, barley, and tobacco (Ibl and Stoger 2012).

In this study, we explored the capacity of two fusion partners ELP and \(\gamma\)-zein in increasing recombinant protein accumulation in soybean seeds, which naturally accumulate higher amounts of endogenous proteins than that of most other crops. In addition, the targeted deposition of the recombinant protein and the influence of recombinant protein accumulation on the total protein and oil content in transgenic soybean seeds were also analyzed. Our aim was to provide a practical and efficient protocol for the production of important pharmaceutical proteins or vaccine candidates with soybean seeds as an expression host.

**Results**

Molecular characterization of transgenic soybean plants

To evaluate the influence of the polypeptide fusion partners ELP and \(\gamma\)-zein on the accumulation of GFP in soybean seeds, three plant expression vectors harboring GFP-ELP, zein-GFP, and GFP were constructed and used for Agrobacterium-mediated genetic transformation (Fig. 1a). The soybean seed-specific promoter BCSP, which is specifically expressed in the cotyledon and embryonic axis of soybean seeds (Imoto et al. 2008), was used to drive the expression of the GFP fusion genes or unfused GFP. A series of transformation experiments resulted in a total of 56 transgenic events with an average transformation efficiency of 2.4\%, including 15 zein-GFP, 18 GFP-ELP, and 23 GFP lines, as confirmed using the LibertyLink strip and polymerase chain reaction (PCR) analysis (Fig. S1A-D). In this study, all the transgenic plants grew normally and showed no differences in morphology compared with that of non-transformed plants (data not showed). Transgenic plants in the subsequent generations (T1–T3) were further screened for their herbicide tolerance and by PCR analysis until homozygous transgenic lines were obtained (Fig. S2). To ensure consistency in comparing accumulation levels of the recombinant proteins, we first analyzed the relative expression of zein-GFP, GFP-ELP, and GFP in the immature seeds at 45 days after flowering by reverse transcription quantitative PCR (RT-qPCR) (Fig. 1b). The results revealed...
extensive variability in the transcript levels across the independent transgenic soybean events (Fig. 1b), which might be caused by positional effects of transgene insertion within the genome, transgene copy number, and/or silencing mechanisms suppressing transgene expression (Phan et al. 2014). However, in some transgenic events, similar expression levels of the foreign genes at the mRNA levels were also observed at the same time point during seed development (Fig. 1b), likely due to the same BCSP promoter used in the study. Integration and copy number of the transgenes were further determined by Southern blot analysis. Hybridization signals were observed in the selected transgenic plants, while no positive signal was detected in the non-transformed plants, confirming the insertion of the foreign genes in the soybean genome (Fig. 2a). In these transgenic events, the copy number ranged from 1 to 3. To avoid influence of copy number and differential expression at the mRNA level of the transgenes on recombinant protein accumulation in different transgenic events, only the transgenic plants with a low-copy (single or double) of insertion and similar expression of the transgenes, including 3 zein-GFP plants DG-12, DG-44 and DG49, 3 GFP-ELP plants DE-1, DE-9 and DE-28, 3 GFP plants DF-4, DF-8 and DF-43, were selected for protein quantification analysis.

---

**Fig. 1** Plant expression constructs and generation of transgenic plants. a Schematic representation of the vector constructs used for Agrobacterium-mediated genetic transformation. BCSP, soybean seed-specific promoter of β-conglycinin alpha subunit; GFP, green fluorescent protein; Zein, maize 27 kDa γ-zein which possesses an intrinsic protein body (PB)-forming property; ELP, 13.56 kDa elastin-like polypeptide; D, enterokinase cleavage site (DDDDK); KDEL, ER localization signal which is located at the C-terminus of ELP to direct retention of the recombinant protein in the ER lumen. The selectable marker BAR driven by the constitutive promoter CaMV 35S was used to select glufosinate-resistant transgenic plants. b Quantification of the GFP fusion genes or GFP alone at the mRNA level in immature seeds of representative transgenic soybean plants. RNA was extracted from the immature seeds of T3 transgenic plants and GFP mRNA expression was determined using the 2-ΔΔCt method after normalization to the expression of the internal control GmACTIN6. The data are presented as the mean of three biological replicates with error bars indicating the standard error. Different letters in each group of GFP-, zein-GFP- and GFP-ELP-expressing plants indicate significant difference at P < 0.05
Expression and quantification of the recombinant GFP protein

Western blotting was conducted to analyze the expression of zein-GFP, GFP-ELP, and unfused GFP with the monoclonal anti-GFP antibodies. The band with an apparent molecular weight of 26.8 kDa representing GFP was observed in the five GFP transgenic plants tested (Fig. 2b). Comparatively, the zein-GFP and GFP-ELP fusion proteins were mostly detected as soluble forms of approximately 51.8 and 40.5 kDa, respectively, which represented the monomer of each recombinant protein of the expected size (Fig. 2b). However, a 26.8-kDa band was also detected in both zein-GFP and GFP-ELP seed samples, which might result from the cleavage of the fusion proteins when they are translated and stayed in the ER before deposited in PBs or PSVs (Fig. 2b). These results indicated that the GFP antibody can effectively bind to the γ-zein- or ELP-fused GFP protein in proper configurations and confirmed the translation of both recombinant proteins in the transgenic soybean plants. Recombinant protein accumulation in soybean seeds was further quantified using ELISA. As shown in Fig. 3a, significant differences were observed in the accumulation of zein-GFP or GFP-ELP and unfused GFP in immature soybean seeds. All three zein-GFP lines DG-12, DG-44, and DG49 displayed significantly higher protein accumulation with an average of 0.99% TSP (0.84%–1.2%), than that of both GFP-ELP and GFP proteins, which showed 0.29% TSP (0.25%–0.34%) and 0.07% TSP (0.06%–0.08%), respectively (Fig. 3a, Table S1). Although lower than that of zein-GFP in immature seeds, the average expression of GFP-ELP was still considerably higher than that of the unfused GFP. Consistent results were also observed in
mature seeds (Fig. 3b). In this case, the average accumulation of zein-GFP and GFP-ELP proteins reached 1.8% (1.42%–2.08%) and 0.84% TSP (0.79%–0.85%), an increase of 3.91- and 1.82-fold, respectively, compared with that of the unfused GFP (0.46% TSP) (Fig. 3b, Table S1). Furthermore, zein-GFP also showed higher accumulation than GFP-ELP in mature seeds. Overall, these observations demonstrated that \( \gamma \)-zein and ELP polypeptide fusions significantly enhance the expression and accumulation of GFP in soybean seeds, and \( \gamma \)-zein appears to have a greater influence on GFP accumulation than ELP under our experiment conditions.

Accumulation of the GFP of GFP with different fusion partners in soybean seeds

Confocal laser scanning analysis was performed to analyze the accumulation of the GFP fused with \( \gamma \)-zein or ELP in immature soybean seeds. Compared with that in the untransformed plants, green fluorescence signal was observed in all transgenic lines, which was attributed to the expression of GFP. In zein-GFP- and GFP-ELP-expressing plants, many small spherical particles were observed in immature soybean seeds (Fig. 4), while there were fewer such fluorescence signals in the same cellular compartments of the unfused GFP-expressing seeds, demonstrating abundant deposition of zein-GFP and GFP-ELP in soybean seeds.

Influence of recombinant protein accumulation on soybean seed protein and oil content

Soybean seeds are characteristically high in proteins. To evaluate the effects of accumulation of the recombinant proteins on the total protein and oil content, we analyzed the changes in the content of total protein and oil in transgenic soybean seeds expressing zein-GFP, GFP-ELP, and GFP proteins. The results showed that there were no significant differences in seed protein and oil content between the transgenic and non-transformed soybean seeds (Fig. 5a, b, Table S2). Increased recombinant protein accumulation does not appear to influence the total seed protein (native and recombinant) content, implying that there appears to be a threshold limit on the total level of soybean protein, due to the competition for resources in protein biosynthesis, trafficking, and deposition between the recombinant and native proteins. We suppose that increased recombinant protein accumulation in soybean seeds is likely accompanied by the displacement of native storage proteins. The influence of recombinant protein accumulation on total protein and oil content needs to be further investigated in a larger field experiment.
Discussion

There are two major limitations to the potential production system of recombinant proteins: low accumulation levels and lack of efficient purification (Gutiérrez et al. 2013). Sufficient accumulation of recombinant proteins in host plant is a major requirement for their commercial-scale production. Currently, polypeptide fusion strategy is an efficient way to enhance recombinant protein yield and also provides a simple means for the purification of recombinant proteins (Conley et al. 2011; Gutiérrez et al. 2013; Hofbauer et al. 2014; Phan et al. 2014). In this study, two polypeptide partners γ-zein and ELP were evaluated for their capability to enhance recombinant GFP accumulation in soybean seeds. To avoid the influence of transgene copy number, which might have a positive or negative association with expression level, only the transgenic lines with low-copy insertion (single or double) were utilized in the expression analysis. Moreover, the transcription levels of the transgenes were determined before quantification of the recombinant proteins. The results revealed that the accumulation of zein-GFP and GFP-ELP was significantly increased compared with that of unfused GFP in both immature and mature soybean seeds. These results are consistent with those of previous studies, which reported that the expression of γ-zein or ELP-fused recombinant proteins was higher than that of proteins without fusion partners in plant leaves or seeds (Scheller et al. 2006; Virgilio et al. 2008; Hofbauer et al. 2014; Phan et al. 2014; Ceresoli et al. 2016; Saberianfar and Menassa 2017). Enhanced accumulation of the recombinant proteins might result from increased transportation and deposition in cell compartments such as PBs or PSVs, which can protect recombinant proteins from proteolytic digestion.
Moreover, accumulation enhancement by the polypeptide fusion is also dependent on the specific fusion partner. In contrast to ELP, γ-zein appears to have a greater influence on GFP accumulation in both immature and mature transgenic soybean seeds. The reason for the difference of accumulation levels may result from the intrinsic PB-forming property of γ-zein.

Accumulation of the recombinant proteins is regulated by several steps at the transcriptional (promoters and transcription factors), translational, and post-translational levels (modification, processing trafficking, and deposition). At the endpoint of the plant secretory pathway, PSVs present the major site for the accumulation and long-term storage of protein in the seeds (Phan et al. 2014). The temporal abundance of PSVs is also important for the accumulation of high-level endogenous proteins in soybean seeds. Soybean seeds are capable of producing a different class of functional recombinant proteins correctly targeted to the PSVs at appropriate levels (Viana et al. 2011). In this study, The PBs triggered by γ-zein or ELP may be sequestered in the PSVs independent of the PSV-sorting determinants, probably via an autophagy-like process for long-term storage (Ibl and Stoger 2012). Actually, in dicots such as soybean seeds, endogenous storage proteins are mainly deposited in PSVs. It is possible that the recombinant proteins fused with γ-zein or ELP partner may behave in the same way as the native storage proteins. Compartmentalization of the recombinant protein in PBs or PSVs not only avoids proteolysis, but also increases the yield and stability of recombinant proteins (Cunha et al. 2011a, b).

In this study, we also evaluated the effects of recombinant protein accumulation on the total seed protein content in the host plant. This is especially important for soybean seeds, which possess a high level of endogenous storage protein. Our results showed that recombinant protein accumulation has no detectable effect on total protein storage, suggesting that an intrinsic threshold mechanism exists to control total protein accumulation. Competition in biosynthesis, transportation, and deposition between the recombinant and native proteins for resources may affect the deposition of amount of foreign proteins in soybean seed. Previous studies have 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 (Kinney and Herman 2001). The accumulation of recombinant proteins might simultaneously result in decreased native protein storage in soybean seeds. This provides a cue that it might further increase the accumulation of desired recombinant proteins in soybean seeds by suppressing native storage protein via RNA interference or gene editing.

Conclusions

Overall, our results demonstrated the practicability of ELP and γ-zein as fusion partners to enhance the accumulation of recombinant protein and exhibit no significant influence on the total protein and oil content in soybean seeds.

Methods

Vector construction and Agrobacterium-mediated transformation of soybean

All cloning steps were carried out using the binary vector pTF101 which contains the phosphinothricin acetyl transferase (BAR) resistance gene as the selection marker (Hardegger et al. 1999). The 1413-bp seed-specific promoter of β-conglycinin alpha subunit (BCSP) was amplified from soybean as previously described (Yang et al. 2018) and sub-cloned into the HindIII/XbaI sites of the compatible pTF101 plasmid. The genes encoding maize 27 kDa γ-zein (Genbank No. AF371261.1) and 13.56 kDa ELP with 30 peptide repeats “VPGXG” were synthesized and optimized based on soybean codon usage (GenScript USA Inc., NJ, USA) (Fig. S3), and then inserted into the binary vector under the control of the promoter BCSP. The genes encoding maize 27 kDa γ-zein (Genbank No. AF371261.1) and 13.56 kDa ELP with 30 peptide repeats “VPGXG” were synthesized and optimized based on soybean codon usage (GenScript USA Inc., NJ, USA) (Fig. S3), and then inserted into the binary vector under the control of the promoter BCSP. The coding sequence of the green fluorescent protein (GFP, Genbank No. X83959.1) from the plasmid pGm-GFP was amplified and cloned upstream of ELP or downstream of γ-zein sequences to generate the transformation vectors pTF101-GFP-ELP and pTF101-zein-GFP. In construct pTF101-GFP-ELP, the ER localization signal KDEL sequence was added at the C-terminus of the codon-optimized ELP sequence, while in construct pTF101-zein-GFP, the enterokinase cleavage site (DDDDK) was located.
between γ-zein and GFP instead (Fig. 1a). For the control vector pTF101-GFP, the GFP gene was cloned into the XbaI/SacI sites between BCSP and nopaline synthase (NOS) terminator. All three vector constructs contained a phosphinothricin acetyltransferase selectable marker gene (BAR) for the selection of glufosinate-resistant transgenic events.

The soybean genotype P03, provided by Prof. Bao Peng from the Jilin Academy of Agricultural Sciences, China, was used for Agrobacterium-mediated transformation as previously described by Yang et al. (2017). Briefly, the imbibed soybean seeds were then cut along the hilum and the explants were immersed in an Agrobacterium suspension for 30 min, and then placed on solid co-cultivation medium. After 3–5 d of co-cultivation in the dark at 25 °C, the explants were transferred onto shoot induction mediums in a growth room at 28 °C, under an 18-h light/6-h dark photoperiod. The green shoots were induced and then transferred to rooting medium until the regenerated resistant plantlets were produced. The T₁ transgenic seeds were produced by self-pollination. The primary transformants were transferred into a greenhouse for subsequent molecular screening.

Molecular analysis of the transgenic plants

The T₀ transgenic plants were first screened using the LibertyLink® strip according to the instructions of the manufacturer (EnviroLogix Inc., ME, USA). The appearance of two red lines simultaneously in a sample indicated expression of the BAR gene at the translational level. PCR was performed to confirm the presence of the transgenes using the primers BG-F1/BG-R1 (5’-CTCACCATCGCTCAACACATTTC-3’), 5’-TGCTTTTGATGTTCCCGTCCGCTCC-3’), which anneal with GFP and its upstream BCSP sequences, respectively. For the subsequent generations, the transgenic plants were further screened by both glufosinate (1500 mg/L) spraying and PCR analysis. Southern hybridization was conducted to determine transgene copy numbers in soybean plants. Total genomic DNA was extracted from the young leaves of T₃ transgenic plants with a modified procedure using cetyltrimethylammonium bromide at a high salt concentration as described previously (Telzur et al. 1999). Approximately 30 μg of genomic DNA was digested overnight with XbaI, separated on 1% agarose gel, and transferred onto a positively charged nylon membrane (GE Amershams, USA) according to the standard protocols. The 714-bp GFP probe was obtained by PCR amplification using the primers GFP-F1/GFP-R1 (5’-ATGAGTAAAGGAGAGAAGACTTTTCA-3’, 5’-TTTGTATAGTTCCATGCAGGCATGT-3’) and labeled using digoxigenin (DIG)-high prime (Roche, Germany). Hybridization, membrane washing, and subsequent chemical staining with BCIP/NBT were conducted as described previously (Yang et al. 2017).

Reverse transcription quantitative PCR

Forty-five days after flowering, immature transgenic soybean seeds were sampled to determine the relative expression of the GFP transcripts by RT-qPCR. Total RNA was extracted from the immature seeds (~ 10 mm in length) of T₃ transgenic plants using the EasyPure Plant RNA Kit (TransGen Biotech, Beijing, China), according to the manufacturer’s protocol. Genomic DNA removal and first-strand cDNA synthesis were carried out using TransStart® Top Green qPCR SuperMix (TansGen Biotech). Amplifications were performed on ABI 7900HT (Applied Biosystems, USA) using TransStart® Top Green qPCR SuperMix (TansGen Biotech). The reaction conditions were as follows: 50 °C for 2 min, 95 °C for 10 min, and 35 cycles of 95 °C for 2 min, 60 °C for 30 s, and 72 °C for 30 s. The specific primers GFP-F2/GFP-R2 (5’-ACGTGCGAGAGGAGGACCAT-3’, 5’-ACTTTGTTGGCCGAGGATGTTT-3’) were designed based on the GFP sequence. The constitutively expressed native soybean gene GmACT6 (GenBank No. NM_001289231) was amplified with the primers GmACT-F1/GmACT-R1 (5’-ATCTTGACTGAGCGTGGTTATTCC-3’, 5’-ATCTTGACTGAGCGTGGTTATTCC-3’), which anneal with GmACT6 and its upstream BCSP sequences, respectively. For subsequent generations, the expression levels were calculated using the relative quantification (2^ΔΔCt) method and the data were normalized using the expression level of the internal control GmACT6. All experiments were conducted with three biological replicates.

Western blotting and GFP quantification analysis

The T₃ transgenic lines carrying a low-copy (single or double) transgene insertion and with similar mRNA
expression of the transgenes in the immature seeds at 45 days after flowering were used for western blotting and enzyme-linked immunosorbent assay (ELISA). Immature seeds of similar size (~10 mm in length) were harvested from the transgenic plants expressing zein-GFP, GFP-ELP, or unfused GFP. Total soluble protein was extracted with extraction buffer (100 mM NaCl, 10 mM EDTA, 200 mM Tris–HCl, 0.05% Tween-20, 0.1% SDS, 14 mM β-mercaptoethanol, 400 mM sucrose, and 2 mM phenylmethylsulfonyl fluoride). Protein quantification was performed using the Bradford method (Bradford 1976). The protein samples were then separated by 12% sodium dodecyl sulfate–polyacrylamide gel electrophoresis and electrotransferred on to a PVDF membrane (GE Healthcare, USA). After blocking with 5% (w/v) fat-free milk powder, the membranes were blotted with monoclonal anti-GFP antibodies (1:500 dilution, Abcam, UK) for 2 h at room temperature. Antibody binding was then detected by the addition of horse-radish peroxidase (HRP)-conjugated goat-anti-rabbit IgG (1:5000 dilution, Abcam). After extensive washing, the signal was visualized using the BiodlightTM Western Chemiluminescent HRP substrate (Bioworld Technology, Inc. USA). To quantify GFP, the protein samples from immature or mature soybean seeds were subjected to ELISA with purified GFP (Abcam) as the standard protein. The wells of the ELISA plates were coated with diluted protein extracts and standard protein in the coating buffer (15 mM Na 2CO 3, 3 mM NaHCO 3, and 3 mM NaNO 3; pH 9.6). After blocking with 5% dried skimmed milk in phosphate buffer containing 0.1% Tween 20, the monoclonal anti-GFP antibodies (1:500 dilution) were added into the wells and incubated for 1 h at room temperature. The HRP-conjugated secondary antibodies were then added at a dilution of 1:5000 and incubated for 1 h. Optical density (OD) of each reaction mixture was measured at 450 nm (ELx800; BioTek, USA) with tetramethylbenzidine (TMB) as the substrate.

Confocal laser scanning

Immature seeds were collected at 45 days after flowering and tissue sections were cut with scalpel for confocal laser scanning. The sections were immediately imaged using a Leica TCS SP2 confocal laser scanning inverted microscope (Leica Microsystems, Germany) equipped with a 63× water immersion objective. To visualize GFP fluorescence, excitation with 488-nm argon laser was performed and the emission was detected at 500–530 nm. The collected images were analyzed using Leica Application Suite for Advanced Fluorescence (LAS AF, V2.3.5; Leica Microsystems).

Protein and oil content analyses in transgenic seeds

Protein and oil content in the transgenic and non-transformed soybean seeds were analyzed as described previously (Yang et al. 2017). Briefly, mature soybean seeds (200 mg) were ground in a grinder after drying at 60 °C. The level of total nitrogen was determined using LECO CHN 2000 analyzer (LECO, St. Joseph, MI) as described by Hwang et al. (2014). The seed protein content was calculated based on the dry mass basis. The seed oil content was determined by pulsed NMR (Maran Pulsed NMR, Resonance Instruments, Oxfordshire, UK), followed by the field induction decay-spin echo procedure (Rubel 1994).

Statistical analysis

All statistical analyses were performed using SPSS software v.17.0 (SPSS Inc., Chicago, IL, USA). The least significant differences (LSD) between the average values of samples were assessed by the t test at P = 0.05 or 0.01 using the analysis of variance (ANOVA).

Acknowledgements

We thank that Prof. Bao Peng from the Jilin Academy of Agricultural Sciences, China, kindly provided the soybean cultivar P03. We also would like to thank Editage (www.editage.cn) for English language editing.

Author Contribution

X.D. and J.L. designed the experiments. J.Y., H.X. and L.N. performed the experiments and drafted the manuscript. H.H., Q.Z. and G.X. conducted the Agrobacterium-mediated transformation experiments. Y.C. conducted the confocal laser scanning analysis. X. Z. analyzed the experimental data. All authors have read and approved the final version of the manuscript.

Funding

This study was financially supported by the National Natural Science Foundation of China (31671764) and the China National Novel Transgenic Organisms Breeding Project (2016ZX08004-003) and Technology Innovation Program of Jilin province-Postdoctoral foundation (c92070520). These funds played an important role in experimental reagents, consumables, and also provided strong support for preparing 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.

Compliance with ethical standards

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

References

Boothe J, Nykiforuk C, Shen Y, Zaplachinski S, Szarka S, Kuhlman P, Murray E, Morck D, Moloney MM (2010) Seed-based expression systems for plant molecular farming. Plant Biotechnol J 8:588–606

Bradford MM (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254

Ceresoli V, Mainieri D, Fabbro MD, Weinstein R, Pedrazzini E (2016) A fusion between domains of the human bone morphogenetic protein-2 and maize 27 kDa γ-zein accumulates to high levels in the endoplasmic reticulum without forming protein bodies in transgenic tobacco. Front Plant Sci 7:358. https://doi.org/10.3389/fpls.2016.00358

Conley AJ, Joensuu JJ, Richman A, Menassa R (2011) Protein body-inducing fusions for high-level production and purification of recombinant proteins in plants. Plant Biotechnol J 9:419–433

Cunha NB, Murad AM, Ramos GL, Maranhão AQ, Kuhlman P, Murray E, Morck D, Moloney MM (2010) Seed-based expression systems for plant molecular farming. Plant Biotechnol J 8:588–606

He Y, Ning T, Xie T, Qiu Q, Zhang L, Sun Y, Jiang D, Fu K, Yin F, Zhang W (2011) Large-scale production of functional human serum albumin from transgenic rice seeds. Proc Natl Acad Sci USA 108(47):19078–19083

Heidari-Japelagh R, Valizadeh M, Haddad R, Dorani-Uliaie E, Jalali-Javaran M (2019) Elastin-like polypeptide fusions enhance transient expression of human iFn-γ in tobacco leaves. S Afr J Bot 125:321–328

Hofbauer A, Peters J, Arcaleis 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. https://doi.org/10.3389/fbioe.2014.00067

Hudson LC, Garg R, Bost KL, Pillar KJ (2014) Soybean seeds: a practical host for the production of functional subunit vaccines. Biomed Res Int 2014:340804. https://doi.org/10.1155/2014/340804

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 Genom 15:1. https://doi.org/10.1186/1471-2164-15-1

Ibl V, Stoger E (2012) The formation, function and fate of protein storage compartments in seeds. Protoplasma 249(2):379–392

Imoto Y, Yamada T, Kitamura K, Kanazawa A (2008) Spatial and temporal control of transcription of the soybean beta-conglycinin alpha subunit gene is conferred by its proximal promoter region and accounts for the unequal distribution of the protein during embryogenesis. Genes Genet Syst 83:469–476

Joensuu JJ, Conley AJ, Lienemann M, Brandle JE, Linder MB, Menassa R (2010) Hydrophobin fusions for high-level transient protein expression and purification in Nicotiana benthamiana. Plant Physiol 152:622–633

Jolliffe NA, Caddock CP, Frigerio L (2005) Pathways for protein transport to seed storage vacuoles. Biochem Soc Trans 33(5):1016–1018

Joseph M, Ludevid MD, Torrent M, Rofidal V, Tauzin M, Rossignol M, Pedrznzzini E, Vitale A, Stoger E, Ibl V, 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. https://doi.org/10.3389/fbioe.2014.00067

Kawakatsu T, Takaiwa F (2010) Cereal seed storage protein synthesis: fundamental processes for recombinant protein production in cereal grains. Plant Biotechnol J 8:939–953

Kim WS, Krishnan HB (2004) Expression of an 11-kDa methionine-rich delta-zein in transgenic soybean results in the formation of two types of novel protein bodies in transitional cells situated between the vascular tissue and storage parenchyma cells. Plant Biotechnol J 2(3):199–210

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(5):1165–1178

Kusnadi AR, Nikolov ZL, Howard JA (2015) Production of recombinant proteins in transgenic plants: practical considerations. Biotechnol Bioeng 56:473–484

Lau OS, Sun SS (2009) Plant seeds as bioreactors for recombinant protein production. Biotechnol Adv 27:1015–1022

Leite A, Kemper EL, da Silva MJ, Luchessi AD, Siloto RMP, Bonaccorsi (2000) Expression of correctly processed
human growth hormone in seeds of transgenic tobacco plants. Mol Breed 6(1):47–53
Mainieri D, Rossi M, Archinti M, Bellucci M, Marchis FD, Vavassori S, Pompa A, Arcioni S, Vitale A (2004) Zeolin. A new recombinant storage protein constructed using maize γ-zein and bean phaseolin. Plant Physiol 136:3447–3456
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
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
Müntz K (1998) Deposition of storage proteins. Plant Mol Biol 38:77–99
O’Keefe BR, Murad AM, Vianna GR, Ramessar K, Saucedo CJ, Wilson J, Buckett KW, Cunha NB, Araújo ACG, Lacorte CC (2015) Engineering soya bean seeds as a scalable platform to produce cyanoovirin-N, a non-ARV microbicide against HIV. Plant Biotechnol J 13:884–892
Phan HT, Hause B, Hause G, Arcalis E, Stoger E, Maresch D, Altman M, Joensuu J, Conrad U (2014) Influence of elastin-like polypeptide and hydrophobin on recombinant hemagglutinin accumulations in transgenic tobacco plants. PLoS ONE 9:e99347. https://doi.org/10.1371/journal.pone.0099347
Reza S, Amirali S, Joensuu JJ, Kohalmi SE, Rina M (2016) Protein bodies in leaves exchange contents through the endoplasmic reticulum. Front Plant Sci 7:693. https://doi.org/10.3389/fpls.2016.00693
Rubel G (1994) Simultaneous determination of oil and water contents in different oil seeds by pulsed nuclear magnetic resonance. J Am Oil Chem Soc 71:1057–1062
Saberi.anfar R, Menassa R (2017) Protein bodies—how the ER deals with high accumulation of recombinant proteins. Plant Biotechnol J 15:671–673
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
Telzur N, Abbo S, Myslabodski D, Mizrahi Y (1999) Modified CTAB procedure for DNA isolation from Epiphytic acti of the Genera Hylocereus and Selenicereus (Cactaceae). Plant Mol Biol Rep 17:249–254
Torrnt M, Llompart B, Lasserre-Ramassamy S, Llop-Tous I, Bastida M, Marzabal P, Westerholm-Parvinen A, Salo-heimo M, Heifetz PB, Ludevid MD (2009) Eukaryotic protein production in designed storage organelles. BMC Biol 7:5. https://doi.org/10.1186/1741-7007-7-5
Vianna GR, Cunha NB, Murad AM, Rech EL (2011) Soybeans as bioreactors for biopharmaceuticals and industrial proteins. Genet Mol Res 10(3):1733–1752
Virgilio MD, Marchis FD, Bellucci M, Mainieri D, Rossi M, Benvenuto E, Arcioni S, Vitale A (2008) The human immunodeficiency virus antigen Nef forms protein bodies in leaves of transgenic tobacco when fused to zeolin. J Exp Bot 59(10):2815–2829
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
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
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 NiB RNA. Transgenic Res 26:665–676
Zhang D, Lee HF, Pettit SC, Zaro JL, Ning H, 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. https://doi.org/10.1186/1472-6750-12-92
Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.