RECEIVED
By hdjohnson for BRS Document Control Officer at 9:34 am, Jun 17, 2020

NO CBI

16 June, 2020

Bernadette Juarez
APHIS Deputy Administrator
Biotechnology Regulatory Services
4700 River Rd, Unit 98
Riverdale, MD 20737

Re: Confirmation of Regulatory Status of CRISPR-Cas9 Soybean with Increased Oil and Protein Content.

Dear Ms. Juarez,

Corteva Agriscience respectfully requests confirmation from the USDA-APHIS's Biotechnology Regulatory Services (BRS) regarding the regulatory status of a soybean line developed using CRISPR-Cas9 gene editing technology to introduce targeted base pair substitutions (edits) in the diacylglycerol acyltransferase *(DGAT1B)* gene which is designed to result in soybeans with increased oil and protein content. As described below, we do not consider the final soybean line to be "regulated article" subject to APHIS oversight under 7 C.F.R. Part 340[1] with respect to introduction into and/or through the U.S. because it does not contain any inserted genetic material from (i) a donor organism, recipient organism, or vector or vector agent listed in Section 340.2 and meeting the definition of "plant pest," or (ii) an unclassified organism and/or an organism whose classification is unknown. Nor is there a basis to believe that this soybean line is or will become a plant pest within the meaning of the Plant Protection Act. Moreover, the amino acid substitution (edits) in the *DGAT1B* gene could have been achieved using conventional breeding techniques and therefore, the final soybean line is indistinguishable from soybean plants that could have resulted from native genome variability or be developed in a conventional breeding program.

The diacylglycerol acyltransferase *(DGAT1)* gene encodes the type 1 diacylglycerol acyltransferases which catalyzes the final step in oil (triacylglycerol) production and is found in many plant species, including soybean (Roesler et al., 2016). Various strategies have been devised to increase the seed oil content via *DGAT1* gene, one of which relies on improving

[1] On May 18th, 2020, USDA revised its regulations at 7 C.F.R. part 340 with an effective date of August 17, 2020 (85 FR 29790-29838, Docket No. APHIS–2018–0034), with the exception of permitting and RSR processes under Sections 340.4 and 340.5, which are effective April 5, 2021. Unless otherwise indicated, the references to "Part 340" herein are as of June 16, 2020.

kinetic properties of the DGAT1 enzyme. By analyzing the existing variations in the DGAT1 protein sequences across various plant species we established that the presence of certain amino acid residues at position 479 of soybean DGAT1B protein (for example, Serine (S)) contributed to increased oil and protein content in yeast and tobacco leaf transient assays. The soybean DGAT1B protein has Isoleucine (I) at that position. The amino acid substitution I479S in the soybean DGAT1B protein is expected to similarly result in plants with increased oil and protein content.

CRISPR-Cas9 gene editing was used to introduce the targeted base pair substitutions (edits) that lead to the I479S amino acid substitution using the "gene edit" (SDN2) technique (Podevin et al., 2013). The DNA repair template consisted of the sequence homologous to the targeted sequence near the double strand break (DSB) except for the specific base pair substitutions and flanked on each side by homology arms (sequences identical to the sequences immediately upstream and downstream of the genomic sequence to be edited). The I479S substitution was accomplished by changing three base pairs to change the Isoleucine coding triplet (ATA) to the Serine coding triplet (TCT). Additionally, one more base pair substitution (T to C) was introduced into the DNA repair template sequence. This is a typical approach to protect the DNA repair template from being cut by Cas9 as its sequence would otherwise be identical to the sgRNA target sequence[2]. This additional base pair substitution was intentionally designed to result in a silent mutation – i.e., a mutation that does not cause the amino acid substitution. Homology arms induced the homology-directed recombination (HDR) DNA repair and resulted in repairing the DSB in a desired manner. The figure below depicts targeted base pair substitutions made to the coding sequence of the soybean *DGAT1B* gene.

ATATTCAGTATCCTTGGTCAACCTATGTGTGTACTGCTATACTA
TCTTTCAGTATCCTTGGTCAACCTATGTGCGTACTGCTATACTA

Figure. Targeted base pair substitutions in the protein coding sequence of soybean DGAT1B gene. ATA to TCT change results in changing Isoleucine to Serine. T to C (in red font) is a silent mutation to protect the DNA repair template from cutting. The Protospacer Adjacent Motif (PAM) sequence, immediately following the DNA sequence targeted by the Cas9 nuclease, is shown in box.

Additional Information on the Edited Lines:

Taxonomic description of organism
Soybean, *Glycine max,* L.

Description of intended phenotype(s)
Plants with increased oil and protein content.

Description of intended activity (movement or release)

[2] https://horizondiscovery.com/en/applications/crispr-cas9/homology-directed-repair-with-a-dna-donor-oligo

Field trials to evaluate the efficacy of introduced mutations.

Description of intended genetic change in final product
Targeted base pair substitutions resulting in one amino acid substitution, as described above.

Description of vector or vector agent
Ochrobactrum haywardense strain H1-8 mediated transformation with one plasmid (patent application US 2018/0216123 Al).

Description of construct
Elements of the CRISPR-Cas9 technology are indicated by bold font.

Plasmid		
Genetic Element	**Origin**	**Function**
PUC ORI	*Escherichia coli*	Origin of replication to facilitate plasmid propagation in bacterial cells.
OVERDRIVE	*Escherichia coli*	tDNA transfer enhancer.
RB (OCTOPINE)	*Agrobacterium tumefaciens*	Right border sequence.
SPACER2	Maize *(Zea mays)*	Sequence based on 19kDa B1 zein terminator used as a buffer sequence to insulate expression cassettes from potential degradation.
MINI-ALLSTOPS3	synthetic	Synthetic sequence of stop codons designed to stop downstream translation of open reading frames.
PSA2	synthetic	A synthetic sequence designed to facilitate PCR analysis of recombined FRT sites.
ATTB4	*Escherichia coli*	Gateway™ cloning system recombination site to facilitate vector construction process.
GM-U6-13.1 PRO-V1	Soybean *(Glycine max)*	Variant of promoter from U6 small nuclear RNA (snRNA) polymerase III gene.
GM-DGAT1B-CR69	Soybean *(Glycine max)*	Sequence fragment complementary to the Cas9 endonuclease recognition site in *DGAT1B* gene.
GUIDE RNA	*Streptococcus pyogenes*	Encodes crRNA-tracrRNA fusion transcript that binds and directs Cas9 endonuclease to the target site. GM-DGAT1B-CR69 and GUIDE RNA together constitute chimeric guide RNA.
GM-U6 TERM	Soybean *(Glycine max)*	Terminator from U6 small nuclear RNA gene.
PSN2	synthetic	A synthetic sequence designed to facilitate PCR analysis of recombined FRT sites.
ATTB1	*Escherichia coli*	Gateway™ cloning system recombination site to facilitate vector construction process.
GM-EF1A2 PRO-V1	Soybean *(Glycine max)*	Variant of promoter from the eEF-1a elongation factor gene.
GM-EF1A2 5UTR	Soybean *(Glycine max)*	5' untranslated region from the eEF-1a elongation factor gene.
GM-EF1A2 INTRON1-V1	Soybean *(Glycine max)*	Variant of intron from the eEF-1a elongation factor gene.

GM-EF1A2 5UTR (2)-V1	Soybean *(Glycine max)*	Variant of 5' untranslated region from the eEF-1a elongation factor gene.
AT-NLS (CO)	*Arabidopsis thaliana*	Nuclear localization signal, which directs a linked polypeptide into the nucleus of a eukaryotic cell.
CAS9 (SP) (SO2) EXON1	*Streptococcus pyogenes*	Exon 1 of Cas9 endonuclease. Cas9 endonuclease introduces a double-strand break in the target endogenous DNA sequence.
ST-LS1 INTRON2	Potato *(Solanum tuberosum)*	Intron introduced for plant-optimized expression of Cas9 endonuclease.
CAS9 (SP) (SO2) EXON2 (TR1)	*Streptococcus pyogenes*	Exon 2 of Cas9 endonuclease. Cas9 endonuclease introduces a double-strand break in the target endogenous DNA sequence.
VIRD2 NLS (CO)	*Agrobacterium tumefaciens*	Nuclear localization signal, that is a bipartite nuclear localizing signal that targets a linked polypeptide back to the nucleus of the expressing cell.
UBQ14 TERM	*Arabidopsis thaliana*	Ubiquitin 14 3' untranslated region and terminator.
PSN1	synthetic	A synthetic sequence designed to facilitate PCR analysis of recombined FRT sites.
ATTB2	*Escherichia coli*	Gateway™ cloning system recombination site to facilitate vector construction process.
GM-DGAT1B-CR69 TARGET SITE	Soybean *(Glycine max)*	Cas9 endonuclease recognition site in *DGAT1B* gene
HDR-GM-DGAT1B-5'	Soybean *(Glycine max)*	5’ (upstream) homology-directed recombination (HDR) arm to facilitate replacement via the native cellular mechanism.
GM-DGAT1B EXON16 (ALT5) (TR1)	Soybean *(Glycine max)*	Part of the DGAT1B coding sequence containing targeted base pair substitutions.
HDR-GM-DGAT1B-3'	Soybean *(Glycine max)*	3’ (downstream) homology-directed recombination (HDR) arm to facilitate replacement via the native cellular mechanism.
GM-DGAT1B-CR69 TARGET SITE	Soybean *(Glycine max)*	Cas9 endonuclease recognition site in DGAT1B gene
ATTB3	*Escherichia coli*	Gateway™ cloning system recombination site to facilitate vector construction process.
AT-UBIQ10 PRO (0.9 KB)	*Arabidopsis thaliana*	Polyubiquitin 10 promoter.
AT-UBQ10 5UTR	*Arabidopsis thaliana*	5' untranslated region of the polyubiquitin 10 promoter
AT-UBIQ10 INTRON1	*Arabidopsis thaliana*	Intron from the ubiquitin promoter
CTP (SYNTHETIC) (MOD D)	Maize (*Zea mays),* Rice (*Oryza sativa)*	Chimeric chloroplast targeting peptide which targets the downstream protein to the chloroplast via binding of the chloroplast membrane and resulting interaction of the targeted protein with the chloroplast import machinery.
SPCN (SO) (TR1)	*Streptomyces spectabilis*	Spectinomycin gene used as a selectable marker during transformation.
UBQ14 TERM	*Arabidopsis thaliana*	Ubiquitin 14 3' untranslated region and terminator.

GM-MTH1 PRO-V2	Soybean *(Glycine max)*	Variant of metallothionein gene promoter.
DS-RED2 (ALT1)	*Discosoma sp.*	Modified fluorescent protein used as a visual marker.
UBQ3 TERM-V1	*Arabidopsis thaliana*	Variant of ubiquitin 3 terminator.
PSB1	synthetic	A synthetic sequence designed to facilitate PCR analysis of recombined FRT sites.
SPACER3	Maize *(Zea mays)*	Sequence based on 19kDa B1 zein terminator used as a buffer sequence to insulate expression cassettes from potential degradation.
MINI-ALLSTOPS4	synthetic	Synthetic sequence of stop codons designed to stop downstream translation of open reading frames.
LB (OCTOPINE)	*Agrobacterium tumefaciens*	Left border sequence.
SC-CEN6 ARS	*Saccharomyces cerevisiae*	Centromere-6 and autonomously replicating sequence used as an origin of replication during vector construction.
SC-URA3 PRO	*Saccharomyces cerevisiae*	Uracil 3 native promoter which drives expression of the URA3 selectable marker gene for transformation during vector construction.
SC-URA3	*Saccharomyces cerevisiae*	Uracil 3 gene. URA3 is the third gene identified that is required for uracil biosynthesis and is used as a selectable marker for transformation during vector construction.
EC-RRNB TERM	*Escherichia coli*	Ribosomal RNA operon B terminator.
PVS1 ORI (3KB)	*Pseudomonas aeruginosa*	Origin of replication to facilitate plasmid propagation in bacterial cells.
SPC (VER2)	*Escherichia coli*	Spectinomycin resistance gene used as a selectable marker to facilitate identification of the plasmid-containing bacterial clones during the vector construction process.
NPTIII	*Streptococcus faecalis*	Neomycin phosphotransferase gene conferring resistance to aminoglycoside antibiotics. Used as selectable marker in plant transformation.

Description of scientific methodology that you used or intend to use to confirm that the intended genetic changes were achieved.

The intended base pair substitutions are confirmed by Next Generation Sequencing (NGS) analysis. Southern-by-sequencing (SbS) is used to confirm absence of unintentionally integrated DNA from the transformation plasmid in the final soybean line.

SbS is an efficient NGS-based sequencing tool described in detail in Zastrow-Hayes et al. (2015) and Brink et al. (2019). SbS analysis covers the sequences of the plasmid used in transformation and detects unique junctions that would be created between the plant genomic DNA and unintended sequences derived from the transformation plasmid, if unintended plasmid DNA integration has occurred. A plant with no detected unintended plasmid-derived DNA is selected and advanced for further development.

As described above, CRISPR-Cas9 soybean with improved oil and protein content is designed to not contain any unintended genetic elements from the transformation plasmid and - more specifically - plant pest genetic material. Soybean is not a plant pest and there is also no reason to believe that the amino acid substitution in the *DGAT1B* gene would generate a plant pest, since no plant pest DNA is inserted into the plant genome, nor are improved oil and protein content phenotypes associated with or likely to result in plant pest risk. Thus, we understand that APHIS would not consider CRISPR-Cas9 soybean with improved oil and protein content as described herein to be a regulated article under Part 340.

Additionally, mutations in the *DGAT1B* gene leading to the amino acid substitution could have been achieved using conventional breeding techniques. For example, the variability of amino acid residues at position 479 (in reference to soybean DGAT1B protein) and in particular the occurrence of S, are known to exist in several plant species. Single Nucleotide Polymorphism (SNP) is the most prevalent type of mutations in plants and the most common source of genetic variation between alleles of the species' gene pool (Glenn et al. 2017; Morgil et al., 2020). Furthermore, allelic differences between different varieties of the same species are often not limited by a single SNP or a single amino acid substitution (Tenaillon et al., 2001; Ching et al., 2002; Gore et al., 2009; Sun et al., 2018). Plant genomes are inherently variable and conventional breeding has been utilizing both spontaneous and induced genetic variability to develop new varieties (Mba, 2013; Glenn et al., 2017; Sun et al., 2018; Custers et al., 2019; Graham et al., 2020). Therefore, coupled with the fact that the final gene-edited soybean line will not contain unintentionally integrated DNA from the transformation plasmid, this soybean line is indistinguishable from soybean plants that could have resulted from native genome variability or be developed in a conventional breeding program.

We appreciate your review of this inquiry and are ready to address any questions you may have.

Sincerely,

Annie Tang Gutsche, Ph.D.
US Seeds Regulatory Policy Lead
Corteva Argiscience
7100 NW 62nd Avenue
Johnston, IA 50131-1000
(302) 415-9434 (Mobile)
annie.gutsche@corteva.com

References

Brink K. et al. (2019) Comparison of Southern-by-Sequencing (SbS^{TM}) Technology and Southern Blot Analysis for Molecular Characterization of Genetically Modified Crops. Journal of Regulatory Science 7: 1.

Ching A. et al. (2002) SNP frequency, haplotype structure and linkage disequillibrium in elite maize inbred limes. BMC Genetic 3: 19.

Custers R. et al. (2019) Genetic alterations that do or domnot occur naturally; consequences for genome edited organisms in the context of regulatory oversight. Frontiers in Bioengineering and Biotechnology 6. Article 213.

Glenn K.C. et al. (2017) Breeding new plant varieties to market: plant breeding and selection practiecs advance beneficial characteristics while minimizing unintended changes. Crop Science 57: 2906.

Gore M.A. et al. (2009) A first-generation haplotype map of maize. Science 326: 1115.

Graham N. et al. (2020) Plant genome editing and the relevance of off-target changes. Plant Physiology. doi: 10.1104/pp.19.01194.

Mba C. (2013) Induced mutations unleash the potentials of plant genetic resources for food and agriculture. Agronomy 3: 200.

Morgil H. et al. (2020) Single nucleotide polymorphisms (SNPs) in plant genetics and breeding. In: Plant Genetics and Breeding, The Recent Topics in Genetic Polymorphisms, Mahmut Çalışkal et al. eds. IntechOpen, DOI: 10.5772/intechopen.91886.

Podevin N. et al. (2013) Site-directed nucleases: a paradigm shift in predictable, knowledge-based plant breeding. Trends in Biotechnology 31(6): 375.

Roesler et al. (2016) An improved variant of soybean type 1 diacylglycerol acyltransferase increases the oil content and decreases the soluble carbohydate content of soybeans. Plant Physiology 171: 878.

Sun S. et al. (2018). Extensive intaspecific gene order and gene structural variations between MO17 and other maize genomes. Nature Genetics 50: 1289.

Tenaillon M.I. et al. (2001) Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.) PNAS USA 98: 9161.

Zastrow-Hayes G.M. et al. (2015) Southern-by-Sequencing: A Robust Screening Approach for Molecular Characterization of Genetically Modified Crops. The Plant Genome 8: 1.