Risk assessment of hybrids between genetically modified soybean and its wild relatives

Li Zhang[1], Laipan Liu[2], Zhixiang Fang[2], Wenjing Shen[2], Ying Dai[2], Rui Zong Jia[3], jingang liang[2], and Biao Liu[2]

[1]Nanjing Institute of Environmental Sciences
[2]Affiliation not available
[3]Institute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agriculture Sciences

January 20, 2022

Abstract

In the present study, a 3-year pot experiment was conducted to investigate the agronomic performance of different generation hybrids between genetically modified (GM) soybean and wild soybeans as well as inheritance of the CP4-EPSPS transgene and its effects on the seed germination rate, aboveground biomass, and fecundity in F1, F2 and F3 populations. Furthermore, the expression of transgenic proteins in various hybrids was also investigated. The results showed that the F1 hybrids had higher germination rates (weaker dormancy) and lower pod and seed numbers than the wild soybean. The F2 and F3 populations also had higher germination rates than wild soybean, but the F2 and F3 populations had nearly the same biomass, pod and seed yield as their maternal parents across the whole life cycle; while the seed germination rate, biomass, and fecundity were similar in EPSPS negative, homozygous and heterozygous plants of F2 and F3 populations. Furthermore, EPSPS proteins were detectable in F1, F2 and F3 progeny at different growth stages. While EPSPS genes had little effect on crop growth and reproduction, hybridization between GM soybean to wild soybean may have more impact on hybrid growth and fecundity, especially the seed germination rate and fecundity. F1, F2 and F3 had lower seed germination but higher pod and seed production than GM soybeans, and these parameters were close to those of wild soybean. Such characteristics acquired by gene flow have the potential to promote the adaptability of hybrids and may increase the possibility of dispersal of transgenes through seed systems.

Risk assessment of hybrids between genetically modified soybean and its wild relatives

Running head: Risk evaluation of transgene escape

Li Zhang[1,2], Laipan Liu[1,2], Zhixiang Fang[1,2], Wenjing Shen[1,2], Ying Dai[3][?]Ruizong Jia[4], Jingang Liang[5,*,b], Biao Liu[1,2,4,*,a]

[1]Key Laboratory on Biosafety of Nanjing Institute of Environmental Sciences, Ministry of Ecology and Environment, Nanjing, 210042, China

[2]State Environmental Protection Scientific Observation and Research Station for Ecology and Environment of Wuyi Mountains, Nanjing Institute of Environmental Sciences, Ministry of Ecology and Environment, Nanjing, China

[3]College of Pharmacy, Nanjing University of Chinese Medicine Hanlin College, Taizhou 225300, China

[4]Hainan Key Laboratory for Biosafety Monitoring and Molecular Breeding in off-Season Reproduction Regions, Sanya, China

[5]Development Center of Science and Technology, Ministry of Agriculture and Rural Affairs, Beijing 100176, China

*Corresponding author.

[a] Biao Liu: liubiao@nies.org

[b] Jingang Liang:*liangjingang@agri.gov.cn*

Abstract

In the present study, a 3-year pot experiment was conducted to investigate the agronomic performance of different generation hybrids between genetically modified (GM) soybean and wild soybeans as well as inheritance of the *CP4 -EPSPS* transgene and its effects on the seed germination rate, aboveground biomass, and fecundity in F_1, F_2 and F_3populations. Furthermore, the expression of transgenic proteins in various hybrids was also investigated.

The results showed that the F_1 hybrids had higher germination rates (weaker dormancy) and lower pod and seed numbers than the wild soybean. The F_2 and F_3populations also had higher germination rates than wild soybean, but the F_2 and F_3 populations had nearly the same biomass, pod and seed yield as their maternal parents across the whole life cycle; while the seed germination rate, biomass, and fecundity were similar in EPSPS negative, homozygous and heterozygous plants of F_2 and F_3 populations. Furthermore, EPSPS proteins were detectable in F_1, F_2 and F_3 progeny at different growth stages.

While *EPSPS* genes had little effect on crop growth and reproduction, hybridization between GM soybean to wild soybean may have more impact on hybrid growth and fecundity, especially the seed germination rate and fecundity. F_1, $F_2$and F_3 had lower seed germination but higher pod and seed production than GM soybeans, and these parameters were close to those of wild soybean. Such characteristics acquired by gene flow have the potential to promote the adaptability of hybrids and may increase the possibility of dispersal of transgenes through seed systems.

Keywords

Genetically modified soybean, Hybrids, Aboveground biomass, Fecundity, Ecological risk.

Introduction

Since the first approval of commercial genetically modified (GM) maize in 1996, the extent of global GM plant cultivation has increased rapidly over the past two decades, reaching 191.7 million hectares in 2018. The principal GM crops in 2018 were soybean, occupying 50% or 95.9 million hectares of the global GM area (ISAAA 2019). Most GM soybeans have been modified with agronomically valuable traits, such as herbicide tolerance, insect resistance, and disease resistance (Kamthan et al. 2016), with herbicide tolerance consistently being the dominant trait. Cultivation of GM crops not only provides enormous economic and social benefits but also has positive effects on the environment in terms of decreasing pollution.

Although GM crops have great benefits, some ecological risks are emerging, such as the evolution of resistance in target insects, the parallel movement of foreign genes, and a decrease in the biodiversity of nontarget organisms (Romeis et al. 2008; Lazebnik et al. 2017). One of the main issues in the safety assessment of GM crops is pollen-mediated gene flow between the GM crop and its wild relatives (Song et al. 2010; Devos et al. 2018).

Annual wild soybean (*Glycine soja*) is widely distributed in China, Japan, Korea and northeastern Russia, and China is one of the main distribution areas of wild soybean. It has been reported that wild soybean has excellent characteristics, such as high protein, high yield, and tolerance to salt (Peng et al. 2013), which are valuable genetic resources for cultivated soybean breeding. Since wild soybean is the wild ancestor of cultivated soybean and its chromosome number is the same as that of soybean (2n = 40), outcrossing between

cultivated soybean and wild soybean can frequently occur under natural field conditions, although some studies have shown that gene flow between cultivated and wild soybean occurred at very low frequencies (Mizuguti et al. 2009; Nakayama and Yamaguchi 2002; Liu et al. 2020). However, wild soybean commonly grows throughout almost all of China, and their distributions largely overlap with the distributions of cultivated soybean fields, especially in northeastern and southeastern China (Wang and Li 2012). While more favorable conditions, such as flowering synchrony and certain climatic conditions, are available, greater gene flow may be observed (Yook et al. 2020).

Transgene flow from GM crops to their wild relatives may have the potential to exacerbate weed problems by providing novel traits that allow these plants to compete better, produce more seeds, and become more abundant (Snow, 2002), which could result in changing variations in wild populations (Lu and Xia 2011), and will probably cause unwanted ecological consequences.

Models to predict consequences of gene flow from GM crops to wild relatives require a measure of hybrid fitness (Hails et al. 2005; Weis et al. 2005; Allainguillaume et al. 2006), which is mainly reflected by the ability of individuals to survive and reproduce in their new environment (Burke and Rieseberg 2003; Snow et al. 2003; Halfhill et al. 2005; Yang et al. 2011). Individuals with a higher fitness level might be associated with a higher ability to adapt to the environment and may therefore be beneficial for the establishment of a new plant generation.

At present, a few studies were conducted to evaluate the growth performance of hybrids between GM soybean and wild soybean under greenhouse or field conditions (Guan et al. 2015; Kan et al. 2015; Yook et al. 2020). Kan et al. (2015) measured the F_1 and F_2 hybrids of four wild soybeans and glyphosate-resistant soybean in a greenhouse and found that hybrids had similar pod and seed numbers per plant as their wild relatives. Field experimental studies conducted by Yook et al. (2020) and Guan et al. (2015) also found that F_1 and F_2 hybrids (especially F_2 hybrids) exhibited intermediate characteristics of their parental soybeans in their vegetative and reproductive stages, and all parameters were close to those of wild soybean. The results of previous studies (Guan et al. 2015; Kan et al. 2015; Yook et al. 2020) indicated that hybrids acquired by gene flow from GM soybean to wild soybean were not associated with a fitness cost when there was no glyphosate selection pressure, hybrids resulting from gene flow from GM soybean to wild soybean possess a higher potential of persistence in ecosystem.

The expression of endogenous genes such as *Bt* and*CP4-EPSPS* could improve resistance to insects or herbicides; if the transgene is normally expressed in crop-wild hybrids and progenies and inherited between different generations, the transgene may change a certain trait of wild plants, possibly leading to further undesired environmental consequences (Lu 2008). Therefore, to evaluate the risk of GM soybean and its hybrids resulting from gene flow, it is necessary to investigate protein expression data for assessing and monitoring the biosafety of GM crops and hybrids; however, previous studies mainly focused on vegetative and reproductive hybrids (Guan et al. 2015; Kan et al. 2015; Yook et al. 2020), and the protein levels in hybrid plants were not well investigated. In addition, wild soybean seeds have strong dormancy, while cultivated soybean seeds do not (Kubo et al. 2013). The seed dormancy of the progeny of the hybrid obtained from a cross of wild soybean and GM soybean is still not clear, especially in higher hybrid generations, such as the F_2 and F_3 generations, the hybrid populations will segregate as homozygous resistant plants (RR), heterozygous resistant plants (RS) and homozygous susceptible plants (SS) based on endogenous genes, and the seed dormancy, vegetative growth and fecundity of these three groups are unknown.

As a developing country, China has always attached great importance to the application of GM technology to improve agricultural productivity. After over 20 years of development, GM soybean in China is now closer to the commercialization stage. It is now necessary to monitor the possible gene transfer from GM crops to wild soybean and investigate the characterization of such hybrids before large-scale commercial production of GM soybean. In the present study, we generated and characterized F_1, F_2 and F_3 hybrids between wild soybean and GM soybean lines 40-3-2, and seed dormancy, vegetation and fecundity between herbicide-resistant (homozygous and heterozygous resistant) and susceptible plants in hybrids were measured under greenhouse conditions. The results of this study will provide valuable scientific information for assessing the

environmental risk of GM soybeans.

Materials and Methods

Plant materials

Glyphosate-resistant (GR) transgenic soybean 40-3-2 (labeled GM) and wild soybean Jiang pu were used in this study. The GM soybean expressing synthetic CP4-EPSPS gene confer tolerance against glyphosate herbicide, and wild soybean accessions used in these crosses were obtained in Jiang pu (32.05°N, 118.62°E, Nanjing, China). The experiment was carried out over four consecutive years, 2017-2020, in a greenhouse at the Key Laboratory on Biosafety of Nanjing Institute of Environmental Sciences. In 2017, using GM plants as the male parent (the pollen donor) and wild soybean as the female parent (the pollen recipient), crosses were performed in July by artificial pollination, and 54 F_1hybrid seeds were collected in mid-October. The next year (2018), 23 out of 54 F_1 seeds germinated and were subsequently transplanted, and the hybridity of F_1 plants was double checked by spraying with glyphosate (14.4 g L^{-1}) and PCR analysis as reported by previous studies (Sumarji and Suparno 2017). Among these 23 F_1 individuals, 11 plants were glyphosate resistant and were grown in a greenhouse with their parental lines to examine their characteristics and produce second filial generation (F_2) seeds by self-pollination. F_2hybrids were examined for their characteristics and harvested separately to obtain third filial generation F_3 seeds by self-pollination in 2019, and F_3 individuals were planted to evaluate their traits in 2020.

Seed germination and plant management

In 2018, the 300 seeds from GM and wild soybean samples were grouped into three replicates, and three replicates of 100 seeds each and all 54 F_1 seeds were then placed individually in a 12-well cell culture plate (Corning Costar, New York, USA) with two layers of filter paper. Finally, 400 µL sterile distilled water was added to each well, and plates were kept in climate chambers (Binder model KBF 720, Tuttlingen, Germany) under 55% RH, 25 ± 2 °C and continuous dark conditions for 21 days. The number of germinated seeds per day was recorded (i.e., radicle protrusion > 5 mm) and was expressed as a percentage of the total number of tested seeds (germination percentage). The germinated seeds were removed immediately once they were counted to prevent any counting errors. Germination experiments of GM soybean, wild soybean, hybrid F_2 and hybrid F_3 seeds were carried out under the same experimental conditions in 2019 and 2020.

The germinated seeds were each transplanted and grown in a plastic pot (730 mm × 560 mm × 230 mm) filled with mixed soil containing farmland soil and soil composite (25% peat, 25% compost, 25% perlite, 25% vermiculite) at a ratio of 1:1. A bamboo pole (diameter of 1.5 cm and height of 230 cm) was inserted into the pot and carefully fixed to allow for plant climbing. During the plant growing season, weeds were manually removed from cultivation pots, and agricultural agents such as plant growth regulators, insecticides, and fertilizers were not applied.

Genotyping assay for F_2 and F_3generations

Fluorescence quantitative PCR and digital PCR were used to identify homozygous (RR), GR heterozygous (RS) or homozygous sensitive (SS) hybrids across different plants of the F_2 and F_3 populations, and the sequences of the primers and probes used in the present study are shown in Table 1.

Thirty days after transplanting, fresh leaf samples from GM, wild, F_2 and F_3 were used for total DNA extraction. Approximately 200 mg of leaf tissue was ground to a fine powder in liquid nitrogen, and DNA was extracted using a DNeasy® Plant Mini Kit (Qiagen, Hilden, Germany). The purified DNA was then used as a template for qPCR to amplify the soybean lectin gene and EPSPS gene, and each reaction was repeated three times. To obtain reliable results, Ct values and the [?]Ct between the Ct for the transgene and the Ct for the endogenous control were used to determine which plants contained the transgene. Samples were considered positive for amplicon production when the lectin gene Ct values and EPSPS gene Ct values were both < 35, and the amplification plot clearly demonstrated an exponential increase in the reporter signal in duplicate PCRs. A negative result was assigned when lectin gene Ct values < 35 and no amplification of

the EPSPS gene occurred. Samples with lectin gene Ct values < 35 and EPSPS gene Ct values > 35 were considered indeterminant and required repeat testing.

Samples testing positive by real-time PCR were analyzed for*lectin* and *EPSPS* copy number by digital PCR. The copy number ratio, which is expressed as a ratio between target and reference*lectin* genes for each DNA sample, was calculated and directly used as an indicator for identifying heterozygous and homozygous individuals. A copy number ratio close to 1 would suggest that the sample is a homozygous individual, and a copy number ratio close to 0.5 would suggest a heterozygous individual.

Investigation of plant characteristics

Recycling of ungerminated seeds for viability tests

At the end of the germination test, all ungerminated seeds were carefully collected and dried again at 25 degC to evaluate the seed's ability to germinate. All normal-shaped seeds were selected, partial seed coats were removed by scraping a small portion of the seed coat with a knife, and seed germination was examined as previously described.

Aboveground biomass

Before the pod color turned from green to brown or blank, each plant was bagged loosely with a 1-mm nylon mesh to prevent seeds from splashing. At maturity, 10 plants of each material were randomly selected and then dried for one week at 70 degC, and total biomass was recorded using a balance (PB602-N, Mettler Toledo, Nanikon-Uster, Switzerland).

Fecundity

Ten competitive plants were selected from each genotype at random for recording the number of pods per plant, number of seeds per plant, number of full seeds per plant and 100-seed weight.

CP4-EPSPS protein expression levels in samples

Leaf samples were collected at the vegetative growth stage (V2) (Fehr et al. 1971), flowering stage (R1), podding stage (R3) and mature stage (R7). All samples were quick-frozen with liquid nitrogen and stored at -70 degC for the estimation of EPSPS protein.

The protein was determined by ELISA using EPSPS detection kits (Envirologix, Portland, USA). Ten milligrams of each leaf sample were suspended in 1 mL of phosphate-buffered saline containing Tween 20 (PBST buffer), which was supplied as part of the kit, and all procedures were performed according to the manufacturer's instructions. The absorption was measured on a microplate reader (Infinite M200, Tecan Group Ltd., Mannedorf, Switzerland) at 450 nm.

Statistical analysis

SPSS 20.0 software was used for statistical analyses. One-way ANOVA was used for comparison of the groups. Tukey's multiple comparison test was used to determine the significance of the differences between groups, which were considered significant at $P < 0.05$.

Results

Genotyping assay for F_2 and F_3generations

QPCR and dPCR assays were used to identify RR, RS and SS genotypes of the F_2 and F_3 populations, a total of 168 F_2 and 123 F_3 populations were submitted to genotype analysis. In this research, the identification and screening of 168 F_2 individuals showed that the F_2 population had 45 RR, 86 RS and 37 SS plants. Among the 123 F_3 population, 38 plants were genotyped as RR, 63 RS, and the remaining 22 SS and F_3 populations were shown to segregate at a ratio of 1:2.86:1.73. the genotype test results are shown in Table S1 and Table S2.

Seed germination and vitality of ungerminated seeds

The 21-day seed germination of GM soybean, wild soybean, and hybrids are shown in Fig. 1. In 2018, GM soybean had the highest total germination rate (93.05%), while wild soybean had the lowest germination rate (14.67%). The F_1 hybrid germination rate was intermediate between that of wild and GM soybeans, and there were significant differences in seed germination between the F_1 hybrid and parental lines.

The germination of GM and wild soybean was similar in different years, and the germination of F_2 and F_3 hybrids was 53.67% and 38.33%, respectively. Both F_2 and F_3 hybrids exhibited intermediate germination between wild and GM soybeans, and significant differences were observed in seed germination between F_2 and F_3 hybrids and the parental lines.

At the end of the germination test, all nongerminated GM seeds were found to be mildewed and rotten; in contrast, most ungerminated wild soybean, F_1, F_2, F_3seeds seemed normal and did not change in shape, size or color. These seeds were evaluated for their ability to germinate after partial seed coat removal, and the results showed that the five-day seed germination of all observed wild soybean and hybrids was above 87.5% (Table 2).

Aboveground biomass

The average aboveground biomass of the F_1 plants and their hybrid female parent, wild soybean, was not significantly different, 85.35 g versus 82.28 g, respectively. In contrast, the average aboveground biomass of GM plants was significantly higher than that of F_1 and wild soybean (Fig. 2).

The aboveground biomass of the F_2 population and GM soybean was significantly higher than that of wild soybean, and there were no significant differences between the F_2population and GM soybean in aboveground biomass. Our results also did not show any significant difference among the F_2population in aboveground biomass.

The total aboveground biomass of the F_3 population ranged from 144.06 g, 138.18 g and 131.54 g for the RR, RS and SS plants, respectively, and no significant differences were found in biomass within the F_3 population. However, F_3 had a significantly higher aboveground biomass ($P < 0.05$) than wild soybean, while F_3had slightly higher aboveground biomass than GM soybean, but the difference was not significant.

Fecundity

Pod number per plant

The pod number per plant of wild soybean and F_1 hybrids was significantly higher than that of GM soybean in 2018. Wild soybean produced 565 pods per plant, while F_1 hybrids produced 353 pods, and a significant difference was observed in the wild soybean and F_1 hybrid groups (Table 3).

F_2 homozygous resistant, heterozygous and homozygous susceptible plants produced 561, 545, and 686 pods per plant, respectively, and no significant differences were found between RR, RS and SS. Analysis of variance also showed that the pod numbers were not significantly different between wild soybean and F_2hybrids (Table 3).

Pod numbers were compared between different genotypes of the F_3 population, and no significant of the parameters differed among RR, RS, and SS. The pod number of the F_3hybrids was also intermediate between that of wild and GM soybeans and was more similar to that of the wild soybean. In addition, both F_3 hybrid and wild soybean pod numbers were much higher than those of GM soybeans.

Number of seeds per plant

The seed number and full seed number per plant for wild, GM and hybrid soybeans are shown in Table 3. The seed number and full seed number of wild soybean and the hybrids were significantly higher than those of GM soybean. F_1 hybrids produced 491 seeds and 426 full seeds, which were significantly different from wild soybean, while F_2 and F_3 hybrids had a similar number of seeds as wild soybean. The details of seed number and full seed number per plant are shown in Table 3.

100-seed weight

In three years, GM soybeans had higher 100-seed weights than wild and hybrid soybeans. The F_1 hybrid 100-seed weight was more similar to that of wild soybean, no differences were recorded for 100-seed weight between the F_1 hybrid and wild soybean, however, both F_2 and F_3 hybrids had higher 100-seed weights than their wild soybean counterparts; no parameters significantly differed among RR, RS, and SS in the F_2 or F_3 population.

CP4-EPSPS protein expression levels in samples

The expression levels of the CP4-EPSPS gene in plant leaf samples were assessed during different growth stages of soybean. The results showed that all wild soybean samples and SS plants were negative for EPSPS expression. In contrast, EPSPS was detectable at different stages in F_1, F_2 RR, F_2 RS, F_3 RR, F_3 RS, and GM plants, and the protein levels in both hybrids and GM soybean were influenced by growth stages. The expression of EPSPS declined significantly as plants matured (Fig. 3). The levels of EPSPS in GM plants ranged between 364.28 and 747.79 μg g^{-1}; in F_2 plants, they ranged between 124.15 and 247.89 μg g^{-1}, and the protein levels were highest in the R2 stage (230.07 μg g^{-1}) and lowest in the R7 stage (74.29 μg g^{-1}) in F_3. In addition, similar levels of relative EPSPS expression were observed for $F_2$and F_3 RR and RS plants, and there were no significant differences between RR and RS in different plant generations.

Discussion

Crop-wild/weed hybridization has generated great concerns simply because gene flow can be an avenue for transgene escape, which could alter the genetic make-up of both populations (Ellstrand and Holfman 1990; Song et al. 2004). Evaluation of the fitness of crop-wild hybrids and their parents, especially the wild parents, is a direct way to investigate the potential consequences of crop-to-wild gene flow (Snow et al. 1998). In this study, agronomic comparisons were evaluated among GM soybean, wild soybean and their F_1, F_2 and F_3 progenies, with implications for transgene escape from GM soybean varieties, resulting in a better understanding of its consequences.

Seed germination and the viability of ungerminated seeds

Seed dormancy is an important component of plant fitness that causes a delay in germination until the arrival of a favorable growth season (Graeber et al. 2012). The high seed dormancy of wild soybean is very important for seed survival in the soil and delays seed germination until environmental conditions are correct (Wang and Li 2012). If seeds of hybrids between wild and GM soybean could have stronger dormancy like their wild relatives, it may favor the formation of a longer-lived seed bank enriched with the transgenic seeds.

In the present study, 21-day seed germination results showed that, although F_1, F_2 and F_3hybrid seed germination was significantly higher than that of wild soybean, about half of the F_1, F_2 and F_3 seeds did not germinate, and most ungerminated F_1, F_2, F_3 seeds were deemed normal, i.e., not changed in shaped, size or color. After their partial seed coats were removed, almost all the ungerminated seeds were viable, and the five-day seed germination of all observed hybrid seeds was above 87.5%, suggesting that hybrid derivatives between wild and GM soybean had similar germination characteristics as wild soybean, F_1, F_2 and F_3 hybrid seeds can persist for considerable time in the soil seed bank.

Aboveground biomass

In this study, no significant differences were found in the aboveground biomass between wild soybean and F_1 plants, while the aboveground biomass of the F_2 population and F_3 population was significantly higher than that of their female parent; thus, hybrids between wild soybean and GM soybean, especially F_2 and F_3, had stronger vegetative growth vigor than wild soybean. A similar conclusion was reported by Guan et al. (2015), who measured the F_1 and F_2 hybrids of two wild soybeans and glyphosate-resistant soybean under greenhouse and field conditions and found that GM F_1 hybrids had similar dry weights of the aboveground biomass, while both GM and non-GM F_2 had a significantly higher aboveground biomass than wild soybean in the field. Hybrid plants with higher biomass were also observed in *Brassica rapa* (Moon et al. 2007;

Stewart et al. 2010), rice (Langevin et al. 1990) and sunflower (Mercer et al. 2007), and the increased growth of hybrids compared with that of wild plants might be due to the paternal parents enhancing the plant performance of the hybrid (Mercer et al. 2007).

Fecundity

F_1, F_2, and F_3 hybrids mainly grew well and produced more pods and more seeds than GM soybean, and their fertility increased similarly to that of wild soybeans. Importantly, there were no significant differences in $F_2$and F_3 RR, RS or SS plants in pod and seed numbers per plant, the *EPSPS* gene and its copy number did not significantly affect the fecundity of the hybrids. In a two-year field experiment, Yook et al. (2020) reported that F_1 and F_2 hybrids had a similar number of pods and seeds as wild soybean, and no differences were found for 100-seed weight between F_1 hybrid and wild soybeans; however, F_2hybrids had a higher 100-seed weight than their wild soybean counterparts. Similar to those of the studies described previously (Yook et al. 2020), our results also showed that F_2 and F_3 hybrids had significantly higher 100-seed weights than their wild soybean counterparts, and no parameters significantly differed among RR, RS and SS in the F_2 or F_3 population. Our findings and those of previous studies suggest that hybrids possess higher seed production potential (Guan et al. 2015; Kuan et al. 2015; Yook et al. 2020), the fecundity benefit could confer hybrid more competitive for natural selection than their wild parents.

CP4-EPSPS protein expression levels

To facilitate the biosafety assessment of transgene escape to populations of wild relative species, it is important to conduct scientific research to properly estimate the expression levels of transgenes in wild individuals as well as the inheritance of the transgenes in wild populations (Zhu et al. 2004; Lu and Snow 2005; Lu 2008). In the present study, ELISA was used to detect CP4-EPSPS protein expression levels at different soybean growth stages, while EPSPS protein was detectable in F_{1}-, F_{2}- and F_3-resistant plants. This observation was consistent with the findings of Kubo et al. (2013), who also showed that the gene was stably expressed and produced functional protein. Our study and the studies mentioned previously (Kubo et al. 2013) suggest that the transgene will be able to confer tolerance to glyphosate in the new host wild population. In addition, it is worth noting that a significant decline in the total protein content in both hybrids was observed compared to that in GM soybean. Since tolerance to glyphosate is very dependent on EPSPS protein expression levels in plant tissues, a reduction in the amount of endotoxin proteins in hybrids may contribute to the variability in tolerance. Consistent with our results, a similar conclusion was reported by Zhu et al. (2004), who also reported a decrease in Bt protein content in transgenic *Brassica rapa* and crop-weed hybrids. This difference may be associated with a weedy genetic background, positional effects, and the number of transgenes inserted per event (Moon et al. 2007).

Our study confirmed that the CP4-EPSPS protein was stably expressed in the hybrid soybean line, endowing these hybrid soybeans with herbicide tolerance, and the RR, RS and SS of F_2 or F_3 populations had similar seed germination, aboveground biomass, pod and seed number per plant and 100-seed weight, which indicated that the presence and absence of *EPSPS* or the copy number of the *EPSPS* gene were not significantly correlated with vegetative growth and fecundity. In contrast, heterosis between GM and wild soybean raises new competitive advantages for hybrids, allowing hybrids to obtain some similar growth characteristics as female wild soybean, such as seed dormancy, a higher stable grain weight, and greater pod and seed numbers per plant; these growth characteristics could increase the possibility of dispersal of transgenes through seed systems and may adversely affect genetic and species diversity of wild soybean. Thus, it is critical to build effective risk management and control measures for the gene flow of transgenes from GM soybean to wild soybean before commercial planting of GM soybean in China.

Acknowledgments

The present research was supported by the National Natural Science Foundation of China (32171656) and the National Special Transgenic Project of China (2021ZX08013001-001-006, 2016ZX08012005).

References

Allainguillaume, J., Alexander, M., Bullock, J. M., Saunders, M., Allender, C. J. G., King, C. S. F., & Wilkinson, M. J. (2010). Fitness of hybrids between rapeseed (*Brassica napus*) and wild*Brassica rapa* in natural habitats. *Molecular Ecology* ,*15* (4), 1175-1184. https://doi.org/10.1111/j.1365294X.2006.02856.x

Burke, J. M., & Rieseberg, L. H. (2003). Fitness effects of transgenic disease resistance in sunflowers. *Science* , *300* (5623), 1250.

Devos, Y., Ortiz-García, S., Hokanson, K. E., & Raybould, A. (2018). Teosinte and maize x teosinte hybrid plants in Europe-Environmental risk assessment and management implications for genetically modified maize. *Agriculture Ecosystems and Environment* , *259* (1), 19-27. https://doi.org/10.1016/j.agee.2018.02.032

Ellstrand, N. C., & Holfman, C. A. (1990). Hybridization as an avenue of escape for engineered genes. *BioScience* , *40* (6), 438-442. https://doi.org/10.2307/1311390

Fehr, W. R., Caviness, C. E., Burmood, D. T., & Pennington, J. S. (1971). Stage of development description for soybeans, Glycine max (L.) Merrill. *Crop Science* , *11* , 929-931. https://doi.org/10.2135/cropsci1971.0011183X001100060051x

Graeber, K., Nakabayashi, K., Miatton, E., Leubner-Metzger, G., & Soppe, W. J. J. (2012). Molecular mechanisms of seed dormancy.*Plant Cell and Environment* , *35* , 1769-1786. https://doi.org/10.1111/j.13653040.2012.02542.x

Guan, Z. J., Zhang, P. F., Wei, W., Mi, X. C., Kang, D. M., & Liu, B. (2015). Performance of hybrid progeny formed between genetically modified herbicide-tolerant soybean and its wild ancestor. Aob Plants 7, plv121. https://doi.org/10.1093/aobpla/plv121

Hails, R. S., & Morley, K. (2005). Genes invading new populations: a risk assessment perspective. *Trends in Ecology and Evolution* ,*20* (5), 245-252. https://doi.org/10.1016/j.tree.2005.02.006

Halfhill, M. D., Sutherland, J. P., Moon, H. S., Poppy, G. M., Warwick, S. I., Weissinger, A. K., Rufty, T. W., Raymer, P. L., & Stewart, C. N. (2005). Growth, productivity, and competitiveness of introgressed weedy*Brassica rapa* hybrids selected for the presence of Bt*cry1AC* and *gfp* transgenes. *Molecular Ecology* ,*14* (10), 3177-3189. https://doi.org/10.1111/j.1365-294X.2005.02649.x

ISAAA (The International Service for the Acquisition of Agribiotech Applications). (2019). Global status of commercialized biotech/GM crops: 2018. *China Biotechnology* , *39* (8), 1-6.

Kamthan, A., Chaudhuri, A., Kamthan, M., & Datta, A. (2016). Genetically modified (GM) crops: milestones and new advances in crop improvement. *Theoretical and Applied Genetics* , *129* (9), 1639-1655. https://doi.org/10.1007/s00122-016-2747-6

Kan, G. Z., Tong, Z. F., Hu, Z. B., Ma, D. Y., Zhang, G. Z., & Yu, D. Y. (2015). Fitness of hybrids between wild soybeans (Glycine soja) and the glyphosate-resistant transgenic soybean (Glycine max). *Soybean Science* , *034* (2), 177-184.

Kubo, A., Aono, M., Nakajima, N., Nishizawa, T., Tamaoki, M., & Saji, H. (2013). Characterization of hybrids between wild and genetically modified glyphosate-tolerant soybeans. *Plant Biotechnology* ,*30* (4), 335-345. https://doi.org/10.5511/plantbiotechnology.13.0314a

Langevin, S. A., Clay, K., & Grace, J. B. (1990). The incidence and effects of hybridization between cultivated rice and its related weed red rice (*Oryza sativa* L.). *Evolution* , *44* (4), 1000-1008. https://doi.org/10.1111/j.1558-5646.1990.tb03820.x

Lazebnik, J., Dicke, M., Braak, C. T., & Loon-Van, J. J. A. (2017). Biodiversity analyses for risk assessment of genetically modified potato. *Agriculture Ecosystems and Environment* , *249* , 196-205. https://doi.org/10.1016/j.agee.2017.08.017

Liu, B., Xue, K., Liu, L. P., Shen, W. J., & Guo, H. (2020). Research on the gene flow from transgenic EPSPS+PAT soybean S4003.14 to non-transgenic soybeans. *Journal of Ecology and Rural Environment* ,

36 (7), 367-373.

Lu, B. R. (2008). Transgene escape from GM crops and potential biosafety consequences: an environmental perspective. International Centre for Genetic Engineering and Biotechnology, Collection of Biosafety Reviews 4, 66-141.

Lu, B. R., & Snow, A. A. (2005). Gene flow from genetically modified rice and its environmental consequences. *Bioscience 55* (8), 669-678. https://doi.org/10.1641/00063568(2005) 055[0669: GFFFGMR]2.0.CO;2

Lu, B. R., & Xia, H. (2011). Environmental biosafety of transgenic plants: research and assessment of transgene escape and its potential ecological impacts. *Chinese Science Bulletin* , *23* (2), 186-194.

Mercer, K. L., Andow, D. A., Wyse, D. L., & Shaw, R. G. (2007). Stress and domestication traits increase the relative fitness of crop-wild hybrids in sunflower. *Ecology Letters* , *10* (5), 383-393. https://doi.org/10.1111/j.1461-0248.2007.01029.x

Mizuguti, A., Yoshimura, Y., & Matsuo, K. (2009). Flowering phenologies and natural hybridization of genetically modified and wild soybeans under field conditions. *Weed Biology and Management* , *9*(1), 93-96. https://doi.org/10.1111/j.1445-6664.2008.00324.x

Moon, H. S., Halfhill, M. D., Good, L. L., Raymer, P. L., & Stewart, C. N. (2007). Characterization of directly transformed weedy *Brassica rapa* and introgressed *B. rapa* with Bt *cry1Ac* and*gfp* genes. *Plant Cell Reports* , *26* (7), 1001-1010. https://doi.org/10.1007/s00299-007-0328-5

Nakayama, Y., & Yamaguchi, H. (2002). Natural hybridization in wild soybean (*Glycine max* ssp. soja) by pollen flow from cultivated soybean (*Glycine max* ssp. max) in a designed population.*Weed Biology and Management* , *2* (1), 25-30. https://doi.org/10.1046/j.1445-6664.2002.00043.x

Peng, C., Yan, K., Shao, H., & Zhao, S. J. (2013). Physiological mechanisms for high salt tolerance in wild soybean (*Glycine soja*) from yellow river delta, China: photosynthesis, osmotic regulation, ion flux and antioxidant capacity. *PLoS One* , *8* , e83227. http://dx.doi.org/10.1371/journal.pone.0083227

Romeis, J., Bartsch, D., Bigler, F., Candolfi, M. P., Gielkens, M. M. C., Hartley, S. E., Hellmich, R. L., Huesing, J. E., Jepson, P. C., & Layton. R. (2008). Assessment of risk of insect-resistant transgenic crops to nontarget arthropods. *Nature Biotechnology* , *26* , 203-208. http://dx.doi.org/10.1038/nbt1381

Snow, A. A. (2002). Transgenic crops-why gene flow matters. *Nature Biotechnology* , *20* , 542.

Snow, A. A., Moran-Palma, P., Rieseberg, L. H., Wszelaki, A., & Seiler, G. J. (1998). Fecundity, phenology, and seed dormancy of F_1 wild-crop hybrids in sunflower (*Helianthus annuus* , Asteraceae). *American Journal of Botany* , *85* (6), 794-801. https://doi.org/10.2307/2446414

Snow, A. A., Pilson, D., Rieseberg, L. H., Paulsen, M. J., Pleskac, N., Reagon, M., Wolf, D. E., & Selbo, S. M. (2003). A *Bt* transgene reduces herbivory and enhances fecundity in wild sunflowers.*Ecological Applications* , *13* (2), 279-286. https://doi.org/10.1890/10510761(2003)013[0279: ABTRHA]2.0.CO;2

Song, X. L., Wei, Z. H., Zuo, J., Huangfu, C. H., & Qiang, S. H. (2010). Potential gene flow of two herbicide-tolerant transgenes from oilseed rape to wild *B. juncea* var. *gracilis.Theoretical and Applied Genetics* , *120* (8), 1501-1510. https://doi.org/10.1007/s00122-010-1271-3

Song, Z. P., Lu, B. R., Bin, W., & Chen, J. K. (2004). Fitness estimation through performance comparison of F_1 hybrids with their parental species *Oryza rufipogon* and *O. sativa* . Annals of Botany, *3* , 311-316. https://doi.org/10.1093/aob/mch036

Stewart, C., All, J. N., Raymer, P. L., & Ramachandran, S. (2010). Increased fitness of transgenic insecticidal rapeseed under insect selection pressure. *Molecular Ecology* , *6* (8), 773-779. https://doi.org/10.1046/j.1365-294X.1997.00239.x

Sumarji, & Suparno. (2017). Detektion soybean (*Glycine Max* L Merrill) transgenic GTS 40-3-2 herbiside resistant active based glyphosate PCR using. *International Journal of Applied Environmental Sciences* , *12* (4), 563-576.

Wang, K. J., & Li, X. H. (2011). Inter specific gene flow and the origin of semi-wild soybean revealed by capturing the natural occurrence of introgression between wild and cultivated soybean populations.*Plant Breeding* , *130* (2), 117-127. https://doi.org/10.1111/j.1439-0523.2010.01815.x

Wang, K. J., & Li, X. H. (2012). Fundamental strategies and methods for collection of wild soybean germplasm resources in China. *Journal of Plant genetic resources* , *013* (3), 325-334.

Weis, A. E. (2005). Assessing the ecological fitness of recipients. In G. M. Poppy & M. J. Wilkinson (eds), Gene flow from GM plants (pp. 143-168). Oxford, Blackwell Publishing.

Yang, X., Xia, H., Wang, W., Wang, F., Su, J., Snow, A. A., & Lu, B. R. (2011). Transgenes for insect resistance reduce herbivory and enhance fecundity in advanced-generations of crop-weed hybrids of rice.*Evolutionary Applications* , *4* , 672-684. https://doi.org/10.1111/j.1752-4571.2011.00190.x

Yook, M. J., Park, H. R., Zhang, C. J., Lim, S. H., & Kim, D. S. (2020). Environmental risk assessment of glufosinate-resistant soybean by pollen-mediated gene flow under field conditions in the region of the genetic origin. *Science of the Total Environment* , *762* (2), 143073. https://doi.org/10.1016/j.scitotenv.2020.143073

Zhu, B., Lawrence, J. R., Warwick, S. I., Mason, P., & Stewart, C.N. (2004). Stable *Bacillus thuringiensis* (Bt) toxin content in inter specific F_1 and backcross populations of wild*Brassica rapa* after *Bt* gene transfer. *Molecular Ecology* , *13* (1), 237-241. https://doi.org/10.1046/j.1365-294X.2004.02018.x

Data Availability Statement

Some or all data used during the study are available in a repository or online in accordance with funder data retention policies.

Author Contribution Statement

Author Contribution Statement Li Zhang conducted the experiments, analyzed the data, and wrote the manuscript. Li Zhang, and Biao Liu designed research. Laipan Liu, Zhixiang Fang, Wenjing Shen and Ying Dai participated in the experiments and assisted in analyzing the data. Ruizong Jia, Jingang Liang and Biao Liu reviewed and revised the manuscript. All authors read and approved the final manuscript.

Conflict of Interest Statements

The authors declare that there are no conflicts of interest.

Supplemental Data

Table S1 Detection of EPSPS by real-time quantitative PCR and EPSPS test strip.

Table S2 Identification of two different genotypes,*EPSPS* -resistant heterozygotes and homozygotes, in F_2 or F_3 population plants based on droplet digital PCR.

Tables and Figures

Table 1 Sequences of primers and probes used for the*lectin* and *EPSPS* genes

Target gene	Primer	Primer sequence(5'-3')	Product size
lectin	lectin-F	GCCCTCTACTCCACCCCCA	118 bp
	lectin-R	GCCCATCTGCAAGCCTTTTT	
	lectin-P	FAM AGCTTCGCCGCTTCCTTCAACTTCAC-BHQ1	
EPSPS	EPSPS-Q-1F	TTCATTCAAAATAAGATCATACATACAGGTT	84 bp
	EPSPS-Q-2R	GGCATTTGTAGGAGCCACCTT	
	EPSPS-Q-1P	FAM-CCTTTTCCATTTGGG-BHQ	

Table 2 Germination rate of ungerminated seeds after the seed coat was manually removed

Year	Material	Average artificial treatment seed no.	Average no. of germinated seeds	Germination rate (%)
2018	Wild	70 ± 0	68 ± 2	97.14 ± 2.86
	F_1	32	28	87.5
2019	Wild	70 ± 0	65 ± 3	93.33 ± 3.60
	F_2	39 ± 9	35 ± 8	89.08 ± 2.18
2020	Wild	70 ± 0	67 ± 4	95.24 ± 5.95
	F_3	48 ± 4	44 ± 4	91.67 ± 3.22

Note: Data presented are means ± SD.

Table 3 Fecundity of GM, wild, and hybrid soybeans

Year	Material	No. of pods per plant	No. of seeds per plant	No. of full seeds per plant	100-seed weight (g)
2018	GM	122 ± 43c	234 ± 76b	189 ± 66b	17.62 ± 2.85a
	Wild	565 ± 203a	1138 ± 406a	1010 ± 400a	1.91 ± 0.04b
	F_1	353 ± 136b	491 ± 190b	426 ± 174b	1.81 ± 0.22b
2019	GM	116 ± 33b	222 ± 58b	199 ± 54b	18.80 ± 2.00a
	Wild	631 ± 136a	1099 ± 219a	949 ± 230a	1.89 ± 0.05c
	F_2-RR	561 ± 179a	881 ± 274a	829 ± 258a	10.93 ± 0.33b
	F_2-RS	545 ± 122a	923 ± 287a	869 ± 277a	10.81 ± 0.63b
	F_2-SS	686 ± 123a	1009 ± 231a	953 ± 208a	10.47 ± 0.34b
2020	GM	115 ± 40b	252 ± 76b	217 ± 69b	18.34 ± 0.74a
	Wild	564 ± 144a	1122 ± 225a	944 ± 406a	1.98 ± 0.12c
	F_3-RR	549 ± 152a	1053 ± 305a	995 ± 322a	8.95 ± 0.31b
	F_3-RS	524 ± 113a	985 ± 261a	931 ± 251a	9.19 ± 0.26b
	F_3-SS	527 ± 158a	1008 ± 271a	951 ± 273a	9.02 ± 0.21b

Note: Data presented are means ± SD, means with different superscripts in the same column are significantly different (Tukey's test, $P< 0.05$).

Figures

Hosted file

`image1.emf` available at https://authorea.com/users/456352/articles/553439-risk-assessment-of-hybrids-between-genetically-modified-soybean-and-its-wild-relatives

Figure 1 Seed germination rate of GM, wild, and hybrid soybeans. Error bars indicate standard error.

Hosted file

`image2.emf` available at https://authorea.com/users/456352/articles/553439-risk-assessment-of-hybrids-between-genetically-modified-soybean-and-its-wild-relatives

Figure 2 Aboveground biomass of GM, wild, and hybrid soybeans. GM genetically modified, RR homozygous resistant, RS heterozygous resistant, SS homozygous susceptible. Different lowercase letters denote statistical differences between treatment groups at the 5% level according to Tukey's test.

Hosted file

image3.emf available at https://authorea.com/users/456352/articles/553439-risk-assessment-of-hybrids-between-genetically-modified-soybean-and-its-wild-relatives

Figure 3 ELISA detection of CP4-EPSPS protein in GM and hybrid soybeans. GM genetically modified, RR homozygous resistant, RS heterozygous resistant. Different lowercase letters denote statistical differences between treatment groups at the 5% level according to Tukey's test.