InnerPlant, Inc.

Regulatory Status Review Request for the Determination of Nonregulated Status for Soybean Expressing a Fluorescent Marker Protein

The purpose of this Regulatory Status Review (RSR) is to request a determination that the article should not be regulated under 7 CFR Part 340

Submitting Company

InnerPlant, Inc.
202 Cousteau Place, Suite 150,
Davis, California, 95618

Submitted by:

Randall Shultz, Ph.D.
Vice President of Research
InnerPlant, Inc.
202 Cousteau Place, Suite 150,
Davis, California, 95618

Contact Information:

Email: randy.shultz@innerplant.com
Telephone: (510) 543-1818

Table of Contents

List of Figures

List of Tables

Abbreviations and Definitions

[	]	**CBI-deleted**
APHIS	Animal and Plant Health Inspection Service	
CaMV	Cauliflower Mosaic Virus	
CFIA-PBO	Canadian Food Inspection Agency-Plant Biosafety Office	
[	]	**CBI-deleted**
cv.	Cultivar	
[	]	**CBI-deleted**
FOIA	Freedom of Information Act	
GE	Genetically Engineered	
GFP	Green Fluorescent Protein	
HSP	Heat Shock Protein	
IFB	InnerPlant Fluorescent Biosensor	
[	]	**CBI-deleted**
MAP	Mitogen-Activated Protein	
NDVI	Normalized Difference Vegetation Index	
OECD	Organization for Economic Cooperation and Development	
PR	Pathogenesis Related	
[	]	**CBI-deleted**
RSR	Regulatory Status Review	
TBS	Transformation Booster Sequence	
[	]	**CBI-deleted**
USDA-NASS	United States Department of Agriculture- National Agricultural Statistics Service	
UTR	Untranslated Region	

1 Confidential Business Information (CBI) Statement

This RSR request does contain CBI.

2 CBI Justification

The Freedom of Information Act (FOIA) exempts federal agencies from releasing information that is trade secret and commercial or financial information that is privileged or confidential (5 U.S.C. 552(b)(4)). InnerPlant considers certain information in this application as trade secret. Disclosure of this information would cause substantial competitive harm to InnerPlant by allowing other companies to unfairly compete with InnerPlant. InnerPlant must keep its research confidential: what it is doing and how it is doing it. Disclosure of this information would enable competitors to duplicate InnerPlant research and development without incurring the investment of time and money expended by the company. Moreover, InnerPlant must protect its intellectual property. InnerPlant must keep research information strictly confidential because in some cases, patent applications have not been filed or patents are pending and have not been published.

Specifically, InnerPlant designates the following as Confidential Business Information:

Genetic Elements Identity. InnerPlants biotechnology traits consist of vectors transferred into plants, which comprise genes for the expression of traits and regulatory sequences such as promoters, enhancers, signaling peptides and terminators. Disclosure of this information may also reveal the origin of these genes and genetic elements and the specific modifications the company made in assembling the DNA constructs and enhance their usefulness. It is in InnerPlants commercial interest that these trade secrets not be publicly disclosed.

Transformation Methodology. InnerPlant has developed a novel method for transformation and selection of transformed plants. The key components of this highly efficient transformation method are the selectable marker and specific genetic elements associated with expression of the selectable marker in soybean tissues, as well as the transformation method itself. Disclosure of this information would enable competitors to duplicate InnerPlants research and development without the investment InnerPlant has made in developing this method.

3 Product Description and Rationale

InnerPlant is developing a new data stream for agricultural producers that is fueled by the creation of genetically engineered (GE) crops that produce an optical fluorescence signal that rapidly and specifically indicate the presence of various biotic and abiotic stresses (optical biosensors).

Importantly, InnerPlant has also developed methodology to detect these optical signals in daylight using remote sensing devices that enable detection from tractors, drones, airplanes, and satellites. The combination of biosensors with scalable remote detection capabilities presents an opportunity to provide the industry with vastly superior information about biological pressures such as pathogen infection or insect damage as well as abiotic stresses such as macro and micro-nutrient deficiencies in the soil. These new data streams will enable producers to reduce pesticide usage by targeting only infected areas of the field and will also increase yields by ensuring that pathogens or insect pests are controlled very early in the infection cycle. In addition, nutrient biosensors will enable a step change in precision agriculture unlocking the opportunity to not only reduce over-application of fertilizers but to optimize inputs on a plant-by-plant level.

The basic concept for each of the biosensors InnerPlant is developing is the same, we identify the genetic pathways that respond specifically to a particular stress using transcriptomic and genomic analyses, we then clone the regulatory elements from those endogenous genes and use them to drive the expression of a fluorescent protein that produces an optical signal that can be detected remotely. This approach does not alter the plant's endogenous metabolic pathways, agronomic characteristics or interactions with the environment, but simply adds the new fluorescence gene in combination with the copy of the regulatory elements. Different biosensor designs are built in the same way using a different set of promoter and regulatory elements that respond to the desired stress condition, and in some cases, we use a different fluorescent protein with unique optical properties (excitation, emission, etc.) to enable multiplexing of biosensor signals. In the present application InnerPlant is requesting a Regulatory Status Review (RSR) of GE soybeans that are designed to express a fluorescent protein specifically in response to very early pathogen infection. Hereafter these soybeans are referred to as InnerPlant Fungal Biosensor soybeans or IFB soybeans.

Soybeans are an important part of United States agriculture with approximately 90 million acres planted in 2022 (USDA-NASS, 2022). Like all crops, soybean yields are negatively impacted by various biotic and abiotic stresses with soybean diseases being particularly detrimental representing an estimated $95 billion total economic loss in the US between 1996 to 2016 (Bandara et al. 2020a). The use of foliar-applied fungicides is an important management strategy to mitigate key diseases including Septoria brown spot, frogeye leaf spot as well as other yield-reducing fungal pathogens (Bandara et al. 2020b), however early detection and intervention is critical to minimize the losses. Unfortunately, current practices including field scouting and remote sensing measurements such as normalized difference vegetation index (NDVI) do not provide early, specific detection of disease pressure thereby limiting the effectiveness of intervention practices.

The IFB soybeans are being produced by [ ] transformation of soybean tissues from non-transgenic cultivar (cv.) Williams 82 using [ ]. The DNA transferred from the [ ] contains two gene constructs. The first gene is *lanFP1* encoding the Green Fluorescent Protein (GFP). IFB soybeans **CBI-deleted** **CBI-deleted** **CBI-deleted**

produce GFP specifically at the onset of pathogen infection and this results in rapid production of an optical fluorescence signal that can be detected in the field. This allows early detection of pathogen infection in the soybean crop and enables application of control measures at the outset of infection to mitigate significant damage to the crop. GFP and other fluorescent protein biosensors represent a step change in disease management by utilizing the plants natural disease response pathways, which are activated within hours post-infection (Westrick *et al*. 2019; Cabre *et al*. 2021; Bueno *et al*. 2022).

The second gene transferred to the soybean genome encodes [ ]. Expression of this protein in plant cells [ ] and serves as a selectable marker for plant transformation. The mechanism of action and safety of the GFP and [ ] proteins are reviewed in this RSR application.

CBI-deleted
CBI-deleted **CBI-deleted**
CBI-deleted
CBI-deleted

4 Description of Comparator Plant

The biology of soybean described herein is based upon the consensus document for *Glycine max* (L.) Merr. prepared by the Organization for Economic Co-operation and Development (OECD, 2000), as well as a biology document published by CFIA-PBO (CFIA, 1996), and recent literature on the topic.

A The Taxonomy and Genetics of Soybean

The cultivated soybean, *Glycine max* (L.) Merr., a diploidized tetraploid (2n=40) that belongs to the family Fabaceae and is further classified taxonomically as follows:

Kingdom	Plantae -- Plants
Subkingdom	Tracheobionta -- vascular plants
Division	Magnoliophyta -- angiosperms, flowering plants
Class	Magnoliopsida -- dicots
Subclass	Rosidae
Order	Fabales
Family	Fabaceae
Genus	Glycine Willd. -- soybean
Species	Glycine max (L.) Merr. -- soybean

The above taxonomic information for soybean was obtained from the Integrated Taxonomic Information System (http://www.itis.gov/) and soybean is assigned the taxonomic serial number 26716.

The genus *Glycine* Willd. contains two subgenera, *Glycine* and *Soja* (Moench) F.J. Herm. The subgenus *Glycine* comprises 22 wild perennial species that are indigenous to Australia, islands in the west, central and southern Pacific Ocean, China, Russia, Japan, Indonesia, Korea, Papua New Guinea, the Philippines, and Taiwan (Hymowitz, 2004) The cultivated soybean, *G. max* (L.) Merr. and its wild annual relatives from Asia, *G. soja* Sieb. and Zucc. are classified in the subgenus *Soja*. *Glycine soja* is an annual that grows in the wild in fields, hedgerows, roadsides, and riverbanks in many countries of East Asia.

In addition to *G. max* and *G. soja*, the subgenus *Soja* also contains a form known as *G. gracilis*. This semi-cultivated or weedy plant is found only in Northeast China and is intermediate in morphology between *G. max* and *G. soja*. *G. gracilis* is a variant of *G. max* (Hermann, 1962; Wang, 1976; Shoemaker *et al*., 1986). The three species of the subgenus *Soja* are capable of cross-pollination and the hybrid seed that is produced can germinate normally and produce plants with

fertile pollen and seed (Singh and Hymowitz, 1989). The wild, weedy relatives of *G. max, G. soja* and *G. gracilis* are indigenous to Asia and do not occur in the U.S. (USDA-APHIS, 2008). Therefore, there is no potential for outcrossing of *G. max* to weedy relatives in the U.S.

B Reproductive Biology and Hybridization with Cultivated Soybean

Soybean is a self-pollinating species that is propagated by seed (OECD, 2000). Due to the strong propensity for self-fertilization, the frequency of soybean cross-pollination is very low. For example, plants grown in close proximity to each other (15 cm) were found to have average outcrossing rates of 1.8%, while plants separated by distances of 0.9 m and 5.4 m had outcrossing rates of 0.41 and 0.03%, respectively (Ray *et al*., 2003). Soybeans are generally not a preferred plant for insect pollinators and insect activity has been found not to increase the outcrossing rate (Erickson, 1975; Erickson, 1984). The regulations governing the production of certified Foundation soybean seed are consistent with the low outcrossing rate recognized for soybean. These regulations place no restriction on the separation distance between different cultivars in the field provided that it is sufficient to prevent mechanical mixing during harvest (USDA-APHIS, 2008).

C Weediness Potential of Cultivated Soybean

Soybean plants are not weedy and are not found outside of cultivation. Soybeans are annuals that reproduce solely from seeds. Cultivated soybean rarely displays any dormancy characteristics (a desirable trait that is selected for in commercial varieties) (TeKrony *et al*., 1987) and are sensitive to cold temperatures (Raper and Kramer, 1987), their potential to survive in the U.S. from one growing season to the next is very low. Soybean seeds normally germinate quickly under the appropriate environmental conditions that include adequate moisture and moderate temperatures and could potentially grow as a volunteer. However, any volunteers that grow after harvest would be destroyed by the low and freezing temperatures encountered during the following winter. In the event that volunteers were to become established, they would not compete well with succeeding crops and they could be controlled by either mechanical or chemical means (OECD, 2000). The low weediness potential for soybeans is reflected in the fact that soybeans are not listed on the USDA Federal Noxious Weed List (USDA, 2006).

D Characteristics of the Recipient Soybean Cultivar

The recipient soybean cultivar Williams 82 (PI518671) was genetically modified to express the fluorescent protein. Cultivar Williams 82 is a Phytophthora-resistant variety and is an F_3-derived selection from the cross Williams x Kingwa (Haun *et al*. 2011). It was released in the United States in 1988. It is a group III maturity cultivar.

5 Genotype of the Modified Plant

As presented above, IFB soybeans are being produced by [ ] CBI-deleted transformation of soybean tissues from non-transgenic cultivar (cv.) Williams 82 using [ ] CBI-deleted. [ ] is presented in Figure 1. The nucleotide sequence of the DNA [ ] to the soybean genome [ ] is presented below in Figure 2. The order of the genetic components in the [ ], the donor organism of each genetic element, a description of the function of the genes and genetic elements as well as the GenBank number of the DNA sequences of the genes and genetic elements are presented in Table 1.

CBI-deleted CBI-deleted CBI-deleted CBI-deleted CBI-deleted CBI-deleted

A [ ] CBI deleted

Figure 1. [CBI-deleted
] CBI-deleted
[CBI-deleted

CBI-deleted

] CBI-deleted

[ ] contains CBI-deleted
two gene constructs. The first gene is *lanFP1* encoding the Green Fluorescent Protein (GFP). The second gene transferred to the soybean genome encodes [CBI-deleted
]. Expression of this protein in plant cells [CBI-deleted CBI-deleted
] and serves as a selectable marker for plant CBI-deleted
transformation.

B Sequence of the [ ] insert in IFB soybeans CBI-deleted

Figure 2. Nucleotide sequence of the [ ] insert in IFB soybeans CBI-deleted

```
   1 [                         ]AAATT GACGCTTAGA CAACTTAATA
  51 ACACATTGCG GACGTTTTTA ATGTACTGAA TTAACGCCGA ATTAATTCGG
 101 GGGATCTGGA TTTTAGTACT GGATTTTGGT TTTAGGAATT AGAAATTTTA
 151 TTGATAGAAG TATTTTACAA ATACAAATAC ATACTAAGGG TTTCTTATAT
 201 GCTCAACACA TGAGCGAAAC CCTATAGGAA CCCTAATTCC CTTATCTGGG
 251 AACTACTCAC ACATTATTAT GGAGAAACTC GAG[
 301
 351
 401
 451
 501
 551
 601
 651
 701
 751
 801
 851
 901
 951
1001
1051
1101
1151
1201
1251
1301    ]TGTTTAG TATACTAGAA TACCAGCGTG TCTCGAGAGA GATAGATTTG
1351 TAGAGAGAGA CTGGTGATTT CAGCGTGTCC TCTCCAAATG AAATGAACTT
1401 CCTTATATAG AGGAAGGTCT TGCGAAGGAT AGTGGGATTG TGCGTCATCC
1451 CTTACGTCAG TGGAGATATC ACATCAATCC ACTTGCTTTG AAGACGTGGT
1501 TGGAACGTCT TCTTTTTCCA CGATGCTCCT CGTGGGTGGG GGTCCATCTT
1551 TGGGACCACT GTCGGCAGAG GCATCTTGAA CGATAGCCTT TCCTTTATCG
1601 CAATGATGGC ATTTGTAGGT GCCACCTTCC TTTTCTACTG TCCTTTTGAT
1651 GAAGTGACAG ATAGCTGGGC AATGGAATCC GAGGAGGTTT CCCGATATTA
1701 CCCTTTGTTG AAAAGTCTCA ATAGCCCTTT GGTCTTCTGA GACTGTATCT
1751 TTGATATTCT TGGAGTAGAC GAGAGTGTCG TGCTCCACCA TGTTATCACA
1801 TCAATCCACT TGCTTTGAAG ACGTGGTTGG AACGTCTTCT TTTTCCACGA
1851 TGCTCCTCGT GGGTGGGGGT CCATCTTTGG GACCACTGTC GGCAGAGGCA
1901 TCTTGAACGA TAGCCTTTCC TTTATCGCAA TGATGGCATT TGTAGGTGCC
1951 ACCTTCCTTT TCTACTGTCC TTTTGATGAA GTGACAGATA GCTGGGCAAT
2001 GGAATCCGAG GAGGTTTCCC GATATTACCC TTTGTTGAAA AGTCTCAATA
2051 GCCCTTTGGT CTTCTGAGAC TGTATCTTTG ATATTCTTGG AGTAGACGAG
2101 AGTGTCGTGC TCCACCATGT TGGCAAGCTG CTCTAGCCAA TACGCAAACC
2151 GCCTCTCCCC GCGCGTTGGC CGATTCATTA ATGCAGCTGG CACGACAGGT
2201 TTCCCGACTG GAAAGCGGGC AGTGAGCGCA ACGCAATTAA TGTGAGTTAG
2251 CTCACTCATT AGGCACCCCA GGCTTTACAC TTTATGCTTC CGGCTCGTAT
2301 GTTGTGTGGA ATTGTGAGCG GATAACAATT TCACACAGGA AACAGCTATG
2351 ACCATGATTA CGAATTGTTG ACGTCAAAAC GCGTAAGTTA ACAGGTACCA
2401 AGTTGTAATG AGTTGCTGGC CTCTCTTGGG ATTGTTCATC AAAGTAGCTG
2451 TCCCTACACA CCACAACAGA ATGGTGTTGT GGAGAGAAAG CACAGACACA
2501 TCCTTGAGAT GGCAAGGGCA CTTAAGTTTC AGAGTGGTGT ACCTACCAGG
```

CBI-deleted

```
2551 TTTTGGGGAG ACTGTGTCAA GACTGTTGTA TACCTGATCA ATAGACTCCC
2601 TACTCCAATC TTACAAGGCA AATGTCCATA TGAACTACTC TATCAGAAAC
2651 ATGCCAAGCT TGATCACTTG AGAGTGTTTG GGTGTCTTGC TTTTGCAGCA
2701 ACACTGCCCA AAGGTGACAA ACTTGCACCT AGAGCCAAGC GAACCATCTT
2751 CATTGGTTAT TCTGAGACTC AAAAGGGTTA CAGGTTGTAT GACTTGGATA
2801 ATAAGGTGAT CATAGTGAGC AGGGATGTAG TTTTCAGAGA GTTTCAGTTC
2851 CCTTTCAAAG AGTCCATCTC CCATGAACCT GATATGTTCA CTCGAGCAGG
2901 TTCACTTAGC TCTGAAGATA CAGCTTTGCA GCTAGTGTAT AATGATATTT
2951 TTCCTTCCAG CTTTCATACT GATGATGATG ATACTCCAGT TCTTATTCCT
3001 GACATTGTAG AGGCCAATGA GGATATCACT TTGGGGACTC ATTCTACCCA
3051 AAATCATATC ACTTCAGCTG ATGCTGCACC TTCAGGCCAC CTGCATGCTC
3101 CTGCAGAGTC AGTTGAGTCA GCTGAGCCTG ACCTAGCAAA TGCTGAATTT
3151 GAGCACCATA CTGCTGTAGC TGATCCCTCT CTAGTTCCAG CAAACCTGCA
3201 CACTAGACCT AAACGCAATG CAGGTCCTCC CATCTGGCTC AAGGACTTTG
3251 TGACACTGAA CAAAGGCTCT AGGGATGTTC CATACCCTAT AGCCAACTAT
3301 GTTTCCTATG ATCACTTGCC TGTTCACTAT CAAGCTTATT TGAGTGCTTT
3351 CTCCACTTAT ACTGAACCTA AATCCTTCAA AGAAGCAGCT CAAGATGAGA
3401 AATGGATGGA GGCCATGTCC CTTGAGATAC AGGCTCTTGA GGATAATAAC
3451 ACCTGGGAGA TTGTCCCTTT ACCCCCTGGT AAACAGCCTA TAGGGTCCAA
3501 ATGGGTGTAC AAAATTAAAT ACAAAGCTAA TGGTGAAGTT GACAGGTTTA
3551 AGGCAAGGCT AGTGGCCAAG GGATACACTC AGCAAGAAGG CCTTGACTAC
3601 CATGAAACTT TTTCTCCAGT GGCCAAAATG GTCACTGTAA AAGCTGTCAT
3651 ATCTGTTGCT GCTTCCAAGG GCTGGTTCCT TTTCCAGATG GATGTCAACA
3701 ATGCTTTTCT ACAAGGTGAC CTCATGGAAG AAGTTTATAT GTCTCTGCCT
3751 CAGGGTTTTC ACAGCCAAGG GGAGACCAAA GTGTGCAGGC TGTTGAATCC
3801 CTCTATGGTT TAAAGCAAGC ATCAAGGCAG TGGAACATCA AGCTTACCAC
3851 TGTCCTTATG CAGGCTGGTT TTATGCAAAG TGCTTATGAT CACTCCTTGT
3901 TTACCAAAAG AGAGGGGACT GACCTTGTCA TAATCCTGAT TTATGTGGAT
3951 GATCTGTTAA TAACAGGCAG CAGCAATGTT TTGATTTCAG AAGCAAAGGC
4001 AACCCTGCAT CAGCATTTCA AGATGAAAGA TTTGGGAGAA CTAAAATACT
4051 TCTTAGGCAT TGAGGTGCTG AGATCAGAAA AGGGAATCCT ACTGAACCAG
4101 AGGAAGTATG CACTGGAACT GATATCAGGT GTTGGTCTGG GAGGCTGCAG
4151 ACCAGTGTCA ACCCCAATGG AACAGAACCA AAGGCTGACT ACTGTAGAAT
4201 ATGACAAGCA CTTAGGAAAG ACTGATGATG CAGAGTTAGA GGATGTGGGA
4251 TCCTATCAGA GACTGGTGGG AAAGCTTCTC TACTTGACAA TCACAAGGCC
4301 AGATATCAGC TTTGCAGTGC AGGTGCTGAG TCAATTCATG CAGCAACCCA
4351 AACAGTCACA CCTGGAAGCA GCATTGAGGG TTGTGAAGTA CATAAAAGGT
4401 TCTCCAGGTG TTAGGAGAGC TCAACTAATT CTAGACAGGG CCCG[
4451
4501
4551
4601
4651
4701
4751
4801
4851
4901
4951
5001
5051
5101
5151
5201
5251
5301
5351                                                    ]
```

CBI-deleted

```
5401 [                                                                                  CBI-deleted
5451                                                                                    CBI-deleted
5501                                                                                    CBI-deleted
5551                                                                                    CBI-deleted
5601                                                                                    CBI-deleted
5651                                                                                    CBI-deleted
5701                                                                                    CBI-deleted
5751                                                                                    CBI-deleted
5801                                                                                    CBI-deleted
5851                                                                                    CBI-deleted
5901                                                                                    CBI-deleted
5951                                                                                    CBI-deleted
6001                                                                                    CBI-deleted
6051                                                                                    CBI-deleted
6101                                                                                    CBI-deleted
6151    ]CCTAGG CGGCGATCCA CAGGGAAGGA GCAGCATCTC CACAAAGACG                             CBI-deleted
6201 CACTACAGAA GACTAAAGAG AGCTTTTTCA TACCAAAGAA GTACAACAAA
6251 AGATTTGCTC CTCATTTTCT GAATCCTGGG ACTCTCTAGC CTGTAGAAGA
6301 AGAAAGGCAG GAATTTCAGC TCAAGAGAAC AGATCACAAT ATTTACCCAC
6351 GGCACTGTCT CGCAATCCAT GGCATTGCCC GCCACCCATG ACATTCATCT
6401 GCACGGTTCC ATAAATGGCC ACGAGTTCGA TATGGTCGGA GGGGGGAAGG
6451 GAGATCCTAA CGCAGGCTCA CTGGTAACAA CAGCAAAGTC AACTAAAGGC
6501 GCTCTGAAGT TCTCACCTTA CTTGATGATA CCCCACCTTG GATACGGGTA
6551 TTATCAATAT CTTCCATATC CCGACGGACC CAGCCCTTTC CAAACCTCTA
6601 TGCTTGAAGG CAGTGGGTAT GCTGTGTATC GCGTCTTTGA CTTTGAGGAC
6651 GGGGGAAAGC TCACAACAGA ATTTAAGTAT TCATACGAAG GCTCACACAT
6701 AAAGGCTGAC ATGAAGTTGA TGGGGAGTGG ATTCCCAGAC GATGGTCCAG
6751 TGATGACTAG CCAGATCGTG GACCAGGACG GCTGCGTGAG CAAAAAGACC
6801 TATCTCAATA ACAATACAAT AGTTGACAGT TTTGACTGGT CATATAACTT
6851 GCAAAACGGA AAAAGATACC GTGCTCGTGT CAGTTCACAT TACATCTTTG
6901 ACAAGCCATT TAGTGCTGAT CTCATGAAAA AACAGCCCGT ATTTGTCTAC
6951 CGCAAATGCC ATGTAAAGGC TAGTAAGACA GAGGTTACCT TGGACGAACG
7001 TGAGAAGGCA TTCTACGAAT TGGCATGA]GG CGCGCCATAT GAAGATGAAG
7051 ATGAAATATT TGGTGTGTCA AATAAAAAGC TTGTGTGCTT AAGTTTGTGT
7101 TTTTTTCTTG GCTTGTTGTG TTATGAATTT GTGGCTTTTT CTAATATTAA
7151 ATGAATGTAA GATCTCATTA TAATGAATAA ACAAATGTTT CTATAATCCA
7201 TTGTGAATGT TTTGTTGGAT CTCTTCTGCA GCATATAACT ACTGTATGTG
7251 CTATGGTATG GACTATGGAA TATGATTAAA GATAAGCCTG CAGGCATGCA
7301 AGCTTGGCGC GCGCGGTGTC ATCTATGTTA CTAGATCGGG AATTAAACTA
7351 TCAGTGTTTG ACA[                       ]                                         CBI-deleted
```

C Annotation of the [ ] inserted genetic material

CBI-deleted

Table 1. Genetic material inserted into the genome of IFB soybeans

Genetic Element	Position in the genetic insertion	Donor Organism	Function	GenBank No.
[ ]	[1-25]	[ ]	[ ]	[ ]
Intervening sequence	26-102	Not applicable. Synthetic sequence, polylinker	Sequence used for DNA cloning.	
CaMV polyadenylation signal	103-277	Cauliflower Mosaic Virus	The 3' Untranslated Region (UTR) sequence of the 35S RNA of cauliflower mosaic virus (CaMV) (Mogen *et al*., 1990) that directs polyadenylation in plant cells and terminates transcription of the [ ] gene inserted in the soybean genome.	GenBank: KY703615.1 (7460..7634)
Intervening sequence	278-283	Not applicable. Synthetic sequence, polylinker	Sequence used for DNA cloning.	

CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted

Table 1 continued. Genetic material inserted into the soybean genome

Genetic Element	Position in the genetic insertion	Donor Organism	Function	GenBank No.
[ ]	284-1075	*Escherichia coli*	Bacterial coding sequence [ ] and serves as a selectable marker for plant transformation [ ]	[ ]
[ ]	1076-1303	*Arabidopsis thaliana*	[ ]	[ ]
Intervening sequence	1304-1369	Not applicable. Synthetic sequence, polylinker	Sequence used for DNA cloning.	
Enhanced CaMV 35S promoter	1370-2047	Cauliflower Mosaic Virus (CaMV)	The cauliflower mosaic virus promoter (Odell *et al*., 1985) with the duplicated enhancer region (Kay *et al*., 1987) used to drive transcription of the [ ] gene.	GenBank: V00140.1 (7016..7441)
Intervening sequence	2048-2362	Not applicable. Synthetic sequence, polylinker	Sequence used for DNA cloning.	

CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted

Table 1 continued. Genetic material inserted into the soybean genome

Genetic Element	Position in the genetic insertion	Donor Organism	Function	GenBank No.
TBS insulator	2363-4444	*Petunia hybrida*	A transformation booster sequence (TBS) from *Petunia hybrida* functions as an enhancer-blocking insulator between promoter and enhancer genetic elements (Hily *et al*., 2009).	GenBank: EU864306.1
[ ] promoter	4445-6154	*Glycine max*	The [ ] promoter driving expression of the GFP was selected by analyzing RNA-sequence data for soybean genes that were specifically upregulated in response to pathogen attack (Innerplant unpublished data). [ ]	[ ]
Intervening sequence	6155-6160	Not applicable. Synthetic sequence, polylinker	Sequence used for DNA cloning.	
OsMac3 5'UTR	6161-6368	*Oryza sativa*	A 5' UTR leader DNA sequence derived from *Oryza sativa* that enhances protein expression in plants. (Aoki *et al*., 2014)	GenBank:NC_029260.1 (25631045..25631240)

CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted

Table 1 continued. Genetic material inserted into the soybean genome

Genetic Element	Position in the genetic insertion	Donor Organism	Function	GenBank No.
lanFP1: Coding sequence for the Green Fluorescent Protein (GFP) from Lancelet	6369-7028	*Branchiostoma floridae*	Expression of the GFP is activated by pathogen attack of the soybean plant and when the GFP receives excitation light it emits green fluorescent light at 509 nm (Baumann *et al*., 2008).	GenBank: XP_035658893.1
Intervening sequence	7029-7032	Not applicable. Synthetic sequence, polylinker	Sequence used for DNA cloning.	
AtHSP terminator	7033-7309	*Arabidopsis thaliana*	The 3'UTR DNA sequence of the terminator for heat shock protein HSP18.2, that terminates transcription of the *GFP* gene (Nagaya *et al*., 2010).	GenBank: NC_003076.8 (24063118..24063367)
Intervening sequence	7310-7358	Not applicable. Synthetic sequence, polylinker	Sequence used for DNA cloning.	
[ ]	7359-7383	[ ]	[ ]	[ ]

CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted

6 Description of New Trait

A Intended trait

GFP:
The soybean contains a fluorescence biosensor protein that produces a fluorescence signal when the plant is experiencing a fungal infection. The fluorescence trait is only visible when the protein absorbs the correct excitation light and optical filters are used to visualize the fluorescent light.

[ ] **CBI-deleted**
The soybean contains a protein [**CBI-deleted**
] serves as a selectable marker for plant transformation. **CBI-deleted**

B Intended phenotype

GFP:
The phenotype of IFB soybean is green fluorescence with an emission peak at approximately 509 nm when the plant is subject to pathogen infection.

[ ] **CBI-deleted**
IFB soybeans are resistant to the antibiotics [ ] **CBI-deleted**
protein serves as a selectable marker for plant transformation.

C Description of the Mechanism of Action (MOA)

Plants possess defense mechanisms against pathogen infection. These defense mechanisms involve highly coordinated and complex molecular mechanisms from detection of the pathogen to signaling pathways and release of pathogenesis-related (PR) proteins by the plant to control pathogen infection (Mazarei *et al*., 2008). Specifically, upon infection, the pathogen releases effector or elicitor molecules into the plant that include glycoproteins, peptides, carbohydrates and lipids (Nürnberger *et al*., 2004). Specific and nonspecific elicitors trigger signal transduction cascades involving protein kinases, elements of the mitogen-activated protein (MAP) kinase pathway, and protein phosphatases ((Nürnberger *et al*., 2004; Desikan *et al*., 1999). Further, the pathways leading to resistance against a particular pathogen are mainly intertwined signaling pathways among which small signaling molecules namely salicylate, jasmonate, and ethylene play key roles which enable the plant to direct the defense responses to both local and systemic tissues (Feys and Parker, 2000). One result of the response to pathogen elicitors and signaling molecules is the activation of expression of PR proteins which have a key role in plant defense mechanisms. In this aspect, promoter sequences are the key elements. The role of a promoter in gene expression

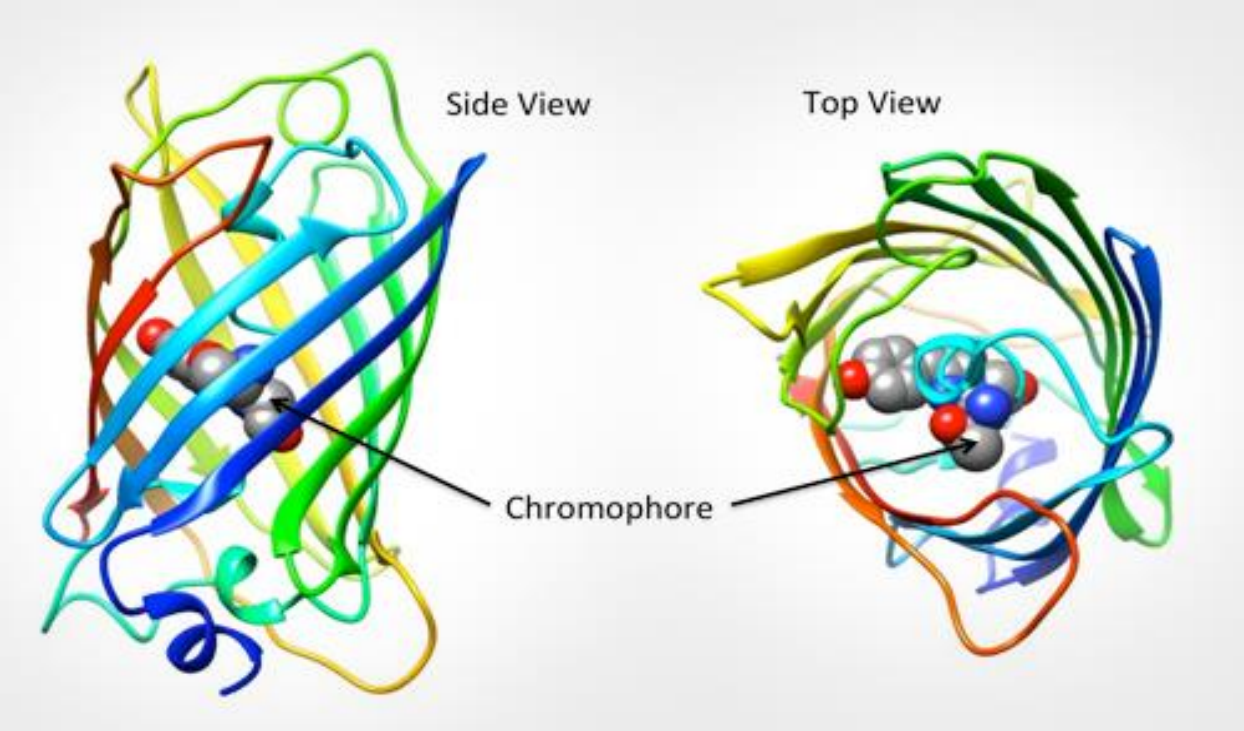
and regulation is well known. Inducible plant defense is a result of the interaction of transcription factors, various related cis-regulatory elements, inducible promoters, and transcription of PR genes (Baruah *et al*., 2020; Mazarei *et al*., 2008).

InnerPlant has utilized the understanding of the molecular pathways for plant defense against pathogens to develop IFB soybeans. An identified inducible promoter was fused to the coding sequence of GFP to produce soybeans that emit a fluorescent signal at the onset of pathogen infection. The promoter driving expression of the GFP was selected by analyzing RNA-sequence data for soybean genes that were specifically upregulated in response to pathogen attack (InnerPlant unpublished data). InnerPlant identified [**CBI-deleted**
] Pathogenesis-Related Protein [ ] **CBI-deleted**
Pathogenesis-related proteins are known to be a central part of the plant defense system and are transcriptionally activated by various pathogens (Pape *et al*., 2010; Fu and Dong, 2013).

Expression of GFP utilizes the plant's natural disease response pathways, which are activated within hours post-infection. Therefore, IFB soybeans produce GFP specifically at the onset of pathogen infection and this results in rapid production of an optical fluorescence signal that can be detected in the field. This allows early detection of pathogen infection in the soybean crop and enables application of control measures at the outset of infection to mitigate significant damage to the crop. The mechanism of action from pathogen infection to activation of the plant's disease response has been described above. The following describes the mechanism of action of GFP.

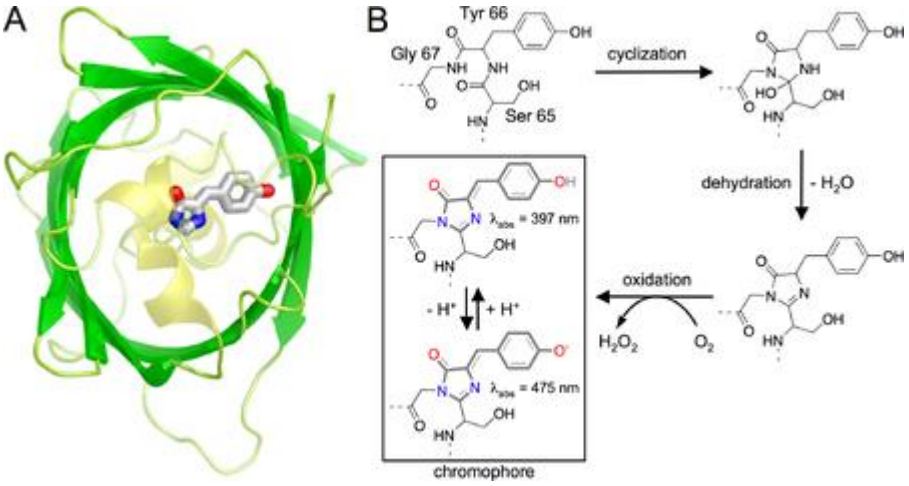
Two independent reports of the x-ray crystal structure of GFP (Ormo *et al*., 1996; Yang *et al*., 1996) revealed that the protein has a unique overall fold comprised of an 11-stranded β-sheet wrapped into a cylindrical β-barrel protein that is 42 amino acids in height and 24 amino acids in diameter (Figure 3). The chromophore is located near the center of the protein, attached to a helical segment of the protein that threads through the center of the β-barrel along its long axis.

Figure 3. The structure of GFP from the side and top. GFP is a hollow barrel shape with a chromophore in the center (the fluorescent portion). Image reproduced from Protein Database Bank, PDB (2022)



The chromophore is spontaneously formed in GFP within the folded β-barrel protein structure. It has been proposed that formation of the chromophore must necessarily involve at least three key steps: cyclization of the main chain, loss of a molecule of water (dehydration), and oxidation with molecular oxygen (Campbell, 2008). An early, and still generally accepted, proposed mechanism is shown in Figure 4B (Heim *et al*. 1994). In this mechanism, chromophore formation starts with the nucleophilic glycine 67 amide nitrogen attacking the electrophilic serine 65 carbonyl carbon to form a 5-membered ring in the main chain of the protein. The resulting tetrahedral hemiaminal intermediate undergoes an elimination of water to form a second intermediate. In the final step, the Cα-Cβ bond of tyrosine 66 is oxidized to a double bond with consumption of molecular oxygen and generation of hydrogen peroxide (Zhang *et al*. 2006). The installation of this double bond simultaneously converts the 5-membered ring into an aromatic system and puts it into conjugation with the aromatic phenol ring of the tyrosine side chain. Chromophore formation is spontaneous only within the context of the fluorescent protein β-barrel structure where steric constraints force the peptide into a tight turn conformation (Branchini *et al*. 1998) and the side chains of highly conserved residues, such as glutamate 222 and arginine 96, are positioned to facilitate the reaction.

Figure 4. A. Top view of the GFP structure with barrel shaped protein and central chromophore. **B**. A proposed mechanism for the series of post-translational modifications that converts the serine 65, tyrosine 66, glycine 67 tripeptide sequence into the fluorescent chromophore (Heim *et al*. 1994). Reproduced from Campbell (2008).



The GFP chromophore exists as an equilibrating mixture of the neutral phenol (absorbance λmax = 397 nm) and anionic phenolate (absorbance λmax = 475 nm) (Morise *et al*. 1974; Heim *et al*. 1994; Patterson *et al*. 1997). Regardless of whether excitation is at 397 nm or 475 nm, the fluorescence emission occurs from the anionic phenolate species (fluorescence λmax = 504 nm) with a quantum yield of 0.79 (Patterson *et al*. 1997).

The safety of GFP has been demonstrated in peer-reviewed literature as well as studies conducted by InnerPlant. Pure GFP and diets containing transgenic canola expressing GFP were fed to young male rats for 26 days to evaluate the potential toxicity and allergenicity of GFP (Richards *et al*., 2003). Ingestion of GFP did not affect growth, food intake, relative weight of intestine or other organs, or activities of hepatic enzymes in serum. It was concluded that GFP does not present a risk of toxicity. Further, the GFP amino acid sequence was analyzed for potential homologies to known protein toxins following the method described in Sharma *et al*. (2022). When the database, ToxinPred2, was searched using default parameters, no significant amino acid sequence homologies between GFP to known protein toxins were found.

The second gene transferred to the soybean genome [ ] Expression of this protein in plant cells [ ] serves as a selectable marker for plant transformation. The following describes the mechanism of action [ ]

[ ]

Figure 5. [ ]

[ ]

]

CBI-deleted

[CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
CBI-deleted
]

[ ] expressed in IFB soybeans is not CBI-deleted
expected to exhibit any enzymatic activity. Therefore, [ ] is not expected to have any effect CBI-deleted
on other soybean metabolic pathways. In addition, the [ ] amino acid sequence was analyzed CBI-deleted
for potential homologies to known protein toxins following the method described in Sharma *et al*. (2022). When the database, ToxinPred2, was searched using default parameters, no significant amino acid sequence homologies between [ ] and known protein toxins were found, CBI-deleted
supporting the safety of the [ ] protein. CBI-deleted

7 Proposed plant-trait-MOA language for website

Plant: Glycine max (soybean)
Trait: Fluorescent marker gene
Phenotype: green fluorescence
MOA: Stress inducible expression of a fluorescent biomarker for stress

Literature Cited

Aoki, H., Teramura, H., Schepetilnikov, M., and Ryabova, L.A. (2014) Enhanced translation of the downstream ORF attributed to a long 5^|^#8242; untranslated region in the OsMac1 gene family members, OsMac2 and OsMac3. Plant Biotechnology 31(3):221-228.

Bandara, A.Y., Weerasooriya, D.K., Bradley, C.A., Allen, T.W., and Esker, P.D. (2020a) Dissecting the economic impact of soybean diseases in the United States over two decades. PLoS ONE 15(4):e0231141. https://doi.org/10.1371/journal.pone.0231141

Bandara, A.Y., Weerasooriya, D.K., Conley, S.P., Bradley, C.A., Allen, T.W., and Esker, P.D. (2020b) Modeling the relationship between estimated fungicide use and disease-associated yield losses of soybean in the United States I: Foliar fungicides vs foliar diseases. PLoS One. 15(6):e0234390. https://doi.org/10.1371/journal.pone.0234390

[**CBI-deleted**
CBI-deleted
]
CBI-deleted

Baruah, I., Baldodiya, G.M., Sahu, J., Baruah, G. (2020) Dissecting the role of promoters of pathogen-sensitive genes in plant defense. Current Genomics 21:491-503.

Baumann, D., Cook, M., Ma, L., Mushegian, A., Sanders, E., Schwartz, J., and Yu, C.R. (2008) A family of GFP-like proteins with different spectral properties in lancelet Branchiostoma floridae. Biol Direct. 3:28.

Branchini, B.R., Nemser, A.R., and Zimmer, M. (1998) A computational analysis of the unique protein-induced tight turn that results in posttranslational chromophore formation in green fluorescent protein. J. Am. Chem. Soc. 120:1-6.

Bueno, T.V., Fontes, P.P., Abe, V.Y., Saito, A.U., Senra, R.L., Oliveira, L.S., Dos Santos, A.B., Capote Ferreira, E.G., Darben, L.M., de Oliveira, A.B., Abdelnoor, R.V., Whitham, S.A., Fietto, L.G., and Marcelino Guimarães, F.C. (2022) A *Phakopsora pachyrhizi* effector suppresses PAMP-triggered immunity and interacts with a soybean glucan endo-1,3- β -glucosidase to promote virulence. Molecular plant-microbe interactions: MPMI, 10.1094/MPMI-12-21-0301-R. Advance online publication. https://doi.org/10.1094/MPMI-12-21-0301-R

Cabre, L., Peyrard, S., Sirven, C., Gilles, L., Pelissier, B., Ducerf, S., and Poussereau, N. (2021) Identification and characterization of a new soybean promoter induced by *Phakopsora pachyrhizi*, the causal agent of Asian soybean rust. BMC Biotechnol. 21(1):27

Campbell, R.E. (2008) Fluorescent proteins. Scholarpedia, 3(7):5410.

CFIA (1996) The biology of *Glycine max* (L.) Merr. (soybean). Biology Document B101996-10. Canadian Food Inspection Agency, Ottawa, Ontario, Canada. http://www.inspection.gc.ca/english/plaveg/bio/dir/t11096e.shtml

Desikan, R., Clarke, A., Atherfold, P., Hancock, J.T., and Neill, S.J. (1999) Harpin induces mitogenactivated protein kinase activity during defense responses in *Arabidopsis thaliana* suspension cultures. Planta 210:97-103.

Erickson, E.H. (1975) Variability of floral characteristics influence honey bee visitation to soybean blossom. Crop Sci. 15:767-771.

Erickson, E.H. (1984) Soybean pollination and honey production: A research progress report. Am. Bee J. 145:775-779.

Feys, B.J., and Parker, J.E. (2000) Interplay of signaling pathways in plant disease resistance. Trends Genet. 16(10):449-455. http://dx.doi.org/10.1016/S0168-9525(00)02107-7 PMID: 11050331

Fu, Z.Q., and Dong, X. (2013) Systemic acquired resistance: turning local infection into global defense. Annual Review of Plant Biology, 64:839–863. https://doi.org/10.1146/annurev-arplant-042811-105606

Haun, W.L., Hyten, D.L., Xu, W.W., Gerhardt, D.J., Albert, T.J., Richmond, T., Jeddeloh, J.A., Jia, G., Springer, N.M., Vance, C.P., and Stupar, R.M. (2011) The composition and origins of genomic variation among individuals of the soybean reference cultivar Williams 82. Plant Physiol. 155(2) 645-655.

Heim, R., Prasher, D.C., and Tsien, R.Y. (1994) Wavelength mutations and posttranslational autoxidation of green fluorescent protein. Proc. Natl. Acad. Sci. U.S.A. 91:12501-12504.

Hermann, F.J. (1962) A revision of the genus *Glycine* and its immediate allies, USDA Tech. Bull. 1268:1-79.

Hily, J-M., Singer, S.D., Yang, Y., and Liu, Z. (2009) A transformation booster sequence (TBS) from *Petunia hybrida* functions as an enhancer-blocking insulator in Arabidopsis thaliana. Plant Cell Rep. 28(7):1095-104.

[CBI-deleted
CBI-deleted
]

Hymowitz, T. (2004) Speciation and cytogenetics. *In:* Soybeans: Improvement, Production, and Uses. 3rd edition, Boerma, H.R. and Specht, J.E. (Eds.), Agron. Monogr. 16. ASA, CSA, and SSSA Publishers, Madison, WI, USA., pp. 97-136.

Kay, R., Chan, A., Daly, M., and McPherson, J. (1987) Duplication of CaMV35S promoter sequences creates a strong enhancer for plant genes. Science 236(4806):1299-1302.

[CBI-deleted
CBI-deleted
]
CBI-deleted

Mazarei, M.; Teplova, I.; Hajimorad, M.R.; Stewart, C.N. (2008) Pathogen phytosensing: plants to report plant pathogens. Sensors (Basel) 8(4):2628-2641. http://dx.doi.org/10.3390/s8042628 PMID: 27879840

Mogen, B.D., MacDonald, M.H., Graybosh, R., and Hunt, A.G. (1990) Upstream sequences other than AAUAAA are required for efficient messenger RNA 3'-end formation in plants. The Plant Cell 2:1261-1272.

Morise, H., Shimomura, O., Johnson, F.H., and Winant, J. (1974) Intermolecular energy transfer in the bioluminescent system of Aequorea. Biochemistry 13:2656-2662.

Nagaya, S., Kawamura, K., Shinmyo, A., and Kato, K. (2010). The HSP terminator of *Arabidopsis thaliana* increases gene expression in plant cells. Plant and Cell Physiol. 51(2):328-332.

Nürnberger, T., Brunner, F., Kemmerling, B., and Piater, L. (2004) Innate immunity in plants and animals: striking similarities and obvious differences. Immunol. Rev. 198:249-266.

Odell, J.T., Nagy, F., and Chua, N.H. (1985). Identification of DNA sequences required for activity of the Cauliflower Mosaic Virus 35S promoter. Nature 313:810-812.

OECD (2000) Organization for Economic Cooperation and Development, Consensus document on the biology of *Glycine max* (L.) Merr. (Soybean). Series on Harmonization of Regulatory Oversight in Biotechnology No. 15. ENV/JM/MONO(2000)9. http://www.oecd.org.

Ormo, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y., and Remington, S.J. (1996) Crystal structure of the *Aequorea victoria* green fluorescent protein. Science 273:1392-1395.

Pape, S., Thurow, C., and Gatz, C. (2010) The Arabidopsis PR-1 promoter contains multiple integration sites for the coactivator NPR1 and the repressor SNI1. Plant physiology 154(4):1805–1818. https://doi.org/10.1104/pp.110.165563

Patterson, G.H., Knobel, S.M., Sharif, W.D., Kain, S.R., and Piston, D.W. (1997) Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. Biophys. J. 73:2782-2790.

Protein Database Bank (PDB) (2022) Molecular models: Exploring the structure of fluorescent proteins. PDB 101

Raper, C.D., and Kramer, P.J. (1987) Stress physiology. In: Soybeans: Improvement, Production, and Uses. 2nd ed., Wilcox, J.R. (Ed.), ASA, CSSA, SSSA, Madison, WI., pp 589-641.

Ray, J.D., Kilen, T.C., Abel, A.C., and Paris, R.L. (2003) Soybean natural cross-pollination rates under field conditions. Environ. Biosafety Res. 2:133-138.

Richards, H.A., Han, C-T., Hopkins, R.G., Failla, M.L., Ward, W.W., and Stewart, Jr. C.N. (2003) Safety assessment of recombinant green fluorescent protein orally administered to weaned rats. J. Nutr. 133:1909–1912.

Sharma, N., Naorem, L.D., Jain, S., and Raghava, G. (2022). ToxinPred2: an improved method for predicting toxicity of proteins. Briefings in bioinformatics, bbac174. Advance online publication. https://doi.org/10.1093/bib/bbac174

Shoemaker, R.C., Hatfield, P.M., Palmer, R.G., and Atherly, A.G. (1986) Chloroplast DNA variation in the genus *Glycine* subenus *Soja*. J. Hered. 77:26-30.

Singh, R.J., and Hymowitz, T. (1989) The genomic relationships among *Glycine soja* Sieb. and Zucc., *G. max* (L.) Merr. and '*G. gracilis*' Skvortz. Plant Breed. 103:171-173.

[

CBI-deleted

CBI-deleted

]

CBI-deleted

TeKrony, D.M., Egli, D.B., and White, G.M. (1987) Seed production and technology. In Soybeans: Improvement, Production, and Uses. 2nd ed. Wilcox, J.R. (Ed.), American Society of Agronomy. Madison, WI., pp. 295-354.

USDA. (2006) USDA federal noxious weed list. http://www.aphis.usda.gov/plant_health/plant_pest_info/weeds/downloads/weedlist2006.pdf.

USDA-APHIS. (2008) http://www.aphis.usda.gov/brs/soybean.html.

USDA-NASS. (2022) Acreage 2022 (June report). United States Department of Agriculture National Agricultural Statistics Service, Washington, D.C.

Wang, C.L. (1976) Review on the classification of soybeans. (In Chinese) Acta Phytotaxon. Sin. 14:22-30.

[**CBI-deleted**

CBI-deleted

]

CBI-deleted

Westrick, N.M., Ranjan, A., Jain, S., Grau, C.R., Smith, D.L., Kabbage, M. (2019) Gene regulation of *Sclerotinia sclerotiorum* during infection of *Glycine max*: on the road to pathogeness. BMC Genomics 20:157

Yang, F., Moss, L.G., and Phillips, G.N.J. (1996) The molecular structure of green fluorescent protein. Nat. Biotechnol. 14:1246-1251.

Zhang, L., Patel, H.N., Lappe, J.W., and Wachter, R.M. (2006) Reaction progress of chromophore biogenesis in green fluorescent protein. J. Am. Chem. Soc. 128:4766-4772.