InnerPlant, Inc.

Regulatory Status Review Request for the Determination of Nonregulated Status for Tomato 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

	Page
List of Figures	3
List of Tables	3
Abbreviations and Definitions	4
1. Confidential Business Information (CBI) Statement	5
2. Product Description and Rationale	5
3. Description of Comparator Plant	7
4. Genotype of the Modified Plant	9
A. Circular map of plasmid vector V148	10
B. Sequence of the T-DNA insert in ICM tomatoes	11
C. Annotation of the T-DNA inserted genetic material	13
5. Description of New Trait	16
A. Intended trait	16
B. Intended phenotype	16
C. Description of the Mechanism of Action (MOA)	16
6. Proposed plant-trait-MOA language for website	20
7. Literature Cited	21

List of Figures

Page

Figure 1. Circular map of plasmid vector V148 showing the genes and genetic elements components of the T-DNA of the plasmid. 10
Figure 2. Nucleotide sequence of the T-DNA insert in ICM tomatoes 11
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). 17
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 18
Figure 5. Chemical reaction catalyzed by NPTII 19

List of Tables

Page

Table 1. Genetic material inserted into the genome of ICM tomatoes 13

Abbreviations and Definitions

APHIS	Animal and Plant Health Inspection Service
CaMV	Cauliflower Mosaic Virus
cv.	Cultivar
FOIA	Freedom of Information Act
GE	Genetically Engineered
GFP	Green Fluorescent Protein
HSP	Heat Shock Protein
ICM	InnerPlant Constitutive Marker
LB	Left Border
NPTII	Neomycin phosphotransferase II
OECD	Organization for Economic Cooperation and Development
RB	Right Border
RSR	Regulatory Status Review
T-DNA	Transfer DNA
UTR	Untranslated Region

1 Confidential Business Information (CBI) Statement

This RSR request does not contain CBI.

2 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 InnerPlant biosensors 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 and emission wavelengths) to enable multiplexing of biosensor signals. An example of one of InnerPlants optical biosensor plants is soybean expressing a fluorescent marker protein in response to pathogen infection of the plant (Regulatory Status Review submission number RSR: 22-235-01rsr).

In the present application InnerPlant is requesting a Regulatory Status Review (RSR) of GE tomatoes that are designed to constitutively express a green fluorescent protein (GFP). Although not intended to be a stand-alone commercial product, these plants with constant fluorescent protein expression are an important ancillary to InnerPlants optical biosensor plants because they serve as

a critical tool to enable us to calibrate, refine and improve the design of our detection equipment in the field. Hereafter these tomatoes are referred to as InnerPlant Constitutive Marker tomatoes or ICM tomatoes.

The ICM tomatoes are being produced by *Agrobacterium tumefaciens*-mediated transformation of tomato tissues from non-transgenic cultivar (cv.) M82 using plasmid vector V148. The DNA transferred from the plasmid to the tomato genome (T-DNA) contains two gene constructs. The first gene is *lanFP1* encoding the Green Fluorescent Protein (GFP). ICM tomatoes produce GFP constitutively and this results in production of an optical fluorescence signal that can be detected in the field. The strong, constitutive expression of the GFP protein enables the development, refinement, and calibration of our detection equipment.

The second gene transferred to the tomato genome encodes neomycin phosphotransferase II enzyme (NPTII). Expression of this protein in plant cells confers resistance to antibiotics neomycin and kanamycin and serves as a selectable marker for plant transformation. The mechanism of action and safety of the GFP and NPTII proteins are reviewed in this RSR application.

3 Description of Comparator Plant

The biology of tomato described herein is based upon the consensus document for *Solanum lycopersicum* (L.) prepared by the Organization for Economic Co-operation and Development (OECD, 2016), and recent literature on the topic.

The cultivated tomato, *Solanum lycopersicum* (L.) is classified taxonomically as follows:

Kingdom	Plantae -- Plants
Subkingdom	Viridiplantae
Division	Tracheophyta
Class	Magnoliopsida
Superorder	Asteranae
Order	Solanales
Family	Solanaceae
Genus	Solanum
Species	*Solanum lycopersicum* (L.) – garden tomato

The above taxonomic information for tomato was obtained from the Integrated Taxonomic Information System (http://www.itis.gov/) and tomato is assigned the taxonomic serial number 521671.

The cultivated tomato is a member of the genus Solanum within the family Solanaceae. The Solanaceae, commonly known as the nightshade family, also includes other notable cultivated plants such as tobacco, chilli pepper, potato and eggplant. The genus Solanum consists of approximately 1,500 species. The tomato clade includes the cultivated tomato (*Solanum lycopersicum*) and 12 wild relatives, all natives to western South America. *Solanum lycopersicum* is derived from two wild ancestor species, *Solanum pimpinellifolium* and *Solanum cerasiforme*. Other wild species are useful for breeding disease resistance, color improvement and desirable quality traits (Ranc *et al*., 2008).

Tomato has a relatively small genome size (around 950 Mb). About 30% of the genome is composed of repetitive sequences which are mainly located in heterochromatin regions (Van der Hoeven *et al*., 2002). Tomato and its wild relatives have 12 chromosomes (2n=2x=24).

Although some tomato wild species of the genus Solanum are allogamous, all commercial tomato cultivars are considered to be mainly self-compatible and inbreeding, i.e. autogamous (Rick, 1979; Taylor, 1986). As is the case for most self-pollinating plants, the viability of exposed tomato pollen is limited. Pollen viability and the number of pollen grains are reduced by high temperatures above 32/26°C day/night. Natural cross-pollination rates among commercial varieties range from 0.07% to 12% (OECD, 2016). The rate of crossing quickly decreases as the distance from the pollen source increases (Currence and Jenkins, 1942) and little viable pollen is transferred beyond 30 m (~95 feet) from its source (Quiros and Marcias, 1978).

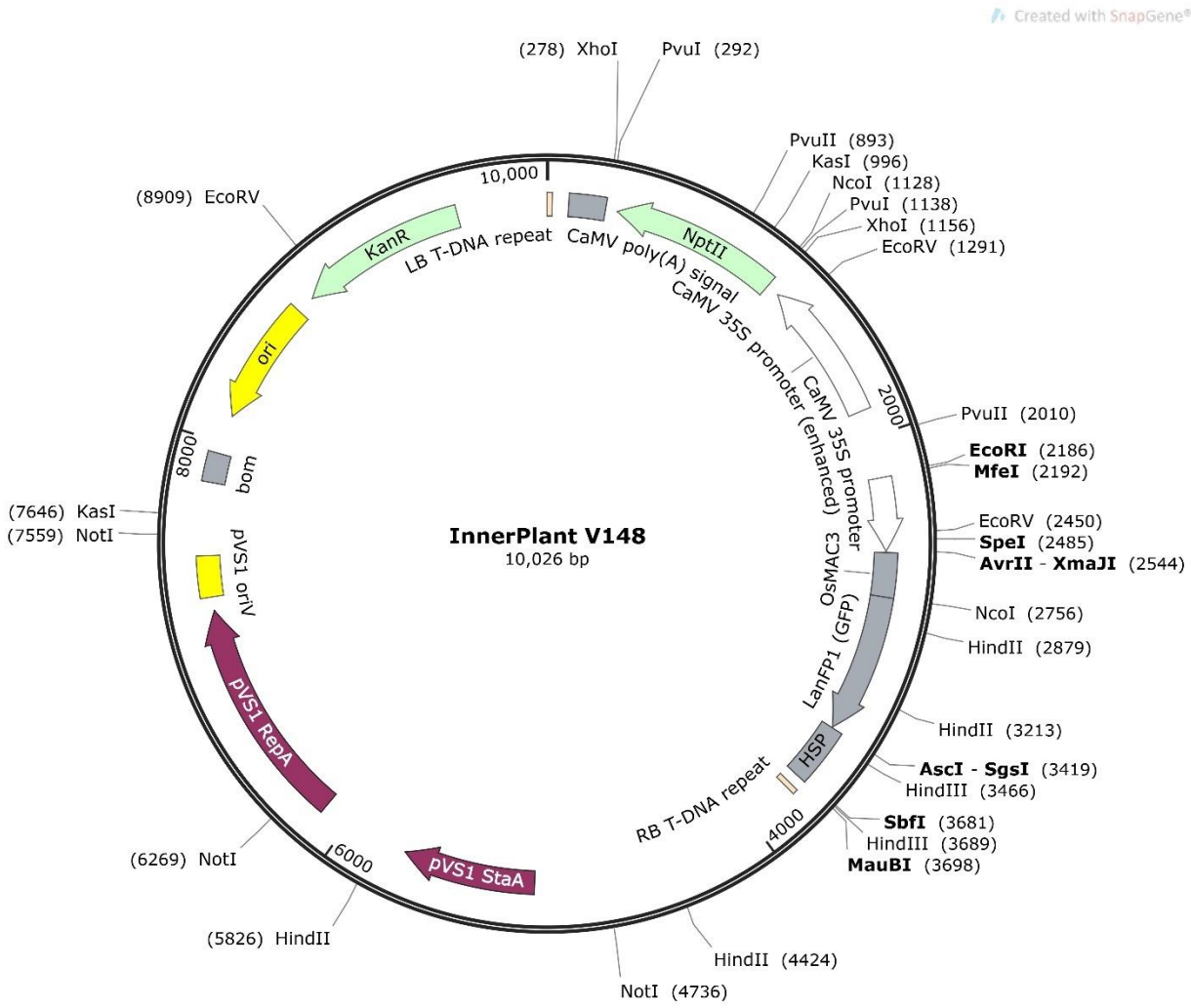
The recipient tomato cultivar M82 (Accession LA3475) is a processing tomato and was released commercially in the mid 1970s (Barrios-Masias *et al*., 2014).

4 Genotype of the Modified Plant

As presented above, ICM tomatoes are being produced by *Agrobacterium tumefaciens*-mediated transformation of tomato tissues from non-transgenic cultivar (cv.) M82 using plasmid vector V148. A circular map of plasmid vector V148 is presented in Figure 1. The nucleotide sequence of the DNA transferred from the plasmid to the tomato genome (T-DNA) is presented below in Figure 2. The order of the genetic components in the T-DNA of V148, 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.

A Circular map of plasmid vector V148

Figure 1. Circular map of plasmid vector V148 showing the genes and genetic elements components of the T-DNA of the plasmid.



The DNA transferred from the plasmid vector V148 to the tomato genome (T-DNA) contains two gene constructs. The first gene is *lanFP1* encoding the Green Fluorescent Protein (GFP). The second gene transferred to the tomato genome encodes the neomycin phosphotransferase II enzyme (NPTII). Expression of NPTII in plant cells confers resistance to the antibiotics neomycin and kanamycin and serves as a selectable marker for plant transformation.

B Sequence of the T-DNA insert in ICM tomatoes

Figure 2. Nucleotide sequence of the T-DNA insert in ICM tomatoes

```
   1 TGGCAGGATA TATTGTGGTG TAAACAAATT GACGCTTAGA CAACTTAATA ACACATTGCG
  61 GACGTTTTTA ATGTACTGAA TTAACGCCGA ATTAATTCGG GGGATCTGGA TTTTAGTACT
 121 GGATTTTGGT TTTAGGAATT AGAAATTTTA TTGATAGAAG TATTTTACAA ATACAAATAC
 181 ATACTAAGGG TTTCTTATAT GCTCAACACA TGAGCGAAAC CCTATAGGAA CCCTAATTCC
 241 CTTATCTGGG AACTACTCAC ACATTATTAT GGAGAAACTC GAGCTTGTCG ATCGACTCTA
 301 GCTAGAGGAT CGATCCGAAC CCCAGAGTCC CGCTCAGAAG AACTCGTCAA GAAGGCGATA
 361 GAAGGCGATG CGCTGCGAAT CGGGAGCGGC GATACCGTAA AGCACGAGGA AGCGGTCAGC
 421 CCATTCGCCG CCAAGCTCTT CAGCAATATC ACGGGTAGCC AACGCTATGT CCTGATAGCG
 481 GTCCGCCACA CCCAGCCGGC CACAGTCGAT GAATCCAGAA AAGCGGCCAT TTTCCACCAT
 541 GATATTCGGC AAGCAGGCAT CGCCATGTGT CACGACGAGA TCCTCGCCGT CGGGCATGCG
 601 CGCCTTGAGC CTGGCGAACA GTTCGGCTGG CGCGAGCCCC TGATGCTCTT CGTCCAGATC
 661 ATCCTGATCG ACAAGACCGG CTTCCATCCG AGTACGTGCT CGCTCGATGC GATGTTTCGC
 721 TTGGTGGTCG AATGGGCAGG TAGCCGGATC AAGCGTATGC AGCCGCCGCA TTGCATCAGC
 781 CATGATGGAT ACTTTCTCGG CAGGAGCAAG GTGAGATGAC AGGAGATCCT GCCCCGGCAC
 841 TTCGCCCAAT AGCAGCCAGT CCCTTCCCGC TTCAGTGACA ACGTCGAGCA CAGCTGCGCA
 901 AGGAACGCCC GTCGTGGCCA GCCACGATAG CCGCGCTGCC TCGTCCTGGA GTTCATTCAG
 961 GGCACCGGAC AGGTCGGTCT TGACAAAAAG AACCGGGCGC CCCTGCGCTG ACAGCCGGAA
1021 CACGGCGGCA TCAGAGCAGC CGATTGTCTG TTGTGCCCAG TCATAGCCGA ATAGCCTCTC
1081 CACCCAAGCG GCCGGAGAAC CTGCGTGCAA TCCATCTTGT TCAATCCCCA TGGTCGATCG
1141 ACAGATCTGC GAAAGCTCGA GAGAGATAGA TTTGTAGAGA GAGACTGGTG ATTTCAGCGT
1201 GTCCTCTCCA AATGAAATGA ACTTCCTTAT ATAGAGGAAG GTCTTGCGAA GGATAGTGGG
1261 ATTGTGCGTC ATCCCTTACG TCAGTGGAGA TATCACATCA ATCCACTTGC TTTGAAGACG
1321 TGGTTGGAAC GTCTTCTTTT TCCACGATGC TCCTCGTGGG TGGGGGTCCA TCTTTGGGAC
1381 CACTGTCGGC AGAGGCATCT TGAACGATAG CCTTTCCTTT ATCGCAATGA TGGCATTTGT
1441 AGGTGCCACC TTCCTTTTCT ACTGTCCTTT TGATGAAGTG ACAGATAGCT GGGCAATGGA
1501 ATCCGAGGAG GTTTCCCGAT ATTACCCTTT GTTGAAAAGT CTCAATAGCC CTTTGGTCTT
1561 CTGAGACTGT ATCTTTGATA TTCTTGGAGT AGACGAGAGT GTCGTGCTCC ACCATGTTAT
1621 CACATCAATC CACTTGCTTT GAAGACGTGG TTGGAACGTC TTCTTTTTCC ACGATGCTCC
1681 TCGTGGGTGG GGGTCCATCT TTGGGACCAC TGTCGGCAGA GGCATCTTGA ACGATAGCCT
1741 TTCCTTTATC GCAATGATGG CATTTGTAGG TGCCACCTTC CTTTTCTACT GTCCTTTTGA
```

1801 TGAAGTGACA GATAGCTGGG CAATGGAATC CGAGGAGGTT TCCCGATATT ACCCTTTGTT
1861 GAAAAGTCTC AATAGCCCTT TGGTCTTCTG AGACTGTATC TTTGATATTC TTGGAGTAGA
1921 CGAGAGTGTC GTGCTCCACC ATGTTGGCAA GCTGCTCTAG CCAATACGCA AACCGCCTCT
1981 CCCCGCGCGT TGGCCGATTC ATTAATGCAG CTGGCACGAC AGGTTTCCCG ACTGGAAAGC
2041 GGGCAGTGAG CGCAACGCAA TTAATGTGAG TTAGCTCACT CATTAGGCAC CCCAGGCTTT
2101 ACACTTTATG CTTCCGGCTC GTATGTTGTG TGGAATTGTG AGCGGATAAC AATTTCACAC
2161 AGGAAACAGC TATGACCATG ATTACGAATT CCAATTGTGA GACTTTTCAA CAAAGGGTAA
2221 TATCCGGAAA CCTCCTCGGA TTCCATTGCC CAGCTATCTG TCACTTTATT GTGAAGATAG
2281 TGGAAAAGGA AGGTGGCTCC TACAAATGCC ATCATTGCGA TAAAGGAAAG GCCATCGTTG
2341 AAGATGCCTC TGCCGACAGT GGTCCCAAAG ATGGACCCCC ACCCACGAGG AGCATCGTGG
2401 AAAAAGAAGA CGTTCCAACC ACGTCTTCAA AGCAAGTGGA TTGATGTGAT ATCTCCACTG
2461 ACGTAAGGGA TGACGCACAA TCCCACTAGT CTTCGCAAGA CCCTTCCTCT ATATAAGGAA
2521 GTTCATTTCA TTTGGAGAGA ACACCTAGGC GGCGATCCAC AGGGAAGGAG CAGCATCTCC
2581 ACAAAGACGC ACTACAGAAG ACTAAAGAGA GCTTTTTCAT ACCAAAGAAG TACAACAAAA
2641 GATTTGCTCC TCATTTTCTG AATCCTGGGA CTCTCTAGCC TGTAGAAGAA GAAAGGCAGG
2701 AATTTCAGCT CAAGAGAACA GATCACAATA TTTACCCACG GCACTGTCTC GCAATCCATG
2761 GCATTGCCCG CCACCCATGA CATTCATCTG CACGGTTCCA TAAATGGCCA CGAGTTCGAT
2821 ATGGTCGGAG GGGGGAAGGG AGATCCTAAC GCAGGCTCAC TGGTAACAAC AGCAAAGTCA
2881 ACTAAAGGCG CTCTGAAGTT CTCACCTTAC TTGATGATAC CCCACCTTGG ATACGGGTAT
2941 TATCAATATC TTCCATATCC CGACGGACCC AGCCCTTTCC AAACCTCTAT GCTTGAAGGC
3001 AGTGGGTATG CTGTGTATCG CGTCTTTGAC TTTGAGGACG GGGGAAAGCT CACAACAGAA
3061 TTTAAGTATT CATACGAAGG CTCACACATA AAGGCTGACA TGAAGTTGAT GGGGAGTGGA
3121 TTCCCAGACG ATGGTCCAGT GATGACTAGC CAGATCGTGG ACCAGGACGG CTGCGTGAGC
3181 AAAAAGACCT ATCTCAATAA CAATACAATA GTTGACAGTT TTGACTGGTC ATATAACTTG
3241 CAAAACGGAA AAAGATACCG TGCTCGTGTC AGTTCACATT ACATCTTTGA CAAGCCATTT
3301 AGTGCTGATC TCATGAAAAA ACAGCCCGTA TTTGTCTACC GCAAATGCCA TGTAAAGGCT
3361 AGTAAGACAG AGGTTACCTT GGACGAACGT GAGAAGGCAT TCTACGAATT GGCATGAGGC
3421 GCGCCATATG AAGATGAAGA TGAAATATTT GGTGTGTCAA ATAAAAAGCT TGTGTGCTTA
3481 AGTTTGTGTT TTTTTCTTGG CTTGTTGTGT TATGAATTTG TGGCTTTTTC TAATATTAAA
3541 TGAATGTAAG ATCTCATTAT AATGAATAAA CAAATGTTTC TATAATCCAT TGTGAATGTT
3601 TTGTTGGATC TCTTCTGCAG CATATAACTA CTGTATGTGC TATGGTATGG ACTATGGAAT
3661 ATGATTAAAG ATAAGCCTGC AGGCATGCAA GCTTGGCGCG CGCGGTGTCA TCTATGTTAC
3721 TAGATCGGGA ATTAAACTAT CAGTGTTTGA CAGGATATAT TGGCGGGTAA AC

C Annotation of the T-DNA inserted genetic material

Table 1. Genetic material inserted into the genome of ICM tomatoes

Genetic Element	Position in the genetic insertion	Donor Organism	Function	GenBank No.
T-DNA Left Border (LB) region	1-25	*Agrobacterium tumefaciens*	DNA sequence derived from *Agrobacterium* containing the left border (LB) sequence from the C58 Ti plasmid for the efficient transfer of the T-DNA from *Agrobacterium tumefaciens* to the plant genome (Barker *et al*., 1983).	GenBank: AJ237588.1 (3623..3647)
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) of the 35S genes of cauliflower mosaic virus (CaMV) (Mogen *et al*., 1990) that directs polyadenylation in plant cells and terminates transcription of the *nptII* gene inserted in the ICM tomato genome.	GenBank: KY703615.1 (7460..7634)
Intervening sequence	278-333	Not applicable. Synthetic sequence, polylinker	Sequence used for DNA cloning.	
nptII	334-1131	*Escherichia coli*	Coding sequence from Tn5 (Beck *et al*., 1982) in *E. coli* encoding for the neomycin phosphotransferase II enzyme (NPTII). Expression of this protein in plant cells confers resistance to the antibiotics neomycin and kanamycin and serves as a selectable marker for plant transformation (Fraley *et al*., 1983).	GenBank: CP047128.1 (35420..36212)

Table 1 continued. Genetic material inserted into the ICM tomato genome

Genetic Element	Position in the genetic insertion	Donor Organism	Function	GenBank No.
Intervening sequence	1132-1193	Not applicable. Synthetic sequence, polylinker	Sequence used for DNA cloning.	
Enhanced CaMV 35S promoter	1194-1871	Cauliflower Mosaic Virus (CaMV)	The promoter for the 35S genes from the cauliflower mosaic virus (Odell *et al*., 1985) with the duplicated enhancer region (Kay *et al*., 1987) that drives transcription of the *nptII* gene.	GenBank: V00140.1 (7016..7441)
Intervening sequence	1872-2197	Not applicable. Synthetic sequence, polylinker	Sequence used for DNA cloning.	
CaMV 35S promoter	2198-2543	Cauliflower Mosaic Virus (CaMV)	The promoter for the 35S genes from the cauliflower mosaic virus (Odell *et al*., 1985) that drives transcription of the *lanFP1* gene.	GenBank: NC_001497.2 (7093..7438)
Intervening sequence	2544-2549	Not applicable. Synthetic sequence, polylinker	Sequence used for DNA cloning.	
OsMac3 5'UTR	2550-2757	*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)

Table 1 continued. Genetic material inserted into the tomato genome

Genetic Element	Position in the genetic insertion	Donor Organism	Function	GenBank No.
lanFP1: Coding sequence for the Green Fluorescent Protein (GFP) from Lancelet	2758-3417	*Branchiostoma floridae*	Synthesis of the GFP is driven by a constitutive promoter 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	3417-3421	Not applicable. Synthetic sequence, polylinker	Sequence used for DNA cloning.	
AtHSP terminator	3422-3698	*Arabidopsis thaliana*	The 3'UTR DNA sequence of the terminator for heat shock protein HSP18.2, that terminates transcription of the *lanFP1* gene (Nagaya *et al*., 2010).	GenBank: NC_003076.8 (24063118..24063367)
Intervening sequence	3699-3747	Not applicable. Synthetic sequence, polylinker	Sequence used for DNA cloning.	
T-DNA Right Border (RB) region	3748-3772	*Agrobacterium tumefaciens*	DNA sequence derived from *Agrobacterium* containing the right border (RB) sequence from the C58 Ti plasmid for the efficient transfer of the T-DNA from *Agrobacterium tumefaciens* to the plant genome (Wang *et al*., 1984)	GenBank: NZ_CP058528.1 (186813..186837)

5 Description of New Trait

A Intended trait

GFP:
The ICM tomato contains a fluorescence protein that produces a fluorescence signal. The fluorescence trait is only visible when the protein absorbs the correct excitation light and optical filters are used to visualize the fluorescent light.

NPTII:
The tomato contains a protein that confers resistance to the antibiotic neomycin and serves as a selectable marker for plant transformation.

B Intended phenotype

GFP:
The phenotype of ICM tomato is green fluorescence with an emission peak at approximately 509 nm.

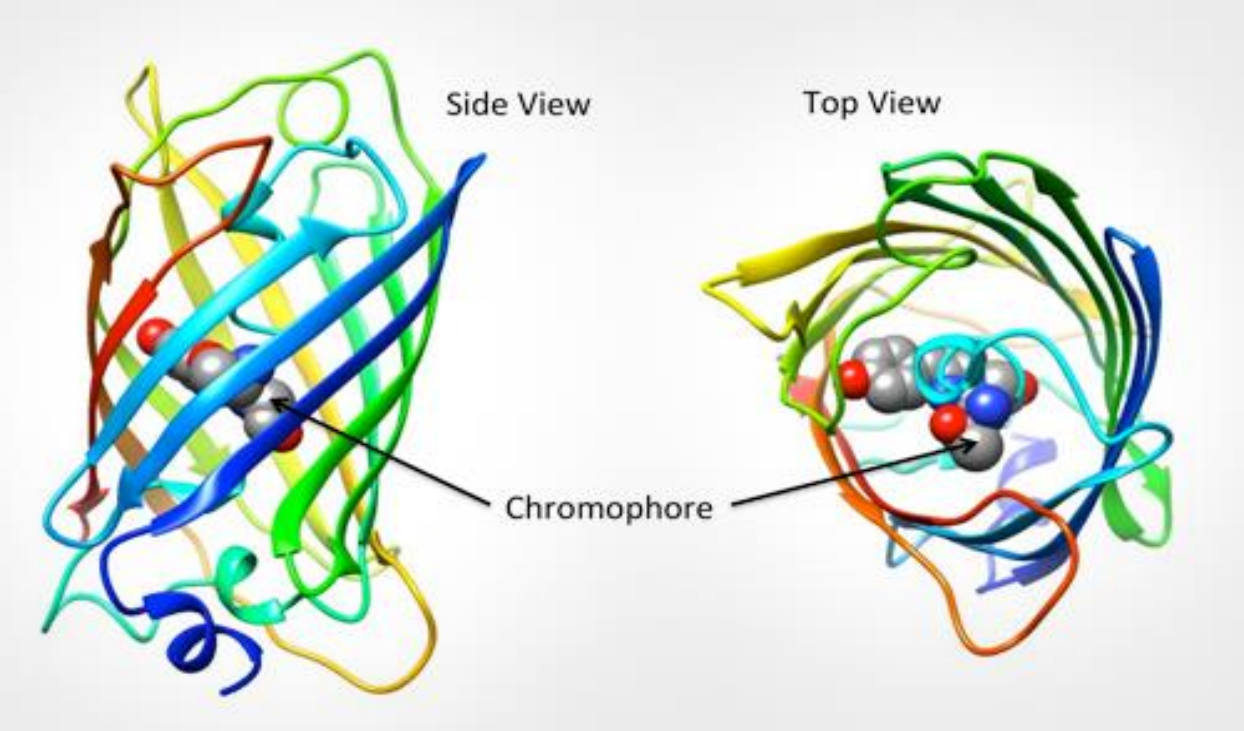
NPTII:
ICM tomatoes are resistant to the antibiotic neomycin and the NPTII protein serves as a selectable marker for plant transformation.

C Description of the Mechanism of Action (MOA)

GFP:
The following describes the mechanism of action of GFP.

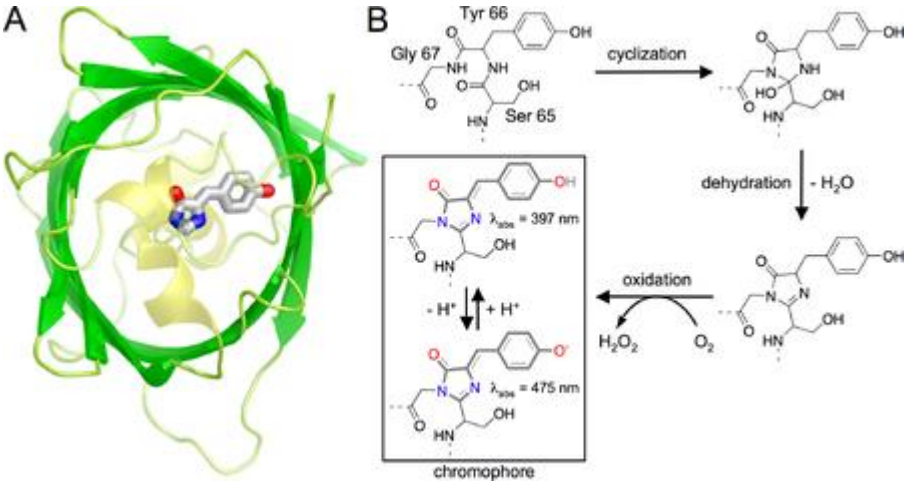
Two independent reports of the x-ray crystal structure of GFP (Ormo *et al*., 1996; Yang *et al*., 1996) revealed that the protein has a unique overall fold comprised of an 11-stranded β-sheet wrapped into a cylindrical β-barrel protein that is 42 amino acids in height and 24 amino acids in diameter (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).



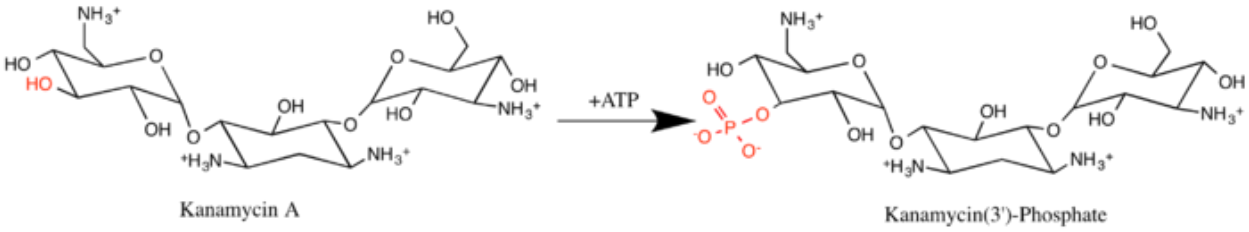
The GFP chromophore exists as an equilibrating mixture of the neutral phenol (absorbance λmax = 397 nm) and anionic phenolate (absorbance λmax = 475 nm) (Morise *et al*. 1974; Heim *et al*. 1994; Patterson *et al*. 1997). Regardless of whether excitation is at 397 nm or 475 nm, the fluorescence emission 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 tomato genome encodes neomycin phosphotransferase II (NPTII). Expression of this protein in plant cells confers resistance to the antibiotics neomycin and kanamycin and serves as a selectable marker for plant transformation. The following describes the mechanism of action of NPTII.

The antibiotics neomycin and kanamycin, bind to the negatively charged backbone of nucleic acids to disrupt protein synthesis, and therefore inhibits bacterial cell growth (Cavallo and Martinetto, 1981). Neomycin phosphotransferase II catalyzes the addition of phosphate from ATP to the 3'-hydroxyl group of the 4,6-disubstituted aminoglycoside neomycin (Figure 5). The NPTII mediated phosphorylation of neomycin/kanamycin introduces a phosphate group on the antibiotic that reduces the binding affinity to nucleic acids due to steric hindrances and unfavorable electrostatic interactions, and thereby disrupts the mechanism of action of the antibiotic (Wright and Thompson, 1999).

Figure 5. Chemical reaction catalyzed by NPTII. Reproduced from Wright and Thompson (1999).



In the absence of neomycin or kanamycin the NPTII expressed in ICM tomatoes is not expected to exhibit any enzymatic activity. Therefore, NPTII is not expected to have any effect on other tomato metabolic pathways. Furthermore, the food, feed and environmental safety of NPTII has been well established. Neomycin phosphotransferase II has been used as a selectable marker in many different commercial genetically-engineered (GE) crops {e.g. Genuity® DroughtGard™ corn (MON 87460), YieldGard® Rootworm corn (MON 863), Bollgard® cotton (MON 531), Bollgard®II cotton (MON 15985), Roundup Ready® cotton (MON 1445)}, and therefore has a history of safe use in the environment as well as in food and feed uses. Furthermore, the NPTII protein has been fully characterized, and the NPTII protein expressed in GE crops has been shown not to pose any discernable environmental, food or feed safety concerns (Fuchs *et al*., 1993a and 1993b; Nap *et al*., 1992; Flavell *et al*, 1992).

7 Proposed plant-trait-MOA language for website

Plant: *Solanum lycopersicum* (tomato)
Trait: Fluorescent marker gene
Phenotype: green fluorescence
MOA: Expression of a fluorescent biomarker

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.

Barker, R. F., Idler, K. B., Thompson, D. V., and Kemp, J. D. 1983. Nucleotide sequence of the T-DNA region from the *Agrobacterium tumefaciens* octopine Ti plasmid pTi15955. Plant Mol. Biol. 2:335-350.

Barrios-Masias, F.H., and Jackson, L.E. (2014) California processing tomatoes: Morphological, physiological and phenological traits associated with crop improvement during the last 80 years. Europ. J. Agronomy. 53:45-55.

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.

Beck, E., G. Ludwig, E.A. Auerswald, B. Reiss and H. Schaller. (1982) Nucleotide sequence and exact localization of the neomycin phosphotransferase gene from transposon Tn5. Gene. 19:327-336.

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.

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

Cavallo G, and Martinetto P. (1981) Meccanismo d'azione degli aminoglicosidi (The mechanism of action of aminoglycosides). G Batteriol Virol Immunol. 74(7-12):335-46. Italian.

Currence, T.M. and Jenkins J.M. (1942) Natural crossing in tomatoes as related to distance and direction. Proceedings of the American Society of Horticultural Sciences. 41:273-276.

Flavell, R.B., Dart, E., Fuchs, R.L. and Fraley, R.T. (1992) Selectable marker genes: Safe for plants? Nature Biotechnology 10: 141-144.

Fuchs, R.L., Heeren, R.A., Gustafson, M.E., Rogan, G.J., Bartnicki, D.E., Leimgruber, R.M., Finn, R.F., Hershman, A. and Berberich, S.A. (1993a) Purification and characterization of microbially

expressed neomycin phosphotransferase II (NPTII) protein and its equivalence to the plant expressed protein. Nature Biotechnology 11: 1537-1542.

Fuchs, R.L., Ream, J.E., Hammond, B.G., Naylor, M.W., Leimgruber, R.M. and Berberich, S.A. (1993b) Safety assessment of the neomycin phosphotransferase II (NPTII) protein. Nature Biotechnology 11: 1543-1547.

Fraley, R.T., Rogers, S.G., Horsch, R.B., Sanders, P.R., Flick, J.S., Adams, S.P., Bittner, M.L., Brand, L.A., Fink, C.L., Fry, J.S., Galluppi, G.R., Goldberg, S.B., Hoffmann, N.L., and Woo, S.C. (1983) Expression of bacterial genes in plant cells. Proceedings of the National Academy of Sciences USA. 80:4803-4807.

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.

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.

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.

Nap, J-P., Bijvoet, J. and Stiekema, W.J. (1992) Biosafety of kanamycin-resistant transgenic plants. Transgenic Research 1: 239-249.

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 (2016) Organization for Economic Cooperation and Development, Consensus document on the biology of tomato. Series on Harmonization of Regulatory Oversight in Biotechnology No. 63. ENV/JM/MONO(2016)53. 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.

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

Quiros, C.F. and Marcias, A. (1978) Natural cross pollination and pollinator bees of the tomato in Celaya, Central Mexico. HortScience 13:290-291.

Ranc, N., Muños, S., Santoni, S., and Causse, M. (2008) A clarified position for Solanum lycopersicon var. cerasiforme in the evolutionary history of tomatoes (Solanaceae), Biomedcentral Plant Biology, Vol. 8, pp. 130.

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.

Rick, C.M. (1979) Biosystematic studies in Lycopersicon and closely related species of Solanum. In: The Biology and Taxonomy of the Solanaceae. Hawkes, J., Lester, R., and Skelding, A.(Eds.). Academic Press, New York, pp. 667-697.

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

Taylor, I.B. (1986) Biosystematics of the tomato. In: The Tomato Crop. A Scientific Basis for Improvement. Atherton, J.G. and Rudich J. (Eds.). Chapman and Hall, New York, pp. 1-34.

Van der Hoeven, R., Ronning, C., Giovannoni, J., Martin G., and Tanksley, S. (2002) Deductions about the number, organization, and evolution of genes in the tomato genome based on analysis of a large expressed sequence tag collection and selective genomic sequencing. Plant Cell, 14:1441-56.

Wang, K., Herrera-Estrella, L., Van Montagu, M., and Zambryski, P. (1984) Right 25 bp terminus sequence of the nopaline T-DNA is essential for and determines direction of DNA transfer from Agrobacterium to the plant genome. Cell 38(2):455-462.

Wright, G.D, and Thompson, P.R. (1999) Aminoglycoside phosphotransferases: proteins, structure, and mechanism. Front Biosci. 4:1-3.

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.