BMC Plant Biology

RESEARCH

Open Access

CRISPR/Cas9 editing of β-Conglycinin subunits reduces IgE binding in soybean [*Glycine max* (L.) Merr.]

Hye Rang Park[1†], Sewon Park[2†], Joo-Mi Jun[2], Yoon Ji Shin[3], Yeojin Hwang[4], Kyoung Yong Jeong[3], Min Young Kim[5], Sun Tae Kim[6], Sangjun Park[1], Yo-Han H. Yoo[1], Eunsoo Lee[1], Girim Park[1], Sang-Gyu Kim[4*] and Soo-Kwon Park[1*]

Abstract

Background Soybean [*Glycine max* (L.) Merr.] is a major source of plant-based protein, yet the seed storage protein β-conglycinin (7 S globulin) is a prominent allergen. The α′, α, and β subunits contain IgE-binding epitopes, and their high sequence similarity enables simultaneous genome editing. The development of soybean lines with reduced β-conglycinin-specific IgE-binding capacity could enhance food safety for individuals with soy allergies.

Results We employed CRISPR/Cas9 to disrupt the α′ (*Glyma.10G246300*) and α (*Glyma.20G148300*, *Glyma.20G148400*), subunit genes and to target the β subunit genes (*Glyma.20G146200*, *Glyma.20G148200*) of β-conglycinin, generating four edited lines: SP1 (α′-null), SP2 (α′α-null), SP3 (β-null), and SP4, which shows an α′α-edited genotype and a β subunit-null protein phenotype. SDS-PAGE and DNA sequencing confirmed complete or near-complete loss of the targeted proteins across the T_0 to T_6 generations, demonstrating stable inheritance of the edited seed protein profiles. IgE immunoblotting and inhibition ELISA using pooled sera from soy-allergic individuals revealed distinct IgE-binding inhibition profiles among the edited lines. At the highest inhibitor concentration, SP4 showed the lowest IgE-binding inhibition (70.0%) compared with the wild type (87.7%), whereas SP1-SP3 exhibited inhibition values similar to or only slightly lower than those of the wild type.

Conclusions CRISPR/Cas9-mediated elimination of β-conglycinin subunits reduces IgE binding to soybean seed proteins and yields lines with stably inherited seed protein phenotypes. These results highlight the potential of targeted genome editing to generate soybean lines with reduced β-conglycinin-specific IgE recognition, supporting the application of precise genome modification in crop improvement for safer soy-based foods.

Keywords Gene editing, β-conglycinin, IgE binding, Seed storage protein, Food allergy

[†]Hye Rang Park and Sewon Park contributed equally to this work.

*Correspondence:
Sang-Gyu Kim
sgkim1@kaist.ac.kr
Soo-Kwon Park
sookwonpark@korea.kr
[1]Department of Upland Crop Research and Development, National Institute of Crop and Food Science, Rural Development Administration, Miryang 50424, Republic of Korea
[2]Department of Crop Foundation Research, National Institute of Crop and Food Science, Rural Development Administration, Wanju 55365, Republic of Korea
[3]Department of Internal Medicine, Institute of Allergy, Yonsei University College of Medicine, Seoul 03722, Republic of Korea
[4]Department of Biological Sciences, KAIST, Daejeon 34141, Republic of Korea
[5]Department of Food Sciences, National Institute of Crop and Food Science, Rural Development Administration, Wanju 55365, Republic of Korea
[6]Department of Plant Bioscience, Life and Industry Convergence Research Institute, Pusan National University, Miryang 627706, Republic of Korea

Introduction

Soybean [*Glycine max* (L.) Merr.] stands as a cornerstone crop in global agriculture, offering a rich source of protein and oil for food and feed [1]. Soy proteins, which constitute roughly 40% of the seed dry weight, are predominantly composed of two storage protein classes, β-conglycinin (7 S globulin) and glycinin (11 S globulin), with β-conglycinin alone representing nearly 30% of the total seed protein content [2–5]. β-conglycinin is a glycoprotein of approximately 150 kDa composed of three subunits, α′ (76 kDa), α (72 kDa), and β (52 kDa), that share high amino acid sequence homology and assemble into diverse trimeric configurations (α′3, α3, α′2β, α2β, α′β2, αβ2, and β3) [4, 6–8]. A total of 15 related genes (*CG-1* to *CG-15*) have been identified, with *CG-1* encoding the α′ subunit, *CG-2* and *CG-3* encoding the α subunit, and *CG-4* encoding the β subunit [9–12], whereas the remaining CG genes are less well characterized [12]. All three major subunits (α′, α, and β) contain IgE-binding epitopes and are clinically recognized soy allergens associated with a broad range of symptoms, including gastrointestinal, respiratory, and cardiovascular manifestations, as well as potentially life-threatening anaphylaxis in sensitive individuals [4, 13–15]. With the increase in global soy consumption, IgE reactivity to β-conglycinin has become a significant food safety concern, making the reduction of β-conglycinin-specific IgE binding an important objective in soybean improvement [4, 14, 16, 17].

Among the more than 16 identified soybean allergens, including Gly m 1 to Gly m 8 [5, 18, 19], β-conglycinin (Gly m 5) is recognized as a major allergen and is implicated in IgE-mediated reactions in a substantial proportion of soy-sensitive patients [20–22]. Furthermore, processed derivatives such as Gly m Bd 30 K (P34, 30 kDa) and Gly m Bd 28 K (P28, 28 kDa) are detected by 65% and 62.5% of soy-allergic individuals, underscoring the central role of β-conglycinin and its related proteins in soy-induced IgE recognition [5, 20–22]. A variety of processing methods, including thermal treatment, fermentation, enzymatic catalysis, high-pressure processing, and glycosylation, have been investigated to attenuate the IgE binding capacity of soy proteins and to reduce overall soy allergenicity [23–27]. These treatments can partially reduce IgE reactivity but do not completely eliminate it [5]. Consequently, combining soybean lines with reduced IgE-binding capacity, generated through genetic modification, with conventional processing technologies has been proposed as a more comprehensive strategy to mitigate IgE-mediated responses to soy [19].

Previous studies utilizing RNA interference (RNAi) [28] targeting β-conglycinin provided proof-of-concept that lowering the accumulation of this storage protein can reduce IgE binding to soybean extracts. However, these knockdown strategies were largely confined to CG1 and typically resulted in incomplete suppression, with residual protein accumulation and limited stability across generations. More recently, a CRISPR/Cas9-based system targeting the intronic region of the long intergenic noncoding RNA *lincCG1* simultaneously reduced the accumulation of the α′, α, and β subunits [29]. Although this approach broadened the impact beyond a single coding gene, it remained centered on the *lincCG1* locus and did not yield a panel of stable, homozygous lines selectively lacking individual β-conglycinin subunits. By contrast, the present study directly targets the coding genes corresponding to CG1 through CG4, enabling the generation of single- and double-null mutants for α′, α, or β, as well as a line in which all three β-conglycinin subunits are undetectable at the protein level across multiple generations. This design provides a set of genetically stable lines with distinct β-conglycinin subunit profiles and allows a more detailed assessment of subunit-specific effects on seed composition.

In contrast to the previous CRISPR/Cas9 approach targeting *lincCG1*, which modulates β-conglycinin accumulation at a regulatory locus, the present study directly edits the coding sequences of the α′, α, and β subunits (CG1–CG4), generating a panel of single-, double-, and triple-null lines (SP1–SP4) with clearly defined subunit combinations. The edited phenotypes are stably inherited from T0 to T5/T6 generations and are accompanied by reproducible changes in seed protein, oil, and free amino acid profiles, allowing a more detailed dissection of subunit-specific contributions to nutritional quality and IgE binding than was possible in the *lincCG1* background.

Earlier efforts to genetically manipulate β-conglycinin have shown that the absence of this storage protein does not impair plant growth or reproduction, indicating that its subunits are largely dispensable for normal development [1, 7, 10, 30–32]. Consistent with these observations, our CRISPR/Cas9-edited soybean lines (SP1–SP4) exhibited normal vegetative and reproductive growth while displaying distinct alterations in seed nutritional traits. Specifically, several lines showed increased total seed protein, reduced oil content, and changes in free amino acid profiles, suggesting subunit-specific roles in regulating seed composition. Notably, one line lacking detectable α′, α, and β subunit proteins showed no coding-sequence mutations in the β subunit genes, indicating that the molecular mechanism underlying β subunit loss in this background is not yet resolved and may involve noncoding, transcriptional, or post-transcriptional regulatory processes that warrant further investigation.

In parallel with these advances, genome editing and integrative molecular breeding are increasingly being applied to modulate allergen-encoding proteins and seed quality traits in soybean. Developing soybean germplasm with reduced β-conglycinin-specific IgE reactivity, while

maintaining normal agronomic performance, represents a promising strategy to support safer soy-based foods and to expand the toolkit for precise, health-oriented crop improvement.

Materials and methods

Plant materials and vector constructions

The Korean soybean cultivar, 'Kwangan' was used for *Agrobacterium*-mediated transformation *via* the half-seed method [33]. To design specific small guide RNAs (sgRNAs) target multiple sites, the *Glycine max* reference genome (*Wm82.a6.v1*) was obtained from the Phytozome database (https://phytozome-next.jgi.doe.gov), and the homology of the β-conglycinin genes *Glyma.20G148300*, *Glyma.20G148400*, *Glyma.20G146200*, *Glyma.20G148200*, and *Glyma.10G246300* was assessed. Sequence alignments of β-conglycinin genes were performed using CLUSTAL W [34], and a phylogenetic tree was constructed with MEGA X [35] (Fig. 1a). To summarized the relationships among these genes, we also compiled the targeted CG loci and other β-conglycinin-related genes together with their genomic position and nucleotide and amino acid sequence similarities (Supplementary Table S1). Three specific sgRNAs targeting both the α′ and α subunits or β subunits were designed in specificity to target exon 1 of the corresponding β-conglycinin genes (Table 1, Supplementary Fig. S1 and S2). Prof. Sang-Gyu Kim (Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea) provided the CRISPR/Cas9 vector constructs pECO201:7 S-α′α and pECO201:7 S-β by leveraging a PCR-free multiplex sgRNA cloning system for plants [36]. DH5α-competent cells were used for cloning and the final vector sequences with sgRNA1, 2, and 3 appear in the supplementary figures (Supplementary Fig. S3 and S4). The plasmids pECO201:7 S-α′α and pECO201:7 S-β were transformed into *Agrobacterium tumefaciens* EHA105 to facilitate soybean transformation.

Agrobacterium preparation and soybean transformation

Agrobactrium-mediated soybean transformation was performed as described by Paz et al. (2006), with slight modifications. Mature soybean seeds of 'Kwangan' were sufcace-sterilized with 70% ethanol (Daejung Chemicals, Siheung, South Korea) for 1 min and subsequently with 2% sodium hypochlorite (Sigma-Aldrich, MO, USA) for 20 min. After rinsing the seeds with deionized water more than five times, the seeds were dried and imbibed in 10 cm diameter Petri dishes (SPL Life Sciences, Pocheon, South Korea) for 24 h at 28 °C in the dark. Imbibed seeds were bisected, retaining the half containing the embryo for transformation. The cotyledonary leaves were excised, and the embryo were wounded axially within the explants. Approximately 100 explants were incubated with an *Agrobacterium* suspension (OD650 = 0.6) for over 30 min in 100 mL of liquid co-cultivation medium supplemented with 1.67 mg/L 6-benzyl amino purine (Duchefa Biochemie, Haarlem, Netherlands), 0.25 mg/L gibberellic acid (MBcell, Seoul, South Korea), 0.1 mg/L dithiothreitol (Duchefa Biochemie, Haarlem, Netherlands), and 0.1 mL/L acetosyringone (Sigma-Aldrich, MO, USA). Explants were then subjected to vacuum infiltration for 30 s, allowed to rest for 30 min at room temperature, and transferred (10 explants per plate) to solid co-cultivation medium containing plant agar (Duchefa Biochemie, Haarlem, Netherlands) and incubated them at 22 °C for 3 days with a 16 h photoperiod. Explants were then transferred to shoot induction medium (SIM1) containing 1 mL/L cefotaxime (MBcell, Seoul, South Korea), 1 mL/L of ticarcillin (MBcell, Seoul, South Korea), and 0.5 mL/L of phosphinothricin (PPT) (Duchefa Biochemie, Haarlem, Netherlands), under a 16 h photoperiod at 25 °C under fluorescent light. Emerging shoots were transferred to fresh SIM2 with 0.5 mL/L DL-phosphinothricin (PPT) every 14 days to control contamination and select transformed shoots. This selection process continued on shoot elongation medium containing 0.3 mL/L PPT for up to 14 days. Shoots longer than 5 cm were transferred to rooting medium containing 1 mL/L indol-3-butyric acid (Sigma-Aldrich, MO, USA) to stimulate root formation. After 30 days, the T_0 plants were acclimatized in a greenhouse and grown to maturity under natural light conditions at 26 ± 3 °C.

Screening of transgenic soybean plants via herbicide resistance

The bialaphos resistance (*bar*) gene in pECO201 vectors served as a selection marker. The *bar* strip test were performed according to the manufacturer's instructions using AgraStrip LL SeedChek (Romer Labs, Getzersdorf, Austria). Herbicide resistance in transgenic (T_0 and T_1) soybean plants was assessed by a PPT-painting assay, in which a 100 mg/L PPT solution (Sigma-Aldrich, MO, USA) mixed with Tween-20 (Sigma-Aldrich, MO, USA) was applied to the midrib of the upper surface of a leaf. Herbicidal response was monitored after three days. PPT-resistant transgenic plants were subsequently cultivated in a greenhouse, and T_5 and T_6 seeds were collected.

Screening of transgenic soybean plants via PCR and deep sequencing

To confirm the mutations introduced by the CRISPR/Cas9 system, transgenic soybean plants (T_0 to T_6) were screened *via* polymerase chain reaction (PCR) and deep sequencing. Genomic DNA was extracted from leaf tissues of both the control 'Kwangan' and the transgenic plants using a DNeasy® Plant mini kit (Qiagen, Hilden, Germany) according to the manufacturer's

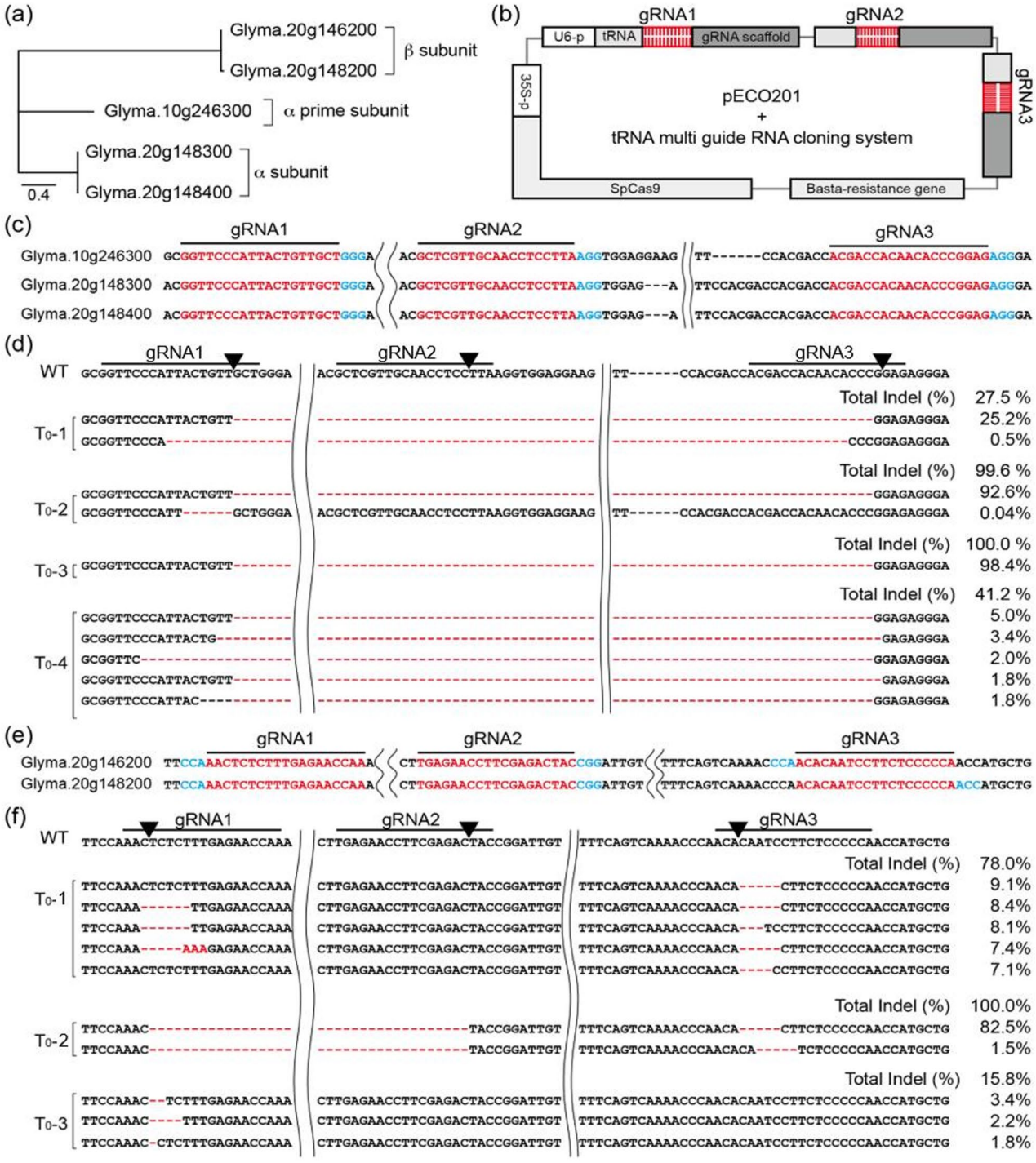

Fig. 1 Generation of T_0β-conglycinin knockout plants. (**a**) A phylogenetic tree drawn for 5 β-conglycinin genes in *Glycine max Wm82.a6.v1*. (**b**) Schematic representation of the three guide RNA(sgRNA) units expressed under a single RNA polymerase III promoter (U6). (**c**) Sequences of the three sgRNAs designed to edit α′ and α subunit genes (*Glyma.10G246300*, *Glyma.20G148300* & *Glyma.20G148400*). (**d**) InDel mutation patterns and frequencies at the target sites of α′ and α subunit genes. The mutation frequencies were calculated by dividing the number of reads containing InDels at the target site. (**e**) Sequences of three sgRNAs designed to target two β subunit genes (*Glyma.20G146200* & *Glyma.20G148200*). (**f**) InDel mutation patterns and frequencies at the target sites of β subunit genes

Table 1 β-conglycinin genes targeted for mutagenesis

Subunit	Gene ID	Chr.	Strand	DNA (bp)	protein (a.a)	predicted protein size(kDa)
CG-1	Glyma.10g246300	10	+	1,866	622	68.3 kDa
CG-2	Glyma.20g148300	20	+	1,818	605	66.6 kDa
CG-3	Glyma.20g148400	20	-	1,818	605	66.6 kDa
CG-4	Glyma.20g146200	20	+	1,320	440	48.3 kDa
CG-4	Glyma.20g148200	20	-	1,320	440	48.3 kDa

instructions. PCR amplification was carried out to detect the presence of transgenes using Prime Taq Premix (GeNetBio, Daejeon, South Korea). Primer pairs were specifically designed (Supplementary Table S2) to span genomic regions encompassing the *7 S* sgRNA target sites, thereby facilitating the detection of mutations within exon 1 of each subunit. To identify the *bar* gene, primers were designed to yield a 412 bp amplicons: forward primer 5′-AAGCACGGTCAACTTCC GTA-3′ and reverse primer 5′-GAAGTCCAGCTGCC AGAAAC-3′. Similarly, insertion of the *SpCas9* gene was confirmed through PCR using primers designed to produce a 1908 bp fragment: forward primer 5′-ATG GATAAGAAGTACTCTATCGG-3′ and reverse primer 5′-GAGCCTTTCTTCAATCATCTCTC-3′.

The insertions and deletions (InDels) at the targeted sites were examined by targeted deep sequencing using an Illumina Miniseq sequencing system (Illumina Inc., CA, USA) at KAIST (Daejeon, South Korea), employing primer sets from the Illumina platform. The CRISPR/Cas9-induced InDels within the β-conglycinin genes for each sgRNA were analyzed using Cas-Analyzer in the CRISPR RGEN tools (https://www.rgenome.net/cas-analyzer/) [37]. The analysis was conducted on raw data FASTQ files using default parameters, including a comparison range of 70 and a minimum frequency threshold of 1. InDel frequency was calculated as the proportion of reads with InDels among the total number of reads at each target site.

Whole genome re-sequencing of 'Kwangan' and SP lines

Whole genome re-sequencing of 'Kwangan' and four lines (SP1-SP4) used in this study was performed using the Illumina NovaSeq 6000 platform at 40x coverage, using previously described mapping and filtering options [38]. The resulting reads were aligned to the reference genome Wm82.a2.v1obtained from the Phytozome(https://phytozome-next.jgi.doe.gov/) database for downstream analysis.

Detection of β-conglycinin subunit proteins via Western blotting

Soybean seeds from pECO201:*7 S* transformats and the 'Kwangan' control were lysed by intermittent vortexing for 1 h in 500 µL of 0.03 M Tris buffer (pH 8.0; Sigma-Aldrich, MO, USA) containing 0.01 M β-mercaptoethanol (Thermo Fisher Scientific, MA, USA), followed by centrifugation at 12,000 × g for 20 min at room temperature to obtain the total protein in the supernatant. Protein concentrations were determined using the Bradford assay [39]. Twenty microliter of total protein extract, diluted to 14 mg/mL with deionized water, was mixed with an equal volume of SDS-sample buffer (0.15 M Tris-HCl, pH 6.8, 4% w/v SDS, 5% v/v β-mercaptoethanol) and heated at 95 °C for 10 min. Ten microliter of this mixture was loaded onto a 10% polyacrylamide gel, and electrophoresis was performed at 50 V for 1 h and then at 100 V until completion. Gels were stained with Coomassie Brilliant Blue R-250 (GE Healthcare, IL, USA) to visualize the overall protein profiles. Proteins were transferred to a membrane, which was blocked with 5% skim milk in phosphate-buffered saline containing 0.1% Tween 20 (PBST) and washed with PBST. The β-conglycinin α′, α, and β subunits were detected using specific antibodies provided by Prof. Kim Sun Tae (Pusan National University, South Korea). Both the primary antibody and the HRP-conjugated goat anti-rabbit IgG secondary antibodies were used at a 1:5000 dilutions [40].

Assessment of IgE Bing in SP lines via Inhibition ELISA

Proteins separated by electrophoresis were transferred from the gel to an Immobilon-P™ PVDF membrane (Millipore, MA, USA) using a semi-dry-transfer system for immunoblotting analysis. To block non-specific binding, the membranes were incubated in Tris-buffered saline with Tween 20 (TBST, Sigma-Aldrich, MO, USA) containing 3% skim milk at 4 °C overnight. Membranes were then incubated with pooled serum from soybean-allergic individuals (serum samples number 2, 8, 9, 14, and 17), diluted 1:4 in PBST with 1% bovine serum albumin (BSA, Sigma-Aldrich, MO, USA), at room temperature overnight. Serum samples were collected and stored at − 20 °C until use. This study was approved by Institutional Review Board of Yonsei University College of Medicine (4-2013-0397), and all patients provided written informed consent blood collection. After incubation with serum, membranes were washed with TBST to remove unbound proteins and then incubated with an alkaline phosphatase-conjugated anti-human IgE secondary

antibody, diluted 1:1000 in TBST containg 1% BSA, for 1 h at room temperature.

The pooled serum was obtained from five patients with physician-diagnosed IgE-mediated soybean allergy, who showed cutaneous and/or respiratory symptoms upon soy exposure and positive specific IgE to soybean extract. According to their medical records, some patients were also sensitized to other common food allergens (e.g., peanut or tree nuts), reflecting the clinical heterogeneity typically observed in soy-allergic individuals.

For the inhibition ELISA, plates were coated with the control protein extract (10 μg/mL) and blocked. Diluted allergic human serum (1:4) was then added to the wells together with varying concentrations of SP line protein extracts as inhibitors (0.08–50 μg/mL). After 1 h of incuation at room temperature, wells were washed four times with TBST. The plates were then incubated with biotinylated goat anti-human IgE (1:1,000-diluted) (Vector, Burlingame, CA, USA) and streptavidin-peroxidase conjugate (1:1,000-diluted) (Sigma-Aldrich, MO, USA). Color development was achieved by adding 3,3′,5,′,-tetramethyl-benzidine (TMB, Kirkegaard & Perry Laboratories, Gaithersburg, MD, USA), and the reaction was stopped with 0.5 M H_2SO_4. Absorbance was measured at 450 nm. The percentage of inhibition of IgE reactivity relative to the control soybean extract was calculated by comparing the absorbance of wells with inhibitors to those without inhibitors.

Crude seed protein quantification

Seed protein content was quantified using the Dumas method [41] with a Rapid N cube analyzer (Elementar Analysen System, Hanau, Germany) according to the manufacturer's protocols [42]. Fifty milligrams of finely ground seed was prepared into pellets for analysis. The combustion, post-combustion, and reduction tube temperatures were set to 955 °C, 750 °C, and 830 °C, respectively, and the detection temperature was maintained at 59.8 °C [43].

Evaluation of seed free amino acid content

Free amino acids in seeds were extracted by adapting the method of Chotekajorn et al. (2021). Free amino acids were analyzed using a triple quadrupole mass spectrometer (Shimadzu, LCMS-8050, Japan). Chromatographic separation was achieved on a Cortecs UPLC C18 column (1.6 μm particle size, 2.1 × 150 mm, Waters, MA, USA) with a sample injection volume set at 1 μL. Precursor ion m/z values and the fragmentation patterns were determined using standard compounds. Mobile phase A was 0.5% (v/v) formic acid in water, and mobile phase B was 0.5% formic acid (Sigma-Aldrich, MO, USA) in acetonitrile (Sigma-Aldrich, MO, USA). The solvent gradient was as follows: 0–1.0 min 0.5% B, 1.0–2.0 min linear gradient 13% B, 2.0–5.5 min 15% B, 5.5 6.5 min 95% B, 6.5–8.5 min 95% B, 8.5–8.6 min 0.5% B, and 8.6–12 min 0.5% B. The flow rate was maintained at 0.5 mL/min.

Evaluation of seed oil content

Seed oil content was determined using the Soxhlet extraction method [44] with a Buchi B-811 extraction system (Buchi, Switzerland). Two grams of powdered seeds was loaded into extraction thimbles (2 × 100 mm) and subjected to n-hexane extraction at 105 °C for 2 h and 40 min. After cooling the extracted oil to room temperature in a desiccator, the oil weight was measured and expressed on a dry mass basis [42].

Statistical analysis

Protein and oil content measurements were performed with three technical replicates, and their average values were used for statistical evaluation. All statistical analyses were conducted using R software (version 4.3.2). Differences in protein and oil content between wild-type and mutant soybean lines were assessed by two-sample Student's t-tests. *P*-values ≤ 0.01 and ≤ 0.001 were considered statistically significant and highly significant and indicated as ** and ***, respectively. *P*-values > 0.05 were considered not significant (N.S.).

Results

Designing SgRNA and creating T_0 plants with mutations in β-conglycinin subunits

The β-conglycinin genes predominantly arose from recent whole-genome duplication events and several tandem duplications [12]. Based on various studies, researchers uncovered specific genes related to *CG-1* to *CG-4* dispersed across soybean chromosomes 2, 10, and 20. Notably, *Glyma.10G246300* is associated with the α′ subunit (*CG-1*), while *Glyma.20G148300* and *Glyma.20G148400* correspond to the α subunit (*CG-2* and *CG-3*), and *Glyma.20G146200* and *Glyma.20G148200* are related to the β subunit (*CG-4*) [12, 45]. To facilitate multi-targeted mutagenesis of β-conglycinin subunits using the CRISPR/Cas9 system, we designed sgRNA based on these homologies and subjected the Korean soybean cultivar 'Kwangan' to *Agrobacterium*-mediated transformation. We constructed a phylogenetic tree using peptide sequence alignments of *Glyma.10G246300*, *Glyma.20G148300*, *Glyma.20G148400*, *Glyma.20G146200*, and *Glyma.20G148200* to assess their homology and identify shared sequences for multi-targeted mutagenesis (Table 1; Fig. 1a). The two genes associated with the α subunit (*Glyma.20G148300* and *Glyma.*20G148400) displayed 100% sequence homology, while two genes associated with the β subunit (*Glyma.20G146200* and *Glyma.20G148200*) had 99.77% sequence homology. The α′ and α subunits sequences showed 88.9% homology,

and the α′ and β subunit sequences showed 73.80%–74.03% homology. Comparatively, α and β subunit sequences displayed 71.98%–74.94% homology (Fig. 1a). To provide an overview of these relationships, we also summarized the four targeted CG loci (CG-1 to CG-4) and other β-conglycinin-related genes, together with their genomic positions and nucleotide and amino acid sequence similarities, in Supplementary Table S1. Using the pECO201:*7 S* CRISPR/Cas9 vector construct (Supplementary Fig. S3 and S4), we produced three sgRNAs targeting three specific sites in exon 1 of the α′ and α subunit genes (Fig. 1b and c). Likewise, three sgRNAs targeted three specific sites in exon 1 of the β subunit genes (Fig. 1b and e). Considering the significant homology of the β-conglycinin subunit genes, the sgRNA sequences were carefully selected to minimize off-target effects while maximizing on-target mutagenesis efficiency (Fig. 1c and e). We then performed *Agrobacterium*-mediated transformation as described in the materials and methods section and carried out targeted deep sequencing to analyze the InDel patterns and assess the gene editing efficiencies on the α′ and α subunit or β subunit genes of T_0 plants (Fig. 1d and f).

Herbicide resistance of pECO201:7 S-α′α induced T_0 plants and targeted deep sequencing

To confirm the presence of the CRISPR/Cas9 construct inserts in the regenerated plants, we conducted a *bar* strip test and PCR analysis targeting the *bar* gene. The *bar* strip test displayed positive results for three of the four transgenic lines (pECO201:7 S-α′α #1, #2, and #3). The PCR analysis further confirmed the presence of the expected 412 bp *bar* gene amplicons in T_0 lines #1, #2, and #3 (Fig. 2a). Additionally, a PPT-painting assay revealed herbicide resistance in T_0 lines #1 and #2, in contrast with the non-transgenic control 'Kwangan' (Fig. 2b). Further PCR amplification using specific primers (Supplementary Table S2) for the 7 S sgRNA region associated with the α′ (*Glyma.10G246300*) and α (*Glyma.20G148300* and *Glyma.20G148400*) subunits revealed induced mutations in the α′ and α genes in T_0 line #2 (Fig. 2c).

Targeted deep sequencing analysis of the T_0 plants (transformed with the pECO201:7 S-α′α vector) demonstrated various deletion mutations in the genes encoding *α′* and *α* subunits. Among the four analyzed T_0 lines (pECO201:7 S-α′α #1 to #4), we identified ten distinct sequence variations, including deletions ranging from (-)6 to (-)18 base pairs at the sgRNA1 target site, either complete deletions or no deletions at the sgRNA2 site, and deletions of (-)28 to (-)32 base pairs or no deletions at the sgRNA3 site (Fig. 1d). The overall recorded frequencies of InDels for T_0 lines #1, #2, #3, and #4 were 27.5%, 99.6%, 100.0%, and 41.2%, respectively. Specifically, the primary InDel occurrences for T_0 lines #1, #2, #3, and #4 were 25.2%, 92.6%, 98.4%, and 5.0%, respectively (Fig. 1d).

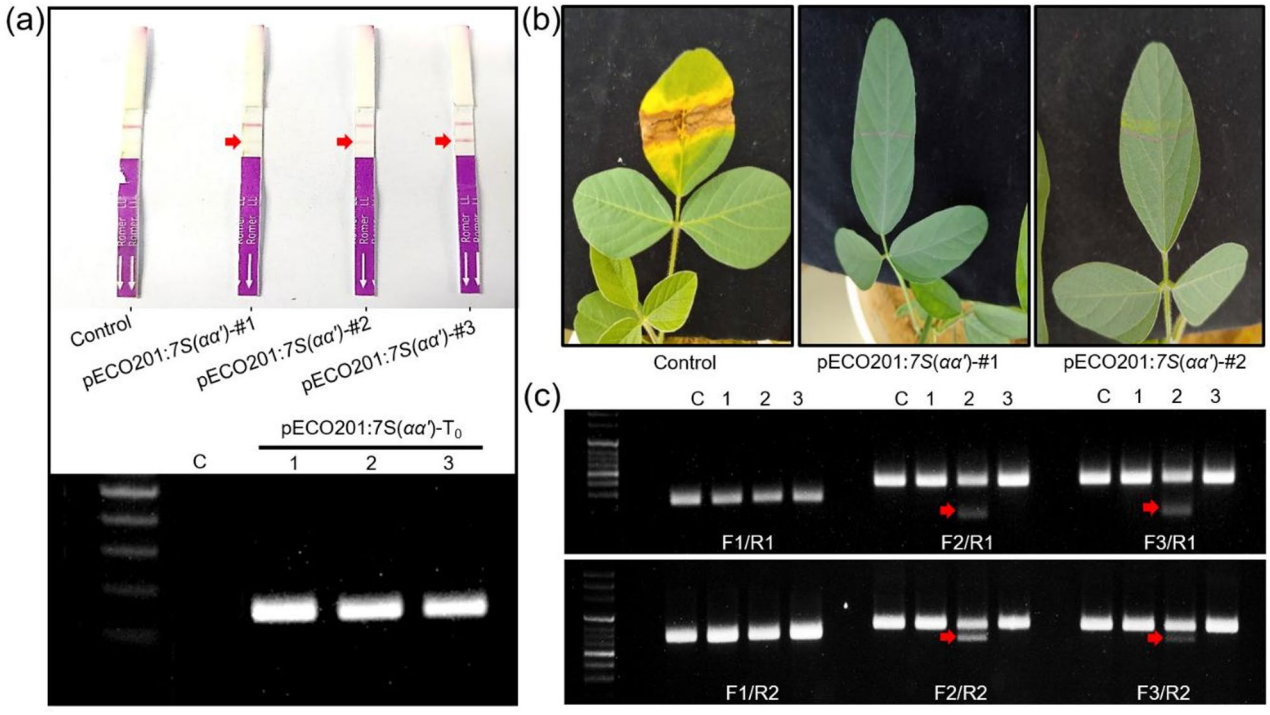

Fig. 2 pECO201:7 S-α′α T_0 plant selection for generational progression. (**a**) *via bar* strip test and PCR amplification of the *bar* gene; (**b**) *via* PPT-painting assay; (**c**) *via* PCR amplification of the target sites of α′ and α genes (Primer information: Supplementary Table S2)

7S protein profile of pECO201:7 S-α′α T_2 plants

Seeds of the T_0 lines (pECO201:7 S-α′α #1 to #3) were collected and propagated to T_1 generation. PCR analysis targeting the α′ subunit gene spanning sgRNA1 to sgRNA3 with the Alpha F2-spR primer set, yielding a 636 bp product, revealed that lines pECO201:7 S-α′α #2–6 and #2–8 possessed homozygous transgenes (Fig. 3a). Further PCR screening of the T_1 generation pECO201:*7 S*-α′α #2 line (pECO201:7 S-α′α #2 – 1 to #2–21) with the Alpha F3-Alpha R1 primers set, specific to the α subunit genes spanning sgRNA1 to sgRNA3 and producing a 362 bp product, revealed induced mutations in 18 out of 21 lines; however, we found no homozygous transgene (Fig. 3b). Therefore, the pECO201:*7 S*-α′α #2–6 line was selected for propagation to the T_2 generation. Subsequent PCR with the 20G.sgRNA3F-20G.sgRNA3R primer set, spanning sgRNA3 of the α subunit genes and yielding a 59 bp product, showed induced mutations in all 18 T_2 lines (pECO201:7 S-α′α #2-6-1 to #2-6-18), with none exhibiting homozygous transgenes (Fig. 3c). SDS-PAGE analysis of the protein profiles from the T_2 generation lines revealed the absence of the 76 kDa α′ subunit protein in lines pECO201:7 S-α′α #2-6-1, #2-6-2, and #2-6-3, as expected from the PCR results (Fig. 3a and d). Notably, line pECO201:7 S-α′α #2-6-2 displayed a significantly lower 72 kDa α subunit protein band intensity than the control "Kwangan," while the band intensity of the β subunit (53 kDa) remained similar across these lines (Fig. 3d).

Herbicide resistance of pECO201:7 S-β induced T_0 plants and targeted deep sequencing

T_0 plants transformed with the pECO201:7 S-β construct underwent *bar* strip testing and were analyzed for mutations induced within exon 1 of the β subunit genes using sgRNA1, sgRNA2, and sgRNA3. Out of six transgenic lines, four (pECO201:7 S-β #2 to #5) tested positive on the *bar* strip test (Fig. 4a). PCR amplification using primers designed for the β subunit genes (*Glyma.20G146200* and *Glyma.20G148200*) (Supplementary Table S2) indicated induced mutations in lines #2 and #4 (Fig. 4b).

Deep sequencing of T_0 lines #1 to #3 revealed various InDel patterns, including deletions from (-)1 to (-)17 bp or a (-)5 bp deletion accompanied by an adenine triad insertion at the sgRNA1 site, (-)18 bp deletion at sgRNA2, and deletions between (-)3 and (-)5 bp at sgRNA3 (Fig. 1f). The total InDel frequencies for lines #1, #2, and #3 were 78.0%, 100.0%, and 15.8%, respectively. Besides, the most prevalent InDels in lines #1, #2, and #3 had frequencies of 9.1%, 82.5%, and 3.4%, respectively (Fig. 1f).

7S protein profile of pECO201:7 S-β T_1 plants

In the T_1 generation propagated from seeds of the T_0 pECO201:7 S-β #2 line, we subjected 28 plants (pECO201:7 S-β #2 – 1 to #2–28) to the PPT-painting assay and PCR analysis for screening (Fig. 4c and d). Among them, 17 T_1 plants exhibited resistance to herbicide (pECO201:7 S-β #2–4, 6, 8, 9, 11, 12, 13, 14, 15, 18, 19, 20, 22, 23, 25, 26, and 28) (Fig. 4c). The plants not showing herbicide resistance (lines #2 – 1, 2, 5, 7, 10, 16, 21, 24, and 27) displayed PCR-confirmed mutations across the regions targeted by sgRNA1, sgRNA2, and sgRNA3 within the β subunit genes. Moreover, these lines lacked the detectable bands for both *bar* and Cas9 sequences in PCR analysis (Fig. 4d). Despite these genetic alterations, however, SDS-PAGE analysis indicated that the T_1 plants and the unaltered control 'Kwangan' had similar β subunit (53 kDa) band intensities (Fig. 4e).

7S protein profile of T_5 or T_6 Storage Protein (SP) plants

Selected seeds from the T_2 generation lines pECO201:7 S-α′α #2-6-1 and pECO201:7 S-α′α #2-6-2 (Fig. 3d) underwent further screenings using PCR analysis and SDS-PAGE (data not shown) while progressing to T_5 or T_6 generations for 7 S protein profile assessment. The T_6 line pECO201:7 S-α′α #2-6-1-4-1-1-15 (designated as SP1) exhibited an absence of the 76 kDa α′ subunit on SDS-PAGE and western blotting; however, PCR analysis revealed the presence of *bar* and Cas9 sequences (Fig. 5a). The T_6 line pECO201:7 S-α′α #2-6-2-9-6-1-11 (designated as SP2) lacked both the 76 kDa α′ and the 72 kDa α subunit proteins, with PCR confirming the presence of *bar* and Cas9 sequences. Western blotting of SP1 and SP2 showed that the β antibody bound to its target protein (Fig. 5a).

The T_5 line pECO201:7 S-α′α #2-6-2-4-2-12 displayed no detectable 76 kDa α′, 72 kDa α, or 53 kDa β subunit proteins, with *bar* and Cas9 sequences identified by PCR, leading to its designation as SP4 (Fig. 5a). Sanger sequencing confirmed that the α′ and α subunits genes in SP4 harbor frameshift-inducing mutations, whereas the β subunit genes retain wild-type coding sequences, indicating the β subunit deficiency in this line is not caused by a direct β-gene knockout (Supplementary Table S2).

The β subunit-null line SP3 was derived from the T_5 generation line pECO201:7 S-β #2-1-20-8-3-11, which originated from a T0 plant transformed with a β-targeting construct. Although the corresponding T1 progeny initially showed no obvious change in the 53 kDa β subunit band intensity on SDS-PAGE, sequence analysis revealed that they carried β subunit edits in a heterozygous state, which likely masked clear band intensity changes at the protein level. Through subsequent selfing and selection, these edits were fixed to homozygosity in later generations, resulting in the SP3 line, which completely lacks the 53 kDa β subunit protein while retaining wild-type levels of the α′ and α subunits (Fig. 5a, Supplementary Table S2).

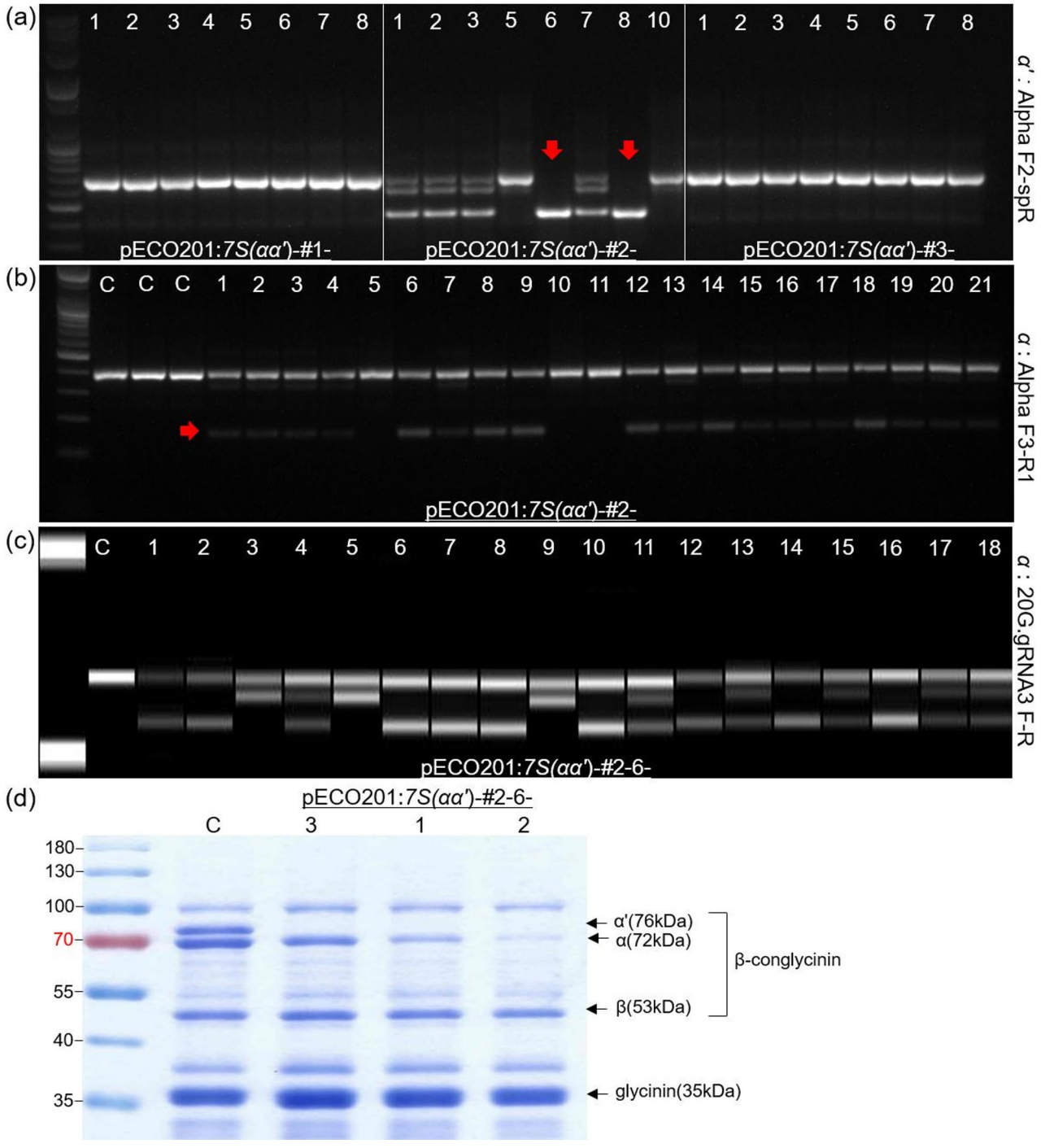

Fig. 3 pECO201:7 S-α'α T_2 plant selection and protein profiles. (**a**) α' target site PCR amplification of T_1 plants. (**b**) α target site PCR amplification of T_1 plants. (**c**) α target site PCR amplification of T_2 plants. (**d**) Protein profile of pECO201:7 S-α'α T_2 plants via SDS-PAGE

InDel patterns in SP plants

We performed Sanger sequencing on the SP lines to identify InDel patterns in the unexpressed β-conglycinin subunit proteins (Fig. 5b and c, and 5d). The InDel patterns in SP lines evolved as they progressed from the T_0 to the T_5 or T_6 generations, aligning with expected digestion sites and demonstrating similar deletion patterns (Figs. 1d and f and 5b and c, and 5d). Within the target range of the two α subunit genes of β-conglycinin (*Glyma.20G148300* and *Glyma.20G148400*), SP1 showed a *Glyma.20G148300* sequence identical to the 'Kwangan' control, and exhibited a (-)4 bp deletion at the sgRNA3 site of *Glyma.20G148400* (Supplementary Table S3 and Fig. 5b). SP2 displayed either an (-)11 bp deletion at the sgRNA3 site in *Glyma.20G148300*

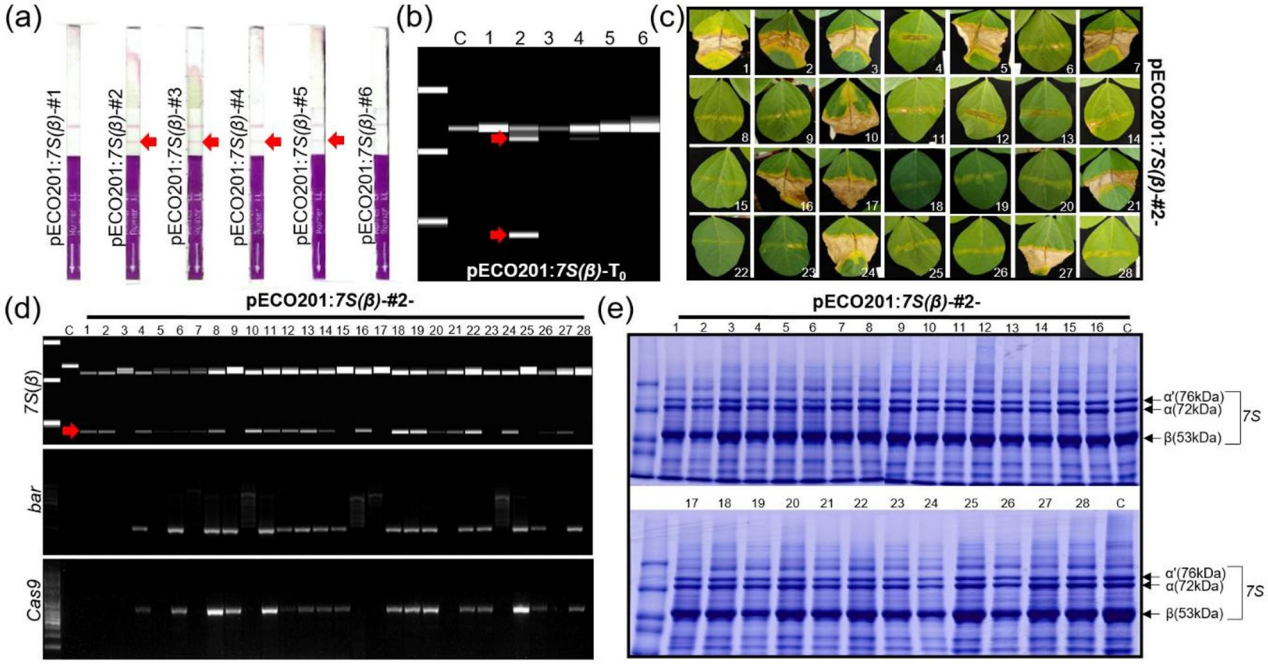

Fig. 4 pECO201:7 S-β T_0 to T_1 plant selection for further generational progression. (**a**) Detection of *bar*-positive T0 plants by *bar* strip test. (**b**) PCR amplification of the target sites of β subunit genes in T0 plants (Primer information: Supplementary table S2); (**c**) PPT-painting assay of T_1 plants derived from a β-construct T0 line. (**d**) PCR amplification of the β subunit target sites, *bar* gene, and Cas9 partial sequences in the same T1 plants. (**e**) Protein profiles of the same set of T1 plants analyzed by SDS-PAGE. In panels (c-e), the numbers 1–28 denote the same individual T1 plants across assays; the T1 plant that later gave rise to the β subunit-null SP3 line is marked in panel (e)

or a (-)4 bp deletion at the sgRNA1 site in *Glyma.20G148400* (Supplementary Table S3 and Fig. 5b). The InDel patterns of SP4 were a 1 bp thymine insertion, (-)10 bp deletion, and (-)4 bp deletion at the sgRNA1, sgRNA2, and sgRNA3 sites of *Glyma.20G148300*, respectively, or a (-)4 bp deletion at the sgRNA3 site of *Glyma.20G148400* (Supplementary Table S3 and Fig. 5b). Within the targeted sites of the α′ subunit (*Glyma.10G246300*), SP1 and SP2 showed identical large deletion patterns (Supplementary Table S3 and Fig. 5c). In SP4, however, we observed heterozygous InDel patterns at the sgRNA sites: one strand with a 1 bp thymine insertion, a (-)1 bp deletion, and a (-)5 bp deletion, and the other strand with a 1 bp thymine insertion, a (-)4 bp deletion, and a (-)5 bp deletion (Supplementary Table S3 and Fig. 5c). Within the target range, the two β subunit genes (*Glyma.20G146200* and *Glyma.20G148200*) in SP3 exhibited identical large deletions across sgRNA1 and sgRNA2 sites and a (-)5 bp deletion at the sgRNA3 site (Supplementary Table S3 and Fig. 5d). Interestingly, despite the unexpressed β subunit protein in SP4 (Fig. 5a), no mutations were detected in either β subunit genes, which remained the same as those of the 'Kwangan' control (Supplementary Table S3 and Fig. 5d).

Off-target analysis

To analyze the effects of any off-target sites related to the unexpressed β-conglycinin subunit phenotype, we performed whole-genome resequencing of the SP lines and utilized these data. Potential off-target sites for the sgRNAs within the soybean genome were assessed using the CRISPR-P website (http://crispr.hazau.edu.cn/CRISPR2), by aligning the 22 bp target sequence with the soybean reference genome (*Glycine max v2.a1*) [29]. Examining the 10 most likely potential off-target positions for each sgRNA revealed alleles identical to those of the 'Kwangan' control, without mutations or InDels (Supplementary Table S4).

Assessment of IgE recognition of SP lines via Immunoblot

IgE immunoblotting was conducted to assess the allergenic potential of proteins extracted from seeds of the SP1 to SP4 soybean lines (Fig. 6a). The analysis aimed to detect the presence of the α′, α, and β subunits by their ability to bind IgE from pooled human serum. The SP1 line, lacking the 76 kDa α′ subunit (Fig. 5a), exhibited no IgE reactivity to the 76 kDa α′ subunit band, suggesting a marked reduction in IgE recognition for this protein fraction (Fig. 6a). Similarly, the SP2 line, which lacks both the 76 kDa α′ and 72 kDa α subunits (Fig. 5a), displayed no bands for the α′ subunit and a significantly lower α subunit band intensity than the 'Kwangan' control (Fig. 6a). The SP3 line, representing the β-null genotype (Fig. 5a), showed a lower 53 kDa β subunit band intensity than the control (Fig. 6a). The SP4 line, expected to lack all three

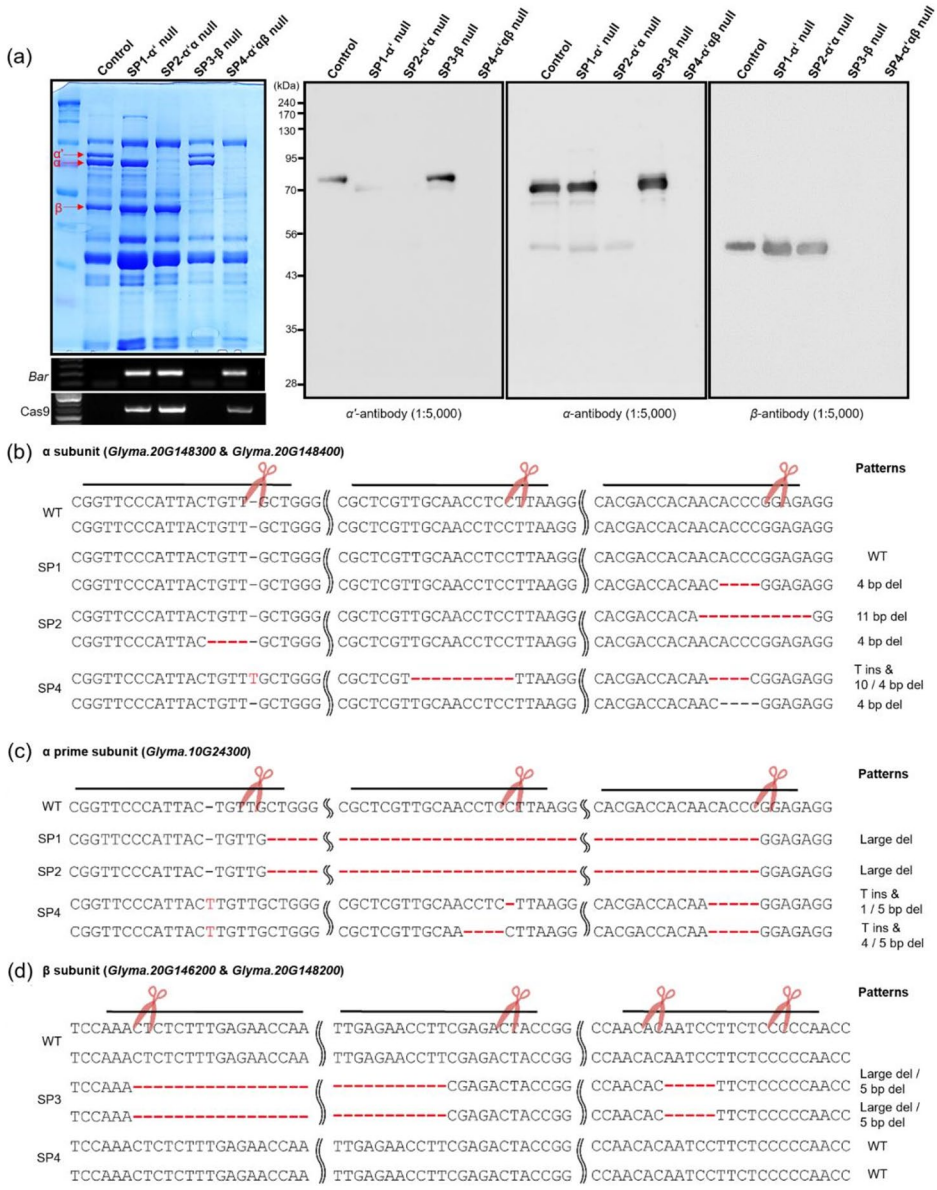

Fig. 5 Protein profiles of SP lines (SP1 to 4) with InDel patterns. (**a**) SDS-PAGE and western blotting of SP lines using specific antibodies against the α′, α, and β subunits (1:5,000). (**b**) InDel patterns within the target sites of α subunit genes (*Glyma.20G148300* and *Glyma.20G148400*). (**c**) InDel patterns within the target sites of α′ subunit gene (*Glyma.10G246300*). (**d**) InDel patterns within the target sites of β subunit genes (*Glyma.20G146200* and *Glyma.20G148200*)

subunits (Fig. 5a), showed a complete absence of bands for the α′ and β subunits and a significantly lower α subunit band intensity than the control, indicating a substantial reduction in IgE-recognized β-conglycinin subunits in this line (Fig. 6a).

Assessment of IgE binding in SP lines via Inhibition ELISA

Subsequent inhibition ELISA assays quantified the capacity of proteins from the modified soybean lines to inhibit IgE binding, compared with proteins extracted from the 'Kwangan' control (Fig. 6b). The ELISA results were generally consistent with the immunoblot findings. The SP1 line demonstrated moderate inhibition of IgE reactivity, with an 83.6% inhibition at 50 μg/mL of inhibitor (Fig. 6a and b). The SP2 line reached 78.6% inhibition at the same concentration, suggesting a reduction in IgE binding that may reflect the removal of multiple *7 S* proteins (Figs. 5a and 6b). In contrast, the SP3 line exhibited 88.9% inhibition at 50 μg/mL compared with 87.7% for the control, and its inhibition curve largely overlapped with that of the wild type across the tested concentrations, indicating that loss of the β subunit alone did not substantially reduce IgE-binding in this assay (Fig. 6b). The SP4 line displayed the lowest level of inhibition, with a 70.0%

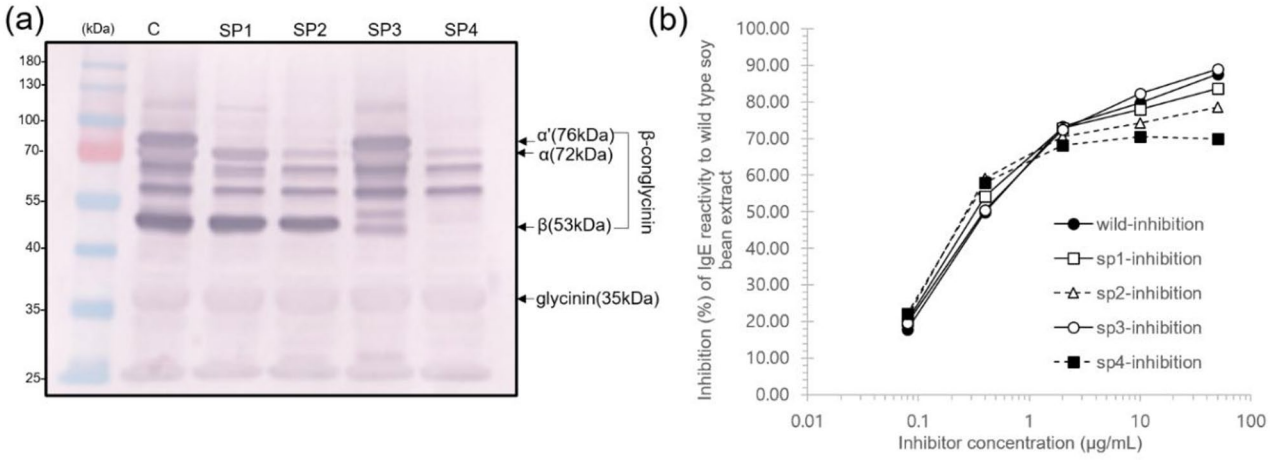

Fig. 6 IgE-binding assessment of SP lines. (**a**) IgE Immunoblotting results. (**b**) Inhibition ELISA results

Table 2 Comparative analysis of total protein and oil content (%) across SP lines

	Control (Kwangan)	SP1	SP2	SP3	SP4
Protein content (%)	39.5	43.56	44.16	43.16	44.59
P-Value[†]	-	2.23E-05	1.19E-07	6.48E-06	4.96E-06
Significance	-	***	***	***	***
Oil content (%)	17.5	15.97	16.39	17.52	16.43
P-Value[†]	-	2.80E-04	1.13E-03	9.11E-01	6.58E-03
Significance	-	***	**	N.S	**

[†]*P*-Values calculated from a two-sample t-test. **: $P \leq 0.01$, ***: $P \leq 0.001$, N.S: not significant

inhibition at 50 μg/mL, indicating a clear decrease in IgE binding under these in vitro assay conditions; however, clinical allergenicity was not directly assessed in this study.

Seed Protein, Oil, and free amino acid contents of SP lines

To assess the impact of altered storage protein subunit expression *via* CRISPR/Cas9 in SP lines on seed quality, we measured total seed protein content, oil content, and free amino acid content, and compared those with the 'Kwangan' control. All SP lines had significantly greater total protein contents than the control, with a *P*-value ≤ 0.001 (Table 2), consistent with a previous study [29]. However, SP1, SP2, and SP4 had significantly lower oil contents than the control, while that of SP3 was similar to the control (*P*-value > 0.05) (Table 2). Interestingly, SP3 exhibited significantly higher levels of 11 out of 15 measured free amino acids (including Asn, Gly, Ala, Ser, Pro, Val, Thr, Leu, Ile, Gln, and His) and lower free Asp levels than the other lines (Fig. 7a). The β-conglycinin-null SP4 line, in particular, showed 1.48-fold and 2.75-fold higher levels of free Asp and Arg, respectively, than the control (Fig. 7a). Regarding plant growth and reproduction, all of the SP lines exhibited similar growth and reproductive traits to the 'Kwangan' control (Fig. 7b). Under our greenhouse and Wagner pot conditions, all SP lines (SP1–SP4) showed vegetative and reproductive growth that was visually similar to the wild-type 'Kwangan', with no obvious differences in plant architecture, flowering, pod set, or seed development.

Discussion

Characteristics of 7 S subunit deficiency in SP lines

β-conglycinin is a vicilin-type protein that significantly impacts the quality of soy-based food products, especially affecting tofu yield and texture [1, 46]. Compared with glycinin, β-conglycinin contains fewer sulfur-containing amino acids (Met and Cys), resulting in lower nutritional quality and diminished gel-forming capacity [17, 47]. It is also less effective in preserving the stability and antioxidant properties of cyanidin-3-glucoside (C3G) [48]. In this study, we utilized CRISPR/Cas9 to generate soybean lines deficient in individual or multiple β-conglycinin subunits (α′, α, and β) in the Korean cultivar 'Kwangan' (Figs.1 and 5). Loss of the β subunit in SP3 line did not significantly affect seed oil content but was accompanied by elevated levels of various free amino acids, including Asn, Gly, Ala, Ser, Pro, Val, Thr, Leu, Ile, Gln, and His (Fig. 7a). All mutant lines (SP1-SP4) exhibited normal growth and reproduction while showing increased total seed protein content, likely reflecting compensatory glycinin accumulation (Fig. 7b) [7, 29, 49]. Consistent with previous reports that some β-conglycinin subunit–deficient lines can complete normal growth and reproduction, the SP lines showed no apparent abnormalities in plant architecture or fertility under greenhouse conditions, although other studies have reported inferior agronomic performance in certain β-conglycinin-deficient germplasm, indicating that the impact of 7 S modification can be genotype dependent. Thus, while our current data suggest that the SP lines maintain normal growth under controlled conditions, replicated multi-environment field

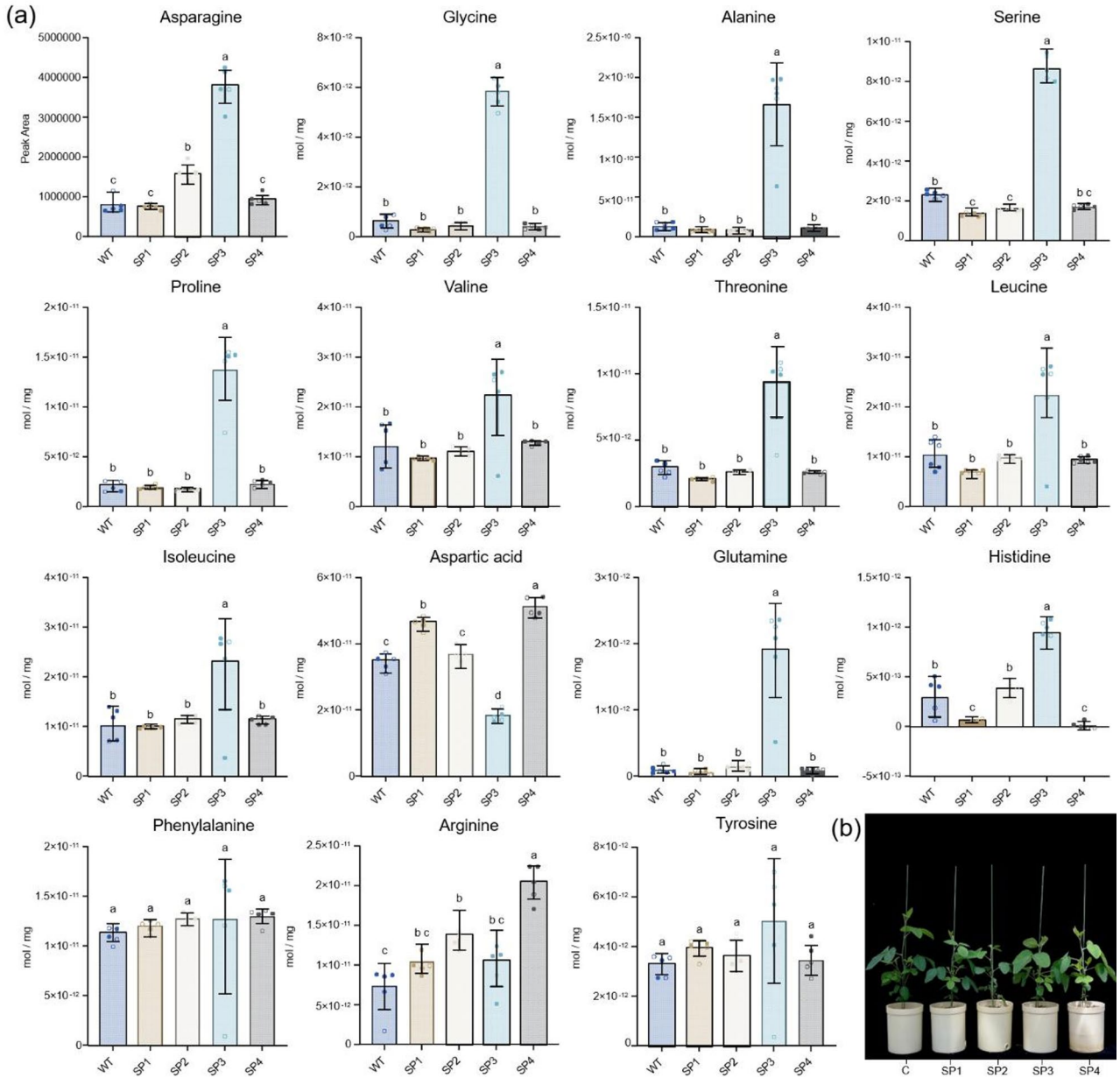

Fig. 7 Agronomic traits of SP lines. (**a**) Content of 15 free amino acids in SP lines and the 'Kwangan' control. Asparagine was measured relative to the peak area and the rest of the free amino acid contents were measured in mol/mg. (**b**) Plant growth patterns of SP lines

trials and targeted stress-response assays will be required to rigorously evaluate yield and stress resistance before large-scale deployment. These edited lines therefore represent valuable germplasm resources for improving processing quality and nutritional properties in soybean-based products.

Notably, the complete β-conglycinin-null SP4 line exhibited increases levels of free arginine [7, 38] and free aspartic acid (Fig. 7), even though Sanger sequencing confirmed no mutations in the β subunit genes in this line (Fig.5b). Regulation of β-conglycinin subunits has been reported to involve transcriptional and post-transcriptional mechanisms [8, 11, 29], and lncRNAs have recently been recognized as critical regulators of gene expression [29]. Loss-of-function mutations in *lincCG1*, generated via CRISPR/Cas9, were shown to cause a deficiency of α′-, α-, and β subunits of β-conglycinin through combined *cis*- and *trans*-regulatory mechanisms and to be associated with increased total protein, increased sulfur-containing amino acids, and elevated free arginine, without agronomic alterations compared with controls [29]. Whole-genome analysis of the SP lines harboring *lincCG1* within *Scg-1* (Suppressor of β-conglycinin) revealed no detectable sequence

variants at the *lincCG1* locus in the SP lines relative to 'Kwangan' (8,38; Supplementary Fig. S5). In contrast, SP4 lacks all three β-conglycinin subunit proteins (α′, α, and β) at the protein level, but Sanger sequencing confirmed that the β subunit genes retain wild-type coding sequences, indicating that the β subunit deficiency in this background is not caused by a CRISPR-induced β-gene knockout. Therefore, the precise regulatory mechanisms responsible for β subunit expression loss in SP4 remain unresolved and warrant further in-depth study.

In contrast to this subunit-targeted approach, the *lincCG1* study used CRISPR/Cas9 to knock out a long intergenic noncoding RNA that regulates the β-conglycinin locus, thereby generating β-conglycinin-deficient lines with increased protein content and sulfur-containing amino acids. Whereas *lincCG1* editing modulated β-conglycinin accumulation through a regulatory RNA, the present study directly targets the coding sequences of the α′, α, and β subunits to create multiple edited genotypes with distinct storage-protein profiles and IgE-binding characteristics (SP1-SP4). This complementary strategy allows fine-tuning of individual subunits and provides germplasm in which seed composition and IgE binding to β-conglycinin can be evaluated separately for each subunit combination, rather than only in a β-conglycinin-null background.

The SP4 line completely lacked detectable β-conglycinin subunit proteins despite retaining wild-type β subunit coding sequences and showing no off-target mutations at predicted sgRNA sites, pointing to the involvement of a regulatory layer beyond simple gene knockout. Given that *lincCG1* mutations can suppress β-conglycinin accumulation through complex cis- and trans- regulatory mechanisms without altering the corresponding coding regions, a similar regulatory scenario may also underlie the SP4 phenotype. As RNA-seq, qRT-PCR, or proteomic analyses were not performed in this study, future work using these approaches will be required to clarify the underlying mechanism.

Allergenicity of Gly m 5 and related subunits

Soybeans contribute to approximately 90% of food allergy reactions and rank among the "big eight" major allergens globally [50]. Soybean allergy affects about 0.5% of the world's population and 0.4% of children in the United States, positioning soybean as a leading allergen in pediatric populations [51, 52]. Importantly, allergenic potential varies among cultivars and cannot be accurately predicted from total protein content alone, as ingestion of as little as 5.3 mg soybean protein can provoke allergic reactions in susceptible individuals [19].

Among soybean allergens, the β-conglycinin (Gly m 5) is particularly significant. It is recognized by most pediatric soybean patients and, together with glycinin (Gly m6), strongly contributes to allergic severity [53]. This study targeted Gly m 5 via CRISPR/Cas9 to disrupt critical β-conglycinin subunits, resulting in reduced IgE-binding capacity to β-conglycinin in the edited lines. At the highest inhibitor concentration, SP4 showed the lowest IgE-binding inhibition among the tested lines (~ 70% compared with 87.7% for the wild-type), whereas SP1-SP3 displayed inhibition values similar to or only slightly lower than the control (Fig. 6b). Across the tested concentrations, the inhibition curve of the β subunit-null line SP3 largely overlapped with that of the wild type, indicating that loss of the β subunit alone did not lead to a large decrease in IgE-binding inhibition in this in vitro assay (Fig. 6b). These data suggest that more extensive modification of β-conglycinin, as in SP4, may be required to achieve a pronounced reduction in IgE binding, although the α′ and α subunits also contribute to IgE recognition.

These findings align with reports that clinical allergenicity differs among 7S protein subunits; for example, the α subunits (Gly m Bd 60 K) have been shown to sensitize approximately 25% of soybean-sensitive dermatitis patients [54]. Collectively, the immunoblot and inhibition ELISA data indicate that CRISPR/Cas9-mediated editing of β-conglycinin subunits alters IgE-binding profiles across the SP lines, with the most pronounced reduction in IgE binding observed in SP4. These genome-edited lines thus represent soybean germplasm with reduced β-conglycinin IgE recognition that may support the development of safer soy-based foods, although clinical allergenicity has not yet been directly evaluated. The application of CRISPR/Cas9 to edit key allergen-encoding proteins in soybean seeds contributes to a growing field of molecular breeding aimed at improving food safety and quality, and these findings provide an example of how modern genome-editing tools can be integrated with phenotypic and molecular analyses to address challenges at the interface of agriculture and human health.

Finally, it should be noted that the IgE-binding assays were performed with sera from only five soy-allergic patients with heterogeneous clinical phenotypes and co-sensitizations, and therefore the results should be interpreted as in vitro IgE-binding data rather than a comprehensive assessment of clinical allergenicity.

Conclusions

In summary, this work provides evidence that CRISPR/Cas9-mediated disruption of the α′, α, and β subunits of β-conglycinin enables the stable development of soybean lines with reduced β-conglycinin-specific IgE binding and markedly reduced IgE reactivity in vitro. The edited genotypes, particularly the α′αβ-null line, demonstrated consistent inheritance across generations and revealed subunit-specific impacts on seed quality traits, underscoring both the precision and robustness

of the genome editing strategy. These findings not only advance our understanding of the molecular basis of IgE recognition of soybean seed proteins but also establish a practical framework for molecular breeding aimed at soybean lines with reduced β-conglycinin IgE binding. Future research should incorporate integrative transcriptomic analyses, molecular marker-assisted breeding, and evaluation across diverse genetic backgrounds to accelerate the application of genome editing for soybean improvement.

Abbreviations

SP	Storage protein
sgRNA	Small guide RNA
InDel	Insertion and Deletion

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-026-08130-8.

Supplementary Material 1: Appendix A. Supplementary data. Supplementary Fig. S1. Alignment of exon1 sequences of the β-conglycinin α′ and α subunit genes with the three sgRNA target sites. The coding sequences of Glyma.10G246300 (α′) and Glyma.20G148300/Glyma.20G148400 (α) are aligned, and the positions of sgRNA1-sgRNA3 are indicated. Supplementary Fig. S2. Alignment of exon1 sequences of the β-conglycinin β subunit genes with the three sgRNA target sites. The coding sequences of Glyma.20G146200 and Glyma.20G148200 (β) are aligned, and the positions of sgRNA1-sgRNA3 are indicated. Supplementary Fig. S3. Final vector map of pECO201:7S-α′α, with the sgRNA target regions highlighted in red. Supplementary Fig. S4. Final vector sequence of pECO201:7S-β, with the sgRNA target regions highlighted in red. Supplementary Fig. S5. Summary of conservation of the lincCG1 locus within Scg-1 (Suppressor of β-conglycinin) among the reference genome Wm82.a2.v1, Kwangan, and the SP1-SP4 lines. Supplementary Table S1. Targeted β-conglycinin genes and related genes with genomic positions and nucleotide and protein sequence similarities Supplementary Table S2. Primer information for amplification of genomic regions around the 7S sgRNA target sitesSupplementary Table S3. Mutagenesis assessment at targeted sites in SP lines through deep sequencing. Supplementary Table S4. Putative 10 off-target sites for each sgRNA.

Acknowledgements
Not applicable.

Authors' contributions
Author contributions: H.R.P. : Data curation, Writing- Review & Editing. S.P.: Methodology, Visualization, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing-Original draft. J.M.J., Y.J.S., Y.H., M.Y.K. , S.P. : Methodology, Data curation. K.Y.J. : Methodology, Data curation, Writing-Review & Editing. S.T.K. : Resources (Providing antibodies). Y.H.Y., E.L., G.P.: Data curation. S.G.K.: Resources, Validation, Visualization, Data curation, Writing-Review & Editing. S.K.P. : Conceptualization, Writing- Review, Supervision, Project administration, Funding acquisition.

Funding sources
This work was carried out with the support of the "Agricultural Research Program (PJ016784022024)" from the Rural Development Administration, Republic of Korea. This study was supported by the 2024 RDA Research Associate Fellowship Program (Postdoctoral fellow) of the National Institute of Crop and Food Science, Rural Development Administration, Republic of Korea.

Data availability
All raw sequence data provided in fastq files used in the study were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under the BioProject accession number PRJNA1345520.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 1 October 2025 / Accepted: 7 January 2026
Published online: 14 January 2026

References

1. Song B, Oehrle NW, Liu S, Krishnan HB. Development and characterization of a soybean experimental line lacking the α′ subunit of β-Conglycinin and G1, G2, and G4 glycinin. J Agric Food Chem. 2018;66(2):432–9. https://doi.org/10.1021/acs.jafc.7b05011.
2. Derbyshire E, Wright DJ, Boulter D. Legumin and Vicilin, storage proteins of legume seeds. Phytochemistry. 1976;15(1):3–24. https://doi.org/10.1016/S0031-9422(00)89046-9.
3. Hayashi M, Nishioka M, Kitamura K, Harada K. Identification of AFLP markers tightly linked to the gene for deficiency of the 7S Globulin in soybean seed and characterization of abnormal phenotypes involved in the mutation. Breed Sci. 2000;50(2):123–9. https://doi.org/10.1270/jsbbs.50.123.
4. Hayashi M, Kitamura K, Harada K. Genetic mapping of Cgdef gene controlling accumulation of 7S Globulin (Beta-Conglycinin) subunits in soybean seeds. J Hered. 2009;100(6):802–6. https://doi.org/10.1093/jhered/esp046.
5. Pi X, Sun Y, Fu G, Wu Z, Cheng J. Effect of processing on soybean allergens and their allergenicity. Trends Food Sci Technol. 2021;118:316–27. https://doi.org/10.1016/j.tifs.2021.10.006.
6. Murphy PA. Soybean Proteins. In *Soybeans*; Vol. 8; Johnson, L. A., White, P. J., Galloway, R., Eds.; AOCS Press, 2008, pp 229–267. https://doi.org/10.1016/B978-1-893997-64-6.50011-1
7. Takahashi M, Uematsu Y, Kashiwaba K, Yagasaki K, Hajika M, Matsunaga R, Komatsu K, Ishimoto M. Accumulation of high levels of free amino acids in soybean seeds through integration of mutations conferring seed protein deficiency. Planta. 2003;217(4):577–86. https://doi.org/10.1007/s00425-003-1026-3.
8. Tsubokura Y, Hajika M, Kanamori H, Xia Z, Watanabe S, Kaga A, Katayose Y, Ishimoto M, Harada K. The Beta-Conglycinin deficiency in wild soybean is associated with the Tail-to-Tail inverted repeat of the Alpha-Subunit genes. Plant Mol Biol. 2012;78(3):301–9. https://doi.org/10.1007/s11103-011-9865-y.
9. da Silva FL, Borém A, Sediyama T, Ludke WH. Soybean breeding. Springer Int Publishing, 2017. https://books.google.co.kr/books?id=lH4nDwAAQBAJ (accessed 23-08-2024).
10. Ha DH, Kim HJ. Mutation of storage protein gene using CRISPR/Cas9 removed α′-Subunit of β-Conglycinin in soybean seeds. Plant Biotechnol Rep. 2023;17(6):939–45. https://doi.org/10.1007/s11816-023-00880-3.
11. Harada JJ, Barker SJ, Goldberg RB. Soybean Beta-Conglycinin genes are clustered in several DNA regions and are regulated by transcriptional and posttranscriptional processes. Plant Cell. 1989;1(4):415–25. https://doi.org/10.1105/tpc.1.4.415.
12. Li C, Zhang YM. Molecular evolution of glycinin and β-Conglycinin gene families in soybean (*Glycine max* L. Merr). Heredity. 2011;106(4):633–41. https://doi.org/10.1038/hdy.2010.97.
13. Krishnan HB, Kim W-S, Jang S, Kerley MS. All three subunits of soybean β-Conglycinin are potential food allergens. J Agric Food Chem. 2009;57(3):938–43. https://doi.org/10.1021/jf802451g.
14. Mulalapele LT, Xi J. Detection and inactivation of allergens in soybeans: A brief review of recent research advances. Grain Oil Sci Technol. 2021;4(4):191–200. https://doi.org/10.1016/j.gaost.2021.11.001.
15. Murakami H, Ogawa T, Takafuta A, Yano E, Zaima N, Moriyama T. Identification of the 7S and 11S globulins as percutaneously sensitizing soybean allergens as demonstrated through epidermal application of crude soybean extract. Biosci Biotechnol Biochem. 2018;82(8):1408–16. https://doi.org/10.1080/09168451.2018.1460573.
16. Krishnan HB. Engineering soybean for enhanced sulfur amino acid content. Crop Sci. 2005;45(2):454–61. https://doi.org/10.2135/cropsci2005.0454.

17. Tsubokura Y, Hajika M, Harada K. Molecular markers associated with β-Conglycinin deficiency in soybean. Breed Sci. 2006;56(2):113–7. https://doi.org/10.1270/jsbbs.56.113.
18. Ballmer-Weber BK, Vieths S. Soy allergy in Perspective. Review. Curr Opin Allergy Clin Immunol. 2008;8(3):270–5. https://doi.org/10.1097/ACI.0b013e3282ffb157.
19. Pi X, Sun Y, Deng X, Xin D, Cheng J, Guo M. Investigation of differences in allergenicity of protein from different soybean cultivars through LC/MS-MS. Int J Biol Macromol. 2022;220:1221–30. https://doi.org/10.1016/j.ijbiomac.2022.08.154.
20. Hanafusa K, Murakami H, Ueda T, Yano E, Zaima N, Moriyama T. Worm wounding increases levels of Pollen-Related food allergens in soybean (*Glycine max*). Biosci Biotechnol Biochem. 2018;82(7):1207–15. https://doi.org/10.1080/09168451.2018.1456319.
21. Havenith H, Kern K, Rautenberger P, Spiegel H, Szardenings M, Ueberham E, Lehmann J, Buntru M, Vogel S, Treudler R, Fischer R, Schillberg S. Combination of two epitope identification techniques enables the rational design of soy allergen Gly m 4 mutants. Biotechnol J. 2017;12(2):1600441. https://doi.org/10.1002/biot.201600441.
22. Selb R, Wal JM, Moreno FJ, Lovik M, Mills C, Hoffmann-Sommergruber K, Fernandez A. Assessment of endogenous allergenicity of genetically modified plants exemplified by Soybean – Where do we stand? Food Chem Toxicol. 2017;101:139–48. https://doi.org/10.1016/j.fct.2017.01.014.
23. Biscola V, de Olmos AR, Choiset Y, Rabesona H, Garro MS, Mozzi F, Chobert JM, Drouet M, Haertlé T, Franco BD. G. M. Soymilk fermentation by Enterococcus faecalis VB43 leads to reduction in the immunoreactivity of allergenic proteins β-Conglycinin (7S) and glycinin (11S). Article. Benef Microbes. 2017;8(4):635–43. https://doi.org/10.3920/BM2016.0171.
24. Bu G, Zhu T, Chen F, Zhang N, Liu K, Zhang L, Yang H. Effects of saccharide on the structure and antigenicity of β-Conglycinin in soybean protein isolate by glycation. Eur Food Res Technol. 2015;240(2):285–93. https://doi.org/10.1007/s00217-014-2326-5.
25. Cabanillas B, Cuadrado C, Rodriguez J, Dieguez MC, Crespo JF, Novak N. Boiling and pressure cooking impact on IgE reactivity of soybean allergens. Int Arch Allergy Immunol. 2018;175(1–2):36–43. https://doi.org/10.1159/000485894.
26. Xi J, He M. Location of destroyed antigenic sites of Gly m Bd 60 K after three processing technologies. Food Res Int. 2020;134:109199. https://doi.org/10.1016/j.foodres.2020.109199.
27. Xia J, Zu Q, Yang A, Wu Z, Li X, Tong P, Yuan J, Wu Y, Fan Q, Chen H. Allergenicity reduction and rheology property of Lactobacillus-Fermented soymilk. J Sci Food Agric. 2019;99(15):6841–9. https://doi.org/10.1002/jsfa.9969.
28. Kim WS, Jez JM, Krishnan HB. Effects of proteome rebalancing and sulfur nutrition on the accumulation of methionine rich δ-zein in Transgenic soybeans. Front Plant Sci. 2014;5:633. https://doi.org/10.3389/fpls.2014.00633.
29. Song B, Luo T, Fan Y, Li M, Qiu Z, Tian Y, Shang Y, Ma C, Liu C, Cao Q, Peng Y, Xu P, Krishnan HB, Wang Z, Zhang S, Liu S. Generation of new β-Conglycinin-Deficient soybean lines by editing the LincRNA lincCG1 using the CRISPR/Cas9 system. J Agric Food Chem. 2024;72(26):15013–26. https://doi.org/10.1021/acs.jafc.4c02269.
30. Li C, Nguyen V, Liu J, Fu W, Chen C, Yu K, Cui Y. Mutagenesis of seed storage protein genes in soybean using CRISPR/Cas9. BMC Res Notes. 2019;12(1):176. https://doi.org/10.1186/s13104-019-4207-2.
31. Schmidt MA, Barbazuk WB, Sandford M, May G, Song Z, Zhou W, Nikolau BJ, Herman EM. Silencing of soybean seed storage proteins results in a rebalanced protein composition preserving seed protein content without major collateral changes in the metabolome and transcriptome. Plant Physiol. 2011;156(1):330–45. https://doi.org/10.1104/pp.111.173807.
32. Wei X, Kim W-S, Song B, Oehrle NW, Liu S, Krishnan HB. Soybean mutants lacking abundant seed storage proteins are impaired in mobilization of storage reserves and germination. ACS Omega. 2020;5(14):8065–75. https://doi.org/10.1021/acsomega.0c00128.
33. Paz MM, Martinez JC, Kalvig AB, Fonger TM, Wang K. Improved cotyledonary node method using an alternative explant derived from mature seed for efficient Agrobacterium-Mediated soybean transformation. Plant Cell Rep. 2006;25(3):206–13. https://doi.org/10.1007/s00299-005-0048-7.
34. Thompson JD, Higgins DG, Gibson TJ, Clustal W. Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, Position-Specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22(22):4673–80. https://doi.org/10.1093/nar/22.22.4673.
35. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35(6):1547–9. https://doi.org/10.1093/molbev/msy096.
36. Oh Y, Lee B, Kim H, Kim S-G. A multiplex guide RNA expression system and its efficacy for plant genome engineering. Plant Methods. 2020;16(1):37. https://doi.org/10.1186/s13007-020-00580-x.
37. Subburaj S, Chung SJ, Lee C, Ryu S-M, Kim DH, Kim J-S, Bae S, Lee G-J. Site-Directed mutagenesis in Petunia × hybrida protoplast system using direct delivery of purified Recombinant Cas9 ribonucleoproteins. Plant Cell Rep. 2016;35(7):1535–44. https://doi.org/10.1007/s00299-016-1937-7.
38. Park S, Kwon H, Park GT, Choi JA, Kim MY, Park S-K, Lee S-H. Selective allele stacking of a novel quantitative trait locus facilitates the enhancement of seed epicatechin content in soybean (*Glycine max* (L.) Merr). Euphytica. 2024;220(6):90. https://doi.org/10.1007/s10681-024-03345-y.
39. Bradford MMA, Rapid. Sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-Dye binding. Anal Biochem. 1976;72(1):248–54. https://doi.org/10.1006/abio.1976.9999.
40. Nishizawa K, Maruyama N, Satoh R, Fuchikami Y, Higasa T, Utsumi SA. C-Terminal sequence of soybean β-Conglycinin α′ subunit acts as a vacuolar sorting determinant in seed cells. Plant J. 2003;34(5):647–59. https://doi.org/10.1046/j.1365-313x.2003.01754.x.
41. Saint-Denis T, Goupy J. Optimization of a nitrogen analyser based on the Dumas method. Anal Chim Acta. 2004;515(1):191–8. https://doi.org/10.1016/j.aca.2003.10.090.
42. Dhungana SK, Seo J-H, Kang B-K, Park J-H, Kim J-H, Sung J-S, Baek I-Y, Shin S-O, Jung C-S, Protein. Amino Acid, Oil, fatty Acid, Sugar, Anthocyanin, Isoflavone, Lutein, and antioxidant variations in colored Seed-Coated soybeans. Plants. 2021;10(9):1765. https://doi.org/10.3390/plants10091765.
43. Park HR, Seo JH, Kang BK, Kim JH, Heo SV, Choi MS, Ko JY, Kim CS. QTLs and candidate genes for seed protein content in two Recombinant inbred line populations of soybean. Plants. 2023;12(20):3589. https://doi.org/10.3390/plants12203589.
44. Lee JH, Cho KM. Changes occurring in compositional components of black soybeans maintained at room temperature for different storage periods. Food Chem. 2012;131(1):161–9. https://doi.org/10.1016/j.foodchem.2011.08.052.
45. Gillman JD, Kim W-S, Song B, Oehrle NW, Tawari NR, Liu S, Krishnan HB. Whole-Genome resequencing identifies the molecular genetic cause for the absence of a Gy5 glycinin protein in soybean PI 603408. G3|Genomes|Genetics. 2017;7(7):2345–52. https://doi.org/10.1534/g3.117.039347.
46. Yagasaki K, Kalzuma N, Kitamura K. Inheritance of glycinin subunits and characterization of glycinin molecules lacking the subunits in soybean (*Glycine max* (L.) Merr). Ikushugaku Zasshi. 1996;46(1):11–5. https://doi.org/10.1270/jsbbs1951.46.11.
47. Ishikawa G, Takada Y, Nakamura T. A PCR-Based method to test for the presence or absence of β-Conglycinin α′- and α-Subunits in soybean. Mol Breed. 2006;17(4):365–74. https://doi.org/10.1007/s11032-006-9010-9.
48. Zheng J, Zheng X, Zhao L, Yi J, Cai S. Effects and interaction mechanism of soybean 7S and 11S globulins on anthocyanin stability and antioxidant activity during in vitro simulated digestion. Curr Res Food Sci. 2021;4:543–50. https://doi.org/10.1016/j.crfs.2021.08.003.
49. Kinney AJ, Jung R, Herman EM. Cosuppression of the α subunits of β-Conglycinin in Transgenic soybean seeds induces the formation of Endoplasmic Reticulum–Derived protein bodies. Plant Cell. 2001;13(5):1165–78. https://doi.org/10.1105/tpc.13.5.1165.
50. Pi X, Wan Y, Yang Y, Li R, Wu X, Xie M, Li X, Fu G. Research progress in peanut allergens and their allergenicity reduction. Trends Food Sci Technol. 2019;93:212–20. https://doi.org/10.1016/j.tifs.2019.09.014.
51. He M, Xi J. Identification of an IgE epitope of soybean allergen Gly m Bd 60K. LWT. 2020;133:110131. https://doi.org/10.1016/j.lwt.2020.110131.
52. Riascos JJ, Weissinger SM, Weissinger AK, Kulis M, Burks AW, Pons L. The seed biotinylated protein of soybean (*Glycine max*): A Boiling-Resistant new allergen (Gly m 7) with the capacity to induce IgE-Mediated allergic responses. J Agric Food Chem. 2016;64(19):3890–900. https://doi.org/10.1021/acs.jafc.5b05873.
53. Fukutomi Y, Sjölander S, Nakazawa T, Borres MP, Ishii T, Nakayama S, Tanaka A, Taniguchi M, Saito A, Yasueda H, Nakamura H, Akiyama K. Clinical relevance of IgE to Recombinant Gly m 4 in the diagnosis of adult soybean allergy. J Allergy Clin Immunol. 2012;129(3):860–e8633. https://doi.org/10.1016/j.jaci.2012.01.031.
54. Ogawa T, Bando N, Tsuji H, Nishikawa K, Kitamura K. α-Subunit of β-Conglycinin, an allergenic protein recognized by IgE antibodies of Soybean-Sensitive patients with atopic dermatitis. Biosci Biotechnol Biochem. 1995;59(5):831–3. https://doi.org/10.1271/bbb.59.831.

Publisher's Note