

RECEIVED

By kldiggs for BRS Document Control Officer at 8:58 am, Jan 13, 2022



DONALD DANFORTH
PLANT SCIENCE CENTER

Institute for International Crop Improvement

January 12, 2022

Bernadette Juarez
APHIS Deputy Administrator
Biotechnology Regulatory Services

RE: Request for Regulatory Status Review

Dear Ms. Juarez,

The Donald Danforth Plant Science Center is requesting an initial regulatory status review (RSR), pursuant to 7 CFR § 340.4, for reduced height teff with improved lodging resistance. Lodging is a major limitation to teff production and for decades the development of lodging resistant varieties has proven difficult with conventional breeding approaches. We used CRISPR/Cas9 to introduce loss-of-function (knockout) mutations in three genes involved in determining plant height.

The enclosed dossier contains all the necessary information to conduct an initial review of the modified teff. This RSR request **does not contain** confidential business information (CBI).

We look forward to addressing any questions you may have on our reduced height teff.

A handwritten signature in black ink, appearing to read 'Donald J. MacKenzie'.

Donald J. MacKenzie
Executive Director, IICI

encl: DPS-22ETAP-US.pdf

Please see the enclosed file.

cc: Getu Duguma, IICI



Reduced Height Lodging Resistant Teff

Information dossier in support of an initial review under the RSR process
described in 7 CFR § 340.4

SUBMITTED TO

Biotechnology Regulatory Services
Animal and Plant Health Inspection Service, USDA
4700 River Road, Unit 98
Riverdale, MD 20737

SUBMITTED BY

Donald Danforth Plant Science Center
975 North Warson Road
St. Louis, MO 63132

REGULATORY DOSSIER NUMBER

DPS-22ETAP-US

© 2022 Donald Danforth Plant Science Center (DDPSC) All rights reserved.

This document is protected under copyright law and is for use only by Biotechnology Regulatory Services in support of the actions requested by the DDPSC. Any other use of this material, without prior written approval of the DDPSC is strictly prohibited. By submitting this document, the DDPSC does not grant any party or entity any right or license to the information or intellectual property described herein.



Summary

Teff is a staple food and a valuable cash crop for millions of people in Ethiopia. Lodging is a major limitation to teff production and for decades the development of lodging resistant varieties has proven difficult with conventional breeding approaches. We used clustered regularly interspersed short palindromic repeats (CRISPR)-CRISPR associated protein 9 (Cas9) to introduce knockout mutations in the teff orthologues of the rice *Semidwarf1* (*Sd1*) and sorghum *Dw1* and *Dw3* genes to confer reduced plant height and ultimately lodging resistance in DPS 22-8-12 teff.

The mutations in progeny null-segregant plants that were free of introduced exogenous DNA were characterized by next generation sequencing (NGS) were comprised of single nucleotide insertions (+A or +T) or dinucleotide deletions (-GC) resulting in codon frameshifts creating premature stop codons in the *sd1*, *dw1*, and *dw3* coding sequences.

The approximate 50 percent reduction in plant height for DPS 22-8-12 teff observed under greenhouse conditions was within the range of natural variability reported for the plant species. Reduced plant height has been one of the main objectives of modern plant breeding, is a familiar trait in many improved crop varieties, and is not associated with a plant pest risk.



Abbreviations

ABCB1	ATP-binding cassette type B1
BIN2	brassinosteroid insensitive 2
BLAST	basic local alignment search tool
BR	brassinosteroid
Cas9	CRISPR associated protein 9
CBI	confidential business information
CRISPR	clustered regularly interspersed short palindromic repeats
DDPSC	Donald Danforth Plant Science Center
GA	gibberellin
gRNA	guide RNA
MUSCLE	MUltiple Sequence Comparison by Log-Expectation
NGS	next generation sequencing
ORF	open reading frame
PAM	protospacer adjacent motif
qPCR	quantitative polymerase chain reaction
RSR	Regulatory Status Review



Contents

1 General Information	6
1.1 Applicant Details	6
1.2 Confidential Information	6
1.3 Brief Identification of the Modified Plant	6
1.4 Purpose of the Application	6
1.5 Development of Reduced Plant Height Teff	6
2 Comparator Plant	7
3 Genotype of the Modified Plant	8
4 The Intended Trait and Mechanism of Action	12
4.1 Intended Trait	12
4.2 Intended Phenotype	12
4.3 Mechanism of Action	12
5 Concluding Comments	14
6 References	15
Appendix A Gene Sequences	18
A.1 <i>Sd1</i> Gene, Chr 3A	18
A.2 <i>Sd1</i> Gene, Chr 3B	19
A.3 <i>Dw1</i> Gene, Chr 4A	20
A.4 <i>Dw1</i> Gene, Chr 4B	21
A.5 <i>Dw3</i> Gene, Chr 8A	23
A.6 <i>Dw3</i> Gene, Chr 8B	25
Appendix B gRNA BLAST Searches	28
B.1 <i>Sd1</i> guide RNA (gRNA)1	28
B.2 <i>Sd1</i> gRNA2	29
B.3 <i>Dw1</i> gRNA	30
B.4 <i>Dw3</i> gRNA	31
Appendix C Nucleotide Sequence Alignments	33



List of Tables

1	Guide RNA target sequences	7
2	Summary of <i>sd1</i> , <i>dw1</i> , and <i>dw3</i> mutations in DPS 22-8-12 teff	9

List of Figures

1	Reduced height teff development	7
2	Guide RNA locations	8
3	SD1 amino acid sequence alignments	10
4	DW1 amino acid sequence alignments	10
5	DW3 amino acid sequence alignments	11
6	DW3 structures	14
7	Comparison of genotype plant heights	15
8	SD1A gene sequence alignments	34
9	SD1B gene sequence alignments	35
10	DW1A gene sequence alignments	36
11	DW1B gene sequence alignments	37
12	DW3A gene sequence alignments (part 1)	38
13	DW3A gene sequence alignments (part 2)	39
14	DW3B gene sequence alignments (part 1)	40
15	DW3B gene sequence alignments (part 2)	41



1. General Information

1.1. Applicant Details

- | | |
|------------------------|--|
| (a) Applicant: | Donald Danforth Plant Science Center |
| (b) Address: | 975 North Warson Road, St. Louis, MO 63132 |
| (c) Contacts: | Getu Beyene and Donald J. MacKenzie |
| (d) Telephone numbers: | (314) 587-1612 (office), (202) 695-0436 (mobile) |
| (e) Email addresses: | GDuguma@DanforthCenter.org,
DMacKenzie@DanforthCenter.org |

1.2. Confidential Information

This request **does not contain** confidential business information (CBI).

1.3. Brief Identification of the Modified Plant

- | | |
|-------------------------|---|
| Designation: | DPS 22-8-12 |
| Plant species: | <i>Eragrostis tef</i> (Zucc.) Trotter, commonly known as teff |
| Phenotype: | Reduced plant height conferring resistance to lodging |
| Introduced mutation(s): | Targeted knockout mutations in the <i>Sd1</i> , <i>Dw1</i> , and <i>Dw3</i> genes |

1.4. Purpose of the Application

This request for a Regulatory Status Review (RSR) is being submitted according to the procedures described under 7 CFR § 340.4.

1.5. Development of Reduced Plant Height Teff

The development of reduced plant height DPS 22-8-12 teff followed a step-wise process that included two plant transformation steps, one to introduce targeted mutations within the *Sd1* gene and a re-transformation of a *sd1* line to simultaneously introduce targeted mutations within the *Dw1* and *Dw3* genes (Figure 1).

Sequences of rice (*Oryza sativa* L.) GA20-oxidases (OsGA20ox1–OsGA20ox4) (Sakamoto et al., 2004) and sorghum (*Sorghum bicolor* L.) *Dw1* (Sobic.009G229800, Chr09:57093313..57095643) and *Dw3* (Sobic.007G163800, Chr07:59821905..59829910) were used as baits to retrieve orthologous sequences from the published teff genome (VanBuren et al., 2020) using CoGeBlast¹ (Lyons et al., 2008), which were then used to design specific gRNAs. The nucleotide sequences of the gRNAs targeting the *Sd1*, *Dw1*, and *Dw3* genes are shown in Table 1 and their relative locations on the target genes are illustrated schematically in Figure 2.

¹ CoGe is a platform for performing Comparative Genomics research. It provides an open-ended network of interconnected tools to manage, analyze, and visualize next-gen data. The CoGe basic local alignment search tool (BLAST) tool can be accessed at: <https://genomevolution.org/coge/CoGeBlast.pl>



COMPARATOR PLANT

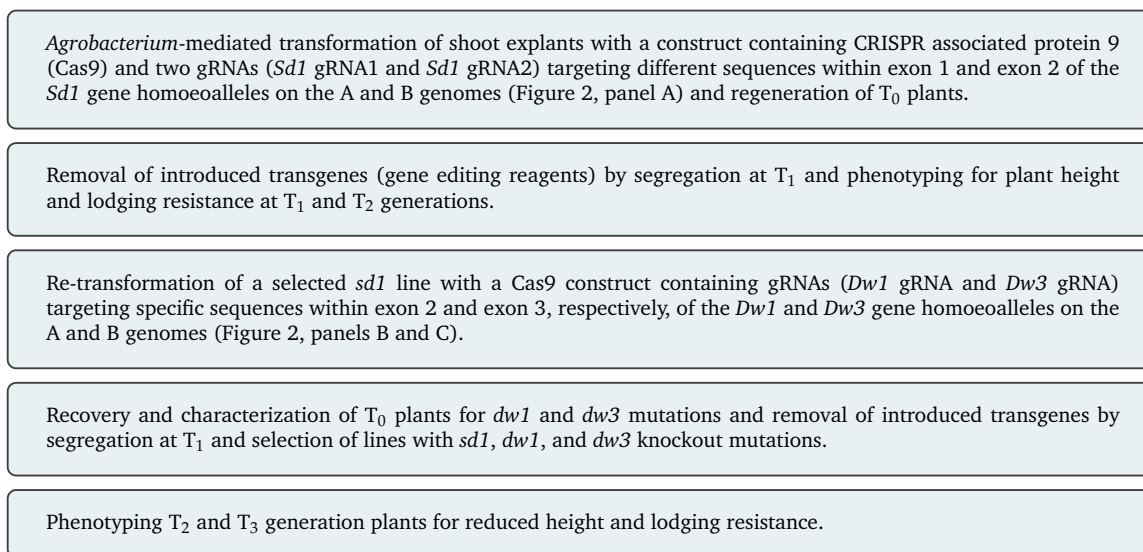


Figure 1. Main steps in the development of reduced height lodging resistant teff. The development process included two transformation steps, one to introduce targeted mutations within the *Sd1* gene and a second to simultaneously introduce targeted mutations within the *Dw1* and *Dw3* genes. In all cases, null-segregants containing only the targeted mutations and without any introduced foreign DNA were confirmed by quantitative polymerase chain reaction (qPCR) analysis and the target site mutations confirmed by next generation sequencing (NGS).

Table 1. Guide RNA target sequences

Name	Target Gene (Exon)	Chromosome Location ^a	Strand	Sequence ^b
<i>Sd1</i> gRNA1	<i>Sd1</i> (Exon 1)	3A (31096526..31096548) 3B (29293138..29293160)	Sense	5'-GGCCCGGACTTCGAGCCAAT <u>GGG</u> -3'
<i>Sd1</i> gRNA2	<i>Sd1</i> (Exon 2)	3A (31096827..31096849) 3B (29293380..29293402)	Sense	5'-GCGACTTCTTCGAGGACAGC <u>CGG</u> -3'
<i>Dw1</i> gRNA	<i>Dw1</i> (Exon 2)	4A (7977763..7977785) 4B (7681439..7681461)	Anti-sense	5'-CCTGGAATATGGTGTGCAC <u>AGG</u> -3'
<i>Dw3</i> gRNA	<i>Dw3</i> (Exon 3)	8A (145900..145878) 8B (134813..134791)	Anti-sense	5'-TCGTGCGCCCGATCATGAAC <u>CGG</u> -3'

^a Chromosome locations are relative to the *E. tef* database (id 50954 PacBio unmasked vV3) located at <https://genomevolution.org/coge/>. BLAST searches against the teff genome database using gRNA sequences as queries are shown in Appendix B, beginning on page 28.

^b Underlined nucleotides correspond to the protospacer adjacent motif (PAM) recognition sites.

Null-segregant plant lines containing targeted mutations without introduced foreign DNA were selected at the T₁ generation using qPCR analysis and NGS was used to confirm the identity of the knockout mutations.

2. Comparator Plant

Teff (*Eragrostis tef* (Zucc.) Trotter) is an allotetraploid ($2n = 4 \times = 40$) C₄ plant belonging to the Poaceae or Grass family (USDA, 2021) and is closely related to finger millet (*Eleusine coracana* Gaerth.) as both are in the subfamily Chloridoideae. The genus *Eragrostis* comprises about 350 species from which only teff is cultivated for human consumption (Assefa et al., 2015).

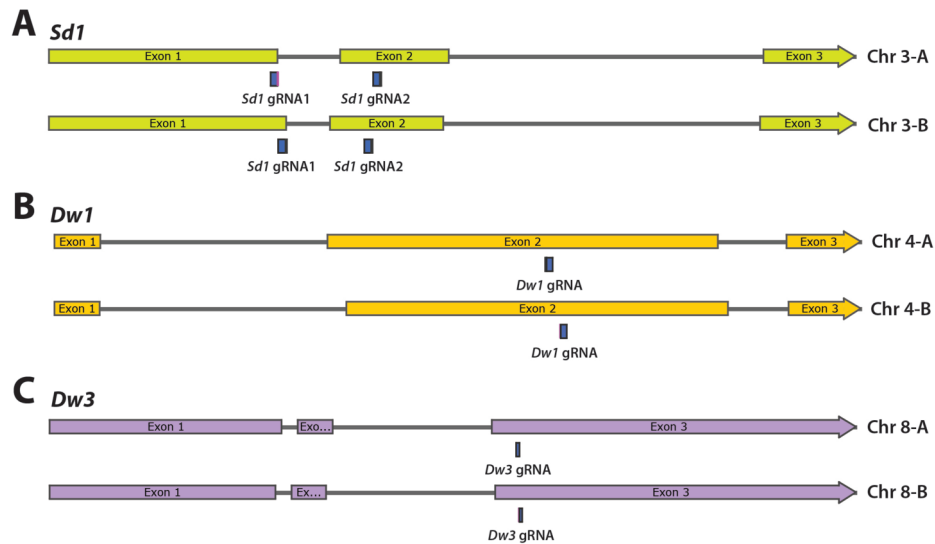


Figure 2. Schematic diagram of the *Sd1* (panel A), *Dw1* (panel B), and *Dw3* (panel C) genes on chromosomes 3, 4 and 8 of the A and B genomes. Relative locations of the gRNA target sites are shown. Nucleotide sequences for each gene, including translated amino acid sequences for each encoded protein, are provided in Appendix A beginning on page 18.

3. Genotype of the Modified Plant

The clustered regularly interspersed short palindromic repeats (CRISPR)-Cas9 targeted mutations were characterized by nucleotide sequencing of the target loci in T_1 generation progeny plants that had been confirmed free of introduced foreign DNA by qPCR. Line DPS 22-8-12 was one of the lines identified that was homozygous at each homoeoallele of the *Sd1*, *Dw1*, and *Dw3* genes for a knockout (loss-of-function) mutation. The *sd1*, *dw1*, *dw3* mutations were limited to single nucleotide insertions (A or T) or dinucleotide (GC) deletions (Table 2). For reference, complete nucleotide sequence alignments between the *Sd1*, *Dw1*, and *Dw3* wild-type sequences and their respective mutant alleles are shown in Appendix C, beginning on page 33.

The wild-type *Sd1* gene homoeologs on genome A and B encode proteins of 420 and 419 amino acids, respectively, with 93.6 percent sequence identity, that share approximately 89 percent identity with the rice GA-20 oxidase (GenBank: BAL03272.1). The *sd1A1* and *sd1A2* alleles, which were homozygous following two generations of selfing of the parental *sd1*-edited line, consisted of a dinucleotide GC deletion in the gRNA1 target site and a single adenine nucleotide insertion (+A) in the gRNA2 target site. These mutations resulted in a codon frameshift and truncation of the coding sequence to 347 amino acids (Table 2). The homozygous *sd1B1* and *sd1B2* alleles contained a single adenine nucleotide insertion (+A) in the gRNA1 target site that resulted in a truncated coding sequence of 245 amino acids. The nucleotide sequence corresponding to gRNA2 target site was unchanged; however, this was of no consequence as the newly created stop codon was located upstream of the gRNA2 target. Alignments of the wild-type *Sd1A* and *Sd1B* encoded sequences and the associated mutated versions are shown in Figure 3.



THE GENETIC MODIFICATION

Table 2. Summary of *sd1*, *dw1*, and *dw3* mutations in DPS 22-8-12 teff

Allele ^a	gRNA1 Target Sequence ^b	Mutation	gRNA2 Target Sequence	Mutation	CDS ^c	Comments
Sd1A	GGCCCCGACTTCGAGCCAAT		GCGACTTCTTCGAGGACAGC		420	Wild-type SD1 (A genome)
<i>sd1A1</i>	GGCCCCGACTTCGA- -CAAT	-GC	GCGACTTCTTCGAGGACAAGC	+A	347	Altered amino acid sequence after F221
<i>sd1A2</i>	GGCCCCGACTTCGA- -CAAT	-GC	GCGACTTCTTCGAGGACAAGC	+A		
Sd1B	GGCCCCGACTTCGAGCCAAT		GCGACTTCTTCGAGGACAGC		419	Wild-type SD1 (B genome)
<i>sd1B1</i>	GGCCCCGACTTCGAGCCAAT	+A	GCGACTTCTTCGAGGACAGC	-	245	Altered amino acid sequence after P222
<i>sd1B2</i>	GGCCCCGACTTCGAGCCAAT	+A	GCGACTTCTTCGAGGACAGC	-		
Allele	gRNA Target Sequence ^b	Mutation	CDS	Comments		
Dw1A	GTGACACACCATATTCCAGG		528	Wild-type DW1 (A Genome)		
<i>dw1A1</i>	GTGACACACCATATTCCAGG	+A	287	Altered amino acid sequence after C276 and premature termination		
<i>dw1A2</i>	GTGACACACCATATTCCAGG	+A				
Dw1B	GTGACACACCATATTCCAGG		528	Wild-type DW1 (B Genome)		
<i>dw1B1</i>	GTGACACACCATATTCCAGG	+A	287	Altered amino acid sequence after C276 and premature termination		
<i>dw1B2</i>	GTGACACACCATATTCCAGG	+A				
Dw3A	GTTTCATGATCGGGCGCACGA		1354	Wild-type DW3 (A Genome)		
<i>dw3A1</i>	GTTTCATGATCGGGCGCACGA	+T	743	Altered amino acid sequence after F630 and premature termination		
<i>dw3A2</i>	GTTTCATGATCGGGCGCACGA	+T				
Dw3B	GTTTCATGATCGGGCGCACGA		1347	Wild-type DW3 (B Genome)		
<i>dw3B1</i>	GTTTCATGATCGGGCGCACGA	+A	736	Altered amino acid sequence after R620 and premature termination		
<i>dw3B2</i>	GTTTCATGATCGGGCGCACGA	+A				

^a The naming convention for wild-type alleles uses three-letter codes for each gene with the first letter capitalized (e.g., *Sd1*, *Dw1*, and *Dw3*) followed by the genome designation, A or B. Mutant alleles are identified with lowercase three-letter codes, followed by the genome designation (A or B), and a number (e.g., *sd1A1*, *sd1A2*, *sd1B1*, etc).

^b All gRNA target sequences are shown as the sense strand (5'-3') without the PAM sites.

^c CDS = coding sequence with size indicated as the number of encoded amino acids.

The unmodified *Dw1* homoeologs encode proteins of 528 amino acids that are 99 percent sequence identical, and share 81 percent sequence identity with the homologous protein from sorghum (UniProtKB–A0A1B6P9X8). The *dw1* mutation consisted of a single adenine nucleotide insertion for each homoeoallele at the same position within the gRNA target sequence (Table 2), which resulted in a truncated open reading frame (ORF) of 287 amino acids for each homoeologous gene. Alignments of the amino acid sequences encoded by the wild-type *Dw1* and mutant *dw1* alleles are shown in Figure 4.

The unmodified *Dw3* homoeologs on the A and B genomes encode proteins of 1354 and 1347 amino acids, respectively, that share 93 percent sequence identity. The teff DW3A and DW3B proteins are, respectively, 87 percent and 88 percent identical to the homologous protein from sorghum (UniProtKB–A0A1Z5RA91). The *dw3* mutation consisted of a single thymine nucleotide insertion in the gRNA target site of the A genome (*dw3A*) and a corresponding single adenine nucleotide insertion in the gRNA target site of the B genome (*dw3B*) (Table 2). The *dw3A* and *dw3B* mutations resulted in truncated ORFs of 743 and 736 amino acids, respectively. Alignments of the wild-type *Dw3A* and *Dw3B* encoded sequences and the associated mutated versions are shown in Figure 5.

THE GENETIC MODIFICATION



Dw3A	.MS	SSDPPEIRGRVVVLGADADELARPELEAFHLPSPTEAQEAGSVGGPGEAAAAPVVAPLETELEPPNAMPSSSSASSN	79
dw3A1	.MS	SSDPPEIRGRVVVLGADADELARPELEAFHLPSPTEAQEAGSVGGPGEAAAAPVVAPLETELEPPNAMPSSSSASSN	79
Dw3B	MSA	SSDPPEIRGRVVVLGADADELARPELEAFHLPSPTEEAAGASVGGPA.AANPFPVVAPLETELEPPNAMPSSSS.SSN	77
dw3B1	MSA	SSDPPEIRGRVVVLGADADELARPELEAFHLPSPTEEAAGASVGGPA.AANPFPVVAPLETELEPPNAMPSSSS.SSN	77
Dw3A	SNSNEONKE	QENTKKKGVSLAPAPLGSLFRFADGLDCVLMVSGTLGALVHGCSLPVFLRFADLVDSFGSHADDPDTMVR	159
dw3A1	SNSNEONKE	QENTKKKGVSLAPAPLGSLFRFADGLDCVLMVSGTLGALVHGCSLPVFLRFADLVDSFGSHADDPDTMVR	159
Dw3B	SNSNEONKE	E...KKKGVSLAPAPLGSLFRFADGLDCVLMVSGTLGALVHGCSLPVFLRFADLVDSFGSHADDPDTMVR	154
dw3B1	SNSNEONKE	E...KKKGVSLAPAPLGSLFRFADGLDCVLMVSGTLGALVHGCSLPVFLRFADLVDSFGSHADDPDTMVR	154
Dw3A	LVA	KYALYFLVVGAAIWASSWAEISCWMTGEROSTRMRIRYLESALRODVSFDDTDVRTSDVIYAINADAVIVQDAISE	239
dw3A1	LVA	KYALYFLVVGAAIWASSWAEISCWMTGEROSTRMRIRYLESALRODVSFDDTDVRTSDVIYAINADAVIVQDAISE	239
Dw3B	LVA	KYALYFLVVGAAIWASSWAEISCWMTGEROSTRMRIRYLESALRODVSFDDTDVRTSDVIYAINADAVIVQDAISE	234
dw3B1	LVA	KYALYFLVVGAAIWASSWAEISCWMTGEROSTRMRIRYLESALRODVSFDDTDVRTSDVIYAINADAVIVQDAISE	234
Dw3A	KLGNLIHYMATFVAGFVVGF	TAAWQLALVTLAVVPLIAVIGGLSAAALAKLSSRSQDALAEASNIAEQVAQIRTVQAFV	319
dw3A1	KLGNLIHYMATFVAGFVVGF	TAAWQLALVTLAVVPLIAVIGGLSAAALAKLSSRSQDALAEASNIAEQVAQIRTVQAFV	319
Dw3B	KLGNLIHYMATFVAGFVVGF	TAAWQLALVTLAVVPLIAVIGGLSAAALAKLSSRSQDALAEASNIAEQVAQIRTVQAFV	314
dw3B1	KLGNLIHYMATFVAGFVVGF	TAAWQLALVTLAVVPLIAVIGGLSAAALAKLSSRSQDALAEASNIAEQVAQIRTVQAFV	314
Dw3A	GEERAMRAYSLALA	AAQRIGYRSGFAGKGLGGTYFTVFCYALLWYGGLLVRRHHTNGGLAIATMFSVMIGGLALGQS	399
dw3A1	GEERAMRAYSLALA	AAQRIGYRSGFAGKGLGGTYFTVFCYALLWYGGLLVRRHHTNGGLAIATMFSVMIGGLALGQS	399
Dw3B	GEERAMRAYSLALG	MAQRIGYRSGFAGKGLGGTYFTVFCYALLWYGGLLVRRHHTNGGLAIATMFSVMIGGLALGQS	394
dw3B1	GEERAMRAYSLALG	MAQRIGYRSGFAGKGLGGTYFTVFCYALLWYGGLLVRRHHTNGGLAIATMFSVMIGGLALGQS	394
Dw3A	APSMAAFARVA	AAAKIFKIIDHKPLISVVHGGDDVQLPSVTGRVEMRGVDFAYPSRPDPVFLRGFSLTVPPGKTIALVG	479
dw3A1	APSMAAFARVA	AAAKIFKIIDHKPLISVVHGGDDVQLPSVTGRVEMRGVDFAYPSRPDPVFLRGFSLTVPPGKTIALVG	479
Dw3B	APSMAAFARVA	AAAKIFRVIDH...LAVVHG...HVVQLPSVTGRVEMRGVDFAYPSRPDPVFLRGFSLTVPPGKTIALVG	470
dw3B1	APSMAAFARVA	AAAKIFRVIDH...LAVVHG...HVVQLPSVTGRVEMRGVDFAYPSRPDPVFLRGFSLTVPPGKTIALVG	470
Dw3A	SSSGKSTVVSLIERFYDPS	AGEILLDGHDLKSNLNRWLROQIGLVSQEPTLFATSIKENLLLGRDSSHATLAEMEEAAR	559
dw3A1	SSSGKSTVVSLIERFYDPS	AGEILLDGHDLKSNLNRWLROQIGLVSQEPTLFATSIKENLLLGRDSSHATLAEMEEAAR	559
Dw3B	SSSGKSTVVSLIERFYDPS	AGEILLDGHDLKSNLNRWLROQIGLVSQEPTLFATSIKENLLLGRDSSQATLAEMEEAAR	550
dw3B1	SSSGKSTVVSLIERFYDPS	AGEILLDGHDLKSNLNRWLROQIGLVSQEPTLFATSIKENLLLGRDSSQATLAEMEEAAR	550
Dw3A	VANAHSFIIKLPDGYDTQV	GERGLQLSGGQKORIAIARAMLNKPGILLLDEATSALDSESEKLVQEAALDRFMIGRTTLVI	639
dw3A1	VANAHSFIIKLPDGYDTQV	GERGLQLSGGQKORIAIARAMLNKPGILLLDEATSALDSESEKLVQEAALDRFMIGRTTLVI	630
Dw3B	VANAHSFIIKLPDGYDTQV	GERGLQLSGGQKORIAIARAMLNKPGILLLDEATSALDSESEKLVQEAALDRFMIGRTTLVI	630
dw3B1	VANAHSFIIKLPDGYDTQV	GERGLQLSGGQKORIAIARAMLNKPGILLLDEATSALDSESEKLVQEAALDRFMIGRTTLVI	621
Dw3A	AHRLSTIRKADLVAVLHGGAV	SEIGTEHEELMGKEDDGAYARLIRMQEQ...AAQEVARRSSARNVSNARNSVSSPIMTRN	717
dw3A1	AHRLSTIRKADLVAVLHGGAV	SEIGTEHEDRAHDPED...RAAVDDPQGRPG...SARNVSSPIMTRN	652
Dw3B	AHRLSTIRKADLVAVLHGGAV	SEIGTEHDEELMAKEDDGAYARLIRMQEQQA...AAQEVARRSSARPS...SARNVSSPIMTRN	709
dw3B1	AHRLSTIRKADLVAVLHGGAV	SEIGTEHEDRAHDPED...RAAVDDPQGRPG...SARNVSSPIMTRN	643
Dw3A	SSYCRSPYSRRLSD	FSNAD.FHYHGGGELPENGTK.KMIHQRVAFRAGASSFLRLAKMNSPEWGYALVGSLSGSMVCGSFS	795
dw3A1	SSYCRSPYSRRLSD	FSNAD.FHYHGGGELPENGTK.KMIHQRVAFRAGASSFLRLAKMNSPEWGYALVGSLSGSMVCGSFS	663
Dw3B	SSYCRSPYSRRLSD	FSNADSHHYHGGGELPENGTK.AHHORRRIAFRAGASSFWRLAKMNSPEWGYALVGSLSGSMVCGSFS	789
dw3B1	SSYCRSPYSRRLSD	FSNAD.FHYHGGGELPENGTK.KMIHQRVAFRAGASSFLRLAKMNSPEWGYALVGSLSGSMVCGSFS	654
Dw3A	AIFAYVLSAVLSVYAPDP	GHMRRERIAKYCYLLMGSSAALVCNTVQHVFWDTVENLTHFRVRERMLGAVLRNEMAWFDA	875
dw3A1	AIFAYVLSAVLSVYAPDP	GHMRRERIAKYCYLLMGSSAALVCNTVQHVFWDTVENLTHFRVRERMLGAVLRNEMAWFDA	674
Dw3B	AIFAYVLSAVLSVYAPDP	GHMRRERIAKYCYLLMGSSAALVCNTVQHVFWDTVENLTHFRVRERMLGAVLRNEMAWFDA	869
dw3B1	AIFAYVLSAVLSVYAPDP	GHMRRERIAKYCYLLMGSSAALVCNTVQHVFWDTVENLTHFRVRERMLGAVLRNEMAWFDA	665
Dw3A	EENASARVAARLALDAQN	VRSIAIGDRISVIVQNSALLLVACTAGFVLOWRLALVLLGVFPLVVAATVLOKMFMKFGSGDI	955
dw3A1	EENASARVAARLALDAQN	VRSIAIGDRISVIVQNSALLLVACTAGFVLOWRLALVLLGVFPLVVAATVLOKMFMKFGSGDI	674
Dw3B	EENASARVAARLALDAQN	VRSIAIGDRISVIVQNSALLLVACTAGFVLOWRLALVLLGVFPLVVAATVLOKMFMKFGSGDI	949
dw3B1	EENASARVAARLALDAQN	VRSIAIGDRISVIVQNSALLLVACTAGFVLOWRLALVLLGVFPLVVAATVLOKMFMKFGSGDI	665
Dw3A	EAHARATQIAGEAVANLRT	VAAFNAEAKIAGLFAANLGRPLRRQLWKGQVAGCGYGVAFQFLLYASYALGLWYAAWLVKH	1035
dw3A1	EAHARATQIAGEAVANLRT	VAAFNAEAKIAGLFAANLGRPLRRQLWKGQVAGCGYGVAFQFLLYASYALGLWYAAWLVKH	674
Dw3B	EAHARATQIAGEAVANLRT	VAAFNAEAKIAGLFAANLGRPLRRQLWKGQVAGCGYGLAFQFLLYASYALGLWYAAWLVKH	1029
dw3B1	EAHARATQIAGEAVANLRT	VAAFNAEAKIAGLFAANLGRPLRRQLWKGQVAGCGYGVAFQFLLYASYALGLWYAAWLVKH	665
Dw3A	GVSDFSRTLRVFMVLMVS	ANGAAETLLTLPDFVRCGRAMSVFETIDRRTEADPDPDAAPLQLPLLTGVELRHVDFCYE	1115
dw3A1	GVSDFSRTLRVFMVLMVS	ANGAAETLLTLPDFVRCGRAMSVFETIDRRTEADPDPDAAPLQLPLLTGVELRHVDFCYE	685
Dw3B	GVSDFSRTLRVFMVLMVS	ANGAAETLLTLPDFVRCGRAMSVFETIDRRTEADPDPDAAPLQLPLLTGVELRHVDFCYE	1108
dw3B1	GVSDFSRTLRVFMVLMVS	ANGAAETLLTLPDFVRCGRAMSVFETIDRRTEADPDPDAAPLQLPLLTGVELRHVDFCYE	677
Dw3A	SRPDMQVLQDLSLRARAGK	TIALVGPSCCCKSSVLALIQRFYEPESGRVLLDGRDARKYNLRALRRAVAVVPOEPLFAA	1195
dw3A1	SRPDMQVLQDLSLRARAGK	TIALVGPSCCCKSSVLALIQRFYEPESGRVLLDGRDARKYNLRALRRAVAVVPOEPLFAA	688
Dw3B	SRPDMQVLQDLSLRARAGK	TIALVGPSCCCKSSVLALIQRFYEPESGRVLLDGRDARKYNLRALRRAVAVVPOEPLFAA	1188
dw3B1	SRPDMQVLQDLSLRARAGK	TIALVGPSCCCKSSVLALIQRFYEPESGRVLLDGRDARKYNLRALRRAVAVVPOEPLFAA	681
Dw3A	SIHDNIAYGRECGATEAEV	LEAATOQANAHKFISALPDGYRTQVGERGVOLSGGQRORIAVARALVKQAALLLDEATSAL	1275
dw3A1	SIHDNIAYGRECGATEAEV	LEAATOQANAHKFISALPDGYRTQVGERGVOLSGGQRORIAVARALVKQAALLLDEATSAL	720
Dw3B	TIHDNIAYGRECGATEAEV	LEAATOQANAHKFISALPDGYRTQVGERGVOLSGGQRORIAVARALVKQAALLLDEATSAL	1267
dw3B1	SIHDNIAYGRECGATEAEV	LEAATOQANAHKFISALPDGYRTQVGERGVOLSGGQRORIAVARALVKQAALLLDEATSAL	712
Dw3A	DAESERSVQOALDRHAKT	RSTTTIVVAHRLATVNRNAHTIAVIDGKVVVEQGSLSHLLNHPDGTYARMLOLORLTSSTS	1354
dw3A1	DAESERSVQOALDRHAKT	RSTTTIVVAHRLATVNRNAHTIAVIDGKVVVEQGSLSHLLNHPDGTYARMLOLORLTSSTS	741
Dw3B	DAESERSVQOALDRHAKT	RSTTTIVVAHRLATVNRNAHTIAVIDGKVVVEQGSLSHLLNHPDGTYARMLOLORLTPSTS	1347
dw3B1	DAESERSVQOALDRHAKT	RSTTTIVVAHRLATVNRNAHTIAVIDGKVVVEQGSLSHLLNHPDGTYARMLOLORLTPSTS	734
Dw3A		1354	
dw3A1	R	743	
Dw3B		1347	
dw3B1	W	736	

Figure 5. Alignment of the wild-type DW3 amino acid sequences encoded by the *Dw3* homeologs on the A and B genomes (*Dw3A* and *Dw3B*) with the amino acid sequences encoded by their respective mutant alleles (*dw3A1* and *dw3B1*) characterized in DPS 22-8-12 teff. The *dw3A2* and *dw3B2* encoded sequences are not shown as they are the same as *dw3A1* and *dw3B1*, respectively. The locations of the codon frameshift for *dw1A1* is indicated by the down-arrow (↓) and for *dw1B1* by the up-arrow (↑). Sequences were aligned with MUSCLE (Edgar, 2004) using the default parameters and displayed using TEXshade (Beitz, 2000).



4. The Intended Trait and Mechanism of Action

4.1. Intended Trait

The intended trait is resistance to lodging. Lodging refers to the permanent displacement of aboveground portions of crops from their vertical stance due to stem bending (characterized by loss of plant elasticity leading to bending of stems while the roots are still secure in the soil) or breaking (breakage of the stem usually near the base of the peduncle), or failure of the root–soil anchorage system (involves uprooting of the whole plant while the stems still appear intact). In practice, bend lodging is by far the most common and economically most important type of lodging in teff (Ketema, 1983, 1993). In teff, lodging reduces grain yield by 11–22 percent, 1000-kernel weight by 35 percent, grain yield per panicle by 51 percent, and percentage and rate of seed germination by 41 and 44 percent, respectively (Ketema, 1983).

4.2. Intended Phenotype

The intended phenotype is reduced plant height.

4.3. Mechanism of Action

Plants contain three major growth-promoting plant hormones: auxin, gibberellins (GAs), and brassinosteroids (BRs), along with other hormones that promote growth in certain circumstances. Over the years, naturally-occurring mutations in genes associated with GA biosynthesis or signalling, BR signalling, or auxin transport have been utilized either individually, or in combination, in conventional breeding programs to develop lodging-resistant cultivars in different crop species.

The most widely deployed and well-studied dwarfing genes are the reduced height-1 (*Rht1*) gene in wheat (Peng et al., 1999), and the *semidwarf1* (*sd1*) mutation in rice (Ashikari et al., 2002; Sasaki et al., 2002; Spielmeyer et al., 2002) both of which were central to the Green Revolution of the 1960s and 1970s (Hedden, 2003). The wild-type alleles of these genes are involved in GA signalling (*RHT1*, a DELLA protein²) and biosynthesis (*SD1*, a GA-20 oxidase). It has also been reported that the semi-dwarfism of barley, *sdw1/denso*, widely introgressed into cultivars grown in Europe, probably results from a defect in an orthologue of the rice *Sd1* gene (Jia et al., 2011).

However, mutations in GA biosynthetic genes causing GA deficiency have not been useful in sorghum because they induce culm bending, which inevitably causes abnormal plant architecture (Ordonio et al., 2014). Rather, four independently inherited non-GA dwarfing mutations (*dw1–dw4*) have been used extensively in commercial grain sorghum breeding mainly in the United States to significantly reduce sorghum plant height to improve lodging resistance and machine harvesting (Quinby and Karper, 1954). As none of the mutations can individually provide sufficient height reduction for machine harvesting, three of them in various combinations are typically used to develop commercial lines. Because of its

² The DELLA proteins are conserved repressors of GA signalling that act immediately downstream of the GA receptor to modulate all aspects of GA-induced growth and development in plants.



ability to improve the harvest index of sorghum, the *dw3* mutation is often included in this combination; however, the commonly-used mutant allele, which contains an 882-bp tandem duplication in exon 5 is subject to a low frequency (0.1–0.5 percent) of reversion back to *Dw3* because of unequal crossing-over during meiosis (Multani et al., 2003).

The genes for three of the four dwarf loci in sorghum have been identified (Hilley et al., 2017; Hirano et al., 2017; Multani et al., 2003; Yamaguchi et al., 2016). *Dw1* encodes for a novel component of BR signalling (Hirano et al., 2017) that positively regulates BR signalling by interacting with brassinosteroid insensitive 2 (BIN2) kinase and inhibiting its nuclear localization (Gruszka, 2020). BRs play important roles in plant growth and development, regulating diverse processes such as cell elongation, cell division, photomorphogenesis, xylem differentiation, and reproduction (Nolan et al., 2020). The recessive *dw1* allele in Dwarf Yellow Milo sorghum, first identified by Quinby and Karper (1954), contains a stop codon in exon 2 that results in protein truncation (Hilley et al., 2016).

Dw2 encodes an AGC kinase³ involved in regulation of stem elongation (Hilley et al., 2017). The *Dw4* locus has been mapped to a region on chromosome 4, but the causal gene for *dw4* is as yet unidentified.

Dw3 is an orthologue to the maize *brachytic2* (*br2*) gene, which encodes for an ATP-binding cassette type B1 (ABCB1)⁴ auxin transporter (Multani et al., 2003). Newer alleles lacking the 882-bp tandem duplication in *dw3* but containing simple nucleotide changes (e.g., *dw3-sd1*, *-sd2*) have been identified that disrupt the DW3 reading frame resulting in a non-functional protein with a 200-amino acid truncation that is expected to confer a stable mutant phenotype (Multani et al., 2003; Barrero Farfan et al., 2012).

In the Arabidopsis auxin-transporting ABCB1, pro-1008 (P1008) is part of a conserved signature D/E-P motif in the C-terminal nucleotide-binding domain that seems to be specific for auxin-transporting ABCB1s (Hao et al., 2020). Mutation of Pro-1008 or the acidic residue (Asp or Glu) at position 1007 abolishes auxin transport activity by Arabidopsis ABCB1. All higher plant ABCB1s for which auxin transport has been conclusively proven carry the conserved D/E-P motif, including the sorghum DW3 protein (Multani et al., 2003). The predicted structural model for teff DW3A protein is analogous to the Arabidopsis auxin-transporting ABCB1 and contains 12 transmembrane helices and the C-terminal D/E-P motif where Pro-1089 corresponds to Pro-1008 in Arabidopsis ABCB1 (Figure 6, panel A). The mutated teff *dw3*-encoded protein lacks 611 amino acids from the C-terminus that include six additional transmembrane helices and the conserved D/E-P motif essential for auxin transporter activity (Figure 6, panel B).

The modes of action of the mutant alleles introduced into DPS 22-8-12 teff are analogous to the those of the *sd1* allele in rice and the *dw1* and the stable *dw3* alleles in sorghum. The

³ AGC kinases are the collective name for cAMP-dependent protein kinase A, cGMP-dependent protein kinase G, and phospholipids-dependent protein kinase C.

⁴ ABCB1 belongs to a large family of ATP-binding cassette transporters and is functionally conserved as an efflux transporter in plants, human, mouse, and *C. elegans* to transport various substrates across the membrane. In plants, ABCB1 is involved in long-distance auxin transport influencing morpho-physiological responses (Mohan et al., 2019).

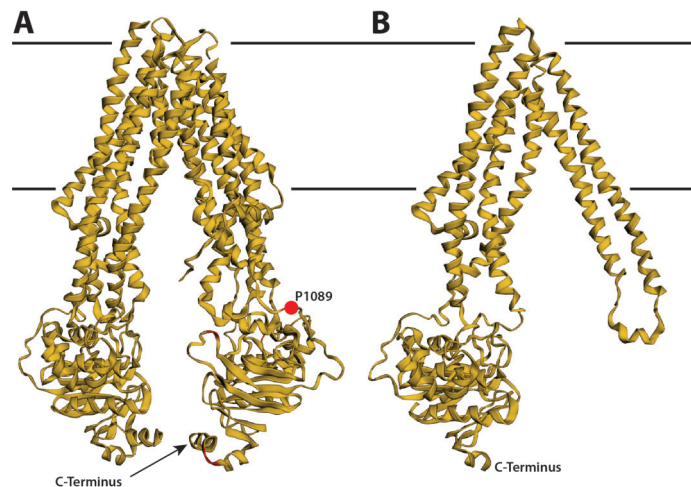


Figure 6. The amino acid sequences of the native teff DW3A protein (panel A) and the *dw3A* encoded protein containing the 611-amino acid C-terminal truncation (panel B) were subjected to protein modelling and structural prediction using the Phyre2 web portal (Kelley et al., 2016). The location of Pro-1089, which is part of the conserved D/E-P motif essential for auxin transporter activity is indicated. The entire C-terminal domain containing six additional transmembrane helices and the D/E-P signature motif is absent from the homeologous *dw3*-encoded sequences.

resulting phenotype is familiar based on a history of use of dwarfing genes in conventional breeding programs for many cereals.

5. Concluding Comments

Reduced plant height has been one of the main objectives of modern plant breeding (Denison et al., 2003) and is a familiar trait in many improved crop varieties. The use of CRISPR-Cas9 genome editing to specifically target *sd1*, *dw1*, and *dw3* knockout mutations, either individually or in various combinations, is a novel approach to introducing lodging-resistance into teff, but it is founded on a long history of experience and familiarity with similar naturally occurring and induced mutations in other crop species.

The approximate 50 percent reduction in plant height for DPS 22-8-12 observed under greenhouse conditions was within the range of natural variability reported for the plant species (Assefa et al., 2017). Furthermore, the reduced height phenotype is unlikely to effect the distribution, density, or development of the plant and its sexually compatible relatives; the production, creation, or enhancement of a plant pest or a reservoir for a plant pest; harm to non-target organisms beneficial to agriculture; or the weedy impacts of the plant and its sexually compatible relatives.

Finally, we are requesting that the RSR determination encompass modified teff lines containing sub-combinations of the *sd1*, *dw1*, and *dw3* mutations carried in DPS 22-8-12 teff, such as those illustrated in Figure 7.



REFERENCES

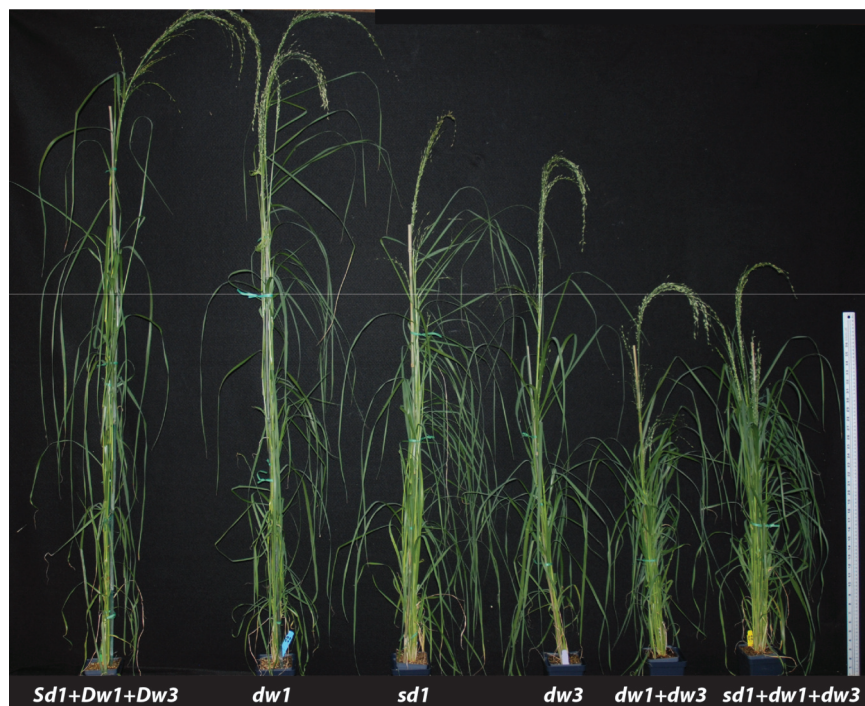


Figure 7. Comparison of plant heights of wild-type (Sd1+Dw1+Dw3) unmodified teff with edited lines containing individual *sd1*, *dw1*, and *dw3* knockout mutations, and lines with combined *dw1+dw3* and *sd1+dw1+dw3* mutations.

6. References

- Ashikari, M., Sasaki, A., Ueguchi-Tanaka, M., Itoh, H., Nishimura, A., Datta, S., Ishiyama, K., Saito, T., Kobayashi, M., Khush, G. S., Kitano, H., and Matsuoka, M. (2002). Loss-of-function of a rice gibberellin biosynthetic gene, GA20 oxidase (*GA20ox-2*), led to the rice 'green revolution'. *Breeding Science*, 52(2):143–150.
- Assefa, K., Cannarozzi, G., Girma, D., Kamies, R., Chanyalew, S., Plaza-Wuthrich, S., Blosch, R., Rindisbacher, A., Rafudeen, S., and Tadele, Z. (2015). Genetic diversity in tef (*Eragrostis tef* (Zucc.) Trotter). *Frontiers in Plant Science*, 6(MAR):1–13.
- Assefa, K., Chanyalew, S., and Tadele, Z. (2017). Tef, *Eragrostis tef* (Zucc.) Trotter. In Patil, J., editor, *Millets and Sorghum: Biology and Genetic Improvement*, pages 226–266. John Wiley & Sons Ltd., New York, first edition.
- Barrero Farfan, I. D., Bergsma, B. R., Johal, G., and Tuinstra, M. R. (2012). A Stable *dw3* Allele in Sorghum and a Molecular Marker to Facilitate Selection. *Crop Science*, 52(5):2063–2069.
- Beitz, E. (2000). TEXshade: shading and labeling of multiple sequence alignments using LATEX2 epsilon. *Bioinformatics (Oxford, England)*, 16(2):135–139.
- Denison, R. F., Kiers, E. T., and West, S. A. (2003). Darwinian agriculture: When can humans find solutions beyond the reach of natural selection? *Quarterly Review of Biology*, 78(2):145–168.
- Edgar, R. C. (2004). MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, 5:1–19.



REFERENCES

- Gruszka, D. (2020). Exploring the brassinosteroid signaling in monocots reveals novel components of the pathway and implications for plant breeding. *International Journal of Molecular Sciences*, 21(1).
- Hao, P., Xia, J., Liu, J., Donato, M. D., Pakula, K., Bailly, A., Jasinski, M., and Geisler, M. (2020). Auxin-transporting ABC transporters are defined by a conserved D/E-P motif regulated by a prolyl isomerase. *Journal of Biological Chemistry*, 295(37):13094–13106.
- Hedden, P. (2003). The genes of the Green Revolution. *Trends in Genetics*, 19(1):5–9.
- Hilley, J., Truong, S., Olson, S., Morishige, D., and Mullet, J. (2016). Identification of Dw1, a regulator of sorghum stem internode length. *PLoS ONE*, 11(3):1–16.
- Hilley, J. L., Weers, B. D., Truong, S. K., McCormick, R. F., Mattison, A. J., McKinley, B. A., Morishige, D. T., and Mullet, J. E. (2017). Sorghum Dw2 Encodes a Protein Kinase Regulator of Stem Internode Length. *Scientific Reports*, 7(1):1–13.
- Hirano, K., Kawamura, M., Araki-Nakamura, S., Fujimoto, H., Ohmae-Shinohara, K., Yamaguchi, M., Fujii, A., Sasaki, H., Kasuga, S., and Sazuka, T. (2017). Sorghum DW1 positively regulates brassinosteroid signaling by inhibiting the nuclear localization of BRASSINOSTEROID INSENSITIVE 2. *Scientific Reports*, 7(1):1–10.
- Jia, Q., Zhang, X.-Q., Westcott, S., Broughton, S., Cakir, M., Yang, J., Lance, R., and Li, C. (2011). Expression level of a gibberellin 20-oxidase gene is associated with multiple agronomic and quality traits in barley. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 122(8):1451–1460.
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2016). The PyMol web portal for protein modeling, prediction and analysis. *Nature Protocols*, 10(6):845–858.
- Ketema, S. (1983). *Studies of Lodging Floral Biology and Breeding Techniques in Tef (Eragrostis tef)*. Ph.d., University of London.
- Ketema, S. (1993). *Tef (Eragrostis Tef): Breeding, Genetic Resources, Agronomy, Utilization and Role in Ethiopian Agriculture*, volume 102. Institute of Agricultural Research, Addis Ababa.
- Lyons, E., Pedersen, B., Kane, J., Alam, M., Ming, R., Tang, H., Wang, X., Bowers, J., Paterson, A., Lisch, D., and Freeling, M. (2008). Finding and comparing syntenic regions among Arabidopsis and the outgroups papaya, poplar, and grape: CoGe with rosids. *Plant Physiology*, 148(4):1772–1781.
- Mohan, A., Dhaliwal, A. K., Nagarajan, R., and Gill, K. S. (2019). Molecular Characterization of Auxin Efflux Carrier- ABCB1 in hexaploid wheat. *Scientific Reports*, 9(1):1–14.
- Multani, D. S., Briggs, S. P., Chamberlin, M. A., Blakeslee, J. J., Murphy, A. S., and Johal, G. S. (2003). Loss of an MDR transporter in compact stalks of maize br2 and sorghum dw3 mutants. *Science*, 302(5642):81–84.
- Nolan, T. M., Vukasinović, N., Liu, D., Russinova, E., and Yin, Y. (2020). Brassinosteroids: Multidimensional regulators of plant growth, development, and stress responses. *Plant Cell*, 32(2):298–318.
- Ordonio, R. L., Ito, Y., Hatakeyama, A., Ohmae-Shinohara, K., Kasuga, S., Tokunaga, T., Mizuno, H., Kitano, H., Matsuoka, M., and Sazuka, T. (2014). Gibberellin deficiency



REFERENCES

- pleiotropically induces culm bending in sorghum: An insight into sorghum semi-dwarf breeding. *Scientific Reports*, 4:1–10.
- Peng, J., Richards, D. E., Hartley, N. M., Murphy, G. P., Devos, K. M., Flintham, J. E., Beales, J., Fish, L. J., Worland, A. J., Pelica, F., Sudhakar, D., Christou, P., Snape, J. W., Gale, M. D., and Harberd, N. P. (1999). 'Green revolution' genes encode mutant gibberellin response modulators. *Nature*, 400(6741):256–261.
- Quinby, J. and Karper, R. (1954). Inheritance of height in sorghum. *Agronomy Journal*, 46:211–216.
- Sakamoto, T., Miura, K., Itoh, H., Tatsumi, T., Ueguchi-Tanaka, M., Ishiyama, K., Kobayashi, M., Agrawal, G. K., Takeda, S., Abe, K., Miyao, A., Hirochika, H., Kitano, H., Ashikari, M., and Matsuoka, M. (2004). An overview of gibberellin metabolism enzyme genes and their related mutants in rice. *Plant Physiology*, 134(4):1642–1653.
- Sasaki, A., Ashikari, M., Ueguchi-Tanaka, M., Itoh, H., Nishimura, A., Swapan, D., Ishiyama, K., Saito, T., Kobayashi, M., Khush, G. S., Kitano, H., and Matsuoka, M. (2002). A mutant gibberellin-synthesis gene in rice. *Nature*, 416(6882):701–702.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7(539).
- Spielmeyer, W., Ellis, M. H., and Chandler, P. M. (2002). Semidwarf (sd-1), "green revolution" rice, contains a defective gibberellin 20-oxidase gene. *Proceedings of the National Academy of Sciences of the United States of America*, 99(13):9043–9048.
- USDA (2021). The PLANTS Database (<https://plants.usda.gov/home/plantProfile?symbol=ERTE>). Technical report, National Plant Data Team, United States Department of Agriculture, Greensboro, NC USA.
- VanBuren, R., Man Wai, C., Wang, X., Pardo, J., Yocca, A. E., Wang, H., Chaluvadi, S. R., Han, G., Bryant, D., Edger, P. P., Messing, J., Sorrells, M. E., Mockler, T. C., Bennetzen, J. L., and Michael, T. P. (2020). Exceptional subgenome stability and functional divergence in the allotetraploid Ethiopian cereal teff. *Nature Communications*, 11(1):1–11.
- Yamaguchi, M., Fujimoto, H., Hirano, K., Araki-Nakamura, S., Ohmae-Shinohara, K., Fujii, A., Tsunashima, M., Song, X. J., Ito, Y., Nagae, R., Wu, J., Mizuno, H., Yonemaru, J. I., Matsumoto, T., Kitano, H., Matsuoka, M., Kasuga, S., and Sazuka, T. (2016). Sorghum Dw1, an agronomically important gene for lodging resistance, encodes a novel protein involved in cell proliferation. *Scientific Reports*, 6(February):1–11.



Appendices

A. Gene Sequences

A.1. *Sd1* Gene, Chr 3A

```

LOCUS       Exported                2364 bp DNA       linear   PLN 02-DEC-2021
DEFINITION  Gibberellin-20 oxidase encoding gene from Eragrostis tef (Chr 3A).
ACCESSION   .
VERSION     .
KEYWORDS    Eragrostis tef Sd1 Gene (A Genome) Encoding GA20 Oxidase
SOURCE      Eragrostis tef
  ORGANISM  Eragrostis tef
REFERENCE   1 (bases 1 to 2364)
  AUTHORS   Donald Danforth Plant Science Center
  TITLE     Direct Submission
  JOURNAL   Exported Dec 2, 2021 from SnapGene 6.0.0
            https://www.snapgene.com
FEATURES             Location/Qualifiers
     source           1..2364
                     /mol_type="genomic DNA"
                     /organism="Eragrostis tef"
     CDS              join(1..674,855..1176,2098..2364)
                     /codon_start=1
                     /product="Gibberellin-20 Oxidase (SD1) from Eragrostis tef
                     (A Genome)"
                     /label=Sd1
                     /translation="MASPVAGPTGLPLSLLAPLLCYKYPTGPDRSPAHSPLAHLMVSQA
                     RQEPRNSGISKLVMDAIPAPLLLRSPAPGIDKIPAPFVWPQDDSRPTSAAAALDVPVVD
                     VGVLRNGGDAAGLRRAAAQVASACATHGFFQVCGHGVDAALARAALDGASDFRLPLAE
                     KQRARRVPGTVSGYTSAHADRFACKLPWKETLSFRFHDAGAASPVVADYFTSTLGPDFE
                     PMGRVYQRYCEEMKELSLTIMELLELSLGVVERGYRDFEFDSRSIMRCNYPPCPEPER
                     TLGTGPHCDPTALTILLQDDVGGLEVLVDGEWRPVRPVPGAMVINIGDTFMALSNGRYK
                     SCLHRAVVNRRQERRSLAFLFCPREDRVVRPPPAAPRRYPDFTWADLMRFTQSHYRA
                     DTRTLDAFTNWLARGPAQQQA "
     misc_feature     652..671
                     /label=Sd1 gRNA1
     misc_feature     672..674
                     /label=PAM
     misc_feature     953..972
                     /label=Sd1 gRNA2
     misc_feature     973..975
                     /label=PAM
ORIGIN
1  atggcgtccc ctgtggcagg ccccacaggg ttgcctttgt cgttgcttgc ccctctcctc
61  tgttacaagt accccaccgg tccggacagg tcccctgcac actcaccctc cgcacatctc
121 atgggtgtccc aagcacggca agagcctcgc aacagcggca tctcgaagct cgtcatggac
181 gccatcccgg ccctcctcct cctgctccgc tctccagctc ccggcatatga caagatcccc
241 gcgccgttgc tgtggccgca ggacgactcg cgcccagctg cggcggcggc gctggacgtc
301 cccgtggtgg acgtgggctg gctccgcaat ggccggcgacg ccgccgggct cgcggcgcgc
361 gcggcgcagg tggcgtcggc gtgcgcgacg cacggcttct tccaggtgtg cgggcacggc
421 gtggacgcgg ccctggcgcg cgccgcgctg gacggcgcac ccgacttctt ccggctgccc
481 ctggccgaga agcagcgcgc ccggcgcgctg ccaggcaccg tgtccgggta cagcagcgcg
541 cacgccgacc ggttcgcgtg caagctgccc tgggaaggaga cgctgtcctt ccgcttccac
601 gacgccggcg ccgcgtcgcc cgtggtcgcc gactacttca ccagcaccct cggcccggac
661 ttcgagccaa tggggtacgt aaccaaccac gttgactggt ttaattaatt actgcatata
721 tagtatccat catctccatg gatatatata tatatatata tatatatata tatatatata
781 tatatatata tatatatctt ggcatatgcat cattggacac gtacacacgg acggctgaaa
841 aaactatatg gcaggcgcgt gtaccagagg tactgcgagg agatgaagga cgtgtccctg
901 acgatcatgg agctgctgga gctgagcctg ggctgggagc gcggctacta ccgcgacttc
961 ttcgaggaca gccggtccat catgctgctg aactactacc cgccgtgccc ggagccggag

```



GENE SEQUENCES

```

1021 cgcacgctgg gcacgggccc gcactgcgac cccaccgcg caccatcct cctccaggac
1081 gacgtcggcg gactcgaggt cctcgtcgac ggcgaatgg gcccgtacg acccgtgccc
1141 ggcccatgg tcatcaacat cggcgacacg ttcatggat ggtattgtat tactctgctg
1201 ctcttggtg cctgcttgc tgcctgccc ctgtttctcc tcctaattgga agtgattgga
1261 cagtcgcaga ttacacacac tcacacgtag ctgcctaggc gtgtcatgtt gcgcaatgta
1321 gacatgcggt cgagttggtt ccatctcgat cgggtgggcta gctcggtagt acaggtctat
1381 caatcgggtg tggcgaacgg cgagccgctt tgatgatttc gccaaagaaa aaaagaaaag
1441 aaagagagag gttgcatgac agtaggagga caaagagaga gagagagaga gagagagaga
1501 gagaagaggt tggcttccgc cgttggcggc ttgacgattg tagaggagtg ggggaggccc
1561 ggaggaggag ggggaggata cgatgcccgc agcgggtggc ctttcgccgg gtggaccxaa
1621 gcccagggtg gcgctcctg cgtcttccct gttttggggc cgtccgggtc ggtggggcca
1681 tctcatctca tcccgcggcg tgactgactg tgcgctgcac gggccgccc gctggatcat
1741 acatgggctt ggctttacca ctggcagctt ctacttggtt gtcgctcgtc ggcttctagc
1801 cgtttccacg gcgtttgtcg cgtttttccg gtgcacgagg cggccgccc cagctcggcc
1861 gtgcgggtgg cagcacacca tcaccccgtc acccgactca cccgtgctg cggcttggtt
1921 ccgccacgcg ggcgaaaac aaaagggttc ctgtgccttg acaggtcgtc acgtgttgtg
1981 ttgtgcaaaa gaactccgga ttgagctaat tgtttgttcg ctactgttg attgcccgc
2041 cgatgcggcg cagcatctct gacttcagca actgactcct ctgctggtt aacgcaggcg
2101 ctgtcgaacg ggcggtatac gagctgcctg caccgcccgc tggatgaacc gcggcaggag
2161 cggcgtcgcg tggccttctt cctgtgcccg cgcgaggacc gcgtcgtcg ccccccccg
2221 gccgccgccc cgcgcgcccg ctaccgggac ttcacgtggg ccgacctcat gcgcttcacg
2281 cagagccact accgcccga cacacgcacc ctcgacgcat tcaccaactg gctcggcccg
2341 ggcccggccc agcagcaggc gtag
//

```

A.2. *Sd1* Gene, Chr 3B

```

LOCUS       Exported                2273 bp DNA       linear       UNA 02-DEC-2021
DEFINITION natural linear DNA
ACCESSION   .
VERSION     .
KEYWORDS    Eragrostis tef Sd1 Gene (B Genome) Encoding GA20 Oxidase
SOURCE      natural DNA sequence
  ORGANISM  unspecified
REFERENCE   1 (bases 1 to 2273)
AUTHORS     IICITI
TITLE       Direct Submission
JOURNAL     Exported Dec 2, 2021 from SnapGene 6.0.0
            https://www.snapgene.com
FEATURES    Location/Qualifiers
  source    1..2273
            /mol_type="genomic DNA"
            /organism="unspecified"
  CDS       join(1..672,794..1114,2007..2273)
            /codon_start=1
            /product="Gibberellin-20 oxidase from Eragrostis tef (B
            genome)"
            /label=Sd1
            /translation="MLVFVVARSPPLFSCYKYPTGPDRSPAHSHLAHLMVSQARQEP
            SGISKLAMDAIPAPPLLRSPAGIDLPKDKIPAFVWPPDDSRPTSAALDVPVVDVG
            VLRNGGDPAGLRRAAAQVASACATHGFFQVRGHGVDAALARAALDGASDFRPLAEKQ
            RARRLPGTVSGYTAHADRFAKLPWKETLSFGFHDGAAASPVVVDYFTSTLGPDFEP
            MGRVYQRYCEEMKDLSTIMELLELSLGVGERGYRDFEDSRSIMRCNYYPPCPEPERT
            LGTGHCDPTALTILLQDDVGGLEVLVDGEWRPVRPVPAMVINIGDTFMALSNGRYKS
            CLHRAVVNRRQERRSLAFFLCPREDRVRPPTTVAAPRRYPDFTWADFMRFTQSHYRAD
            TRTLDAFTTWLARGPPAQA"
  misc_feature 649..668
            /label=Sd1 gRNA1
  misc_feature 669..671
            /label=PAM
  misc_feature 891..910
            /label=Sd1 gRNA2
  misc_feature 911..913
            /label=PAM

```



GENE SEQUENCES

ORIGIN

```

1 atgcttgtgt ttgtcgttgc tcgctcccct cccctgttct cctgttataa ataccccacc
61 ggccccggaca ggtccccctgc acaactcacac ctcgcacatc tcatgggtgtc ccaagcacgg
121 caagagcctc gcagcagcgg catctccaag ctcgccatgg acgcatccc ggccccgcct
181 ctccctgctcc gctccccagc tcccggcatt gacctcccca aagacaagat ccccgcgccg
241 ttcgtgtggc cgccggacga ctcgggccg acgtcggcgg cggcgctgga cgtccccgtg
301 gtggacgtgg gcgtgctgcg caatggcggc gaccccgccg ggctgcggcg cgcggcggcg
361 caggtggcgt cggcgtgcgc gacgcacggc ttcttcagg tgccgggca cggcgtggac
421 gccgccctgg cgcgcgcggc gctggacggc gccagcgact tcttcggct gccgctcgcc
481 gagaagcagc gcgcgcggcg actccccggc accgtgtccg ggtacacgag cgcccacgcc
541 gaccggttcg cgtgcaagct cccgtggaag gagacgctgt ccttcggctt ccacgacggc
601 gccgctgcgt cgccggccgt cgtcgtcgac tacttcacca gcaccctcg cccggacttc
661 gagccaatgg ggtaacatat aaacactact ggccggcctt ctgtggatat agtatcttgt
721 cgcttatgca ttgacaccg aactagctt gggacgactg gaacctcca tcatgtacta
781 cttatatggc aggcgcgtgt accagaggta ctgagaggag atgaaggacc tgccttgac
841 gatcatggag ctgcttgagc tgagcctggg cgtggagcgc ggctactacc gcgacttctt
901 cgaggacagc cggtcacatca tgcggtgcaa ctactaccg ccgtgcccgg agccggagcg
961 cacgctgggc acgggcccgc actgcgaccc cacggcgctc accatcctgc tccaggacga
1021 cgtcggcggga ctcgaggtcc tcgtcgacgg ggaatggcgc ccgtcccgc cctgcccgg
1081 agccatggtc atcaacatcg gcgacacggt catggtataa ttcaatttac tctgttcttt
1141 ggctggcctg tagcttgtcc tgccctgttc ccgatgtcct cctagctaat ggaactaatt
1201 ggacagcgct cagattacac acacacgtag gcgtgtcatg ttgcgcaatg tagacatgcg
1261 ttcgagttgg ttccatctcg atccgtgggc tagctcggta gtaggtctac cgggtgggtg
1321 gaacggtgag ccgccttgat ggttcgcca aagaagaaaa aagagagaga gagagagaga
1381 gagagaggta gcgatgcagt aggaggacaa agagagagag agaagaggtt ggcttgtgcc
1441 gttggcggct tgcgagttgt agaggagtgg gggaggaccg gaggaggagg gggatacgat
1501 gccgcgagcg gtggcgcttt cgccctgggg cccaagccc aggtgcgcgc gcgctcgtt
1561 tcttccctgt ttggggcccg tccggggccc tctcatccgc ggctgactg tgcgccgac
1621 gggccgcccgc gctggagcac aggatcgtca gcttggcgtt acccgctggt atgctctg
1681 taattctact tggctgtcgc tcgcccgggt ctagccggtt ccgcccgtt tggcgggtt
1741 tcccgtgca cgccggccc gcgcccgggc tccagctcgg tgcgggtggc actacgctag
1801 cgtatagtac gcgcgccgca acaaaagggt tcctgtgcct tgacaagcct tcaactcagt
1861 ctagtgtttg cacaaaaggg gggggtttcc ggattgagct gattgttgtt tcgtccagt
1921 gttcactggt gatctgcac tctccctgac tgactgactg actgactgac tgacttctt
1981 actggttgtt tgcacactgg caacaggcgc tgtccaacgg gcggtacaag agctgcctgc
2041 accgcgaggc ggtgaaccgg cggcaggagc ggcggtcgct ggcccttctc ctgtgcccgc
2101 gcgaggaccg cgtcgtgcgc ccgccagca ccgtcgccgc gccgcgccgc taccggact
2161 tcacgtgggc cgacttcacg cgcttcacgc agagccacta ccgcgccgac acacgaccc
2221 tcgacgcctt caccacctgg ctgcccgcg gcccgccggc ccagcaggcg tag

```

//

A.3. *Dw1* Gene, Chr 4A

```

LOCUS      Exported                2501 bp DNA      linear    PLN 02-DEC-2021
DEFINITION Dw1 gene (Chr 4A) encoding protein involved in brassinosteroid
            signalling.
ACCESSION  .
VERSION    .
KEYWORDS   Eragrostis tef Dw1 Gene (A Genome)
SOURCE     Eragrostis tef
ORGANISM   Eragrostis tef
REFERENCE  1 (bases 1 to 2501)
AUTHORS    Donald Danforth Plant Science Center
TITLE      Direct Submission
JOURNAL    Exported Dec 2, 2021 from SnapGene 6.0.0
            https://www.snapgene.com
FEATURES   Location/Qualifiers
            source          1..2501
                        /mol_type="genomic DNA"
                        /organism="Eragrostis tef"
            CDS             join(1..146,848..2060,2274..2501)
                        /codon_start=1
                        /product="DW1 protein involved in brassinosteroid
                        signalling."

```



```

/label=Dw1
/translation="MASAGSSSSGGAGGGSSIRAANGAAAISAAATAVGSADARFHSHP
QDRRSRWAGCLSLGSCFGSKGKRIVPAARTPDGNGSSARGNGHQSGSNSNQVPLN
LSLLAPPPSPASFSNSALPSTVQSPNFLSISANSPGGPNSMFAVGPYANEPQLVSP
VFSTYTTEPSTAPLTPPELAHATTPSSPDVPYARFLSSFMIDIKTASKEHNMAFLSTTY
SGGSLQASYPLYPESPCSSLISPASATPRTGLSSPIPEQEVPTAHWKTSRSACDTPYS
RASPIPEQEATAQWKTSRSACDTPYARASPSNIFGLDSSASRNYLLDGNFFRPAASAQF
YLDQAQQTYPYNGRRSVSRDKQDADEVEAYRASFGFSADEIMQTSYVEIPDALDES
SISPPFGNAPATEVSPFNDPPNEVQKAEKSSPKKTADQISNGSPHRVLHIDIFKGTGG
HPSEYEGIVKDGHPFRKTRDEISLKPIEVRRKSSPPGHSCSDAEIEYRRARSLREANGV
SWRSTLSRQLQ"
misc_feature      1525..1527
                  /label=PAM
misc_feature      1528..1547
                  /label=Dw1  gRNA1
ORIGIN
    1 atggcctcag ctgggagcag cagtggaggt gcaggtggag gcagcagcat cagggcggcg
    61 aatggcgcag ctgccatcag tgcggccgcg acggcagtag gctcagccga cgccagattc
   121 cactcccacc caccacagca ggacagggta tatacactac acgacttctt gttccatggc
   181 gcatgcaatt gctcatgctt tagctgcagc ttttacctt cttgctttag atgtgtggat
   241 gagtaccagc aatttctagt aggccttgac gacctagatt gtcacaggct ttgcgtactt
   301 agtacgaatg gtactgccca gtggtgggtg cagcttgcaa ccccggcatac tttaccccca
   361 ataattagcg ttcagtagtg tctgcttggt tgcttctagt actgaatagt ttggtcagct
   421 gaattttccg ttttgctgtt gatatttgat tattacgcca ttgcaactcc ctctatgttt
   481 ttccccccgg atcctagatg actgtactct taccataaca ctctgtccaga gacctatatg
   541 ttgtgtaact gttgatttga ccttgcatat acttccacac gtcgtgtcag ttgatgtcat
   601 gagcaaaact actatgctgg gttgattcct tcttgcatga tggatctttt tccaaatgaa
   661 ttttcagatt cctgtaaaat gcttggtggg gtagctatta atgtcaacac atttcttata
   721 ttcttaagta aaattggttt agattaatac tccaaggaaa tttcgttggt gaacaactag
   781 tcatacatat tttcccctac caacagctat gtacttgata ctaatatgtg catctctgat
   841 ctaaacagcg gagtagatgg gctggctgct tgcggggtct tcatgtttc ggatctcaga
   901 agggcgggaa gaggattgtt cctgccgcac gtactcctga tgggaatgga tcaagtgtct
   961 gtggaaatgg tcatcagtc ggttctaatt caaaccaaaa tgtgcctta aatctgtctc
  1021 ttctggctcc accatcctca ccggcatcct tctcaaatc tgcacttctc tcaactgttc
  1081 aatcgccctaa taactttctg tcaatctcag caaatctctc tgggtgtcca acatctaata
  1141 tgtttctgtg tgggccatat gctaataaac ctacagctgt ctgcctcct gtcttctcaa
  1201 cttacacaa tgagccttcc acagcaccat tgaccctcct acctgaacta gctcatgcaa
  1261 ccactccctc atctccagat gttccatacg ctcggtttct tcttctttt atggatatca
  1321 aaactgctag taaggagcat aacatggcct tcttatcaac aacatattct ggtggttcag
  1381 gactccaggc atcataccct cttaccctg agagcccttg tagcagcctc atatcacctg
  1441 cttctgctac tccgaggact ggtctttcct caccaatacc tgaacaagag gtccctactg
  1501 cccattggaa gacttctagg tctgcctgtg acacaccata ttccagggtc tcaccatcc
  1561 ctgagcaaga agccactgag cagtgggaaga cttctagatc ggcattgtat acgcttatg
  1621 ccagggtctc accatcaaac atctttgggc tggactcatc tgcttctaga aactatttgt
  1681 tagatggcaa ctcttccgg ccagctgctt ctgctcaatt ctacctggac caggctcagc
  1741 agacatatcc atataatggt gggaggcgta gtgtctcagc ggacaagcaa gatgcagatg
  1801 aagttgaagc ttacagagct tcgtttggtc ttagtgcaga tgaatcatg caaactcaat
  1861 cttatgtgga gataccagat gcacttgatg aatcattcag tatatcgcca tttgaaata
  1921 atgctcctgc tactgaggtg tccccattta atgatccacc caacgaggtt cagaaggctg
  1981 agaagtcaag tccaaagaaa acggcagatc agatttccaa tggctctcca catagagtgc
  2041 tgcacattga catattcaag ggtaaatgtc catactttt gcaccaacca ttcagtgtga
  2101 atctaggcat agtagaatta ctctgttata tccttgctat ctactgcacc taacaagctc
  2161 ttggtaaact tcaatgctaa taacctgaac tctgggccat ttgtcatcag cacttattgg
  2221 tattgtgttt ctatttatgt aatgcttat tatgcttatg tgctcgatt tcaggaacaa
  2281 aaggaggaca tcgctctgag tacgagggta ttgtgaaaga tggccatcct ttcagaaaga
  2341 caagggacga aatatctcta aaaccatag aagtaaggaa gaaatctcca cctgggcatt
  2401 catgctcaga tgctgaaatt gagtacagaa gggcaaggag tctgagggaa gccaatggtg
  2461 tcccctcgtg gcgcagcagc ttgtcaagac agctgcagtg a
//

```

A.4. Dw1 Gene, Chr 4B

LOCUS Exported 2561 bp DNA linear PLN 02-DEC-2021
DEFINITION Dw1 gene (Chr 4B) encoding protein involved in brassinosteroid



GENE SEQUENCES

```

1741 tctgcttcta gaaactatTT gttagatggc aacttcttcc ggccagctgc ttctgctcaa
1801 ttctacctgg accaggctca gcagacatat ccatataatg gtgggaggcg tagtgtctca
1861 cgggacaagc aagacgcaga tgaagttgaa gcttacagag cttcgtttgg ctttagtgca
1921 gatgaaatca tgcaactca atcttatgtg gagataccag atgcacttga tgaatcattc
1981 agtatatcac catttgaaa taatgctcct gctactgagg tgtccccatt taatgatctg
2041 cccaatgagg ttcagaaggc tgagaagtca agtccaaaga aatcagcaga tcagatttcc
2101 aatggttctc cacatagagt gctgcacggt gacatattca agggtaaag tccatacttt
2161 ttgcatcaac cttcagttg taatctaggc gtagtagaat tactcctttt gttatccttg
2221 ctatctactg cacctaactg gtaacttcca atgctaataa cctgaactct gggccattgg
2281 tattgtgttt ctgtttacgt aaatgcttat catgcttatg tgctcacatt tcaggaacaa
2341 aaggaggaca tccgtttgag tacgagggga ttgtgaaaga tggccatcct ttcagaaaga
2401 caagggatga gatatctcta aaacccatag aagtaaggaa gaaatctcca cctggtcatt
2461 catgctcaga tgctgaaatc gactacagaa gggcaaggag tctgagggaa gccaacgggtg
2521 tcccctcgtg gcgcagcacg ttgtcgagac agctgcagtg a

```

//

A.5. Dw3 Gene, Chr 8A

```

LOCUS       Exported                               5181 bp DNA       linear     PLN 02-DEC-2021
DEFINITION  Dw3 gene encoding an ATP-binding cassette type B1 (ABCB1) auxin
            transporter.
ACCESSION   .
VERSION     .
KEYWORDS    Eragrostis tef Dw3 Gene (A Genome)
SOURCE      Eragrostis tef
            ORGANISM  Eragrostis tef
REFERENCE   1 (bases 1 to 5181)
            AUTHORS   Danforth Plant Science Center
            TITLE     Direct Submission
            JOURNAL   Exported Dec 2, 2021 from SnapGene 6.0.0
            https://www.snapgene.com
FEATURES    Location/Qualifiers
            source     1..5181
                    /mol_type="genomic DNA"
                    /organism="Eragrostis tef"
            CDS         join(1..1498,1596..1825,2845..5181)
                    /codon_start=1
                    /product="ATP-binding cassette type B1 (ABCB1) auxin
                    transporter."
                    /label=Dw3
                    /translation="MSSSDPEEIRGRVVVLGADADELARPELEAFHLPSPPTLAQEAGSV
                    GPGPAAAAAPVVAPLETLPNATMPSSSSASSNSNSNEQNKEQENTKKKGVSLAPAPL
                    GSLFRFADGLDCVLMVGTGLGALVHGCSLPVFLRFFADLVDVDFGSHADDDPDMVRLVAK
                    YALYFLVVGAAIWASSWAEISCWMTGERQSTRMRIRYLESALRQDVSSFDDTDVRTSDV
                    IYAINADAVIVQDAISEKLGNIHYMATFVAGFVVGFTAAWQLALVTLAVVPLIAVIGG
                    LSAAALAKLSSRSQDALAEASNIAEQAVAQIRTVQAFVGEERAMRAYSLALAAAQRIGY
                    RSGFAKGLGLGGTYFTVFCCYALLLWYGGLLVRRHHTNGGLAIATMFSVMIGGLALGQS
                    APSMAAFAKARVAAAKIFKIIDHKPLSVVHGDDDVQLPVSVTGRVEMRGVDFAYPSRPD
                    VPVLRGFSLTVPPGKTIALVGSSESGSKSTVVSLIERFYDPSAGEIILDGHDLKSLNLRW
                    LRQQIGLVSQEPTLFATSIKENLLLGRDSSHATLAEMEEAARVANAHSFIIKLPDGYDT
                    QVGERGLQLSGGQKQRIAIARAMLKNPGILLLDEATSALDSESEKLVQEALDRFMIGRT
                    TLVIAHRLSTIRKADLVAVLHGGAVSEIGTHEELMGKGEDGAYARLIRMQEQAQAEVAA
                    RRSSARNSVSARNSVSSPIMTRNSSYGRSPYSRRLSDFSNADPFHYHGGGELPEGNTKKM
                    IHQRVAFRAGASSFLRLAKMNSPEWGYALVGSLSGSMVCGSFAIFAYVLSAVLVSYYAP
                    DPGHMRREIAKYCYLLMGMSAALVCNTVQHVFWDTVGENLTKRVRERMLGAVLRNEMA
                    WFDAENASARVAARLALDAQNVRSAGDRISVIVQNSALLLVACTAGFVLQWRALALVL
                    LGVFPLVVAATVQLKMFMKGFGDLEAAHARATQIAGEAVANLRTVAAFNAEAKIAGLF
                    AGNLRGPLRRCLWKGGQVAGCGYGVAQFLLYASYALGLWYAAWLKXGVSDFSRITRVFM
                    VLMVSANGAAETLTLAPDFVRRGRAMRSVFETIDRRTEADPDDPAAPLQLPLLTGVEL
                    RHVDFCYPSPREPVQLQDLSLRARAGKTLALVGPSPGCGKSSVLALIQRFYEPTSGRVLL
                    DGRDARKYNLRALRRAVAVVPQEPFLFAASIHNDIAYGREGGATEAEVLEAATQANAHK
                    FISALPDGYRTQVGERGVQLSGGQRQRIAVARALVKQAAVLLLDEATSALDAESERSVQ
                    QALDRHAKTRSTTTIIVVAHRLATVRNAHTIAVIDEGKVVEQGSHELLNHPDGTYARM
                    LQLQLTSSTS"

```




GENE SEQUENCES

```
misc_feature      3000..3002
                  /label=PAM
misc_feature      3003..3022
                  /label=Dw3  gRNA
ORIGIN
1 atgtctagca gcgacccgga ggagatccgg gggcgctgg tcgtcctcgg cgccgatgcg
61 gacgagttgg ctgcgccga gctggaggcc ttccacctcc cgtctccgac actagctcag
121 gaggcaggat ctgtcggagg cccaggccct gctgctgctg ctcccgtggt ggcgccattg
181 gagacggagc tgccctccaa tgccaccatg ctttcttctt cctccgcctc ctgcaacagc
241 aacagcaatg agcagaacaa ggagcaggag aatacgaaga agaagggggt gtctttggcg
301 cctgctcctc ttggttcgct gttccgtttc gccgacggtc tggactgctg tctgatgtcg
361 gtgggcacgt tgggcgcgct tgtccacggt tgctccctcc cgggtgttct cgccttcttc
421 gcggacctcg tcgactcgtt cggctcccac gccgacgacc cggacaccat ggtgaggctg
481 gtggccaagt acgcgctcta cttcctgggt gtgggcgcgg ccatctgggc gtcctctggg
541 gcggagatct cgtgctggat gtggaccggc gagcggcagt cgacgcggat gtcgatccgg
601 tacctggagt cggcgctccg gcaggacgtg tccttcttct acaccgactg ccgacctcc
661 gacgtgatct acgcatcaa cgcggacgcc gtgatcgtgc aggacgccat cagcgagaag
721 ctgggcaacc tcatccacta catggccacc ttctgtggcg ggttcgtggt cgggttcacg
781 gcggcatggc agctggcgcct ggtgacgctg gccgtgggtc cgtctatcgc cgtgatccgg
841 ggcctcagcg cggcggcgcct ggcaagctg tcgtcgcgga gccaggacgc cgtcggcggg
901 gcgagcaaca tcgaggagca ggcggtggcg cagatacgga cgggtcaggc cttcgtcggg
961 gaggagcgcg caatgcccgc ctactcgttg gcgctggccg cggcgacagag gatcggatac
1021 cgcagcggct tcgccaaggg cctgggcctg ggcgccacct acttcaccgt cttctgctgc
1081 tacgcgctgc tgctctggta cggcggcctc ctcgtcgcac gccaccacac caacggcggc
1141 ctgcacctcg ccacctggtt ctccgtcatg atcgggggcc tcgctgctgg ccagtcggcg
1201 cccagcatgg ccgcttccgc aaaggcgcgc gtcgcccggg ccaagatctt caaaatcatc
1261 gaccacaagc ccctctcctg cgtcgtccat ggcgacgacg atgtccaact cccttccgtg
1321 acggggcggg tggagatgcg gggcgttgac ttcgcctacc cgtcgcggcc ggagctcccc
1381 gtcctgcgcg gcttctccct caccgtgccc cccggcaaga ccatcgcctc cgtcgggagc
1441 tcaggctccg ggaagagcac ggtcgtctcc ctcatcgaga gtttctacga cccaagtgca
1501 ggtacacaac tacttaacgt cttagaagtg tctgtcagtc aattgtcttc ttcattcaca
1561 atcgtattca ttgattgaaa gaaaaatatg cttctcaggc gagattctgc ttgacgggca
1621 cgacctcaag tcctcaacc tgcggtggct ccggcagcag atcgggctgg tgagccagga
1681 gccgacgctg ttcgccacca gcatcaagga gaacctgctg ctgggcccgg acagccacag
1741 gccaccctc gccgagatgg aggaggccgc aagggtcgc aacgcccact cattcatcat
1801 caagctgccc gacggatacg acaccagggt atgtactagc accgttgctt tcctttcaca
1861 ttctctctc ctctgagca actagatatg tactagcacc attgtccat ctcaatgcca
1921 ccatcctcaa aaaagaagga aagcttttct ccatctcatc tccatctct cattcctggc
1981 gctcgtgccc tgctcagat ccatgggcgc aagtggacgg gcacgcacga gacgtggcta
2041 gtagtactcc agtatagcaa tgcttattag cctcactcac tcaccagtgc tatactagta
2101 gggcagccag cagcaacaca tgacacggga ggagacagca ccaccggccc ggggtccctc
2161 ccattgttag gttaggagcc tgttctgttc acctgctgtt gcgagttaac aacatgtaac
2221 atcctaaaga ttcttcttc cttcctttcc acaacacagg tcgctttctg acttctgtc
2281 ttctgttct cgacacaatg caatgcagca gctagcagcc catcacctag tcctcaccca
2341 tctttctgta gtactgctc cattttcacc atggcttggc tcaccggccg ggggtgtagt
2401 agaaagtaga attcccagga atctagtaag agaacaagtc ccaagttggc attttgtcag
2461 gaatgtcaag ttgocacttg cgagatccca tcccaagaat ttcttgcaat gggagaggag
2521 aggaccatgg ccgttgtgtt tattttactg tcaagcgcag gcatagggcg gttcctcctt
2581 ccttgagtca ttaacacaaa tcctgcgtgg agtgtattaa caagagtgtg tagcactacc
2641 aaaccacatg gccactgac tctgcagctg aggagttaga aatagtgca gttgtggaggc
2701 cggcccatcc tcctatgac atgacatgat ccgacggaga tagtagaata gatccgacat
2761 gatattttt ttaattcatg acaatcatga taaaaaaaaa gcaaagtcaa acaaggtcaa
2821 ttatgataaa aaaaatgcaa ttgacaggtc ggtgagcgcg gctcgcagct gtctggcggt
2881 cagaagcagc gcatcgccat tgcccgcgcc atgctcaaga acccggcat cttgctgctg
2941 gacgaggcga cgagcgcgct ggactctgag tcggagaagc tgggtcagga ggcctggac
3001 cggttcatga tcgggcgcac gacctggtg atcgcgcacc ggctgtcag gatccgcaag
3061 gccgacctgg tggcgggtgct gccggcggc gcctgtcgg agattgggac gcacgaggag
3121 ctgatgggca agggcgaggga cggcgcgtac gcgcgctga tcgggatgca gacgagggcc
3181 gcgaggagg tggccgcccg tcggagcagc gcgcgcaact cgggtgagcg gcgcaactcg
3241 gtgagctcgc ccatcatgac gcgcaactcc tcctacgggc ggtcgcctca ctcgcccgc
3301 ctctccgact tctccaacgc cgacttccac taccacggcg gtggtgaaat accggagggt
3361 aataccaaga agatgattca tcagagggtt gcgttcggg cggggggcag cctttctctg
3421 cgtctggcga agatgaactc gcccgagtgg gggtagcctc tgggtggctc gctgggctcc
3481 atgggtgtcg gtccttcag cgccatcttc gcctacgttc tgagcgcctg gctgagcgtg
```



```

3541 tactacgccc ccgaccggg ccacatgctg cgggagatcg ccaagtactg ctacctgctg
3601 atgggcatgt cttccgctg gctgggtgtg aacacgggtg agcacgtgtt ctgggacacg
3661 gtgggcgaga acctgacgaa ggcgctccgg gagcgcgatg tggcgccgtg gctccgcaac
3721 gagatggcgt ggttcgacgc cgaggagaaac gccagcgcgc ggggtggccgc gaggctggcg
3781 ctggacgcgc agaactgctg ctccgccatc ggcgaccgca tctcgggtgat cgtgcagaac
3841 tcggcgcctg tgctgggtgg gtgcacggcc gggttcgtg tgcagtggcg cctcgccctg
3901 gtgctgctgg gcgtgttccc gctgggtggg gccgccaccg tgctgcagaa gatgttcatg
3961 aagggcttct cgggggacct ggaggcggcc cacgcccgcg ccacgcagat cgccggcgag
4021 gccgtcgcca acctgcgcac cgtcgccgcg ttcaacgccc aggccaagat cgcgggactc
4081 ttcgcgggca acctcggggg cccgctgctg cgggtgctctt ggaaggcca ggtggccggg
4141 tgccgctacg ggggtggcga gttcctgctg tacgctcct acgctggg gctctggtac
4201 gcggcgtggc tggatgaagca cggcgtctcc gacttctccc gcaccatccg cgtcttcatg
4261 gtgctcatgg tctccgcaa cggcgcggcc gagacgctca cgtggcgcc ggacttctgct
4321 agggcgggac gcgccatgct cctcgtcttc gagaccatcg acaggcgac cgaggcggac
4381 cccgacgacc cggacgcggc gccattgcaa ctgccattat tgacagggtt ggagctgctg
4441 cacgtggact tctgctaccg gtcgcccggc gaggtgcagg tgctgcagga cctcagcctg
4501 cgccgcgcgc cggggaagac gctggcgctc gtggggccca gcggctgctg caagagctcc
4561 gtgctggcgc tcatccagcg cttctacgag cccacctcgg gccgctcct cctcgacggc
4621 cgggacgcgc gcaagtacaa cctccggggc ctgcccgcgc cgtcgcctg ggtgccgacg
4681 gagcccttcc tgttcgccc gagcatccac gacaacatcg cctacggcgc cgaggggggc
4741 gccacggagg cggaggtgct ggaggcggcc acgcaggcca acgcgcaaa gttcatctcc
4801 gcgctgccc acggctacag gacgcaggtc ggggagcgcg ggggtgcagct ctccggcggc
4861 cagcggcagc gcatgcctg ggcgcgcgc ctcgtcaagc aggcgcctg tctgctgctg
4921 gacgagcca ccagcgcct ggacgcccag tcggagcgca cgtgcagca ggcgctggac
4981 gccacgcca agaccgcag caccaccacc atcgtcgtc gcaccgcct cgccaccgtg
5041 cgcaacgcc acaccatgc cgtcatcgac gagggcaagg tcgtcgagca gggatcgac
5101 tcgcacctg tcaaccacca cccggacgga acctacgcgc gcatgctcca actacagcga
5161 ctcacctcat ccacttcta a

```

//

A.6. Dw3 Gene, Chr 8B

```

LOCUS       Exported                               5238 bp DNA       linear     PLN 02-DEC-2021
DEFINITION  Dw3 gene encoding an ATP-binding cassette type B1 (ABCB1) auxin
            transporter.
ACCESSION   .
VERSION     .
KEYWORDS    Eragrostis tef Dw3 Gene (B Genome)
SOURCE      Eragrostis tef
ORGANISM    Eragrostis tef
REFERENCE   1 (bases 1 to 5238)
AUTHORS     Donald Danforth Plant Science Center
TITLE       Direct Submission
JOURNAL     Exported Dec 2, 2021 from SnapGene 6.0.0
            https://www.snapgene.com
FEATURES    Location/Qualifiers
     source          1..5238
                   /mol_type="genomic DNA"
                   /organism="Eragrostis tef"
     CDS             join(1..1472,1575..1803,2896..5238)
                   /codon_start=1
                   /product="ATP-binding cassette type B1 (ABCB1) auxin
                   transporter."
                   /label=Dw3
                   /translation="MSASSDPEEIRGRVVVLGADADELARPELEAFHLPSPITLEEAGAS
                   VGGPAAANPPVVAPLETLPNAMPSSSSSSNSNSNEQNKEEKKKGVALAPAPLGLSLFR
                   FADGLDCVLMVGTGLGALVHGC SLPVFLRFFAELVDSFGSHADDPDTMVRLVARYALYF
                   LVVGAAIWASSWAEIS CWMW TGERQSTRMRIRYLESALRQDV SF FDTDVRTSDVIYAIN
                   ADAVIVQDAISEKLNLIHYMATFVAGFVVGFTA AWQLALVTLAVPLI AVIGLSAAA
                   LAKLSSRSQDALAEASNIAEQAVAQIRTVQAFVGEERAMRAYSLALGMAQRIGYRSGFA
                   KGLGLGGTYFTVFCCYALLLWYGGHLVRRHHTNGGLAIATMFSVMIGGLALGQSAPSMA
                   AFAKARVAAAKIFRIIDHLAVVHGDHVQLPSVTGRVEMRGVDFAYPSRPDIPVLRGFSL
                   TVPPGKTIALVGS SSGKSTVSLIERFYDPSAGEILLDGHDLKSLNLRWLRQQIGLVS
                   QEPTLFATS IKENLLLRDSQSATLAEMEEAARVANAHSF I I KLPDGYDTQVGERGLQL

```



SGGQKQRIATARAMLKNPGILLLDEATSALDSESEKLVQEALDRFMIGRTTLVIAHRLS
 TIRKADLVAVLHGGAVSEIGTHDELMAKGEDGAYARLIRMQQEQAAAQVAARRSSARP
 SSARNSVSSPIMTRNSSYGRSPYSRRLSDFSNADSHHYHGGELIESNNKAHHQRRIA
 FRAGASSFWRLAKMNSPEWGYALVGLSGSMVCGSFSIFAAYALSAVLSVYYAPDPGYMR
 RQIGKYCYLLMGMSAALVFNTVQHAFWDTVGENLTKRVRERMFGAVLRNEVAWFDAEE
 NASARVAARLALDAQNVRSAIGDRISVIVQNSALLLVACTAGFVLQWRLALVLLAVFPL
 VVAATVLQKMFMKGFGSDLEAAHARATQIAGEAVANLRTVAAFNAQAKIAGLFAANLRG
 PLRRCFWKQAAGCGYGLAQFLLYASYALGLWYAAWLVKHGVSDFSRAIRVFMVLMVSA
 NGA AETLTLAPDFVKGRAMRSVFETIDRRTEPDDPDAAPLPSDAVSVELRHVDFCY
 PSRPDVRVLQDLSLRARAGKTLALVGPSPGCGKSSVLALIQRFYEPTSGRVLLDGRDARK
 YNLRALRRAIAVVPQEPFLFAATIHNDIAYGREGATEAEVVEAATQANAHKFISALPDG
 YRTQVGERGVQLSGGQRQRIAVARALVKQAAVLLDEATSALDAESERSVQALDRHAK
 TRSTTTIVVAHRLATVRDAHTIAVIDDGKVVVEQGSHELLNHPDGTYARMLHLQRLTA
 PSTS"

misc_feature 3051..3053
 /label=PAM
 misc_feature 3054..3073
 /label=Dw3 gRNA

ORIGIN

1	atgtcagcta	gtagcgaccc	ggaggagatc	cggggccgcg	tgggtggtcct	cggcgccgat
61	gccgacgagt	tggctcgccc	cgagctggag	gccttcacc	tcccatctcc	gacattagag
121	gaggcaggag	cctctgtcgg	aggccctgct	gctgctaatac	ctccggtggt	ggcgccattg
181	gagacggagc	tgcctcccaa	tgcgatgcct	tcttcttct	cctcctcaa	cagcaacagc
241	aatgagcaga	acaaggagga	gaagaagaag	gggggtggcgt	tggcgctgc	tcctcttggg
301	tcgctgttcc	gtttcgccga	cggtcttgac	tgcgttctga	tgtcggtggg	cacgttgggc
361	gcgctgttcc	acggctgctc	cctccccgtg	ttcctcgcct	tcttcgcgga	gctggtggac
421	tcgttcggct	cccacgccga	cgaccggac	accatgggtgc	ggctggtggc	caggtacgcg
481	ttgtattttc	tgggtggtggg	cgcggccatc	tgggctcct	cctgggcgga	gatctcgtgc
541	tggatgtgga	cgggggagcg	gcagtcgacg	cggatgcgga	tccggtacct	ggagtcggcg
601	ctccgtcagg	acgtgtcctt	cttcgacacc	gacgtccgca	cctccgacgt	gatctacgcc
661	atcaacgccg	acgccgtcat	cgtgcaggac	gccatcagcg	agaagctcgg	caacctcatc
721	cactacatgg	ccaccttctg	ggcgggattc	gtgggtggggt	tcacggccgc	ctggcagctg
781	gcgctgggta	cgctggccgt	ggtgccgctc	atcgccgtga	tcggcggcct	cagcgcggcg
841	gcgctggcga	agctgtcgtc	gcggagccag	gacgcgctgg	cggaggcgag	caacatcgcg
901	gagcaggcgg	tggcgcagat	acggacgggtg	caggcgcttcg	tcggggagga	gcgcgcaatg
961	cgggcctact	cgttggctct	aggaaatggcg	cagaggatcg	ggtaccgcag	cggcttcgcc
1021	aagggcctgg	gcctggggcg	cacctacttc	accgtcttct	gctgctacgc	gctgctgctc
1081	tggtagcgag	gtcacctcgt	gcccgccac	cacaccaacg	cggcctcgc	catcgccacc
1141	atgtttcccg	tcatgatcgg	gggcctcgcg	ctgggcccag	ccgcgccag	catggccgcc
1201	ttcgcaaaag	cgcgctcgc	cgccgccaag	atcttcagga	tcacgacca	ccttgccgctc
1261	gtccatggcg	accatgtcca	gctcccttcc	gtgacggggc	gggtggagat	gcggggcggt
1321	gacttcgcct	accgctcgcg	accggacatc	cccgtcctgc	gcggcttctc	cctcacctgt
1381	ccgccaggca	agaccatcgc	gctcgtcggg	agctcaggct	cggaaagag	cacagctcgc
1441	tccttcattg	agaggttcta	cgacccaagt	gcaggtacac	aaaaaaaaactt	ccttggttga
1501	aaatctctct	tgctcgttcg	ttcacacaca	ttcatcgatt	caaaaacata	tgcttcttac
1561	taattactac	tactaggtga	gattctgctt	gacgggacag	acctcaagtc	cctgaacctg
1621	cgctggctcc	ggcagcagat	cgggctgggtg	agccaggagc	cgacgctggt	cgccacgagc
1681	atcaaggaga	acctgctgct	gggcccgcgac	agccagagcg	ccacctcgc	cgagatggag
1741	gaggccgcca	gggtcgccaa	cgcccactca	ttcatcatca	agctgcccga	cggctacgac
1801	accaggtat	gtacagtatg	cttcgctttg	ctccatctca	ctcaatgcca	ccctcggaaa
1861	ggaagaaaag	cttttcctcc	atcatctcat	tccttcctgg	cgctctcaga	tccacgggca
1921	cgcacgagac	tggcggggct	actagagtac	taccagtata	ctagcaatgc	ttgacctcac
1981	tcaccagtac	tagtactcta	gtagggcagc	aacacatgac	acgggagaca	gcaccaccgc
2041	gccgggggtcc	ctccatttgt	taggttagga	ggctgttgct	gtcacctggt	ggtgaggtt
2101	actcctaaca	acatgtaaca	tgctaaagat	tctttcttcc	tttctttcca	caacacaggt
2161	cgctttctga	gttgcttctg	tctcgacaca	atgcaatgca	gcagcagcag	cagctagcag
2221	cccatactc	atctttttgt	gctgcattca	ttttccacct	ggcttgccctc	aaaaaagaag
2281	ctactcactc	cagtaaatca	tcctctcgt	ggtggtagta	ctccaaagta	gaattccag
2341	gaatctagta	acggaacaag	tcccagttg	gcattttgtc	aggaatgtca	agttgcgaga
2401	tcccatccca	agaattttgt	ggagagagag	agagagagag	agggagcaat	gggagaggag
2461	gaggaccatg	gccgttctgt	tgttatttta	gtgtcaagtg	caggcagagt	gcagttaa
2521	aatcctgagt	agtttattaa	caagagtttg	tactagtacc	aaaccacatg	ggcgccact
2581	gactctgcat	tgctagttaa	aaaatattgc	agtgtggtgg	aggccggggc	ggcccagccc
2641	atcctcctgt	gcacatggca	tgatccgacg	gagatagatc	ctgcattagt	ttagtttaa



GENE SEQUENCES

```
2701 cttttctttc tgttctttct tctagtagta gtaggagtat cccagtcaa ctaagcactc
2761 tgctgggtcaa aaaaaaaaaa aaacaggcgg agtactcctg atttaattaa ctgatgatga
2821 tgatgacatt tgccttaagt tagtaattca tgatgagaat aatagaagga aaaacaaagt
2881 caatggatgt gcatgcaggt cggtagcgc ggtctgcagc tgtctggcgg tcagaagcag
2941 cgcacgcca tcgcgcgcg catgctcaag aaccctggga tcttgctgct ggacgaggcg
3001 acgagcgcgc tggactcggg gtccgagaag ctgggtgcagg aggcgctgga ccggttcag
3061 atcgggcgca cgaccctggt gatcgcgcac cggctgtcga cgatccgcaa ggccgacctg
3121 gtggccgtgc tgcacggcgg cgccgtgtcg gagattggga cgacgcaga gctgatggcc
3181 aagggcgagg acggcgcgta cgcgcggctg atccggatgc agcaggagca ggccgcccgc
3241 caggaggtgg ccgcccgtcg gagcagcgcg cggccgtcga gcgcgcgcaa ctcggtcagc
3301 tcgcccataca tgacgcgcaa ctctctctac gggcggctcg cctactcgcg gcgctctcc
3361 gacttctcca acgcccactc ccaccactac taccatggtg tgaacttat tgagagtaac
3421 aacaaggctc atcatcaaag gaggaggatt gcgttcgggg ctggggcgag ctcttctgg
3481 aggttggtcga agatgaactc gcccgagtgg gggtacgcgc tgggtggggtc cctgggctcc
3541 atgggtgtcgc gctcctttag cgccatcttc gcctacgcgc tgagcgcgt gctgagcgtg
3601 tactacgccc ccgacccggg ttacatgcgg cggcagatcg gcaagtactg ctacctgctg
3661 atgggcatgt ctccggcggc gctgggtgtc aacacgggtc agcacgcgtt ctgggacacg
3721 gtgggcgaga acctgacgaa gcgcgtccgg gagcgcagat tcggcgcctg gctccgcaac
3781 gaggtggcgt gtttcgacgc cgaggagaac gccagcgcgc gggtggcgcg caggctggcg
3841 ctggacgcgc agaactgtcg ttccgccatc ggcgacgcga tctcgggtgat cgtgcagaac
3901 tcggcgcgtc tgctgggtgg ctgcacggcg gggttcgtgc tgcagtggcg cctggcgtg
3961 gtgctcctgg ccgtgttccc gctgggtggt gccgccaccg tgctgcagaa gatgttcag
4021 aagggcttct cgggcgacct ggaggcggcc cacgcccgcg ccacgcagat cgccgcccag
4081 gccgtcgcca acctgcgcac cgtcgcgcgc ttcaacgccc aggcctaagat cgcgggctc
4141 ttcgccgcca acctgcgggg cccgctgcgg cgctgcttct ggaagggccca ggccgcccgc
4201 tgcggctacg ggctggcca gttcctgctc tacgcgtcct acgcgtggg cctctggtac
4261 gcggcctggc tcgtcaagca cggcgtctcc gacttctccc gcgccatccg ggtcttcag
4321 gtgctcatgg tctccgcaa cggcgcgcgc gagacgctca ccctggcgc ggacttctg
4381 aaaggcggcc gcgccatgcg atccgtcttc gagaccatcg acaggcgcac cgagacggag
4441 cccgacgacc cggacgcgc gccactgcca tcggacgcgg tctcgtgga gctcggcac
4501 gtggacttct gctaccctc gcggccggac gtgcgggtgc tgcaggacct cagcctgccc
4561 gcgcgcggc ggaaagacgct ggcctcgtg gggcccagcg ggtgcggcaa gagctcctg
4621 ctgcgcctca tccagcgtt ctacgagccc acctcgggc gcgtcctcct cgcggcgc
4681 gacgcgcgca agtacaacct cgggcgctg cggcgcgcca tcgccgtcgt gccgaggag
4741 cccttctct tcgggccac catccacgac aacatcgct acggccgca gggcggcac
4801 gaggcggagg tggtaggagg gccacgcag gccaacgcgc acaagttcat ctccgcctg
4861 cccgacgggt acaggacgca ggtcggcag cgcggggtgc agctctcgg cggcagcga
4921 cagcgcctc ccgtggctc ggcccttgtc aaacaggccg cgttctgct cctgcagag
4981 gccaccagcg cgtggacgc cgagtggag cgcagcgtgc agcaggcgt ggaccgcc
5041 gccaaagacc gtagcaccac caccatcgtc gtcgcgcacc gcctcggcc cgtccgcgac
5101 gcacacacca tcgccgtcat cgacgacggc aaggtcgtcg agcagggatc gcactcacac
5161 ctgctcaacc accaccccga cggaacctac gcgcgcatgc tccacctcca gcgactcacc
5221 gcgccatcca ctctctaa
```

//



B. gRNA BLAST Searches

B.1. *Sd1 gRNA1*

BLASTN 2.2.28+

Reference: Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-3402.

Database: *Eragrostis tef* (tef)
878 sequences; 577,738,711 total letters

Query= seq

Length=23

Sequences producing significant alignments:	Score (Bits)	E Value
lc1 3B	44.9	9e-05
lc1 3A	44.9	9e-05

> lc1|3B
Length=32575812

Score = 44.9 bits (23), Expect = 9e-05
Identities = 23/23 (100%), Gaps = 0/23 (0%)
Strand=Plus/Plus

```
Query 1          GGCCCGGACTTCGAGCCAATGGG 23
                |||
Sbjct 29293138   GGCCCGGACTTCGAGCCAATGGG 29293160
```

> lc1|3A
Length=34643735

Score = 44.9 bits (23), Expect = 9e-05
Identities = 23/23 (100%), Gaps = 0/23 (0%)
Strand=Plus/Plus

```
Query 1          GGCCCGGACTTCGAGCCAATGGG 23
                |||
Sbjct 31096526   GGCCCGGACTTCGAGCCAATGGG 31096548
```

Lambda K H
1.33 0.621 1.12

Gapped
Lambda K H
1.33 0.621 1.12

Effective search space used: 2888614535



GRNA BLAST SEARCHES

Database: Eragrostis tef (tef)
Posted date: Feb 11, 2020 4:50 PM
Number of letters in database: 577,738,711
Number of sequences in database: 878

Matrix: blastn matrix 1 -2
Gap Penalties: Existence: 5, Extension: 2

B.2. *Sd1 gRNA2*

BLASTN 2.2.28+

Reference: Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-3402.

Database: Eragrostis tef (tef)
878 sequences; 577,738,711 total letters

Query= seq

Length=23

Sequences producing significant alignments:

Score (Bits)	E Value
44.9	9e-05
44.9	9e-05

lc1|3B
lc1|3A

> lc1|3B
Length=32575812

Score = 44.9 bits (23), Expect = 9e-05
Identities = 23/23 (100%), Gaps = 0/23 (0%)
Strand=Plus/Plus

```
Query 1          GCGACTTCTTCGAGGACAGCCGG 23
                |||
Sbjct 29293380   GCGACTTCTTCGAGGACAGCCGG 29293402
```

> lc1|3A
Length=34643735

Score = 44.9 bits (23), Expect = 9e-05
Identities = 23/23 (100%), Gaps = 0/23 (0%)
Strand=Plus/Plus

```
Query 1          GCGACTTCTTCGAGGACAGCCGG 23
                |||
Sbjct 31096827   GCGACTTCTTCGAGGACAGCCGG 31096849
```

Lambda K H
1.33 0.621 1.12



GRNA BLAST SEARCHES

Gapped

```
Lambda      K      H
      1.33    0.621  1.12
```

Effective search space used: 2888614535

```
Database: Eragrostis tef (tef)
Posted date: Feb 11, 2020 4:50 PM
Number of letters in database: 577,738,711
Number of sequences in database: 878
```

```
Matrix: blastn matrix 1 -2
Gap Penalties: Existence: 5, Extension: 2
```

B.3. Dw1 gRNA

BLASTN 2.2.28+

Reference: Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

```
Database: Eragrostis tef (tef)
      878 sequences; 577,738,711 total letters
```

Query= seq

Length=23

Sequences producing significant alignments:

	Score (Bits)	E Value
lcl 4B	44.9	9e-05
lcl 4A	44.9	9e-05

```
> lcl|4B
Length=29936223
```

```
Score = 44.9 bits (23), Expect = 9e-05
Identities = 23/23 (100%), Gaps = 0/23 (0%)
Strand=Plus/Plus
```

```
Query 1          CCTGTGACACACCATATTCCAGG 23
      |||
Sbjct 7681439    CCTGTGACACACCATATTCCAGG 7681461
```

```
> lcl|4A
Length=32664196
```

```
Score = 44.9 bits (23), Expect = 9e-05
Identities = 23/23 (100%), Gaps = 0/23 (0%)
Strand=Plus/Plus
```



GRNA BLAST SEARCHES

```

Query 1      CCTGTGACACACCATATTCCAGG 23
          |||
Sbjct 7977763 CCTGTGACACACCATATTCCAGG 7977785

```

```

Lambda      K      H
    1.33    0.621  1.12

```

Gapped

```

Lambda      K      H
    1.33    0.621  1.12

```

Effective search space used: 2888614535

```

Database: Eragrostis tef (tef)
Posted date: Feb 11, 2020 4:50 PM
Number of letters in database: 577,738,711
Number of sequences in database: 878

```

```

Matrix: blastn matrix 1 -2
Gap Penalties: Existence: 5, Extension: 2

```

B.4. Dw3 gRNA

BLASTN 2.2.28+

Reference: Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-3402.

```

Database: Eragrostis tef (tef)
      878 sequences; 577,738,711 total letters

```

Query= seq

Length=23

Sequences producing significant alignments:

	Score (Bits)	E Value
lc1 8B	44.9	9e-05
lc1 8A	44.9	9e-05

```

> lc1|8B
Length=21147804

```

```

Score = 44.9 bits (23), Expect = 9e-05
Identities = 23/23 (100%), Gaps = 0/23 (0%)
Strand=Plus/Minus

```

```

Query 1      CCGGTTTCATGATCGGGCGCACGA 23
          |||
Sbjct 134813  CCGGTTTCATGATCGGGCGCACGA 134791

```




GRNA BLAST SEARCHES

```
> lc1|8A
Length=24151120

Score = 44.9 bits (23), Expect = 9e-05
Identities = 23/23 (100%), Gaps = 0/23 (0%)
Strand=Plus/Minus

Query 1          CCGGTTTCATGATCGGGCGCACGA 23
          |||
Sbjct 145900    CCGGTTTCATGATCGGGCGCACGA 145878
```

```
Lambda      K      H
      1.33  0.621  1.12
```

```
Gapped
Lambda      K      H
      1.33  0.621  1.12
```

Effective search space used: 2888614535

```
Database: Eragrostis tef (tef)
Posted date: Feb 11, 2020 4:50 PM
Number of letters in database: 577,738,711
Number of sequences in database: 878
```

```
Matrix: blastn matrix 1 -2
Gap Penalties: Existence: 5, Extension: 2
```



C. Nucleotide Sequence Alignments

The following figures present full-length nucleotide sequence alignments between the *Sd1*, *Dw1*, and *Dw3* wild-type sequences and their respective mutant alleles. Alignments are provided separately for homoeologs on the A and B genomes.

The remainder of this page is intentionally left blank.



NUCLEOTIDE SEQUENCE ALIGNMENTS

Sd1B	ATGCTTGTGTTTGTCTGCTCGCTCCCTCCCTGTTCTCTGTTTACAAATACCCACCGGCGCCGGACAGGTCCCCCTGC	80
sd1B1_Mutant	ATGCTTGTGTTTGTCTGCTCGCTCCCTCCCTGTTCTCTGTTTACAAATACCCACCGGCGCCGGACAGGTCCCCCTGC	80
Sd1B	ACACTCACACCTCGCACATCTCATGGTGTCCCAAGCACGGCAAGAGCCTCGCAGCAGCGGCATCTCCAAGCTCGCCATGG	160
sd1B1_Mutant	ACACTCACACCTCGCACATCTCATGGTGTCCCAAGCACGGCAAGAGCCTCGCAGCAGCGGCATCTCCAAGCTCGCCATGG	160
Sd1B	ACGCCATCCCGGCCCGCCTCTCCTGTCCGCTCCCAAGCTCCCGGATTGACCTCCCAAAGACAAGATCCCGCGCGCC	240
sd1B1_Mutant	ACGCCATCCCGGCCCGCCTCTCCTGTCCGCTCCCAAGCTCCCGGATTGACCTCCCAAAGACAAGATCCCGCGCGCC	240
Sd1B	TTCGTGTGGCCGCGGACGACTCGCGGCCGACGTCGGCGGGCGGCGTGGACGTCGCCGTTGGTGGACGTGGGCGTGTCTGC	320
sd1B1_Mutant	TTCGTGTGGCCGCGGACGACTCGCGGCCGACGTCGGCGGGCGGCGTGGACGTCGCCGTTGGTGGACGTGGGCGTGTCTGC	320
Sd1B	CAATGGCGGCGACCCGCGGGCTGCGGGCGCGGGCGGCGAGGTGGCGTGGCGTGGCGGACGACAGGCTTCTTCCAGG	400
sd1B1_Mutant	CAATGGCGGCGACCCGCGGGCTGCGGGCGCGGGCGGCGAGGTGGCGTGGCGTGGCGGACGACAGGCTTCTTCCAGG	400
Sd1B	TGCGCGGGCACGGCGTGGACGCGCCCTGGCGCGCGGGCGGCTGGACGGCGGCGAGCATTCTTCCGGTGCCTGCGC	480
sd1B1_Mutant	TGCGCGGGCACGGCGTGGACGCGCCCTGGCGCGCGGGCGGCTGGACGGCGGCGAGCATTCTTCCGGTGCCTGCGC	480
Sd1B	GAGAAGCAGCGCGCGGCGGCGACTTCCCGGGCACCGTGTCCGGGTACACGAGCGCCACGCGGACCGGTTCGCGTGCAAGCT	560
sd1B1_Mutant	GAGAAGCAGCGCGCGGCGGCGACTTCCCGGGCACCGTGTCCGGGTACACGAGCGCCACGCGGACCGGTTCGCGTGCAAGCT	560
Sd1B	CCCGTGAAGGAGACGCTGTCTTCCGCTTCCACGACGGCGCGGCTGCGTTCGCGCGCGTTCGTCGTCGACTACTTACCA	640
sd1B1_Mutant	CCCGTGAAGGAGACGCTGTCTTCCGCTTCCACGACGGCGCGGCTGCGTTCGCGCGCGTTCGTCGTCGACTACTTACCA	640
Sd1B	GCACCCTCGGCCCGGACTTCGAGCC↓AATGGGGTAAACATATAAACTACTGGCGGGCTTGTGTGGATATAGTATCTTG	719
sd1B1_Mutant	GCACCCTCGGCCCGGACTTCGAGCCAAATGGGGTAAACATATAAACTACTGGCGGGCTTGTGTGGATATAGTATCTTG	720
Sd1B	TGCGCCATTGCATTTCGACACCGACACTAGCTTGGGACGACTGGAACCATCCATCATGTACTACTTATATGGCAGGCGCGTG	799
sd1B1_Mutant	TGCGCCATTGCATTTCGACACCGACACTAGCTTGGGACGACTGGAACCATCCATCATGTACTACTTATATGGCAGGCGCGTG	800
Sd1B	TACCAGAGGTACTGCGAGGAGATGAAGGACCTGTCCCTGACGATCATGGAGCTGCTTGAGCTGAGCCTGGGCGTGGAGCG	879
sd1B1_Mutant	TACCAGAGGTACTGCGAGGAGATGAAGGACCTGTCCCTGACGATCATGGAGCTGCTTGAGCTGAGCCTGGGCGTGGAGCG	880
Sd1B	CGGCTACTACCGGCACTTCTTCGAGGACAGCGGCTCCATCAATGCGGTTGCAACTACTACCGCGCTGCGCGGAGCGGAGC	959
sd1B1_Mutant	CGGCTACTACCGGCACTTCTTCGAGGACAGCGGCTCCATCAATGCGGTTGCAACTACTACCGCGCTGCGCGGAGCGGAGC	960
Sd1B	GCACGC TGGGCACGGGCGCGGCGACTGCGACCCACGGCGCTACCATCTCTGCCAGGACGAGTTCGGCGGACTTCGAGGTC	1039
sd1B1_Mutant	GCACGC TGGGCACGGGCGCGGCGACTGCGACCCACGGCGCTACCATCTCTGCCAGGACGAGTTCGGCGGACTTCGAGGTC	1040
Sd1B	CTCGTCGACGGGGAATGGCGCCCGTCCGCCCCGTCGCCGGAGCCATGGTTCATCAACATCGCGGACAGGTTTCATGGTACA	1119
sd1B1_Mutant	CTCGTCGACGGGGAATGGCGCCCGTCCGCCCCGTCGCCGGAGCCATGGTTCATCAACATCGCGGACAGGTTTCATGGTACA	1120
Sd1B	ATTCAATTTACTCTGTCTTTGGCTGGCTGCTAGCTTGTCTTCCGCTTTCGCCGATGTCCTCTAGCTAATGGAACATAA	1199
sd1B1_Mutant	ATTCAATTTACTCTGTCTTTGGCTGGCTGCTAGCTTGTCTTCCGCTTTCGCCGATGTCCTCTAGCTAATGGAACATAA	1200
Sd1B	TGGACAGCGCTCAGATTACACACACAGTAGGCGTGTATGTTGGCAATGTAGACATGCGTTTCGAGTTGGTTCCATCTC	1279
sd1B1_Mutant	TGGACAGCGCTCAGATTACACACACAGTAGGCGTGTATGTTGGCAATGTAGACATGCGTTTCGAGTTGGTTCCATCTC	1280
Sd1B	GATCCGTGGGCTAGCTCGGTAGTAGGCTACCGGTGGTGGTGAACCGTGGAGCGGCTTGCATGGTTTCGCGAAAGAGAAA	1359
sd1B1_Mutant	GATCCGTGGGCTAGCTCGGTAGTAGGCTACCGGTGGTGGTGAACCGTGGAGCGGCTTGCATGGTTTCGCGAAAGAGAAA	1360
Sd1B	AAAGAGAGAGAGAGAGAGAGAGAGAGAGGATAGCGATGACAGTGGAGGACAAAGAGAGAGAGAGAGAGAGGTTGGCTTGTG	1439
sd1B1_Mutant	AAAGAGAGAGAGAGAGAGAGAGAGAGAGGATAGCGATGACAGTGGAGGACAAAGAGAGAGAGAGAGAGAGAGAGGTTGGCTTGTG	1440
Sd1B	CGTTGGCGGCTTGGAGTTGTAGAGGAGTGGGGGAGGACCGGAGGAGGAGGGGGATACGATGCCGCGAGCGGTTGGCGCTT	1519
sd1B1_Mutant	CGTTGGCGGCTTGGAGTTGTAGAGGAGTGGGGGAGGACCGGAGGAGGAGGGGGATACGATGCCGCGAGCGGTTGGCGCTT	1520
Sd1B	TGCGCTGGTGGCCCCAAGCCAGGTGGCGCGCGCTGTTGTCTTCCCTGTTTGGGCCCCGTCGGGGCCGCTCATCCG	1599
sd1B1_Mutant	TGCGCTGGTGGCCCCAAGCCAGGTGGCGCGCGCTGTTGTCTTCCCTGTTTGGGCCCCGTCGGGGCCGCTCATCCG	1600
Sd1B	CGGCGTGACTTGGCGCGCACGGGCGCGCGCTGGAGCACAGGATCGTCAGCTTGGCGTTACCCGCTGGTATGCTCTGC	1679
sd1B1_Mutant	CGGCGTGACTTGGCGCGCACGGGCGCGCGCTGGAGCACAGGATCGTCAGCTTGGCGTTACCCGCTGGTATGCTCTGC	1680
Sd1B	GTAATTTACTTGGCTGTGCTCGCTCGCCGGGTTCTAGCCGTTTCCGCGCGGTTTGGCGCGGTTTCCCGGTGCACGCGCGCC	1759
sd1B1_Mutant	GTAATTTACTTGGCTGTGCTCGCTCGCCGGGTTCTAGCCGTTTCCGCGCGGTTTGGCGCGGTTTCCCGGTGCACGCGCGCC	1760
Sd1B	CGCGCGGGGCTCCAGCTCGGTGCGGGTGGCACTACGCTAGCGTATAGTACGCGCGCGCAACAAAAGGGTTCTCTGTGCC	1839
sd1B1_Mutant	CGCGCGGGGCTCCAGCTCGGTGCGGGTGGCACTACGCTAGCGTATAGTACGCGCGCGCAACAAAAGGGTTCTCTGTGCC	1840
Sd1B	TGACAAGCCTTCACTCACGCTAGTGTGTTGCACAAAAGGGGGGGTTCCGGATTGAGCTGATTGTTGTTTCGTTCCAG	1919
sd1B1_Mutant	TGACAAGCCTTCACTCACGCTAGTGTGTTGCACAAAAGGGGGGGTTCCGGATTGAGCTGATTGTTGTTTCGTTCCAG	1920
Sd1B	TGTTCACTGTTGATCTGCATCTCTCCCTGACTGACTGACTGACTGACTGACTGACTGACTGACTTCTCTACTGGTTGTTTGCACACTG	1999
sd1B1_Mutant	TGTTCACTGTTGATCTGCATCTCTCCCTGACTGACTGACTGACTGACTGACTGACTGACTGACTTCTCTACTGGTTGTTTGCACACTG	2000
Sd1B	GCAACAGGCGCTGTCCAACGGGCGGTACAAGAGCTGCTTGCACCCGCGGTTGGTGAACCGCGGCGAGGAGCGGCGGTTGC	2079
sd1B1_Mutant	GCAACAGGCGCTGTCCAACGGGCGGTACAAGAGCTGCTTGCACCCGCGGTTGGTGAACCGCGGCGAGGAGCGGCGGTTGC	2080
Sd1B	TGGCCTTCTTCTGTGTCGCGCGGAGGACCGGCTGCTGCGCGCGCGGACCCGCTGCGCGCGCGCGCGGCTTACCCGGAC	2159
sd1B1_Mutant	TGGCCTTCTTCTGTGTCGCGCGGAGGACCGGCTGCTGCGCGCGCGGACCCGCTGCGCGCGCGCGCGGCTTACCCGGAC	2160
Sd1B	TTCAGTGGGGCGACTTCATGCGCTTACGACAGAGCCACTACCGCGCGGACACACGACCCCTGCAGCGCTTACCACCTG	2239
sd1B1_Mutant	TTCAGTGGGGCGACTTCATGCGCTTACGACAGAGCCACTACCGCGCGGACACACGACCCCTGCAGCGCTTACCACCTG	2240
Sd1B	GCTCGCCGCGGCGCCCGGGCCAGCAGGCGTAG	2273
sd1B1_Mutant	GCTCGCCGCGGCGCCCGGGCCAGCAGGCGTAG	2274

Figure 9. Alignment of the wild-type *Sd1B* allele nucleotide sequence with the *sd1B1* allele. The *sd1B2* allele sequence is not shown as it was identical to *sd1B1*. The location of the adenine nucleotide insertion at position 666 in *sd1B1* is indicated. Sequences were aligned with Clustal Omega (Sievers et al., 2011) using the default parameters and displayed using TEXshade (Beitz, 2000).



NUCLEOTIDE SEQUENCE ALIGNMENTS

Dw1A	ATGGCTTCAGCTGGGAGCAGCAGTGGAGGTGCAGGTGGAGGCAGCAGCATCAGGGCGGCGAATGGCGCAGCTGCCATCAG	80
dw1A1_Mutant	ATGGCTTCAGCTGGGAGCAGCAGTGGAGGTGCAGGTGGAGGCAGCAGCATCAGGGCGGCGAATGGCGCAGCTGCCATCAG	80
Dw1A	TGCGGGCGGCGAGGCAGTAGGCTCAGCCGACGCCAGATTCCACTCCCACCCACCACAGCAGGACAGGGTATATACACTAC	160
dw1A1_Mutant	TGCGGGCGGCGAGGCAGTAGGCTCAGCCGACGCCAGATTCCACTCCCACCCACCACAGCAGGACAGGGTATATACACTAC	160
Dw1A	ACGACTTCTTGTTCATGGCGCATGCAATTGCTCATGCTTTAGCTGCAGCTTTTACCCTTCTTGCTTTAGATGTGTGGAT	240
dw1A1_Mutant	ACGACTTCTTGTTCATGGCGCATGCAATTGCTCATGCTTTAGCTGCAGCTTTTACCCTTCTTGCTTTAGATGTGTGGAT	240
Dw1A	GAGTACCAGCAATTTCTAGTAGGCCTTGACGACCTAGATTGCACAGGCTTTGCGTACTTAGTACGAATGGTACTGCCCA	320
dw1A1_Mutant	GAGTACCAGCAATTTCTAGTAGGCCTTGACGACCTAGATTGCACAGGCTTTGCGTACTTAGTACGAATGGTACTGCCCA	320
Dw1A	GTGGTGGTTGCAGCTTGCAACCCCGGCATCTTTACCCTCAATAATTAGGCTTCAGTACTGTCTGCTTGTGTTGCTTCTAGT	400
dw1A1_Mutant	GTGGTGGTTGCAGCTTGCAACCCCGGCATCTTTACCCTCAATAATTAGGCTTCAGTACTGTCTGCTTGTGTTGCTTCTAGT	400
Dw1A	ACTGAATAGTTTGTTCAGCTGAATTTCCGTTTTGCTGTTGATATTTGATTATTACGGGATTTGCACCTCCCCTATGTTT	480
dw1A1_Mutant	ACTGAATAGTTTGTTCAGCTGAATTTCCGTTTTGCTGTTGATATTTGATTATTACGGGATTTGCACCTCCCCTATGTTT	480
Dw1A	TTCCCCTCCGGATCCTAGTAGTACTGTACTCTTACCATAACACTCGTCCAGAGACCTATATGTTGTGTAAGTGTGATTGA	560
dw1A1_Mutant	TTCCCCTCCGGATCCTAGTAGTACTGTACTCTTACCATAACACTCGTCCAGAGACCTATATGTTGTGTAAGTGTGATTGA	560
Dw1A	CCTTGCATACACTTCCACACGCTCGTGTGATGTCATGAGCAAACTAATATGCTGGGTTGATTCCCTTCTTGCATGA	640
dw1A1_Mutant	CCTTGCATACACTTCCACACGCTCGTGTGATGTCATGAGCAAACTAATATGCTGGGTTGATTCCCTTCTTGCATGA	640
Dw1A	TGGATCTTTTTGCCAAATGAATTTTCAGATTCCGTGATAAATGCTTGGTGGTGTAGCTATTAATGTCAACACATTTCTTATA	720
dw1A1_Mutant	TGGATCTTTTTGCCAAATGAATTTTCAGATTCCGTGATAAATGCTTGGTGGTGTAGCTATTAATGTCAACACATTTCTTATA	720
Dw1A	TTCTTAAGTAAATTTGTTTGTAGATTAATACTCCAAGGAAATTTCTGTTGTTGAACAACCTAGTCAATACATTTTTTCCCCTAC	800
dw1A1_Mutant	TTCTTAAGTAAATTTGTTTGTAGATTAATACTCCAAGGAAATTTCTGTTGTTGAACAACCTAGTCAATACATTTTTTCCCCTAC	800
Dw1A	CAACAGCTATGTAAGTACTGATACTAATAATGTCATCTCGATCTAAACAGCGGAGTAGATGGGCTGGCTGTTCGGGCTCT	880
dw1A1_Mutant	CAACAGCTATGTAAGTACTGATACTAATAATGTCATCTCGATCTAAACAGCGGAGTAGATGGGCTGGCTGTTCGGGCTCT	880
Dw1A	TTTATGTTTCGGATCTCAGAAGGGCGGGAAGAGGATGTTTCTCGCCACGCTACTCCTGATGGGAATGGATCAAGTGCTC	960
dw1A1_Mutant	TTTATGTTTCGGATCTCAGAAGGGCGGGAAGAGGATGTTTCTCGCCACGCTACTCCTGATGGGAATGGATCAAGTGCTC	960
Dw1A	GTGGAAATGGTTCATCAGTCCGGTCTAATTCAAACAAAATGTGCCCTTAAATCTGTCTCTTCTGGCTCCACCATCTCTCA	1040
dw1A1_Mutant	GTGGAAATGGTTCATCAGTCCGGTCTAATTCAAACAAAATGTGCCCTTAAATCTGTCTCTTCTGGCTCCACCATCTCTCA	1040
Dw1A	CGGGCATCCTTCTCAAATTTCTGCACCTCCTTCAACTGTTCAATCGCCTAATAACTTTCTGTCAATCTCAGCAAATTTCTCC	1120
dw1A1_Mutant	CGGGCATCCTTCTCAAATTTCTGCACCTCCTTCAACTGTTCAATCGCCTAATAACTTTCTGTCAATCTCAGCAAATTTCTCC	1120
Dw1A	TGGTGTTCACACATCTAATATGTTTGTCTGTTGGGCCATATGCTAATGAACCTCAGCTTGTCTCGCCCTCCTGTCTTCFAA	1200
dw1A1_Mutant	TGGTGTTCACACATCTAATATGTTTGTCTGTTGGGCCATATGCTAATGAACCTCAGCTTGTCTCGCCCTCCTGTCTTCFAA	1200
Dw1A	CTTACACAACCTGAGCCTTCCACAGCACCATTGACCCCTCCACCTGAACTAGCTCATGCAACCCTCCCTCATCTCCAGAT	1280
dw1A1_Mutant	CTTACACAACCTGAGCCTTCCACAGCACCATTGACCCCTCCACCTGAACTAGCTCATGCAACCCTCCCTCATCTCCAGAT	1280
Dw1A	GTTCCTACGCTCGGTTTCTTCTCTTTTATGGATATCAAACCTGCTAGTAAAGGAGCATAACATGGCTTTCTTATCAAC	1360
dw1A1_Mutant	GTTCCTACGCTCGGTTTCTTCTCTTTTATGGATATCAAACCTGCTAGTAAAGGAGCATAACATGGCTTTCTTATCAAC	1360
Dw1A	AACATATCTGTGGTTCAGGACTCCAGGCATCATACCCCTTTACCCCTGAGAGCCCTTGTAGCAGCCTCATATCACCTG	1440
dw1A1_Mutant	AACATATCTGTGGTTCAGGACTCCAGGCATCATACCCCTTTACCCCTGAGAGCCCTTGTAGCAGCCTCATATCACCTG	1440
Dw1A	CTTCTGCTACTCCGAGGACTGGTCTTTTCTCACCAATACCTGAACAAGAGGTCCTACTGCCCATTTGGAAGACTTCTAGG	1520
dw1A1_Mutant	CTTCTGCTACTCCGAGGACTGGTCTTTTCTCACCAATACCTGAACAAGAGGTCCTACTGCCCATTTGGAAGACTTCTAGG	1520
Dw1A	TCTGCCGTGTAACACACCATAATCCAGGGCTTCACCCATCCCTGAGCAAGAAGCCACTGCGCAGTGAAGACTTCTAGAT	1599
dw1A1_Mutant	TCTGCCGTGTAACACACCATAATCCAGGGCTTCACCCATCCCTGAGCAAGAAGCCACTGCGCAGTGAAGACTTCTAGAT	1600
Dw1A	CGGCATGTGATACGCCCTTATGCCAGGGCTTCACCATCAAACATCTTTGGGCTGGACTCATCTGCTTCTAGAAAATTTG	1679
dw1A1_Mutant	CGGCATGTGATACGCCCTTATGCCAGGGCTTCACCATCAAACATCTTTGGGCTGGACTCATCTGCTTCTAGAAAATTTG	1680
Dw1A	TTAGATGGCAACTTCTTCCGGCCAGCTGCTTCTGCTCAATTTACCTGGACCAGGCTCAGCAGACATATCCATATAATGG	1759
dw1A1_Mutant	TTAGATGGCAACTTCTTCCGGCCAGCTGCTTCTGCTCAATTTACCTGGACCAGGCTCAGCAGACATATCCATATAATGG	1760
Dw1A	TGGGAGGCTAGTGTCTCAGGGGACAGCAAGATGCAGATGAATTTGAAGCTTACAGAGCTTGGTTTGGCTTTAGTGCAG	1839
dw1A1_Mutant	TGGGAGGCTAGTGTCTCAGGGGACAGCAAGATGCAGATGAATTTGAAGCTTACAGAGCTTGGTTTGGCTTTAGTGCAG	1840
Dw1A	ATGAAATCATGCAAACTCAATCTTATGTTGAGATACAGATGCACCTTGATGAATCATTCAGTATATCGCCATTTGGAAT	1919
dw1A1_Mutant	ATGAAATCATGCAAACTCAATCTTATGTTGAGATACAGATGCACCTTGATGAATCATTCAGTATATCGCCATTTGGAAT	1920
Dw1A	AATGCTCCTGCTACTGAGGTGTCCCATTTAATGATCCACCCAACGAGGTTTCAAGGCTGAGAAGTCAAGTCCAAAAGAA	1999
dw1A1_Mutant	AATGCTCCTGCTACTGAGGTGTCCCATTTAATGATCCACCCAACGAGGTTTCAAGGCTGAGAAGTCAAGTCCAAAAGAA	2000
Dw1A	AACGGCAGATCAGATTCCAATGGCTCTCCACATAGAGTGTGCACATTGACATATTCAGGGTAAATGTCCATACTTTT	2079
dw1A1_Mutant	AACGGCAGATCAGATTCCAATGGCTCTCCACATAGAGTGTGCACATTGACATATTCAGGGTAAATGTCCATACTTTT	2080
Dw1A	TGCACCAACCATTCAGTTGTAATCTAGGCATAGTAGAATTACTCTTGTGTTATCTTGTCTACTGACCTAACCAAGC	2159
dw1A1_Mutant	TGCACCAACCATTCAGTTGTAATCTAGGCATAGTAGAATTACTCTTGTGTTATCTTGTCTACTGACCTAACCAAGC	2160
Dw1A	TTTGGTAACTTTCAATGCTAATAAACCCTGAACCTGGGCCATTTGTCTATCAGCACTTATTGGTATTGTTTCTATTTATG	2239
dw1A1_Mutant	TTTGGTAACTTTCAATGCTAATAAACCCTGAACCTGGGCCATTTGTCTATCAGCACTTATTGGTATTGTTTCTATTTATG	2240
Dw1A	TAAATGCTTATTATGCTTATGTGCTCGCATTTCAGGAACAAAAGGAGGACATCCGCTGAGTACGAGGGTATTGTGAAAG	2319
dw1A1_Mutant	TAAATGCTTATTATGCTTATGTGCTCGCATTTCAGGAACAAAAGGAGGACATCCGCTGAGTACGAGGGTATTGTGAAAG	2320
Dw1A	ATGGCCATCCTTTTCAGAAAGACAAGGGACGAAATATCTCTAAAACCCATAGAAGTAAAGGAAGAAATCTCCACCTGGGCAT	2399
dw1A1_Mutant	ATGGCCATCCTTTTCAGAAAGACAAGGGACGAAATATCTCTAAAACCCATAGAAGTAAAGGAAGAAATCTCCACCTGGGCAT	2400
Dw1A	TCATGCTCAGATGCTGAAATTTGAGTACAGAAGGGCAAGGAGTCTGAGGGAAAGCCAAATGGTGTCCCTCGTGGCGCAGCAC	2479
dw1A1_Mutant	TCATGCTCAGATGCTGAAATTTGAGTACAGAAGGGCAAGGAGTCTGAGGGAAAGCCAAATGGTGTCCCTCGTGGCGCAGCAC	2480
Dw1A	GTGTGCAAGACAGCTGCAGTGA	2501
dw1A1_Mutant	GTGTGCAAGACAGCTGCAGTGA	2502

Figure 10. Alignment of the wild-type *Dw1A* allele nucleotide sequence with the *dw1A1* allele. The *dw1A2* allele sequence is not shown as it was identical to *dw1A1*. The location of the adenine nucleotide insertion at position 1531 in *dw1A1* (+A) is indicated. Sequences were aligned with Clustal Omega (Sievers et al., 2011) using the default parameters and displayed using TEXshade (Beitz, 2000).



NUCLEOTIDE SEQUENCE ALIGNMENTS

Dw1B	ATGGCTTCAGCTGGGAGCAGCAGTGGAGGTGCAGGTGGAGGCAGCAGCATCAGGGCGGCGAATGGCCGAGCTGCTATCAG	80
dw1B1_Mutant	ATGGCTTCAGCTGGGAGCAGCAGTGGAGGTGCAGGTGGAGGCAGCAGCATCAGGGCGGCGAATGGCCGAGCTGCTATCAG	80
Dw1B	TGCGGGCGGCGAGCCAGTAGGCTCAGCCGACGCCAGATTCCACTCCCACCCACCACAACAGGACAGGGTATATACACTAC	160
dw1B1_Mutant	TGCGGGCGGCGAGCCAGTAGGCTCAGCCGACGCCAGATTCCACTCCCACCCACCACAACAGGACAGGGTATATACACTAC	160
Dw1B	ACGACTTCTTGTTCATGGCCGATGCAATTGCTCATGCTTTGGCTGCAGCTTTTACCCTTCTTGCTTTAGATGTATGGAT	240
dw1B1_Mutant	ACGACTTCTTGTTCATGGCCGATGCAATTGCTCATGCTTTGGCTGCAGCTTTTACCCTTCTTGCTTTAGATGTATGGAT	240
Dw1B	GAGTACCAGCAATTTCCAGTAGGCCCTTGACGACCTAGATTGTCACAGGCTTTGCGTACTTAGTACGAATGGTACTGCCCA	320
dw1B1_Mutant	GAGTACCAGCAATTTCCAGTAGGCCCTTGACGACCTAGATTGTCACAGGCTTTGCGTACTTAGTACGAATGGTACTGCCCA	320
Dw1B	GTGGTGGTGCAGCTTGCAACCCCGGCATCTTTACCCCAATAAATTAGGCTTCAGTACTGCTGCTTGTTTGCTTCTAC	400
dw1B1_Mutant	GTGGTGGTGCAGCTTGCAACCCCGGCATCTTTACCCCAATAAATTAGGCTTCAGTACTGCTGCTTGTTTGCTTCTAC	400
Dw1B	TGAATAATTTGTCGGCTGAATTTTTGCTTTGCTGTCGATATTGATTGCACGCCATTTGCACTCCCTCTATGTACTCTT	480
dw1B1_Mutant	TGAATAATTTGTCGGCTGAATTTTTGCTTTGCTGTCGATATTGATTGCACGCCATTTGCACTCCCTCTATGTACTCTT	480
Dw1B	ACCAGTAGACTGTTTGTACTCTTACCATAAACCCTCCACAGAGATCCTAGATGCTGTATAACTGTTGATTTGACCTTGA	560
dw1B1_Mutant	ACCAGTAGACTGTTTGTACTCTTACCATAAACCCTCCACAGAGATCCTAGATGCTGTATAACTGTTGATTTGACCTTGA	560
Dw1B	TACACTTCCACACGTTGTGTCACTGATGTCATGAGCAAACAACTATGCTGGGCTGAAATCTTCTTGCATGATAAATCT	640
dw1B1_Mutant	TACACTTCCACACGTTGTGTCACTGATGTCATGAGCAAACAACTATGCTGGGCTGAAATCTTCTTGCATGATAAATCT	640
Dw1B	TTTTCCATATGAAAATTTCCAGTAGGCCCTGTAATAATCTTGGTGGTGTAGCTATTATTGTCAACAAAATTTCTTATATTCT	720
dw1B1_Mutant	TTTTCCATATGAAAATTTCCAGTAGGCCCTGTAATAATCTTGGTGGTGTAGCTATTATTGTCAACAAAATTTCTTATATTCT	720
Dw1B	TAGTAAACTTGTTTTTATGCTTACCCTGTTATAGTAAGTTTGGCCAGTTTCCCTCCCATGATTTTAAATGAATAAATCA	800
dw1B1_Mutant	TAGTAAACTTGTTTTTATGCTTACCCTGTTATAGTAAGTTTGGCCAGTTTCCCTCCCATGATTTTAAATGAATAAATCA	800
Dw1B	CTTTTTCTGTTTGTAGCGAGTATAGATTAATACTCTGAGGAAAATTTGTTGTTGAACATTTAGTCATACATTTCCCTCCCT	880
dw1B1_Mutant	CTTTTTCTGTTTGTAGCGAGTATAGATTAATACTCTGAGGAAAATTTGTTGTTGAACATTTAGTCATACATTTCCCTCCCT	880
Dw1B	ACCAACAGCTATGTACTTGATACTAATAATGTGCATCTCTGATCTAAACAGCGGAGTAGATGGGCTGGCTGCTTTTCGGGG	960
dw1B1_Mutant	ACCAACAGCTATGTACTTGATACTAATAATGTGCATCTCTGATCTAAACAGCGGAGTAGATGGGCTGGCTGCTTTTCGGGG	960
Dw1B	CTTTCATGTTTCGGATCTCAGAAGGGCGGGAAGAGGATTGTTCCGGCCGACGTTACTCCTGATGGGAATGGATCAAGTGC	1040
dw1B1_Mutant	CTTTCATGTTTCGGATCTCAGAAGGGCGGGAAGAGGATTGTTCCGGCCGACGTTACTCCTGATGGGAATGGATCAAGTGC	1040
Dw1B	TGGTGAATATGGTCACTCAGTCTGGTTCTAATTCAAAACAAAATGTGCCTTTAAATCTGTGCTTCTGGCTCCACCATCGT	1120
dw1B1_Mutant	TGGTGAATATGGTCACTCAGTCTGGTTCTAATTCAAAACAAAATGTGCCTTTAAATCTGTGCTTCTGGCTCCACCATCGT	1120
Dw1B	CACCGGCATCTTCTCAAATTTGCACTTCCCTTCAACTGTTTCACTGCGCAAAATAAATTTCTGTCAAATCTCAGCAAAATCT	1200
dw1B1_Mutant	CACCGGCATCTTCTCAAATTTGCACTTCCCTTCAACTGTTTCACTGCGCAAAATAAATTTCTGTCAAATCTCAGCAAAATCT	1200
Dw1B	CCTGGTGGTCCGACATCTAATATGTTTGTCTGTTGGGCCATATGCTAATGAACCTCAGCTTGTCTCGCTCCTGTCTTCTC	1280
dw1B1_Mutant	CCTGGTGGTCCGACATCTAATATGTTTGTCTGTTGGGCCATATGCTAATGAACCTCAGCTTGTCTCGCTCCTGTCTTCTC	1280
Dw1B	AACCTACACAACCTGAGCCTTCCACAGCACCATTGACCCCTCCACCTGAACCTAGCTCATGCAACCACCTCCCTCGTCCCAG	1360
dw1B1_Mutant	AACCTACACAACCTGAGCCTTCCACAGCACCATTGACCCCTCCACCTGAACCTAGCTCATGCAACCACCTCCCTCGTCCCAG	1360
Dw1B	ATGTTCCATATGCCCCGTTTCTTTCTTCTTATGGATATCAAACCTGCTAGTAAGGAGCATAAATAGGCTTCTTATCA	1440
dw1B1_Mutant	ATGTTCCATATGCCCCGTTTCTTTCTTCTTATGGATATCAAACCTGCTAGTAAGGAGCATAAATAGGCTTCTTATCA	1440
Dw1B	ACAACCTATTTCTGGTGGTTCAGGACTCCAGGCATCATACCTCTTTACCCTGAGAGCCCTTGTAGCAGCCTCATATCACC	1520
dw1B1_Mutant	ACAACCTATTTCTGGTGGTTCAGGACTCCAGGCATCATACCTCTTTACCCTGAGAGCCCTTGTAGCAGCCTCATATCACC	1520
Dw1B	TGCTTCTGCAACTCCGAGGACTGGTCTTTCCCTACCAATACCTGAACAAGAGGTCCCACAGCCCATTTGGAAGACTTCTA	1600
dw1B1_Mutant	TGCTTCTGCAACTCCGAGGACTGGTCTTTCCCTACCAATACCTGAACAAGAGGTCCCACAGCCCATTTGGAAGACTTCTA	1600
Dw1B	GGTCTGCCTGTGACACACCATATTTCCAGGGCTTCCACCATCCCTGAGCAAGAAGCCACTGCGCAGTGGAAAGACTTCTAG	1679
dw1B1_Mutant	GGTCTGCCTGTGACACACCATATTTCCAGGGCTTCCACCATCCCTGAGCAAGAAGCCACTGCGCAGTGGAAAGACTTCTAG	1680
Dw1B	ATCAGCATGTGATACACCTTATGCCAGAGCTTCAACCATCAAACATCTTTGGGCTGGATTATCTGCTTCTAGAAAATATT	1759
dw1B1_Mutant	ATCAGCATGTGATACACCTTATGCCAGAGCTTCAACCATCAAACATCTTTGGGCTGGATTATCTGCTTCTAGAAAATATT	1760
Dw1B	TGTTAGATGGCAACTTCTTCCGGCCAGCTGCTTCTGCTCAATTTCACTGGACCAGGCTCAGCAGACATATCCATATAAT	1839
dw1B1_Mutant	TGTTAGATGGCAACTTCTTCCGGCCAGCTGCTTCTGCTCAATTTCACTGGACCAGGCTCAGCAGACATATCCATATAAT	1840
Dw1B	GGTGGGAGGCGTAGTGTCTCAGGGGACAAGCAAGACGAGATGAAGTTGAAGCTTACAGAGCTTCGTTTGGCTTTAGTGC	1919
dw1B1_Mutant	GGTGGGAGGCGTAGTGTCTCAGGGGACAAGCAAGACGAGATGAAGTTGAAGCTTACAGAGCTTCGTTTGGCTTTAGTGC	1920
Dw1B	AGATGAAATCATGCAAACTCAATCTTATGTGGAGATACCAGATGCACCTTGATGAATCATTCAGTATATCACCAATTTGGAA	1999
dw1B1_Mutant	AGATGAAATCATGCAAACTCAATCTTATGTGGAGATACCAGATGCACCTTGATGAATCATTCAGTATATCACCAATTTGGAA	2000
Dw1B	ATAATGCTCCTGCTACGAGGTGTCGCCATTTAATGATCTGCCCATGAGGTTTCAAGAGGTCAGAAAGTCAAGTCCAAAG	2079
dw1B1_Mutant	ATAATGCTCCTGCTACGAGGTGTCGCCATTTAATGATCTGCCCATGAGGTTTCAAGAGGTCAGAAAGTCAAGTCCAAAG	2080
Dw1B	AAATCAGCAGATCAGATTTCCAATGGTTCTCCACATAGAGTGTGCACGTTGACATATTCAGGGTAAATGTCACATCTT	2159
dw1B1_Mutant	AAATCAGCAGATCAGATTTCCAATGGTTCTCCACATAGAGTGTGCACGTTGACATATTCAGGGTAAATGTCACATCTT	2160
Dw1B	TTTGCAACAACCATTCAGTTGTAATCTAGGGCTAGTAGAATTAAGTCTTTTGTATCCTTGTCTATCTACTGCACCTAACT	2239
dw1B1_Mutant	TTTGCAACAACCATTCAGTTGTAATCTAGGGCTAGTAGAATTAAGTCTTTTGTATCCTTGTCTATCTACTGCACCTAACT	2240
Dw1B	GGTAACCTCCAATGCTAATAAACCCTGAACCTCGGGCCATGGTATTGTTTCTGTTTACGTAATGCTTATCATGCTTAT	2319
dw1B1_Mutant	GGTAACCTCCAATGCTAATAAACCCTGAACCTCGGGCCATGGTATTGTTTCTGTTTACGTAATGCTTATCATGCTTAT	2320
Dw1B	GTGCTCACATTTTCAGGAACAAAAGGAGGACATCCGTTTGTAGTACGAGGGGATTTGTGAAAGATGGCCATCCCTTTCAGAAAG	2399
dw1B1_Mutant	GTGCTCACATTTTCAGGAACAAAAGGAGGACATCCGTTTGTAGTACGAGGGGATTTGTGAAAGATGGCCATCCCTTTCAGAAAG	2400
Dw1B	ACAAGGGATGAGATATCTCTAAAACCCATAGAAGTAAGGAAGAAAATCTCCACCTGGTCAATTCATGCTCAGATGCTGAAAT	2479
dw1B1_Mutant	ACAAGGGATGAGATATCTCTAAAACCCATAGAAGTAAGGAAGAAAATCTCCACCTGGTCAATTCATGCTCAGATGCTGAAAT	2480
Dw1B	CGAGTACAGAAGGGCAAGGAGTCTGAGGGAAAGCCAAACGGTGTCCCTTGTGGCGCAGCACGTTGTGAGACAGCTGCAGT	2559
dw1B1_Mutant	CGAGTACAGAAGGGCAAGGAGTCTGAGGGAAAGCCAAACGGTGTCCCTTGTGGCGCAGCACGTTGTGAGACAGCTGCAGT	2560
Dw1B	GA 2561	
dw1B1_Mutant	GA 2562	

Figure 11. Alignment of the wild-type *Dw1B* allele nucleotide sequence with the *dw1B1* allele. The *dw1B2* allele sequence is not shown as it was identical to *dw1B1*. The location of the adenine nucleotide insertion at position 1613 in *dw1B1* is indicated. Sequences were aligned with Clustal Omega (Sievers et al., 2011) using the default parameters and displayed using TEXshade (Beitz, 2000).

NUCLEOTIDE SEQUENCE ALIGNMENTS



Dw3A	ATGTCTAGCAGCGACCCGGAGGAGATCCGGGGGCGCGTGGTCTCCTCGGGCCGATGCGGACGAGTTGGCTCGCCCCGA	80
dw3A1_Mutant	ATGTCTAGCAGCGACCCGGAGGAGATCCGGGGGCGCGTGGTCTCCTCGGGCCGATGCGGACGAGTTGGCTCGCCCCGA	80
Dw3A	GCTGGAGGCCCTTCCACCTCCCGTCTCCGACACTAGCTCAGGAGGCAGGATCTGTCCGAGGCCAGGCCCTGCTGTCTGTG	160
dw3A1_Mutant	GCTGGAGGCCCTTCCACCTCCCGTCTCCGACACTAGCTCAGGAGGCAGGATCTGTCCGAGGCCAGGCCCTGCTGTCTGTG	160
Dw3A	CTCCCGTGGTGGGCCATTGGAGACGGAGCTGCCTCCCAATGCCACCATGCCTTCTTCTTCTCCGCTCCCGAACAGC	240
dw3A1_Mutant	CTCCCGTGGTGGGCCATTGGAGACGGAGCTGCCTCCCAATGCCACCATGCCTTCTTCTTCTCCGCTCCCGAACAGC	240
Dw3A	AACAGCAATGAGCAGAACAAAGGAGCAGGAGAATACGAAGAAGAAGGGGGTGTCTTTGGCGCCTGCTCCTCTTGGTTTCGT	320
dw3A1_Mutant	AACAGCAATGAGCAGAACAAAGGAGCAGGAGAATACGAAGAAGAAGGGGGTGTCTTTGGCGCCTGCTCCTCTTGGTTTCGT	320
Dw3A	GTTCCGTTTCGCCGACGGTCTGGACTCGCTTCTGATGTCCGTTGGGCACGTTGGGCGCGCTTGTCCACGGTTGCTCCCTCC	400
dw3A1_Mutant	GTTCCGTTTCGCCGACGGTCTGGACTCGCTTCTGATGTCCGTTGGGCACGTTGGGCGCGCTTGTCCACGGTTGCTCCCTCC	400
Dw3A	CGGTGTTCCTCCGCTTCTTCCGGGACCTCGTCGACTCGTTCCGGCTCCACGCGGACGACCCGGACACCATGGTCCGGGTG	480
dw3A1_Mutant	CGGTGTTCCTCCGCTTCTTCCGGGACCTCGTCGACTCGTTCCGGCTCCACGCGGACGACCCGGACACCATGGTCCGGGTG	480
Dw3A	GTGGCAAGTACGCGCTCTACTTCTTGGTGGTGGGCGGGCCATCTGGGCGTCTCTCCGCGGAGATCTCTGCTGGAT	560
dw3A1_Mutant	GTGGCAAGTACGCGCTCTACTTCTTGGTGGTGGGCGGGCCATCTGGGCGTCTCTCCGCGGAGATCTCTGCTGGAT	560
Dw3A	GTGGACCGGCGAGCGGAGTTCGACCGGATCGGGATCCGGTACCTGGAGTCCGCGCTCCGGCAGGACGTGTCTTCTTCC	640
dw3A1_Mutant	GTGGACCGGCGAGCGGAGTTCGACCGGATCGGGATCCGGTACCTGGAGTCCGCGCTCCGGCAGGACGTGTCTTCTTCC	640
Dw3A	ACACCGACGTCCGACCTCCGACGTGATCTAGGCCATCAACGCGGACGCGGTGATCGTGCAGGACGCCATCAGCGAGAAG	720
dw3A1_Mutant	ACACCGACGTCCGACCTCCGACGTGATCTAGGCCATCAACGCGGACGCGGTGATCGTGCAGGACGCCATCAGCGAGAAG	720
Dw3A	CTCGGCAACCTCATCCACTACATGGCCACCTTCCGTGGCCGGTTCGTTGGTCCGGTTACGGCGGCATGGCAGTGGCGCT	800
dw3A1_Mutant	CTCGGCAACCTCATCCACTACATGGCCACCTTCCGTGGCCGGTTCGTTGGTCCGGTTACGGCGGCATGGCAGTGGCGCT	800
Dw3A	GGTGGCTGGCCGTGTTGCCGCTCAFCGCCGTGATCGGCGCCTCAGCGCGCGCGCTGGCGAAGCTGTCTGTCGGGA	880
dw3A1_Mutant	GGTGGCTGGCCGTGTTGCCGCTCAFCGCCGTGATCGGCGCCTCAGCGCGCGCGCTGGCGAAGCTGTCTGTCGGGA	880
Dw3A	GCCAGGACGCGCTGGCGGAGGCGAGCAACATCGCCGAGCAGCGGTTGGCGAGATACGGACGGTGCAGGCCCTTCGTCGGG	960
dw3A1_Mutant	GCCAGGACGCGCTGGCGGAGGCGAGCAACATCGCCGAGCAGCGGTTGGCGAGATACGGACGGTGCAGGCCCTTCGTCGGG	960
Dw3A	GAGGAGCGCGCAATGCGGGCTACTCGTTGGCGCTGGCCGCGCGCAGAGGATCGGATACCGCAGCGGCTTCGCGAAGGG	1040
dw3A1_Mutant	GAGGAGCGCGCAATGCGGGCTACTCGTTGGCGCTGGCCGCGCGCAGAGGATCGGATACCGCAGCGGCTTCGCGAAGGG	1040
Dw3A	CTGGGCTGGCGGACCTACTTCCCGCTTCTGCTGCTACGGCTGCTGCTGCTGTTACGGCGGCTTCTGTCGGAC	1120
dw3A1_Mutant	CTGGGCTGGCGGACCTACTTCCCGCTTCTGCTGCTACGGCTGCTGCTGCTGTTACGGCGGCTTCTGTCGGAC	1120
Dw3A	GCCACCACACCAACGGCGGCCFCGCCATCGCCACCAATGTTTCCGCTCATGATCGGGGGCCTCGCGTGGCCAGTCCGGG	1200
dw3A1_Mutant	GCCACCACACCAACGGCGGCCFCGCCATCGCCACCAATGTTTCCGCTCATGATCGGGGGCCTCGCGTGGCCAGTCCGGG	1200
Dw3A	CCCAGCATGGCGCCTTCGCAAAAGGCGCGCTCGCCGCGGCCAAGATCTTCAAATCATCGACCACAAGCCCTCTCCGT	1280
dw3A1_Mutant	CCCAGCATGGCGCCTTCGCAAAAGGCGCGCTCGCCGCGGCCAAGATCTTCAAATCATCGACCACAAGCCCTCTCCGT	1280
Dw3A	CGTCCGCTGGCGGACGATGTCACCTCCCTCCGTCGACGGGGCGGGTGGAGATGCGGGGCGTTGACTTCCGCTACC	1360
dw3A1_Mutant	CGTCCGCTGGCGGACGATGTCACCTCCCTCCGTCGACGGGGCGGGTGGAGATGCGGGGCGTTGACTTCCGCTACC	1360
Dw3A	CGTCCGCGGCGGACGCTCCCGTCTGCGCGGCTTCTCCCTACCGTGGCGCGGCAAGACATCGCGCTTCGCGGAGC	1440
dw3A1_Mutant	CGTCCGCGGCGGACGCTCCCGTCTGCGCGGCTTCTCCCTACCGTGGCGCGGCAAGACATCGCGCTTCGCGGAGC	1440
Dw3A	TCAGGTCGCGGGAAGAGCAGGTCGCTCCCTCATCGAGAGGTTCTACGACCAAGTGCAGGTACACAACCTACTTAACGT	1520
dw3A1_Mutant	TCAGGTCGCGGGAAGAGCAGGTCGCTCCCTCATCGAGAGGTTCTACGACCAAGTGCAGGTACACAACCTACTTAACGT	1520
Dw3A	CTTAGAAGTGTCTGTCACTCAATTTGCTTCTTCAATTCACAAATCGTATTCATTTGATTGAAAGAAAAATATGCTTCTCAGGC	1600
dw3A1_Mutant	CTTAGAAGTGTCTGTCACTCAATTTGCTTCTTCAATTCACAAATCGTATTCATTTGATTGAAAGAAAAATATGCTTCTCAGGC	1600
Dw3A	GAGATTCTGCTTACGGGCGACGACCTCAAGTCCCTCAACTGGCGTGGCTCCGGCAGCAGATCGGGCTTGTGAGCCAGGA	1680
dw3A1_Mutant	GAGATTCTGCTTACGGGCGACGACCTCAAGTCCCTCAACTGGCGTGGCTCCGGCAGCAGATCGGGCTTGTGAGCCAGGA	1680
Dw3A	GCCGACGCTGTTCCGCCACAGCATCAAGGAGAACCTGCTGCTGGGCGGGGACAGCCACAGCGCCACCTCGCGGAGATGG	1760
dw3A1_Mutant	GCCGACGCTGTTCCGCCACAGCATCAAGGAGAACCTGCTGCTGGGCGGGGACAGCCACAGCGCCACCTCGCGGAGATGG	1760
Dw3A	AGGAGGCGCAAGGGTCGCCAACGCCACTCATTCATCAAGCTGCCCGAGGATACGACACCCAGGTATGTACTAGC	1840
dw3A1_Mutant	AGGAGGCGCAAGGGTCGCCAACGCCACTCATTCATCAAGCTGCCCGAGGATACGACACCCAGGTATGTACTAGC	1840
Dw3A	ACCGTTGCTTTTCTTTTACATTCCTTCCTTCCTGAGCAACTAGATATGTACTAGCACCATTGCTCCATCTCAATGCCA	1920
dw3A1_Mutant	ACCGTTGCTTTTCTTTTACATTCCTTCCTTCCTGAGCAACTAGATATGTACTAGCACCATTGCTCCATCTCAATGCCA	1920
Dw3A	CCATCCCAAAAAAGAAAGAAAGCTTTTCTCCATCTCATCTCCATCTCTCATTCCCTGGCGCTCGTGCCTGCCTCAGAT	2000
dw3A1_Mutant	CCATCCCAAAAAAGAAAGAAAGCTTTTCTCCATCTCATCTCCATCTCTCATTCCCTGGCGCTCGTGCCTGCCTCAGAT	2000
Dw3A	CCATGGGCGCAAGTGGACGGGACGACAGAGACGTGGCTAGTAGTACTCCAGTATAGCAATGCTTATTAGCCTCACTCAC	2080
dw3A1_Mutant	CCATGGGCGCAAGTGGACGGGACGACAGAGACGTGGCTAGTAGTACTCCAGTATAGCAATGCTTATTAGCCTCACTCAC	2080
Dw3A	TCACCAGTGTATATACTAGTAGGGCAGCCAGCAGCAACATGACACGGGAGGAGACAGCACCACCCGCGGGGTCCTCTC	2160
dw3A1_Mutant	TCACCAGTGTATATACTAGTAGGGCAGCCAGCAGCAACATGACACGGGAGGAGACAGCACCACCCGCGGGGTCCTCTC	2160
Dw3A	CCATTGTTAGGTTAGGAGCCTGTTGCTGTCACCTGCTGTTGGGAGTTAACAAACATGTAACATGCTAAAGATTGCTTCTTC	2240
dw3A1_Mutant	CCATTGTTAGGTTAGGAGCCTGTTGCTGTCACCTGCTGTTGGGAGTTAACAAACATGTAACATGCTAAAGATTGCTTCTTC	2240
Dw3A	CTTCTTTCCACAACACAGGTCGCTTCTGACTTGGCTTCTCCTGCTTCTCGACACAAATGCAATGCAGCAGTAGCAGCC	2320
dw3A1_Mutant	CTTCTTTCCACAACACAGGTCGCTTCTGACTTGGCTTCTCCTGCTTCTCGACACAAATGCAATGCAGCAGTAGCAGCC	2320
Dw3A	CATCACCTAGTCTCACCACCTCTTCTGCTAGTACTGCTTCCATTTTACCATGGCTTGGCTCACCGCCGGGGTGGTAGT	2400
dw3A1_Mutant	CATCACCTAGTCTCACCACCTCTTCTGCTAGTACTGCTTCCATTTTACCATGGCTTGGCTCACCGCCGGGGTGGTAGT	2400
Dw3A	AGAAAGTAGAATTTCCAGGAATCTAGTAAGAGAACAAGTCCCAAGTTGGCATTGTTGTCAGGAATGTCAAGTTGCCACTTG	2480
dw3A1_Mutant	AGAAAGTAGAATTTCCAGGAATCTAGTAAGAGAACAAGTCCCAAGTTGGCATTGTTGTCAGGAATGTCAAGTTGCCACTTG	2480
Dw3A	CGAGATCCCATCCCAAGAATTTTCTGCAATGGGAGAGGAGGACCATGGCCGTTGTTTATTTTACTGTCAAGCGGAG	2560
dw3A1_Mutant	CGAGATCCCATCCCAAGAATTTTCTGCAATGGGAGAGGAGGACCATGGCCGTTGTTTATTTTACTGTCAAGCGGAG	2560
Dw3A	GCATAGGGCGGTTTCTCCTTCTGAGTCATTAACAACAATCCTGCGTGGAGTGATTAACAAGAGTTTGTAGCACTACC	2640
dw3A1_Mutant	GCATAGGGCGGTTTCTCCTTCTGAGTCATTAACAACAATCCTGCGTGGAGTGATTAACAAGAGTTTGTAGCACTACC	2640
Dw3A	AAACCACATGGGCCACTGACTCTGCAGCTGAGGAGTAGTAAAAATAGTGCAGTGTGGAGGCCGGCCCATCTCCTATGCAC	2720
dw3A1_Mutant	AAACCACATGGGCCACTGACTCTGCAGCTGAGGAGTAGTAAAAATAGTGCAGTGTGGAGGCCGGCCCATCTCCTATGCAC	2720
Dw3A	ATGACATGATCCGACGGAGATAGTAGAATAGATCCGACATGATATTTTTTTTAAATTCATGACAATCATGATAAAAAAA	2800
dw3A1_Mutant	ATGACATGATCCGACGGAGATAGTAGAATAGATCCGACATGATATTTTTTTTAAATTCATGACAATCATGATAAAAAAA	2800

Figure 12. Part 1—Alignment of the wild-type *Dw3A* allele nucleotide sequence with the *dw3A1* allele. The *dw3A2* allele sequence is not shown as it was identical to *dw3A1*. Sequences were aligned with Clustal Omega (Sievers et al., 2011) using the default parameters and displayed using TEXshade (Beitz, 2000).

NUCLEOTIDE SEQUENCE ALIGNMENTS



Dw3A	GCAAAGTCAAACAAGGTCAATTATGATAAAAAAATGCAATGTGCAGGTCGGTGAGCGCGGCTGCAGCTGTCTGGCGGT	2880
dw3A1_Mutant	GCAAAGTCAAACAAGGTCAATTATGATAAAAAAATGCAATGTGCAGGTCGGTGAGCGCGGCTGCAGCTGTCTGGCGGT	2880
Dw3A	CAGAAGCAGCGCATCGCCATTGCCCGGCCATGCTCAAGAACCCTGGCATCTTGTCTGGACGAGGCGACGAGCGCGCT	2960
dw3A1_Mutant	CAGAAGCAGCGCATCGCCATTGCCCGGCCATGCTCAAGAACCCTGGCATCTTGTCTGGACGAGGCGACGAGCGCGCT	2960
Dw3A	GGACTCTGAGTCGGAGAAGCTGGTGCAGGAGGCGCTGGACCGGTTTCATGATCGGGCGCACACCCCTGGTGTATCGCCGAC	3039
dw3A1_Mutant	GGACTCTGAGTCGGAGAAGCTGGTGCAGGAGGCGCTGGACCGGTTTCATGATCGGGCGCACACCCCTGGTGTATCGCCGAC	3040
Dw3A	CGGCTGTGCAGCATCCGCAAGGCCGACCTGGTGGCGGTGCTGCACGGCGGCCCGTGTCCGAGATTGGGACGCACGAGGA	3119
dw3A1_Mutant	CGGCTGTGCAGCATCCGCAAGGCCGACCTGGTGGCGGTGCTGCACGGCGGCCCGTGTCCGAGATTGGGACGCACGAGGA	3120
Dw3A	GCTGATGGGCAAGGGCGAGGACGGCGCGTACCGCGCGCTGATCCGGATGCAGGAGCAGGCCGCGCAGGAGGTGGCCGCCC	3199
dw3A1_Mutant	GCTGATGGGCAAGGGCGAGGACGGCGCGTACCGCGCGCTGATCCGGATGCAGGAGCAGGCCGCGCAGGAGGTGGCCGCCC	3200
Dw3A	CTCGGAGCAGCGCGCAACTCGGTGAGCGCGCGCAACTCGGTGAGCTCGCCCATCATGACGCGCAACTCCTCCTACGGG	3279
dw3A1_Mutant	CTCGGAGCAGCGCGCAACTCGGTGAGCGCGCGCAACTCGGTGAGCTCGCCCATCATGACGCGCAACTCCTCCTACGGG	3280
Dw3A	CGGTCCGCCTACTCGCGCCGCTCTCCGACTTCTCCAACGCCGACTTCCACTACCACGGCGGTGGTGAACCTACGGGAGG	3359
dw3A1_Mutant	CGGTCCGCCTACTCGCGCCGCTCTCCGACTTCTCCAACGCCGACTTCCACTACCACGGCGGTGGTGAACCTACGGGAGG	3360
Dw3A	TAAATCCAAGAAGATGATTCATCAGAGGGTTGCGTTCCGGGGCGGGGCGAGCTCTTTTCTCGCTCTGGCGAAGATGAAC	3439
dw3A1_Mutant	TAAATCCAAGAAGATGATTCATCAGAGGGTTGCGTTCCGGGGCGGGGCGAGCTCTTTTCTCGCTCTGGCGAAGATGAAC	3440
Dw3A	CGCCCGAGTGGGGGTACGCTCTGGTGGGCTCGTGGGCTCCATGGTGTGCGGCTCCTTCAGCGCCATCTTCGGCTACGTT	3519
dw3A1_Mutant	CGCCCGAGTGGGGGTACGCTCTGGTGGGCTCGTGGGCTCCATGGTGTGCGGCTCCTTCAGCGCCATCTTCGGCTACGTT	3520
Dw3A	CTGAGCGCCGTGCTGAGCGTGTACTACGCCCCCGACCCGGGCGACATGCGCGGGGAGATCGCCAAGTACTGCTACCTGCT	3599
dw3A1_Mutant	CTGAGCGCCGTGCTGAGCGTGTACTACGCCCCCGACCCGGGCGACATGCGCGGGGAGATCGCCAAGTACTGCTACCTGCT	3600
Dw3A	GATGGGCATGTCTTCCGCGCGCTGGTGTGCAACACGGTGCAGCACGTGTTCTGGGACACGGTGGGCGAGAACCTGAGGA	3679
dw3A1_Mutant	GATGGGCATGTCTTCCGCGCGCTGGTGTGCAACACGGTGCAGCACGTGTTCTGGGACACGGTGGGCGAGAACCTGAGGA	3680
Dw3A	AGCGCGTCCGGGAGCGCATGCTCCGCGCGCTGCTCCGCAACGAGATGGCGTGGTTCGACGCGGAGGAGAACCAGCCAGCC	3759
dw3A1_Mutant	AGCGCGTCCGGGAGCGCATGCTCCGCGCGCTGCTCCGCAACGAGATGGCGTGGTTCGACGCGGAGGAGAACCAGCCAGCC	3760
Dw3A	CGGGTGGCCGCGAGGC TGGCGCTGGAAGCGCGAGAAGTGCCTCCGCCATCGGCGACCGCATCTCGGTGATCGTGCAGAA	3839
dw3A1_Mutant	CGGGTGGCCGCGAGGC TGGCGCTGGAAGCGCGAGAAGTGCCTCCGCCATCGGCGACCGCATCTCGGTGATCGTGCAGAA	3840
Dw3A	CTCGCGCTGCTGCTGGTGGCGTGCACGGCCGGGTTCTGCTGTCAGTGGCGCTCGCCCTGGTGTCTGGCGGTGTTC	3919
dw3A1_Mutant	CTCGCGCTGCTGCTGGTGGCGTGCACGGCCGGGTTCTGCTGTCAGTGGCGCTCGCCCTGGTGTCTGGCGGTGTTC	3920
Dw3A	CGCTGGTGGTGGCCGCCACCTGCTGCGAGAAGATGTTCATGAAGGGCTTCTCGGGGACCTGGAGGCGGCCACGCGCCG	3999
dw3A1_Mutant	CGCTGGTGGTGGCCGCCACCTGCTGCGAGAAGATGTTCATGAAGGGCTTCTCGGGGACCTGGAGGCGGCCACGCGCCG	4000
Dw3A	GCCACGAGATCGCCGCGAGGCCGTCCCAACCTGGCCACCGTCCGCCGTTCAACGCCGAGGCCAAGATCGCGGGACT	4079
dw3A1_Mutant	GCCACGAGATCGCCGCGAGGCCGTCCCAACCTGGCCACCGTCCGCCGTTCAACGCCGAGGCCAAGATCGCGGGACT	4080
Dw3A	CTTCGGCGGCAACCTGCGGGGCCGCTGCGCGGGTGCCTCTGGAAGGGCCAGGTGGCGGGTGCGGTACGGGGTGGCGC	4159
dw3A1_Mutant	CTTCGGCGGCAACCTGCGGGGCCGCTGCGCGGGTGCCTCTGGAAGGGCCAGGTGGCGGGTGCGGTACGGGGTGGCGC	4160
Dw3A	AGTTCTGTCTGACCGCTCTACGCGCTGGGGCTCTGGTACCGCGCTGGTGGTGAAGCAGGGCTCTCCGACTTCTCC	4239
dw3A1_Mutant	AGTTCTGTCTGACCGCTCTACGCGCTGGGGCTCTGGTACCGCGCTGGTGGTGAAGCAGGGCTCTCCGACTTCTCC	4240
Dw3A	CGCACCATCCGCGCTTTCATGGTGTCTATGGTCTCCGCCAACGGCGCGCCGAGACGCTCACGCTGGCGCCGACTTGGT	4319
dw3A1_Mutant	CGCACCATCCGCGCTTTCATGGTGTCTATGGTCTCCGCCAACGGCGCGCCGAGACGCTCACGCTGGCGCCGACTTGGT	4320
Dw3A	CAGGGGCGGACGCGCCATGCGCTCCGCTCTTCGAGACCATCGACAGGGCGACCGAGGCGGACCCGACGACCAGGCGGG	4399
dw3A1_Mutant	CAGGGGCGGACGCGCCATGCGCTCCGCTCTTCGAGACCATCGACAGGGCGACCGAGGCGGACCCGACGACCAGGCGGG	4400
Dw3A	CGCCATGCAACTGCCATTATGACAGGGGTGGAGCTGCGGCGACGTGGACTTCTGCTACCCGTCGCGGGCGGAGGTGCAG	4479
dw3A1_Mutant	CGCCATGCAACTGCCATTATGACAGGGGTGGAGCTGCGGCGACGTGGACTTCTGCTACCCGTCGCGGGCGGAGGTGCAG	4480
Dw3A	GTGCTGACGAGCCTCAGCCTGCGCGCGCGCGGGGAAAGCCTGGCGCTCGTGGGGCCAGCGGCTGCGGCAAGAGCTC	4559
dw3A1_Mutant	GTGCTGACGAGCCTCAGCCTGCGCGCGCGCGGGGAAAGCCTGGCGCTCGTGGGGCCAGCGGCTGCGGCAAGAGCTC	4560
Dw3A	CGTGTGGCGCTCATCCAGCGCTTCTACGAGGCCACCTCGGGCGCGTCTCTCTCGACGGCCGGGACGCGGCAAGTACA	4639
dw3A1_Mutant	CGTGTGGCGCTCATCCAGCGCTTCTACGAGGCCACCTCGGGCGCGTCTCTCTCGACGGCCGGGACGCGGCAAGTACA	4640
Dw3A	ACCTCCGGGCGCTGCGCGCGCGCGTGCCTGGTGCAGGAGGCCCTTCCGTTCGCGGCGAGCATCCACGACAACATC	4719
dw3A1_Mutant	ACCTCCGGGCGCTGCGCGCGCGCGTGCCTGGTGCAGGAGGCCCTTCCGTTCGCGGCGAGCATCCACGACAACATC	4720
Dw3A	GCCACGGCCGAGGGGGGGCCACGGAGGCGGAGGTGCTGGAGGCGGCCACGCAGGCCAACGCGCACAAAGTTCATCTC	4799
dw3A1_Mutant	GCCACGGCCGAGGGGGGGCCACGGAGGCGGAGGTGCTGGAGGCGGCCACGCAGGCCAACGCGCACAAAGTTCATCTC	4800
Dw3A	CGCGCTGCCCGACGGCTACAGGACGCAAGTTCGGGGAGCGCGGGGTGCAGCTCTCCGGCGGCCAGCGCCAGGCATCGCC	4879
dw3A1_Mutant	CGCGCTGCCCGACGGCTACAGGACGCAAGTTCGGGGAGCGCGGGGTGCAGCTCTCCGGCGGCCAGCGCCAGGCATCGCC	4880
Dw3A	TGGCGCGCGCTCGTCAAGCAGGGCCGCTTCTGCTGCTGGACGAGGCCACAGCGCGCTGGACGCGGAGTTCGGAGGCG	4959
dw3A1_Mutant	TGGCGCGCGCTCGTCAAGCAGGGCCGCTTCTGCTGCTGGACGAGGCCACAGCGCGCTGGACGCGGAGTTCGGAGGCG	4960
Dw3A	AGCGTGCAGCAGGGCGCTGGACCGCCACGCCAAGACCAGCAGCACCACCATCGTCTGCGCACCGCCTCGCCACCGT	5039
dw3A1_Mutant	AGCGTGCAGCAGGGCGCTGGACCGCCACGCCAAGACCAGCAGCACCACCATCGTCTGCGCACCGCCTCGCCACCGT	5040
Dw3A	GCGCAACGCCACACCATCGCGCTCATCGACGAGGGCAAGTTCGTCGAGCAGGGATCGCACTCGCACTTGTCAACCAC	5119
dw3A1_Mutant	GCGCAACGCCACACCATCGCGCTCATCGACGAGGGCAAGTTCGTCGAGCAGGGATCGCACTCGCACTTGTCAACCAC	5120
Dw3A	ACCCGGACGGAACCTACGCGCGCATGCTCCAACCTACAGCGACTCACCTCATCCACTTCTTA	5181
dw3A1_Mutant	ACCCGGACGGAACCTACGCGCGCATGCTCCAACCTACAGCGACTCACCTCATCCACTTCTTA	5182

Figure 13. Part 2–Alignment of the wild-type *Dw3A* allele nucleotide sequence with the *dw3A1* allele. The *dw3A2* allele sequence is not shown as it was identical to *dw3A1*. The location of the thymine nucleotide insertion at position 3004 in *dw3A1* is indicated. Sequences were aligned with Clustal Omega (Sievers et al., 2011) using the default parameters and displayed using TEXshade (Beitz, 2000).



NUCLEOTIDE SEQUENCE ALIGNMENTS

Dw3B	ATTTAATTAAGTATGATGATGATGACATTTGCTTAAAGTTAGTAATTCATGATGAGAATAATAGAAGGAAAAACAAGT	2880
dw3B1_Mutant	ATTTAATTAAGTATGATGATGATGATGACATTTGCTTAAAGTTAGTAATTCATGATGAGAATAATAGAAGGAAAAACAAGT	2880
Dw3B	CAATGGATGTGCATGCAGGTGCGTGGAGCGCGTCTGCAGCTGTCTGGCGGTGAGAAGCAGCGCATCGCCATCGCGCGCGC	2960
dw3B1_Mutant	CAATGGATGTGCATGCAGGTGCGTGGAGCGCGTCTGCAGCTGTCTGGCGGTGAGAAGCAGCGCATCGCCATCGCGCGCGC	2960
Dw3B	CATGCTCAAGAACCCTGGGATCTTGTGCTGGACGAGGCGACGAGCGCGCTGGACTCGGAGTCCGAGAAGCTGGTGCAGG	3040
dw3B1_Mutant	CATGCTCAAGAACCCTGGGATCTTGTGCTGGACGAGGCGACGAGCGCGCTGGACTCGGAGTCCGAGAAGCTGGTGCAGG	3040
Dw3B	AGGCGCTGGACCGGTT↓CATGATCGGGCGCACGACCCTGGTGTATCGCGCACCGGCTGTGACGATCCGCAAGGCCGACCT	3119
dw3B1_Mutant	AGGCGCTGGACCGGTT↓ACATGATCGGGCGCACGACCCTGGTGTATCGCGCACCGGCTGTGACGATCCGCAAGGCCGACCT	3120
Dw3B	GGTGGCGGTGCTGCACGGCGCGCGCTGTGGAGATTGGGACGCACGACGAGCTGATGGCCAAAGGGCAGGACGGCGCGT	3199
dw3B1_Mutant	GGTGGCGGTGCTGCACGGCGCGCGCTGTGGAGATTGGGACGCACGACGAGCTGATGGCCAAAGGGCAGGACGGCGCGT	3200
Dw3B	ACGCGCGGCTGATCCGGATGCAGCAGGAGCAGGCGCGCGCAGGAGGTGGCCGCCCGTGGAGCAGCGCGCGCGCGCTCG	3279
dw3B1_Mutant	ACGCGCGGCTGATCCGGATGCAGCAGGAGCAGGCGCGCGCAGGAGGTGGCCGCCCGTGGAGCAGCGCGCGCGCGCTCG	3280
Dw3B	AGCGCGCGCAACTCGGTACGCTCGCCCATCATGACGCGCAACTCCTCCTACGGGCGGTGCGCCCTACTCGCGCGCCTCTC	3359
dw3B1_Mutant	AGCGCGCGCAACTCGGTACGCTCGCCCATCATGACGCGCAACTCCTCCTACGGGCGGTGCGCCCTACTCGCGCGCCTCTC	3360
Dw3B	CGACTTCTCCAACGCCGACTCCACCCTACTACCAATGGTGGTGAACCTTATGAGAGTAACAACAAGGCTCATCATCAA	3439
dw3B1_Mutant	CGACTTCTCCAACGCCGACTCCACCCTACTACCAATGGTGGTGAACCTTATGAGAGTAACAACAAGGCTCATCATCAA	3440
Dw3B	GGAGGAGGATTGGCTTCCGGGCTGGGGCGAGCTCGTCTGGAGGCTGGCCAAGATGAACCTCGCCGAGTGGGGGTACGCC	3519
dw3B1_Mutant	GGAGGAGGATTGGCTTCCGGGCTGGGGCGAGCTCGTCTGGAGGCTGGCCAAGATGAACCTCGCCGAGTGGGGGTACGCC	3520
Dw3B	CTGGTGGGGTCCCTGGGCTCCATGGTGTGCGGCTCCTTCACGCCCATCTTCGCCTACGCGCTGAGCGCCGCTGTGAGCGT	3599
dw3B1_Mutant	CTGGTGGGGTCCCTGGGCTCCATGGTGTGCGGCTCCTTCACGCCCATCTTCGCCTACGCGCTGAGCGCCGCTGTGAGCGT	3600
Dw3B	GTAAGCGCCCGACCGGGTTACATGCGCGCGCAGATCGGCAAGTACTGCTACCTGCTGATGGGCATGTCTTCGGGG	3679
dw3B1_Mutant	GTAAGCGCCCGACCGGGTTACATGCGCGCGCAGATCGGCAAGTACTGCTACCTGCTGATGGGCATGTCTTCGGGG	3680
Dw3B	CGCTGGTGTTCACACCGTGCAGCAGCGGTTCTGGGACACGGTGGGCGAGAACCTGACGAAGCGCGTCCGGGACGCGATG	3759
dw3B1_Mutant	CGCTGGTGTTCACACCGTGCAGCAGCGGTTCTGGGACACGGTGGGCGAGAACCTGACGAAGCGCGTCCGGGACGCGATG	3760
Dw3B	TTCCGGCGCGTGTCCGCAACGAGGTGGCGTGGTTCGACGCGAGGAGAACGCCAGCGCCGGGTGGCCCGCAGGCTGGC	3839
dw3B1_Mutant	TTCCGGCGCGTGTCCGCAACGAGGTGGCGTGGTTCGACGCGAGGAGAACGCCAGCGCCGGGTGGCCCGCAGGCTGGC	3840
Dw3B	GCTGGACGCGCAGAACCTGCGTTCCGCCATCGGCGACCGCATCTCGGTGATCGTGCAGAACTCGGCGCTGCTGCTGGTGG	3919
dw3B1_Mutant	GCTGGACGCGCAGAACCTGCGTTCCGCCATCGGCGACCGCATCTCGGTGATCGTGCAGAACTCGGCGCTGCTGCTGGTGG	3920
Dw3B	CCTGCACGGCGGGGTTCTGTGCTGCAGTGGCGGCTGGCGCTGGTGTCTTCCCGCTGGTGGTGGCCGCCACC	3999
dw3B1_Mutant	CCTGCACGGCGGGGTTCTGTGCTGCAGTGGCGGCTGGCGCTGGTGTCTTCCCGCTGGTGGTGGCCGCCACC	4000
Dw3B	GTGCTGCAGAAGATGTCATGAAGGGTCTTCGCGGACCTGGAGGCGGCCACGCGCGCGCCACGAGATCGCCGGGGA	4079
dw3B1_Mutant	GTGCTGCAGAAGATGTCATGAAGGGTCTTCGCGGACCTGGAGGCGGCCACGCGCGCGCCACGAGATCGCCGGGGA	4080
Dw3B	GGCCGTGCGCAACCTGCGCACCGTTCGCGCCCTTCAACGCCCAAGGATCGCGGGCTCTTCGCGCGCAACCTGCGGG	4159
dw3B1_Mutant	GGCCGTGCGCAACCTGCGCACCGTTCGCGCCCTTCAACGCCCAAGGATCGCGGGCTCTTCGCGCGCAACCTGCGGG	4160
Dw3B	GCCCGTGGCGCGTGTCTTGGAAAGGGCCAGGCGCGCGGTGCGCGCTACGGGCTGGCCAGTTCCTGTCTACGCGTCC	4239
dw3B1_Mutant	GCCCGTGGCGCGTGTCTTGGAAAGGGCCAGGCGCGCGGTGCGCGCTACGGGCTGGCCAGTTCCTGTCTACGCGTCC	4240
Dw3B	TACGCGCTGGGCTCTGGTACGCGGCTGGCTCGTCAAGCAGCGGCTCTCCGACTTCTCCCGGCCATCCGGGTCTTTCAT	4319
dw3B1_Mutant	TACGCGCTGGGCTCTGGTACGCGGCTGGCTCGTCAAGCAGCGGCTCTCCGACTTCTCCCGGCCATCCGGGTCTTTCAT	4320
Dw3B	GGTGTCTATGGTCTCCGCCAACGGCGCGCGGAGACGCTCACCTGGCGCGGACTTCGTCAAAGGCGGCGCGCCATGC	4399
dw3B1_Mutant	GGTGTCTATGGTCTCCGCCAACGGCGCGCGGAGACGCTCACCTGGCGCGGACTTCGTCAAAGGCGGCGCGCCATGC	4400
Dw3B	GATCCGCTTCGAGACCATCGACAGGCGCACCGAGACGGAGCCGACCGGACCGCGCCACTGCCATCGGAGCGG	4479
dw3B1_Mutant	GATCCGCTTCGAGACCATCGACAGGCGCACCGAGACGGAGCCGACCGGACCGCGCCACTGCCATCGGAGCGG	4480
Dw3B	GTCTCCGTGGAGCTGCGGCACGTGGACTTCTGCTACCCGCTCGCGGCGGACGTCGCGGTGCTGCAGGACCTCAGCCTGCG	4559
dw3B1_Mutant	GTCTCCGTGGAGCTGCGGCACGTGGACTTCTGCTACCCGCTCGCGGCGGACGTCGCGGTGCTGCAGGACCTCAGCCTGCG	4560
Dw3B	CGCGCGCGGGGAAGACGCTGGCGCTCGTGGGGCCAGCGGCTGGCGCAAGAGCTCCGTGCTCGCGCTCATCCAGCGCT	4639
dw3B1_Mutant	CGCGCGCGGGGAAGACGCTGGCGCTCGTGGGGCCAGCGGCTGGCGCAAGAGCTCCGTGCTCGCGCTCATCCAGCGCT	4640
Dw3B	TCTACGAGCCACCTCGGGCCGCTCTCCTCGACGCGCGGACGCGCGCAAGTACAACCTCGGGGCTGCGGGCGCGCC	4719
dw3B1_Mutant	TCTACGAGCCACCTCGGGCCGCTCTCCTCGACGCGCGGACGCGCGCAAGTACAACCTCGGGGCTGCGGGCGCGCC	4720
Dw3B	ATCGCGCTCGTGGCGCAGGAGCCCTTCTCTTCGCGGCCACCATCCACGACAACATCGCCTACGGCCGCGAGGGCGCCAC	4799
dw3B1_Mutant	ATCGCGCTCGTGGCGCAGGAGCCCTTCTCTTCGCGGCCACCATCCACGACAACATCGCCTACGGCCGCGAGGGCGCCAC	4800
Dw3B	GGAGGCGGAGGTGGTGGAGGCGGCCACGCAAGGCAACAGTTCATCTCCGCGCTGCGCCAGGGGTACAGGAGCG	4879
dw3B1_Mutant	GGAGGCGGAGGTGGTGGAGGCGGCCACGCAAGGCAACAGTTCATCTCCGCGCTGCGCCAGGGGTACAGGAGCG	4880
Dw3B	AGGTCCGCGAGGCGGGGTGCAGCTCTCCGGGGGCGAGCAGCGCATCGCGGTGGTTCGGGCGCTTGTCAAACAGGCG	4959
dw3B1_Mutant	AGGTCCGCGAGGCGGGGTGCAGCTCTCCGGGGGCGAGCAGCGCATCGCGGTGGTTCGGGCGCTTGTCAAACAGGCG	4960
Dw3B	GCCGTTCTGCTCCTCGACGAGGCCACGACGCGCTGGACGCGAGTCGGAGCGCAGCGTGCAGCAGGCGCTGGACCGCA	5039
dw3B1_Mutant	GCCGTTCTGCTCCTCGACGAGGCCACGACGCGCTGGACGCGAGTCGGAGCGCAGCGTGCAGCAGGCGCTGGACCGCA	5040
Dw3B	CGCCAAGACCCGTAGCACCAACACCATCGTCTCGCGCACCGCCTCGCCACCGTCCCGGACGACACACCATCGCCGCTCA	5119
dw3B1_Mutant	CGCCAAGACCCGTAGCACCAACACCATCGTCTCGCGCACCGCCTCGCCACCGTCCCGGACGACACACCATCGCCGCTCA	5120
Dw3B	TCCGACGCGCAAGGTCGTCGAGCAGGGATCGCACTCACACCTGCTCAACCACCCCGACGGAACCTACGCGCGCATG	5199
dw3B1_Mutant	TCCGACGCGCAAGGTCGTCGAGCAGGGATCGCACTCACACCTGCTCAACCACCCCGACGGAACCTACGCGCGCATG	5200
Dw3B	CTCCACCTCCAGCGACTCACCAGCCATCCACTTCCATAA	5238
dw3B1_Mutant	CTCCACCTCCAGCGACTCACCAGCCATCCACTTCCATAA	5239

Figure 15. Part 2—Alignment of the wild-type *Dw3B* allele nucleotide sequence with the *dw3B1* allele. The *dw3B2* allele sequence is not shown as it was identical to *dw3B1*. The location of the adenine nucleotide insertion at position 3057 in *dw3B1* is indicated. Sequences were aligned with Clustal Omega (Sievers et al., 2011) using the default parameters and displayed using TEXshade (Beitz, 2000).