Expression and accumulation of heterologous molecules in the protein storage vacuoles of soybean seeds

Giovanni Vianna (giovanni@cenargen.embrapa.br)
LTG lab
Elíbio Rech (rech@cenargen.embrapa.br)
LTG lab
Nicolau da Cunha
LTG lab

Method Article

Keywords: transgenic soybean, recombinant protein, seed storage promoter

Posted Date: February 1st, 2011

DOI: https://doi.org/10.1038/protex.2011.206

Abstract

The production of recombinant molecules in seeds plant has presented interesting results related to the synthesis, secretion, compartmentalisation, post-translational modifications and purification, scalability, stable expression, the absence of contaminants and human pathogens, and knowledge of the growth, harvest, storage and processing practices. In this context, the soybean plant has emerged as an option to express recombinant proteins. We obtained several transgenic soybean plant lines, expressing different types of recombinant molecules, under the control of the tissue-specific soybean seed storage β-conglycinin promoter. This promoter is the most abundant protein in soybean seeds, directing the expression of the molecules that accumulate in the seed storage tissue. Cotyledonary immunocytochemical analysis of the seeds demonstrated that the targeted proteins effectively drove polypeptide accumulation into the protein storage vacuoles \(PSVs). Here, we provide a detailed protocol to reproduce the results described above and properly direct the expression of recombinant protein into the PSVs.

Introduction

Plants have emerged as a viable alternative to microbial fermentation and mammalian cell culture for the industrial production of biopharmaceuticals. They offer scalability and low production costs and are free of human pathogens and toxins \(Fiedler et al., 1997; Evangelista et al., 1998; Hood et al., 2002; Ma et al., 2003; Stoger et al., 2005; Verma et al., 2008; Orzáez et al., 2009; Woodard et al., 2009; Dreesen et al., 2010). In addition, plant enzymatic machinery can provide a method of biosynthesis, folding and assembly of complex proteins and promote proper post-translational modifications, resulting in high quality recombinant products \(Ma et al., 2003). The genetic transformation of a broad range of vegetable species and the efficient production of more than 100 different recombinant proteins using plant systems has confirmed the potential to utilise transgenic plants as efficient bioreactors of several relevant biomolecules, such as antigens, vaccines, antibodies and growth factors \(Perrin et al., 2000; Ma et al., 2003; Stoger et al., 2005). As a potential, low-cost platform for recombinant protein biosynthesis, soybeans \[_Glycine_ _max_ L. \(Merril)] might constitute one of the least expensive systems for the large-scale production of biopharmaceuticals. Their high biomass capacity, short life cycle, low allogamy frequency and photoperiod sensitivity are intrinsic and unique characteristics \(Abud et al., 2003; 2007; Cunha et al., 2010b). Under greenhouse conditions, utilising a photoperiod of 19 h of light can increase soybean seed production 10-fold when compared with seed production under field conditions as a consequence of an extended vegetative growth and flowering delay \(Cunha et al., 2010b). Soybean seeds are a rich source of protein, which can reach up to 40% of the dry weight of these organs \(Cantoral et al., 1995; Moravec et al., 2007; Boothe et al., 2010), and, unlike leaves, they can be easily transported and stored for extended periods without significant protein degradation and do not require special storage conditions \(Leite et al., 2000; Stoger et al., 2005; Moravec et al., 2007; Lau & Sun, 2009; Cunha et al., 2010a). Because seeds generally express a wide range of genes related to abundant endogenous storage proteins, it is advisable to use a seed-specific promoter with high specificity and strong

transcriptional activity to maximise transgene expression and reach economic scalability for the production of recombinant proteins \(Stoger et al., 2000, 2005). Monocot promoters, such as the commonly used rice glutelin GluA-2 \(Gt-1), rice globulin, barley D hordein, and maize zein promoters, are usually employed to drive the expression of genes related to storage proteins that are localised in the endosperm, the main storage tissue of monocot seeds. For this reason, expression cassettes using monocot promoters are specifically optimised for gene expression in monocot seeds, resulting in expression levels reaching up to 15% of the total seed protein in transgenic seeds \(Takagi et al., 2005, Yang et al., 2006 and Yang et al., 2007). Among the dicot promoters, legume promoters, such as the seed-specific promoters from the _Phaseolus_ _vulgaris_ phaseolin and arcelin-5 genes, soybean lectin, glycinin, the β-conglycinin α' subunit, pea legumin \(legA) and the bean unknown seed protein \(USP), have emerged as a suitable choice for recombinant protein production in dicot seeds \(Boothe et al., 2010). In a study that examined the expression of a functional, active single chain antibody in _Arabidopsis_ seeds, impressive results were achieved when both _Phaseolus_ promoters were used, accumulating up to 36 times more immunoglobulin than was observed in transgenic _Arabidopsis_ seeds that were transformed with the CaMV 35S promoter \(De Jaeger et al., 2002). Following mRNA translation, the accumulation of newly synthesised proteins in seeds is particularly sensitive to subcellular targeting and, for certain proteins, it can have a major effect on the final yields of the recombinant product by avoiding unintended proteolysis and achieving accumulation levels suitable for economic production \(Ma et al., 2003; Stoger et al., 2005; Streatfield, 2007; Semenyuk et al., 2010; Boothe et al., 2010). For example, protein targeting to the secretory pathway components can lead to a considerable increase in the final protein yield, which can be as much as 10-fold higher than the accumulation in the cytosol alone \(Conrad & Fiedler, 1998; Avesani et al., 2003). The most utilised targeting approaches are to retrieve newly synthesised proteins in the endoplasmic reticulum \(ER) by adding N- or C-terminal signal peptides to the nascent proteins or to direct them to other subcellular organelles, such as the mitochondria, the chloroplasts, and the many components of the secretory pathway, especially the main sites for protein accumulation in seeds, namely, the protein storage vacuoles \(PSVs) \(Stoger et al., 2000; Jiang & Sun, 2002; Stoger et al., 2005; Vitale & Pedrazzini, 2005; Boothe et al., 2010). Plant vacuoles are the intracellular endpoints of the plant secretory pathway, the final destination of proteins bypassing the ER and travelling through the Golgi complex \(Jolliffe et al., 2005). They can be divided into two main categories: \(i) LVs \(lytic vacuoles), which are acidic compartments that are rich in hydrolases and can be regarded as the equivalent of mammalian lysosomes, and \(ii) PSVs, the ER-derived cisternae, which are found in the storage organs, mainly seeds, and are the main specialised storage sites of large amounts of protein that will later be used as a source of aminated compounds during seed germination \(Yoo & Chrispeels, 1980; Vitale & Hinz, 2005; Jolliffe et al., 2005). Cotyledonary PSVs constitute an excellent subcellular target for the long-term storage of recombinant proteins as their lumen are characterised by a non-acidic environment with low concentrations of amino peptidases, thus minimising protein degradation \(Zheng et al., 1992; Muntz, 1998; Jolliffe et al., 2005; Takaiwa et al., 2007; Cunha et al., 2010c). PSVs also contain large amounts of the major seed storage globulins 7S and 11S and toxic proteins, such as lectins, protease inhibitors and ribosome inactivating proteins, which probably evolved to protect the seeds from predators \(Herman &

Larkins 1999; Vitale & Hinz 2005). The mechanism related to the transport of recombinant proteins to the PSVs has not been well studied compared to that of the LVs \(Nishawa et al., 2003). However, the addition of N-terminal signal peptides without any additional signal leads the protein of interest to be targeted to the ER, followed by further delivery along the secretory pathway and to the PSVs and apoplast \(Vitale & Pedrazzini, 2005). Another model for protein targeting to plant vacuoles is based on recent advances in understanding the role of a restricted group of cargo proteins, named vacuolar sorting signals \(VSS), that can mediate protein targeting to the PSVs \(Jolliffe et al., 2005). VSSs of barley lectin, common bean phaseolin, and the soybean β-conglycinin α' subunit have already been identified as potential mediators for recombinant protein targeting to the PSV \(Bednarek & Raikhel, 1991; Frigerio et al., 1998; Nishizawa et al., 2003; Robinson et al., 2005). Although they are sophisticated targeting tools, few studies have reported the interference of the high flow of seed storage proteins to the PSVs in the correct targeting mediated by VSSs, which resulted in the unintended accumulation of proteins in non-targeted organelles \(Lau & Sun, 2009). The beta-conglycinins \(7S) and the glycinins \(11S) are the major storage proteins of the soybean, which can account for up to 70% of the total seed protein \(Chen et al., 1986; 1988; Nishizawa et al., 2003). Beta-conglycinin is a multimeric protein and consists of three subunits: alpha prime \(alpha'), alpha and beta \(Derbyshire et al., 1976; Chen et al., 1986; Doyle et al., 1986). The alpha and alpha' subunits are synthesised as pre-proproteins and the beta subunit as a pre-protein \(Nishizawa et al., 2003). The expression of these subunit genes is spatially and temporally regulated, coinciding with the development of the seed and resulting in high accumulation levels of the related proteins during seed maturation. In this work, we describe results showing a highly efficient, reproducible system for recombinant protein expression in seeds that targets the PSVs, where several molecules with different molecular weights and structures, such as the human Growth Hormone \(hGH) \(Cunha et al., 2010b), the human coagulation factor IX \(hFIX) \(Cunha et al., 2010c), two single chain fragment variable \(scFVDIR83D4 and anti-CD18), and a potent microbicide against HIV, Cyanovirin-N \(CVN), were stably accumulated in the PSVs from transgenic soybean seeds using the α' subunit regulatory sequences of the soybean beta-conglycinin. We successfully utilised two seed-specific expression vectors of soybean construction, containing the following: i) the promoter and signal peptide from the alpha' subunit of beta-conglycinin \(with step-by-step directions for how to construct it) \(Fig. 1A); ii) the promoter from the soybean alpha' subunit of beta-conglycinin, along with the monocot signal peptide α-Coixin from _Coix_ _lacrima_-jobi L. \(responsible for directing the polypeptides to the PSVs of _Coix_ _lacrima_-jobi L. seeds) \(Ottoboni et al., 1993) \(Fig. 1B).

Reagents

• DNeasy Plant Mini Kit \(QIAGEN, cat. no. 69106). • dNTPs \(Invitrogen, cat. no. 10297018). • Restriction endonucleases. • A set of primers for sequence amplification \(Table 2). • PCR cloning vector \(Promega pGEM-T Easy Vector Systems I, cat. no. A1360). • T4 DNA ligase \(Invitrogen, cat. no. 15224017). • Platinum Taq DNA Polymerase High Fidelity \(Invitrogen, cat. no. 11304-011) • QIAquick Gel Extraction Kit \(QIAGEN, cat. no. 28704) • Absolute ethanol \(Sigma-Aldrich Chemical Co. Ltd., cat. no. 459844) • Isopropanol \(2-propanol; Sigma-Aldrich Chemical Co. Ltd., cat. no. 67630) • Plasmid DNA isolation Kit \

(QIAGEN Plasmids Kits, cat. no. 12162) • Plant material: young, green, healthy soybean leaves from mature seeds from BR-16 cultivar were used in the protocol \(supplied by Embrapa Soybean, Rodovia Carlos João Strass, Distrito de Warta, C.P. 231, CEP 86001-970, Londrina, PR, Brazil). • Sterile distilled deionised water. • Tris, ultra pure \(Invitrogen, cat. no. 15504-020). • Paraformaldehyde \($C_7H_{14}O$; J.T. Baker, cat. no. 30525894). • Glutaraldehyde \(25% vol/vol) in Milli-Q water \($C_5H_8O_2$; Electron Microscopy Sciences, cat. no. 16530). • Sodium Cacodylate Trihydrate \($C_2H_6AsNaO_2.3H_2O$; Electron Microscopy Sciences, cat. no. 124652). • LR White embedding resin and Benzoyl peroxide accelerator \(Electron Microscopy Sciences, cat. no. 14381). • Gold conjugate protein A 5 nm \(SPI Supplies, cat. no. 04283-AB). • Sodium phosphate, dibasic, anhydrous \(Na_2HPO_4, Sigma-Aldrich Chemical Co. Ltd., cat. no. 7558794). • Potassium phosphate monobasic \(KH_2PO_4, Sigma-Aldrich Chemical Co. Ltd., cat. no. 7778770). • Sodium chloride \(NaCl, Sigma-Aldrich Chemical Co. Ltd., cat. no. 7647145). • Potassium chloride \(KCl, Sigma-Aldrich Chemical Co. Ltd., cat. no. 7447407). • Bovine serum albumin \(BSA, Sigma-Aldrich Chemical Co. Ltd., cat. no. 9048-46-8).

Equipment

• PTC-100 Peltier thermal cycler \(Bio-Rad, cat. no. PTC-1196). • Thin wall 0.2 ml PCR tubes \(Midwest Scientific, cat. no. AVTW2). • Refrigerated microcentrifuge \(Eppendorf 5415R, Fisher, cat. no. 05-401-05). • 1.5-ml Polypropylene microcentrifuge tubes \(Fisher Scientific Ltd., cat. no. 05-669-32). • Wide-mouth bottles, 250 ml \(Fisher Scientific Ltd., cat. no. 02911916). • Filters, 0.22 um \(Fisher Scientific Ltd., cat. no. MPGL02GH2). • NanoDrop 3300 \(Thermo Scientific). • Ultramicrotome \(model EM UCT, Leica Microsystems, Vienna, Austria). • Transmission electron microscope \(model DSM 962, Carl Zeiss, Oberkochen, Germany).

Procedure

A. Plant material Genomic DNA of the variety BR-16 \(Embrapa) was isolated from young leaves with the DNeasy Plant Kit \(QIAGEN), following the manufacturer's instructions, and quantified with a NanoDrop 3300 \(Thermo Scientific). The concentration was confirmed by analysis with a 1% agarose gel stained with ethidium bromide and adjusted to 10 ng/µl with distilled water. The DNA was kept at -20 °C until it was analysed using polymerase chain reaction \(PCR). ■ PAUSE POINT Genomic DNA and PCR products can be stored at -20 °C for several days without a loss of quality. B. Cloning of the regulatory sequences For amplification of the signal peptide and the promoter region of the alpha' subunit from the beta-conglycinin soybean protein, the primers CongF and CongR were constructed. These primers were chosen based on the alpha' subunit promoter and signal peptide sequence available in GenBank \(accession number: M13759). The primer sequences can be viewed in Table 1 and include sites for the restriction enzymes _Xho_ I and _Hin_d III. The primers anneal to the start of the promoter region and to the end of the signal peptide, respectively, resulting in the amplification of a 1,148-bp fragment when using Platinum _Taq_ DNA Polymerase High Fidelity, according to the manufacturer's \(Invitrogen) instructions, and a PTC-100 Peltier thermal cycler \(Tables 2 and 3). The PCR fragment was purified using a QIAquick Gel

Extraction Kit \(QIAGEN) and quantified using a NanoDrop. The final concentration was adjusted with nuclease free water to 5 ng/ul and utilised for cloning into the pGEM-T Easy cloning vector \(Promega) \(Table 4). The vector pGEM-T Easy, containing the amplified fragment, was digested with the restriction enzymes _Xho_ I and _Hin_d III, and the resulting fragment was cloned into the pBluescript KS^+ \(Stratagene), utilising the same sites \(Table 4). After generation of the intermediate vector, two other primers \(named 35STF and 35STR) were designed to amplify a 228-bp fragment containing the CaMV 35S terminator \(Table 1), utilising the pCAMBIA 2300 vector \(http://www.cambia.org/daisy/cambia/585.html) as a template and creating the sites _Eco_ RI and _Bam_ HI. The PCR fragment was generated, purified from 1% agarose gel as described above, cloned into a pGEM-T Easy vector and digested by _Bam_ HI and _Eco_ RI. After purification, the insert was cloned into the previously obtained vector to generate the vector pbetaCong1. The vector pbetaCong1 was utilised as a model to clone ten different genes that encoded various molecules \(Table 5). We always utilised _Hin_d III and _Eco_ RI sites to clone these genes \(Fig. 2). Each construct was utilised at the concentration of 1 $ug.ul^{-1}$ in the soybean transformation experiments. All of the vectors obtained were completely sequenced by Macrogen Corp. \(USA). We also tested the vector pbetaCong3 \(Fig. 1B, kindly provided by Roger N. Beachy) to express four different genes in the PSVs of soybean seeds. The pbetaCong3 vector contains the promoter from the soybean alpha' subunit of beta-conglycinin and the CaMV35S poly A sequence. Below, we describe each construct in greater detail: a) pbetaCong3hGH – The sequence that contains the human growth hormone and the coix signal peptide was digested from pRT-PGK-PShGH \(Leite et al., 2000) with the restriction enzymes _Nco_ I and _Bam_ HI and cloned into the pbetaCong3 utilising the same sites. b) pbetaCong3scFv – The sequence that contains the scFv and the coix signal peptide was digested from pRT-PGK-PSscFv \(kindly provided by Adilson Leite, Unicamp), with the restriction enzymes _Nco_ I and _Bam_ HI and cloned into the pbetaCong3 vector utilising the same sites. The fragment containing the alpha-coixin signal peptide and the 805-bp fragment _scfv_ \(Oppezzo et al., 2000) was cloned into the vector pbetaCong3, utilising the restriction enzymes _Nco_ I and _Bam_ HI, to generate the 4,826-bp vector pbetaCong3scFv \(Fig. 1). The recombinant protein included a carboxy terminal 6xHis-tag. The plasmids pbetaCong3scFv and pAC 321 were utilised for soybean genetic co-transformation experiments in a 1:1 ratio. c) pbetaCong3CD18 – We constructed primers to add the monocot signal peptide α-coixin from _Coix_ _lacrima_-jobi L., which permitted the addition of _Nco_ I and _Bam_ HI sites before and at the terminus of each molecule of CD-18 \(Table 6). d) pbetaCong3FIX – We constructed primers to add the monocot signal peptide α-coixin from _Coix_ _lacrima_-jobi L., which permitted the addition of _Nco_ I and _Bam_ HI sites before and at the terminus of FIX \(Table 6). The construction scheme is shown in Fig. 3. C. Soybean transformation The vector that was constructed was utilised to obtain transgenic soybean plants that expressed several genes of interest to molecular farming \(Table 5) by bombardment of the somatic embryonic axes from mature soybean seeds cv. BR-16, utilising a particle bombardment procedure. The genetic transformation process, greenhouse plant cultivation and molecular analysis were conducted according to an article published in _Nature_ _Protocols_ \(http://www.nature.com/nprot/journal/v3/n3/full/nprot.2008.9.html). D. Immunocytochemical analysis For immunolocalisation of the cited heterologous molecules mature transgenic and nontransgenic seeds were sliced \(2 mm thick) and fixed for 4 h at 4 °C in a 0.05 M

sodium cacodylate buffer solution containing 2% paraformaldehyde and 0.5% glutaraldehyde at pH 7.2. After fixation, the seed slices were washed three times in fixation buffer and dehydrated for 5 h in a series of ethanol solutions \(30, 50, 70, 95 and 100%, 1 h each, under -20 °C and partial vacuum). Following ethanol desitration, the samples were infiltrated with increased concentrations \(30 to 100%) of ethanol used to dilute LR White resin \(SPI Supplies, USA) in 3 h increments, followed by an 8 h incubation period in pure LR White resin. Inclusion was done by transferring the samples to 1.98% benzoyl peroxide in LR White resin and incubating at 4 °C under UV light for 72 h. Ultra-thin sections \(50 nm thick) were collected in 400 mesh nickel nets. The nets were then incubated for 1 h at room temperature with 1X PBS-T \(10 mM Na_2HPO_4, 1.7 mM KH_2PO_4, 140 mM NaCl, 2.7 mM KCl, 0.5% Tween 20 and 2% bovine serum albumin \(BSA)) and for 2 h with rabbit polyclonal anti-specific molecule \(1:500, 24 ng/ul) diluted in PBS-T. The samples were washed for 1 h in 1X PBS and incubated in gold conjugate protein A \(SPI Supplies) in PBS solution \(20 nM) for 2 h at room temperature. After being washed in PBS and dried for 24 h, the samples that were contrasted with 1% uranile acetate in a 0.1 M PBS solution were observed under a Zeiss DSM 962 transmission electron microscope \(Carl Zeiss, Germany).

Anticipated Results

Seeds have the intrinsic capability for stable protein accumulation that makes them a suitable bioreactor platform for the manufacture of different recombinant molecules. Exploiting the capacity of soybean seeds regulatory sequences it is possible to manipulate and induce the synthesis of functional proteins. The results presented in this protocol allowed the introduction of genes for expressing different biomolecules of pharmaceutical or industrial interest. The construction of a seed-specific transformation vector should result in a fully functional plasmid to express different recombinant proteins in the soybean seeds protein storage vacuoles. Transgenic shoots harboring the seed-specific expression vector should regenerate under the selectable conditions described by Rech et al., 2008. Analytical assays demonstrated that the recombinant proteins presented in total soluble protein extracts from transgenic soybean seeds are expected to present the correct molecular weights in western blot and be detected only in the seeds, being absent in roots, stem, flowers and leaves \(Cunha et al., 2010b and 2010c). The subcelullar compartimentalization obtained by the utilization of the promoter and the signal peptide from the α' portion of the β-conglycinin or the α-coixin cotyledonary vacuolar signal peptide from _Coix lacryma_-jobi \(Poaceae), resulted in the stable accumulation of the recombinant proteins in the transgenic seeds, even when these seeds remained stored for six years under room temperature. Also, the sequencing of the protein fragments by trypsin digestion followed by nanoLC-MS assay can confirm the expected monoisotopic masses and positions of the portions covered by the spectrum assay, without the PSV targeting signal peptide detection \(Cunha et al., 2010b and c).

References

1. Abud, S. et al. Dispersão de pólen em soja transgênica na região do Cerrado. Pesq. Agropec. Bras. 38, 1229-1235 \(2003). 2. Abud, S. et al. Gene flow from transgenic to nontransgenic soybean plants in the

Cerrado region of Brazil. Genet. Mol. Res. 6, 345-352 \(2007). 3. Avesani, L. et al. Improved in planta expression of the human islet autoantigen glutamic acid decarboxylase \(GAD65). Transgenic Res. 12, 203–212 \(2003). 4. Bednarek, S.Y. & Raikhel, N.V. The barley lectin carboxyl-terminal propeptide is a vacuolar protein sorting determinant in plants. Plant Cell. 3, 1195–1206 \(1991). 5. Boothe, J. et al. Seed-based expression systems for plant molecular farming. Plant Biotechnol. J. 8, 588-606 \(2010). 6. Cantoral, R., Fernández-Quintela, A., Martínez, J.A. & Macarulla, M.T. Comparative study of the composition and nutritional value of the seeds and protein concentrations in legumes. Arch. Latinoam. Nutr. 45, 242-248 \(1995). 7. Chen, Z.L., Pan, N.S. & Beachy, R.N. A DNA sequence element that confers seed-specific enhancement to a constitutive promoter. EMBO J. 7, 297 – 302 \(1988). 8. Chen, Z.L., Schuler, M.A. & Beachy, R.N. Functional analysis of regulatory elements in a plant embryo-specific gene. Proc. Natl. Acad. Sci. USA. 83, 8560-8564 \(1986). 9. Conrad, U. & Fiedler, U. Compartment-specific accumulation of recombinant immunoglobulins in plant cells: an essential tool for antibody production and immunomodulation of physiological functions and pathogen activity. Plant Mol. Biol. 38, 101–109 \(1998). 10. Cunha, N.B. et al. Correct targeting of proinsulin in protein storage vacuoles of transgenic soybean seeds. Genet. Mol. Res. 9, 1163-1170 \(2010a). 11. Cunha, N.B. et al. Expression of functional recombinant human growth hormone in transgenic soybean seeds. Transgenic Res. \(2010b). 12. Cunha, N.B. et al. Accumulation of functional recombinant human coagulation factor IX in transgenic soybean seeds. Transgenic Res. \(2010c). 13. De Jaeger, G. et al. Boosting heterologous protein production in transgenic dicotyledonous seeds using _Phaseolus vulgaris_ regulatory sequences. Nat. Biotechnol. 20, 1265–1268 \(2002). 14. Derbyshire, E., Wright, D.J. & Boulter, D. Legumin and vicilin, storage proteins of legume seeds. Phytochemistry. 15, 3-24 \(1976). 15. Doyle, J.J., Schuler, M.A., Godette, W.D., Zenger, V. & Beachy, R.N. The glycosylated seed storage proteins of _Glycine max_ and _Phaseolus vulgaris_. J. Biol. Chem. 261, 9228-9238 \(1986). 16. Dreesen, I.A.J., Hamri, G.C. & Fussenegger, M. Heat-stable oral alga-based vaccine protects mice from _Staphylococcus aureus_ infection. J. Biotechnol. 145, 273-280 \(2010). 17. Evangelista, R.L., Kusnadi, A.R., Howard, J.A. & Nikolov, Z.L. Process and economic evaluation of the extraction and purification of recombinant β-glucuronidase from transgenic corn. Biotechnol. Prog. 14, 607–614 \(1998). 18. Fiedler, U., Phillips, J., Artsaenko, O. & Conrad, U. Optimization of scFv antibody production in transgenic plants. Immunotechnology. 3, 205-216 \(1997). 19. Frigerio, L., de Virgilio, M., Prada, A., Faoro, F. & Vitale, A. Sorting of phaseolin to the vacuole is saturable and requires a short C-terminal peptide. Plant Cell. 10, 1031–1042 \(1998). 20. Herman, E.M. & Larkins, B.A. Protein storage bodies and vacuoles. Plant Cell. 11: 601-613 \(1999). 21. Hood, E.E., Woodard, S.L. & Horn, M.E. Monoclonal antibody manufacturing in transgenic plants - myths and realities. Curr. Opin. Biotechnol. 13, 630-635 \(2002). 22. Jiang, L. & Sun, S.S. Membrane anchors for vacuolar targeting: application in plant bioreactors. Trends Biotechnol. 20, 99-102 \(2002). 23. Jolliffe, N.A. Craddock, C.P. & Frigerio, L. Pathways for protein transport to seed storage vacuoles. Biochem. Soc. Trans. 33, 1016–1018 \(2005). 24. Lau, O.S. & Sun, S.S. Plant seeds as bioreactors for recombinant protein production. Biotechnol. Adv. 27, 1015-1022 \(2009). 25. Leite, A. et al. Expression of correctly processed human growth hormone in seeds of transgenic tobacco plants. Mol. Breed. 6, 47-53 \(2000). 26. Ma, J.K.C., Drake, P.M.W. & Christou, P. The production of recombinant pharmaceutical proteins in plants. Nat. Rev. Genet. 4, 794-805 \(2003). 27. Moravec, T., Schmidt, M.A., Herman, E.M. & Woodford-Thomas, T. 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 \(2007). 28. Muntz, K. Deposition of storage proteins. Plant. Mol. Biol. 38, 77-99 \(1998). 29. Nishizawa, K. et al. A C-terminal sequence of soybean beta-conglycinin alpha' subunit acts as a vacuolar sorting determinant in seed cells. Plant J. 34, 647–659 \(2003). 30. Orzáez, D., Granell, A. & Blázquez, M.A. Manufacturing antibodies in the plant cell. Biotechnol. J. 4, 1712-1724 \(2009). 31. Ottoboni, L.M.M. et al. Sequence analysis of 22kDa-like alpha-coixin genes and their comparison with homologous zein and kafirin genes reveals highly conserved protein structure and regulatory elements. Plant Mol. Biol. 21, 765-778 \(1993). 32. Oppezzo, P. et al. Production and functional characterization of two mouse/human chimeric antibodies with specificity for the tumor-associated Tn-antigen. Hybridoma 19, 229-239 \(2000). 33. Perrin, Y. et al. Transgenic pea seeds as bioreactors for the production of a single-chain Fv fragment \(scFv) antibody used in cancer diagnosis and therapy. Mol. Breed. 6, 345–352 \(2000). 34. Rech, E.L., Vianna, G.R. & Aragão, F.J.L. High-efficiency transformation by biolistics of soybean, common bean and cotton transgenic plants. Nat. Protoc. 3, 410-418 \(2008). 35. Robinson, D.G., Oliviusson, P. & Hinz, G. Protein sorting to the storage vacuoles of plants: a critical appraisal. Traffic. 6, 615–624 \(2005). 36. Semenyuk, E.G., Schmidt, M.A., Beachy, R.N., Moravec, T. & Woodford-Thomas, T. Adaptation of an ecdysone-based genetic switch for transgene expression in soybean seeds. Transgenic Res. 19, 987–999 \(2010). 37. Stoger, E. et al. Cereal crops as viable production and storage systems for pharmaceutical scFv antibodies. Plant Mol. Biol. 42, 583-590 \(2000). 38. Stoger, E., Ma, J.K., Fischer, R. & Christou, P. Sowing the seeds of success: pharmaceutical proteins from plants. Curr. Opin. Biotechnol. 16, 167-173 \(2005). 39. Streatfield, S.J. Approaches to achieve high-level heterologous protein production in plants. Plant Biotechnol. J. 5, 2-15 \(2007). 40. Takagi, H. et al. A rice-based edible vaccine expressing multiple T cell epitopes induces oral tolerance for inhibition of Th2-mediated IgE responses. Proc. Natl. Acad. Sci. USA. 102, 17525-17530 \(2005). 41. Takaiwa, F., Takagi, H., Hirose, S. & Wakasa, Y. Endosperm tissue is good production platform for artificial recombinant proteins in transgenic rice. Plant Biotechnol. J. 5, 84-92 \(2007). 42. Verma, D., Samson, N.P., Koya, V. & Daniell, H. A protocol for expression of foreign genes in chloroplasts. Nat. Protoc. 3, 739 – 758 \(2008). 43. Vitale, A. & Hinz, G. Sorting of proteins to storage vacuoles: how many mechanisms? Trends Plant Sci. 10, 316-323 \(2005). 44. Vitale, A. & Pedrazzini, E. Recombinant pharmaceuticals from plants: the plant endomembrane system as bioreactor. Mol. Interv. 5, 216–225 \(2005). 45. Woodard, S.L., Wilken, L.R., Barros, G.O., White, S.G. & Nikolov, Z.L. Evaluation of monoclonal antibody and phenolic extraction from transgenic Lemna for purification process development. Biotechnol Bioeng. 104, 562-571 \(2009). 46. Yang, L., Suzuki, K., Hirose, S., Wakasa, Y. & Takaiwa, F. Development of transgenic rice seed accumulating a major Japanese cedar pollen allergen \(Cry j 1) structurally disrupted for oral immunotherapy. Plant Biotechnol. J. 5, 815–826 \(2007). 47. Yang, L. et al. A transgenic rice seed accumulating an anti-hypertensive peptide reduces the blood pressure of spontaneously hypertensive rats. FEBS Lett. 580, 3315–3320 \(2006). 48. Yoo, B.Y. & Chrispeels, M.J. The origin of protein bodies in developing soybean cotyledons: a proposal. Protoplasma. 103: 201-204 \(1980). 49. Zheng, Y., He, M., Hao, S. & Huang, B. The ultrastructural evidence on the origin of protein bodies in the rough endoplasmic reticulum of developing cotyledons of soybean. Ann. Bot. 69, 377-383 \(1992).

Acknowledgements

We are grateful to Dr. Roger N. Beachy from the Donald Danforth Plant Science Center \(St. Louis, MO, USA) for providing the alpha' subunit of the beta-conglycinin promoter and terminator sequence and to Ana C.M.M. Gomes from Bio Image Laboratory from Embrapa Genetic Resources and Biotechnology for assistance with scanning electron microscopy. This study was supported in part by the National Council for Scientific and Technological Development \(CNPq), Fundação de Apoio a Pesquisa \(FAP-DF) and Embrapa Genetic Resources and Biotechnology.

Figures

Figure 1

Figure 1 A) Sequence of the vector pbetaCong1 that contains the alpha' subunit promoter and signal peptide from soybean beta-conglycinin and CaMV35S poly A. B) Sequence of the vector pbetaCong3 that contains the alpha' subunit promoter and the signal peptide α-coixin from _Coix lacrima_-jobi L and alpha'

subunit beta-conglycinin poly A. The promoters are in black, the signal peptides are in red, and the terminators are in blue. Underlining denotes the sites included in the sequences to clone all the regulatory and coding sequences. CTCGAG, _Xho_ I; AAGCTT, _Hin_d III; GAATTC, _Eco_ RI ; GGATCC, _Bam_ HI; CCATGG, _Nco_ I.

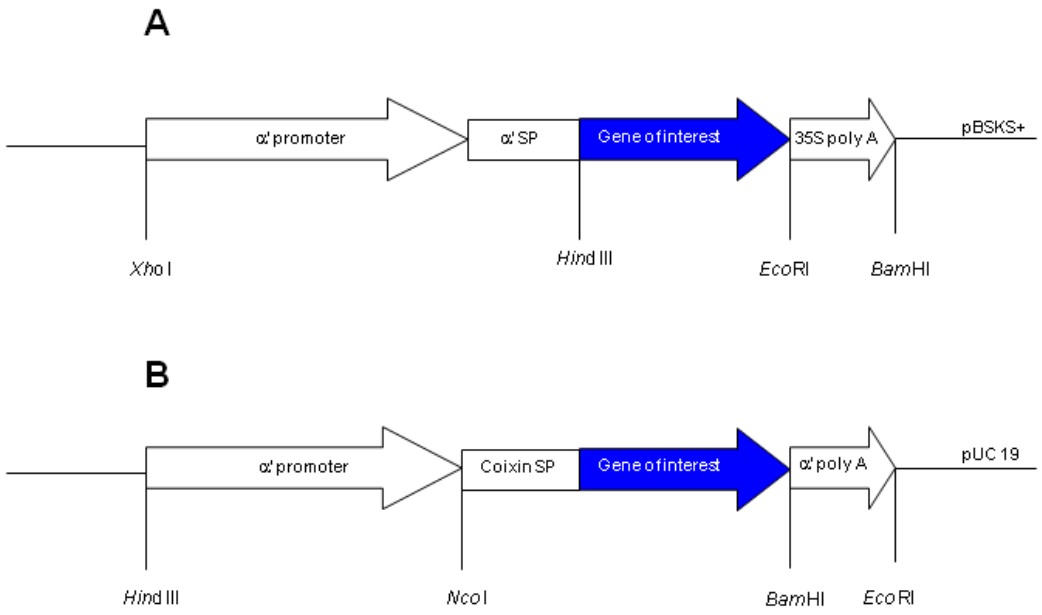

Figure 2

Figure 2 Schematic representations of the vectors used for soybean seed heterologous protein expression. These maps show the alpha' subunit promoter and signal peptide from soybean beta-conglycinin and CaMV35S poly A. The sites _Hin_d III and _Eco_ RI indicate the site for cloning the gene of interest. A) PβCong1; B) pβCong3.

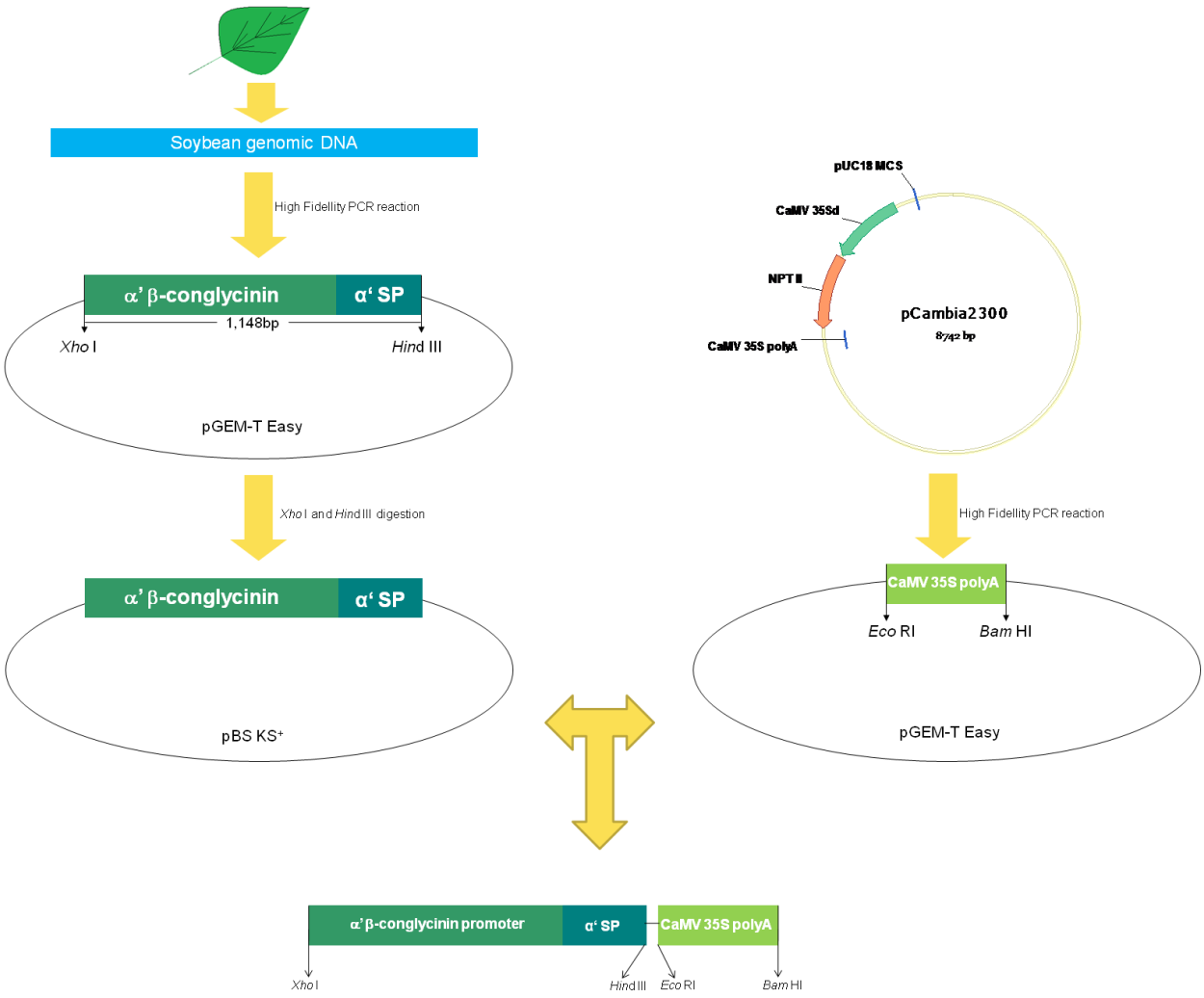

Figure 3

Figure 3 Scheme for construction of the vector pβCong1 used for heterologous expression of biomolecules in PSVs from soybean seeds, based on the use of the promoter and signal peptide portion of the α' subunit from β-conglycinin from soybean.

Primer name	Sequence (5' to 3')	GenBank Accession No.	Restriction site	Fragment size (bp)
CongF	ACTCGAGGTTTTCAAATTTGAA	M13759	*Xho* I	1,148
CongR	AAAGCTTAAGGAGGTTGCAACG		*Hind* III	
35STF	CGGATCCCTGAATTAACGCCGAA	AF234315	*Bam* HI	228
35STR	CGAATTCTCGACAAGCTCGAGT		*Eco* RI	

Figure 4

Table 1 Table 1 Primer sequences used for cloning the regions of interest in the soybean beta-conglycinin and CaMV 35S poly A genes, as well as their accession numbers in GenBank (http://www.ncbi.nlm.nih.gov). Underlining denotes the restriction sites utilised for cloning.

Component	Quantity (µl)	Initial concentration	Final concentration
PCR Buffer	2.5	10X	1X
dNTP mix	0.5	10 mM (each)	0.2 mM
$MgSO_4$	1.0	50 mM	2 mM
Primer 1	0.5	10 µM	200 nM
Primer 2	0.5	10 µM	200 nM
Taq Platinum high fidelity	0.2	5 U/µl	1 U/reaction
Template (Genomic DNA)	2.0	10 ng/µl	20 ng/reaction
Nuclease free water	17.8		
Total	**25.0**		

Figure 5

Table 2 Table 2 Platinum Taq DNA polymerase high fidelity PCR reaction conditions to amplify the sequences utilised in this work.

Objective	Temperature (°C)	Time (minutes)	No. cycles
Initial denaturation	94	2.0	01
Denaturation	94	0.5	35
Annealing	55	0.5	35
Elongation	68	2.0	35
Final elongation	68	5.0	01
Conservation	04	∝	

Figure 6

Table 3 Table 3 PCR Platinum high-fidelity reaction conditions utilised to amplify the fragments.

Component	Quantity (µl)	Final concentration
Rapid Ligation Buffer 2X	5.0	1X
pGEM®-T Easy Vector (50 ng/µl)	1.0	50 ng/reaction
T4 DNA Ligase (5 U/µl)	1.0	5 U/reaction
PCR product (5 ng/µl)	3.0	15 ng/reaction
Nuclease free water	0.0	
Total	**10.0**	

Figure 7

Table 4 Table 4 Ligation conditions to clone fragments into the pGEM-T Easy vector (Promega).

Molecule	Vector construct	Size (bp)	Characteristic
hGH	pβCong3hGH and pβCong1hGH	576	Human Growth Hormone
CD-18	pβCong3CD18 and pβCong1CD18	768	*Mus musculus* immunoglobulin gamma heavy chain variable (nucleotide sequence of the variable region of an Anti-CD18 antibody) and *Mus musculus* immunoglobulin kappa light chain variable region (IgK) (nucleotide sequence of the variable region of an anti-CD18 antibody)
ScFvDIR83D4	pβCong3scFv	720	*Mus musculus* anti-Tn immunoglobulin heavy chain monoclonal antibody VH) gene, and *Mus musculus* anti-Tn immunoglobulin light chain monoclonal antibody (Vk)
hFIX	pβCong3FIX and pβCong1FIX	1,248	Human coagulation factor IX (plasma thromboplastic component, Christmas disease, haemophilia B)
SVN	pβCong1SVN	285	Synthetic construct scytovirin gene, a potent novel anti-HIV protein from the cultured *Cyanobacterium scytonema* varium
CVN	pβCong1CVN	306	Synthetic gene cyanovirin-N fragment anti-virus protein
CVNm	pβCong1CVNm	306	Synthetic gene cyanovirin-N fragment anti-virus protein – mutated
CTAG 1B	pβCong1CTAG	543	Cancer/testis antigen
PLAC 1	pβCong1PLAC	639	Placenta-specific 1 [*Homo sapiens*]. Plac1 is a tumour-specific antigen capable of eliciting spontaneous antibody responses in human cancer patients
GAGE 2B	pβCong1GAGE	351	G antigen 2B – Cancer/testis antigen
GRFT	pβCong1GRFT	420	Synthetic construct His-tagged griffithsin – a novel HIV-inactivating protein, from the red alga *Griffithsia* sp

Figure 8

Table 5 Table 5 Different heterologous molecules expressed into the PSVs of soybean seeds and their respective constructs.

Primer name	Sequence (5' to 3')	Restriction site	Fragment size (bp)
FIXF	ACCATGGCTACCAAGATATTTGCCCTCCTTGTGCTCCTTGCTCTTTCAGCGAGC GCTACAACTGCGACAGTTTTTCTTGATCATGA	*Nco* I	1,321
FIXR	AGGATCCTTAAGTGAGCTTTGTTTTTT	*Bam* HI	
CD18F	AGATATCATGGCTACCAAGATATTTGCCCTCCTTGTGCTCCTTGCTCTTTCAGC GA GCGCTACAACTGCGCGGGCTCAAGTTCAGTTG	*Eco* RV	1,038
CD18R	AGGATCCGAGCTCGGTACCTTAAGC	*Bam* HI	

Figure 9

Table 6 Table 6 Primer sequences utilised for cloning four different molecules and the signal peptide from alpha-coix. Underlining denotes the restriction sites utilised for cloning.

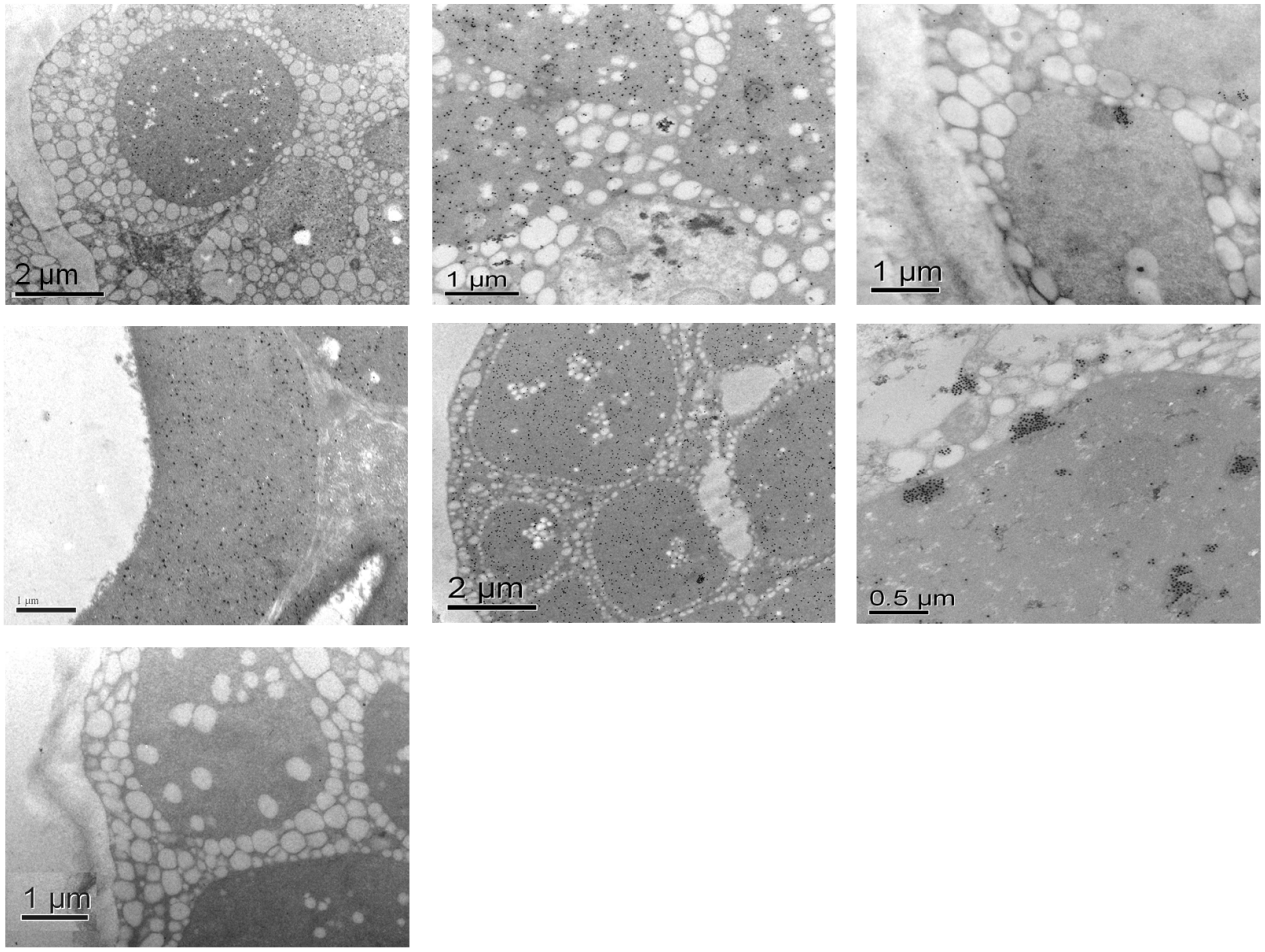

Figure 10

Figure 4 Figure 4 Evaluation by ultrastructural immunocytochemistry of the accumulation of different recombinant proteins in protein storage vacuoles (PSVs) in ultra-thin sections of soybean cotyledon of transgenic seeds. (A) Microbicide cyanovirin-N; (B) human Growth Hormone; (C) single chain fragment variable of the anti-Tn monoclonal antibody; (D) mutated cyanovirin-N; (E) human blood coagulation factor IX; (F) single chain fragment variable of an Anti-CD18 antibody, and (G) negative control.