Supplementary Information

Title: Epigenetic Regulation of Intronic Transgenes in *Arabidopsis* Authors: Kenji Osabe, Yoshiko Harukawa, Saori Miura, and Hidetoshi Saze.

Supplementary Figs. S1-S11. Supplementary Table S1.





Supplementary Figure 1. Gene and T-DNA structures inserted into the AGAMOUS locus. (A) T-DNA sequences determined by sequencing analysis, and expected sizes of DNA fragments after digestion with EcoRI and HindIII. The red, double bar represents the 5' junction of the T-DNA insertion site analysed by bisulfite-sequencing. (B) Southern analysis of DNAs from wild-type Columbia (WT Col) and segregating progen of ag T-DNA mutant. DNAs were digested with either EcoRI (upper two panels) or HindIII (the lower panel) and hybridised with probes indicated in (A). Note that the left border probe contained part of a flanking sequence from the AG locus and therefore detected a wild-type DNA fragment from the AG locus. * represents a band of unknown origin. (C) A gel image of PCR-genotyping for AG. M: 100bp DNA ladder marker.



Supplementary Figure 2. Gene and T-DNA structures inserted into the *LEAFY* locus. (A) A T-DNA sequence determined by sequencing analysis, and expected sizes of DNA fragments after digestion with *EcoRI*. The double bar represents the 5' junction of the T-DNA insertion site analysed by bisulfite-sequencing. (B) Southern analysis of DNAs from wild-type Columbia (WT Col) and segregating progeny of a *lfy* T-DNA mutant. DNAs were digested with *EcoRI* and hybridised with probes indicated in (A). Note that the left border probe contained part of the flanking sequence from the *lfy* locus; therefore, it detected a wild-type DNA fragment from the *LFY* locus. (C) A gel image of PCR-genotyping for *LFY*. M: 100bp DNA ladder marker.



Supplementary Figure 3. Gene and T-DNA structures inserted into the *GLABRA2* locus. (A) T-DNA sequence determined by sequencing analysis, and expected sizes of DNA fragments after digestion with *EcoR*I. The gray box represents a sequence that was not determined by sequencing analysis. The double bar represents the 5' junction of the T-DNA insertion site analysed by bisulfite-sequencing. (B) Southern analysis of DNAs from wild-type Columbia (WT Col) and segregating progeny of *gl2* T-DNA mutant. DNAs were digested with either *EcoR*I (upper two panels) or *Hind*III (lower panel) and hybridised with probes indicated in (A). Note that the left border probe contained part of the flanking sequence from the *GL2* locus; therefore it detected a wild-type DNA fragment from the *GL2* locus. * represents bands of unknown origin. (C) A gel image of PCR-genotyping for*GL2*. M: 100bp DNA ladder marker.



В

Cross	F1 (epi) genotype	F2 phenotype	n	p-value*
<i>lfy*/lfy*</i> #1 x <i>lfy/</i> +	<i>lfy*/+</i> #1-1	WT: 34 <i>lfy</i> : 0	34	-
	<i>lfy*/+</i> #1-2	WT: 33 <i>lfy</i> : 0	33	-
	<i>lfy*/lfy</i> #1-1	WT: 27 <i>lfy</i> : 0	27	0.0027
	<i>lfy*/lfy</i> #1-2	WT: 25 <i>lfy</i> : 0	25	0.0039
<i>lfy*/lfy* #2 x lfy/</i> +	<i>lfy*/+</i> #2-1	WT: 29 <i>lfy</i> : 0	29	-
	<i>lfy*/+</i> #2-2	WT: 31 <i>lfy</i> : 0	31	-
	<i>lfy*/lfy</i> #2-1	WT: 33 <i>lfy</i> : 0	33	0.0009
	<i>lfy*/lfy</i> #2-2	WT: 31 <i>lfy</i> : 0	31	0.0013

*Chi-square test with the expectation of 3:1 segregation ratio.

Supplementary Figure 4. A paramutation-like effect of the suppressed lfy^* allele. (A) Crossing scheme of lfy and lfy^* mutants. (B) A summary table of the lfy phenotyping in the F2 population. Results from two independent crosses are shown.

AGtdm : Accession accessio	ACGAATCGTC ACCAATCGTC ACCAATCGTC ACCAATCGTC ACCAATCGTC ACCAATCGTC ACCAATCGTC ACCAATCGTC ACCAATCGTC CTCTAGCCG TCTCTAGCG TCTCTAG
Type1-1: ICCCCCC IN ACCANCEGARCIA GENERAL NUMBER OF CONTENDED AND CONCERNANCE AND CONTENDED AND	ACCAATCGTC ACCAATCGTC ACCAATCGTC ACCAATCGTC ACCAATCGTC ACCAATCGTC TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCC CCCATTTAGC CCCATTTAGC CCCATTTAGC CCCATTTAGC GGGTAGCATT GGGTAGCATT
Type1-2: ACGGCGTACCANTCGCARGCARGENGENATTCCTCTCCTCTGAGGAAATCTGGCARGENAGTCGAATCCAANCCGARGCAACCAACAC Type2-2: ACGGCGTACCANTCGCAGCTAGGAGGAGATTCCTCTCCTC	ACCAATCGTC ACCAATCGTC ACCAATCGTC ACCAATCGTC ACCAATCGTC TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG CCCATTTAGCG CCCATTTAGCG CCCATTTAGCG CCCATTTAGCG CCCATTTAGCG CCGATTTAGCG CCGATTTAGCG CCGATTTAGCG CCGATTTAGCG CCGATTTAGCG CCGATTTAGCG CCGATGGCATT CCCGATTTAGCG
Type2-2: ICCGCGTIACCANTCGCAGCIAGCAGGAGATTCCTCTCCTCTCCTCGAGAAATCTGGGAAGCAAAGCGAAATCCAGAAATCCAAACGAAATCAGAAATCCGAAATCAAAATCAAAACGAAATCAAAACGAAATCAAAACGAAATCAAAACGAAATCAAAACGAAATCAAAATCAAAATCAAAACGAAATCAAAACGAAATCAAAATCAAAACGAAATCAAAACGAAATCAAAATCAAAACGAAATCAAAATCAAATCAAAACGAAATCAAAACGAAATCAAAATCAAAACGAAATCAAAATCAAAACGAAATCAAAATCAAAATCAAAACGAAAATCAAAATCAAAACGAAAATCAAAATCAAAACGAAAATCAAAATCAAAATCAAAACGAAATCAAAATCAAAATCAAAACGAAAATCAAAATCAAATCAAAACGAAATCAAAATCAAATTGGAAAGCTGGAAATCGAAATCCAAAATCAAAATCAATTGGAAAGCTGGAAGTCGAAATCGGAACTCAACGGTCAATGGAAGTCGAAATCGAAATCGAAATCACAATTGGAAAGTCGAAAGTCGAAATCGAAATCGAAATCGAAATCACAATTGGAAAGTCGAAAGTCGAAATCGAAATCGAAATCACAATTTGGAAAGCTGAAAGTCGAAAGTCGAAATCGAAAGTCGAAATTGGAAAGTCGAAAGTCGAAATTGGAAAGTCGAAAGTCGAAATTGGAAAGTCGAAGTCGAAAGTCGAAAGTCGAAAGTCGAAAGTCGAAAGTCGAAAGTCGAAAGTCGAAAGTCGAAAGTCGAAAGTCGAAAGTCGAAGTCGAAAGTCGAAAGTCGAAGTTGGAAGTGGAAAATTAAATTAAATTAAATTAAATTAAATTAAAATTAAATTAAATTAAATTAAAA	ACGAATCGTC ACCAATCGTC ACCAATCGTC ACCAATCGTC CTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG CCCGATTTAGCG CCGATTTAGCG CCGATTTAGCG GGGTAGGATT GGGTAGCATT GGGTAGCATT
Type-2-2: NCCGCCTINCCANCCENCENCENCENCENCENCENTCOCTECCENTENCENANTCIGENANCENCENANCEMACCENCENALACEN Type-3: NCCGCCTINCCANTCEGAGETAGEAGEAGATTCCTCTCCCTTECCENTENGEAAATCCTGGCAAAGEANATCAGAANCENANCEGAAACEAACACA Type-3: NCCGCCTINCCANTCEGAGETAGEAGEAGATTCCTCTCCCTTCCCTGGCAAGEAAAATCCAGAAAATCCAAATCCAAATCGAAATCCAAATTGCAAAGTCGAAAGTCGAAGTCGAAGTCGAAGTCGAAGTCGAAGTCGAAGTCGAAGTCGAAGTCGAAGTCGAAGTCGAAGTCGAAGTCGAAGTCGAAGTCGAAATCGAAATTGCAAAATTGAAATTGAATTGCAAAAGTTAACTAAAATTGAATTGCAAAGTCGAAAGTCGAAATTGCAAAATTGAAATTGGATTGCTAAAGTAAAAGTTAACTAAAAGTTAACTAAAAGTTAACAAAAGGTTTCCTCTCTCT	ACGAATCGTC ACCAATCGTC ACCAATCGTC 200 TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG CCGATTTAGC CCGATTTAGC CCGATTTAGC GGGTAGCATT GGGTAGCATT
Type3 INCEGEGETACEAATECEGAECTAGEAEGAANTECTCCCCCTCAEGEAAAATECEGEAEAACTECEAAAACCEAATECEAATECEAATECEAATECEAATECEAATECCTCCCCCTCAEGEAAAATECTGEAEAACTECEAAAACCEAATECEAATECEAATECEAATECEAATECCTCCCCTCCCCCTCAEGEAAAATECTCCEAAAACCEAAAACCEAATECEAATECEAATECEAATECCTCCCCTCCCCTCCCCCAEAAACCACACAAATECEAAACGEAATECEAATECEATECTCCCCCTCCCTCCTCTCTTCTTCTCAATECEGAEAGEAATECEAATECEAATECTTTCTCCCAEAAAAGETTACCAEGEATECTCTTTCTCTTTCTTCCAAACGEAATECEAATECEATECCCAEAAATECETTTCTTTCTTTC	ACCAATCGTC ACCAATCGTC 200 TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG CCCATTTAGG CCCATTTAGG CCCGATTTAGG 400 GGGTAGCATT GGGTAGCATT GGGTAGCATT
MT_AG NCGCCGIACCAATCCGAGCTAGGAGGAGATTCCCTCTCCCCTTAGGAGAAACTCGAGGAAAGATCCAAAACGGATCGACAACAACAA * 120 * 140 * 160 * 180 AGCdma AAGTCACTTTTGCAAAGGACAAATGGTTTCCCCAACGAAAGCTTACCAGCTCTGTTCTTGTGTAGGCGAAGAGCGACTAACGGGCATCAACGGCGACTACGCGGACTACGCGGACTACGCGGACTACGCGGACTACGCGGACTACGCGGACTCACGGCGCTCCGTGTTCTTGTGTGGGAGGCGACTCAAGGGACGACGACGGCGACTCAAGGGAGCTTACCAGCTCCGTGTCTTTGTGGAGGCGACGACGACGCACGGCGCTCTGTGTTTGGGAGGGA	ACCAATCGTC 200 TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG CCCATTTAGG CCCGATTTAGG CCCGATTTAGG CCCGATTTAGG GGGTAGGATT GGGTAGGATT GGGTAGGATT
* 120 * 140 * 160 * 180 * AGtdma * AAGTCACTTTTGCAAAGGTACAAATGGTTCCCCAACAAAGCTTACCAGGTCTCGTGTCTTGTGCCGAAGTCGCACTACGGCGCTTT Type1-2 * AAGTCACTTTTGCAAAGGTACAAATGGTTGCCCAACAAAGCTTACCAGCTCTCGTGTCTTGTGCGAGGCCAGGCCACTCACGGCCT Type1-2 * AAGTCACTTTTGCAAAGGTACAAATGGTTGCCCAACAAAGCTTACCAGCTCTGTGTCTTGTGCGAGGCGAGGCCACTCACGGCCT Type2-2 * AAGTCACTTTTGCAAAGGTACAAATGGTTGCCCAACAAAGGTTACCAGCTCTGTGTCTTGTGTGCGAGGCGAGGCCACTCATCGTCTT Type2-2 * AAGTCACTTTTGCAAAGGTACAAATGGTTGCCCAACAAAGGTTACCAGCTCTGTGTCTTGTGTGCGAGGCGACTCATCGTCGTT Type2-2 * AAGTCACTTTTGCAAAGGTACAAATGGTTGCCCAACAAAGGTTACCAGCGCTCTGTGTCTTGTGTGGTGGCGAGGCACTCATCGTCGTT Type2-2 * AAGTCACTTTTGCAAAGGTACAAATGGTTGCCCAACAAAGGTTACCAGGTCTGTGTCTTGTGTGTG	200 TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG CCCATTTAGG CCCATTTAGG CCCATTTAGG CCCATTTAGG 400 3GGTAGGATT GGGTAGGATT
* 120 * 140 * 160 * 160 * 180 * AGtdma : AAGTCACTTTTGCAAAGGTAGAAATGGTTGGTCAAGAAGGTTACGAGGTCTGTGTTTTGTGATGGTGAAGTCGAAGTCGATGGTGT Type1-1 : AAGTCACTTTTGCAAAGGTAGAAATGGTTGGTCAAGAAGGTTACGAGGTCTGTGTTTTGTGATGGTGAAGTCGAACTCATGGTGTT Type2-2 : AAGTCACTTTTGCAACGTAGAAATGGTTGGTCAAGAAGGTTACGAGGTCTCTGTTCTTGGTAGTGGCAAGTCGCACTCATCGTCGTT Type2-2 : AAGTCACTTTTGCAACGTAGAAATGGTTTGGTCAAGAAGGTTACGAGGTCTCGTTCTTGTGATGGTGAAGTCGCACTCATCGTCTT Type2-2 : AAGTCACTTTTGCAACGTAGAAATGGTTTGGTCAAGAAAGGTTACGAGGTCTCGTTCTTGTGATGGTGAAGTCGCACTCATCGTCTT Type2-2 : AAGTCACTTTTGCAACGTAGAAATGGTTTGGTCAAGAAAGGTTACGAGCTCTCGTTCTTGTGATGGTCAAGTCGCACTCATCGTCTT Type2-2 : AAGTCACTTTTGCAACGTAGAAATGGTTTGGTCAAGAAAGGTTACGAGCTCTCGTTCTTGTGATGGTCAAGTCGCACTCATCGTCTT Type3 : AAGTCACTTTTGCAACGTAGAAATGGTTTGGTCAACAAAGGTTACGCAGCTCTCGTTCTTGTGTAGTGGCAAGTCGCACTCATCGTCTT Type2-1 : TGGTCGTCTTATCGAACACTTAACAACAC Type2-2 : TGGTCGTCTTATCGTAACAACAC Type3 : TGGTCGTCTTATCGAGTACTCTAACAACAC Type2-1 : TGGTCGTCTTATCGAGTACTCTAACAACAC Type3 : TGGTCGTCTTATCGAGTACTCTAACAACAC Type3 : TGGTCGTCTTATCGAGTACTCTAACAACAC Type2-2 : TGGTCGTCTTATCGAGTACTCTAACAACAC Type2-2 : TGGTCGTCTTATCGAGTACTCTAACAACAC Type2-2 : TGGTCGTCTTATCGAGTACTCTAACAACAC Type2-2 : TGGTCGTCTTATCGAGTACTCTAACAACAC Type2-2 : TGGTCGTCTTATCGAGTACTCTAACAACAC Type2-2 : TTTTTCCACTTTTTTTTTTTTTTTTAAGGCATAAATTTTTTACCAGGGTTTTCCACTCTTCTAAGTAGTTGGTAGTTGGTTAATTTTTTAAAGTTTTTTTAAGTTTATTTTAAAGTTTATTTTAAAGTTTATTTTTAAGTTTATCTTTAAGCATAAATTTTTTCCACACACA	200 TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG CCGATTTAGG CCGATTTAGG CCGATTTAGG 400 GGGTAGGATT GGGTAGGATT GGGTAGGATT GGGTAGGATT
* 120 * 140 * 160 * 180 * AGtdaa: * AAGTCACTTTTGCAAAGGTAGAAAGGTTTGGTCAAGAAAGGTTAGCAGCGCTCTGTGTTTTGGAAGTGGAAGTGGCACTCAACGTCTTTGTTGGAAGTGGCAAGTGGCAACGTGCACTCATCGTCTTTGTP121 * AAGTCACTTTTGCAAACGTAGAAATGGTTTGGTCAAGAAAGGTTACCAGCGTCTGTGTTTTGGAAGTGCGAAGTGGCACTCATCGTCTTTTGGAAGTGCCAACGTGCACTCATCGTGTTTTGGAAGTGCAAGTGCGACTGCAGTGTTGGTCAAGAAGGTTAGCAAGGTTGGCAAAGTGCGAAGTGCGAAGTGCGAAGTGCGACTCATCGTGTTTTTGGAAGGCGAAGTGCGAAGTGCGACTCATCGTGTTTTTGGAAGGCGAAGTGCGAAGTGCGACTCATCGTGTTTTTGGAAGGCGAAGTGCGAAGTGCGACTCATCGTGTTTTTGGAAGGCGAAGTGCGAAGTGCGACTCATCGGCTTTTGTGAAGGGCGAAGTGCGAAGTGCGACTCATCGGCGACTCATCGTCTTTTGGAAGGCGAAGTGCGAAGTGCGACTCATCGGCGCTCTTTGTGAAGGGCGAAGTGCGAAGTGCGACTCATCGGCGCTCTTTGGAAGTGCGAAGTGCGAAGTGCGACTCATCGGCGCTCTTTGGAAGTGCGAAGTGCGAAGTGCGAAGTGCGAAGTGCGAAGTGCGAAGTGCGAAGTGCGAAGTGCGAAGTGCGCACTCATCGGCGCTCTTTGGAAGTGCGAAGTGCGAAGTGCGAAGTGCGAAGTGCGCACTCATCGGCGCTCTTTGGAAGTGCGAAGTGCGAAGTGCGAAGTGCGCACTGATGGGAAGTTTGCAAGAGTTTGGAAGTGCGAAGTGCGCAAGTGGCGCAAGTGGGGGGCGCCCAAGTGGCGCAAGTGGCGCGAAGTGGCGGAAGTGGGGAAGTGGGGAG	200 TCTCIAGCCG TCTCIAGCCG TCTCIAGCCG TCTCIAGCCG TCTCIAGCCG CCGATTAGCG CCGATTAGGG CCGATTAGGG CCGATTAGG GGGTAGGATT GGGTAGGATT GGGTAGGATT GGGTAGGATT
AGtdna : AAGTCACTTTTGCAAAGTACAAAGGTTAGCTCAACAAAGGTTACCAGGCTCCTGTTCTTGTCATGCTAAGTGCCCATGCGCGTCT Typel-2 : AAGTCACTTTTGCAAAGTAAAGTAGAAAGGTTAGCTAAGAAAGCTTACGAGGTCTCGTGTCTTTGTGATGCTCAAGGAGGCACTCATCGTCTT Type2-2 : AAGTCACTTTTGCAAAGTAGAAAGGTTAGCTAAGAAAGCTTACGAGGTCTCGTGTCTTTGTGATGCTCAAGGTGCCGACTCATCGTCTT Type2-2 : AAGTCACTTTTGCAAAGTAGAAAGGTTAGCTAACAAAGCTTACGAGGTCCTGTGTCTTGTGATGCTCAAGGTGCCCATGGTCTT Type2 : AAGTCACTTTTGCAAAGTAGAAAGGTTGCTCAACAAAGCTTACGAGGTCCTGTGTCTTGTGATGCTCAAGGTGCCCATGGTCTT Type3 : AAGTCACTTTTGCAAAGTAACAAAGGTTGCTCAACAAAGCTTACGAGGTCCTGTGTTGTTGTAGGTGAAGGGCACTCATCGTCTT Type3 : AAGTCACTTTTGCAAAGTACGAAAGGTTGCTCAACAAAGCTTACGAGGTCCTGTGTTTTGTCAAGGCAAGTCGACGTACGAGGCACTCATGGTCTT Type3 : AAGTCACTTTTGCAAAGTACGAAAGGTTGCTCAACAAAGCTTACGAGGTCCTGGTCTTTGTGATGGCCAAGTCGCCCATGGTCGT AGGCGCTCTATGAGTACGTAACAACAC C C C C C C C 	TCTCIAGCCG TCTCIAGCCG TCTCIAGCCG TCTCIAGCCG TCTCIAGCCC TCTCIAGCCC CCCATTIAGC CCCATTIAGC CCCATTIAGC CCCATTIAGC 400 SGGTAGCATT SGGTAGCATT
Type1-1: AAGTCACTTTTTGCAAAGTACAAAGGTTAGCAAAGGCTTACCAGGCTCCTGTTCTTGTCAAGTCGCACAGGCCACTCATGGCTCTT Type2-1: AAGTCACTTTTTGCAAAGGTAGAAAGGTTTGCTCAAGAAGCTTACGAGGCTCCTGTTCTTGTGATGCTGAAGTGGCACTCATGGACGACTTTTGCAAGAGGACTATGGCACTCATGGCACTCATGGACACTAGGAGGACTATGGCACTCATGGAGGACTATGGCACTCATGGAGACTATGGCACTCATGGAGGACTCATGGACTCATGGCACTCATGGAGGACTCATGGAGGACTCATGGACGACTATGGCACTCATGGAGGACTCATGGAGGACTCATGGAGGACTGACAGACTTTGGCACTCATAGGAGGACTGACAGACTTTGGCACTCATGGAGGCACTGACGACGACTTTGGCACTCATGGCACTGACGACGACTGGCACTATGGACGACTGACGACGACTGACT	TCTCIAGCCG TCTCIAGCCG TCTCIAGCCG TCTCIAGCCG TCTCIAGCCG TCTCIAGCCG CCCATTIAGG CCCATTIAGG CCCGATTIAGG 400 GGGTAGCATT GGGTAGCATT GGGTAGCATT
Type1-2: AAGTCACTTTTTGCAAAGTACATAGGTTTGCTCAACAAAGCTTLACAGGCTCCTGTTCTTTGCAAGTCGCACAGGCCACTCATCGTCTTT Type2-2: AAGTCACTTTTTGCAAAGGTAGAATGGTTTGCTCAACAAAGCTTLACAGGCTCCTGTTCTTTGCAAGTGCAAGTCGCACGCAGTCATCGTCAT Type3: AAGTCACTTTTTGCAAAGGTAGAATGGTTTGCTCAACAAAGCTTLACAGGCTCCTGTTCTTTGCAAGTGCCAAGTGCCACTCATCGTCTT Type3: AAGTCACTTTTGCAAAGGTAGAATGGTTTGCTCAACAAAGCTTLACGAGCTCCTGTTCTTTGCAAGTGCCAAGTGCCACTCATCGTCTT Type3: AAGTCACTTTTGCAAACGTAGAATGGTTTGCTCAACAAAGCTTLACGAGCTCCTGTTCTTTGCAAGTGCCAAGTGCGCACTCATCGTCTT YT AAGTCACTTTTGCAAACGTAGAATGGTTTGCTCAACAAAGCTTLACGAGCTCCTGTTCTTTGCAAGTGCGCAAGTGCCCAACGCACTCATCGTCTT YT AAGTCACTTTTGCAAACGTACAAGTGTTTGCTCAACAAAGCTTACGAGCTCCTGTTCTTTGCAAGTGCCCAAGTGCACTCATCGTCATCGCTCAACGACTT YT Y AG AGGCCTCTTTGCAACAACAACGTTCTTCTCTCTTCTTCTCTCTC	TCTCIAGCCG TCTCIAGCCG TCTCIAGCCG TCTCIAGCCC 3000 CCGATTIAGCG CCGATTIAGG CCGATTIAGG CCGATTIAGG GGGTAGGATT GGGTAGGATT GGGTAGGATT
Type2-1: AAGTCACTTTTGCAAAGTAGAAAGGTAGAAAGGTTTGCCAACAAAGGTTACGAGGTCCTGTTCTTGTAGTGCAAGAAGGTGCCAAGGGCGCCAAGGGCGCCAAGGGCGCAAGGGCGCAAGGGCGCACGACG	TCTCTAGCCG TCTCTAGCCG TCTCTAGCCG 300 CCGATTTAGG
Type2-2: AAGTCACTTTTGCAAAGTAAAGTIACAAAGGTTAGCAAAGCTTAACAGGCTCOCGTCGTTGTTGTCAAAGTGCACAAGTCGACGCAAGCGCACCAAGGCGCCAAGGCGCCAAGGCGCCAAGGCGCACGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCACGCAAGGCGCACGCACGCGAGGCGCCGGCGCGGCG	TCTCIAGCCG TCTCTAGCCC 3000 CCGATTTAGG CCGATTTAGG CCGATTTAGG GGGTAGGATT GGGTAGGATT GGGTAGGATT GGGTAGGATT
Type3 : AAGTCACTTTTTGCAAAGGTAGAAAGGTAGAAAGGTTAGGAAGGTCTTGGTGGTGTTTGTGAAGGTGCCAAGGCGCACTCATGGTCTT YTAG <td: aagtcactttttgcaaaggtaggaaaggttaggaaggttctggtggtgttgtggaaggcgcactcatggtctt<="" td=""> YTAG <td: aagtcacttttgcaaaggtaggaaaggttggtcaagaaaggttaggaggtctctggtgttgtgaaggcgcactcaaggcactcatggtctt<="" td=""> YTAG <td: aagtcacttttgcaaaggtaggtaggtaggaaagggtttggtcaagaaggtctctggtcttaggaggcaggc<="" td=""><td>TCTCIAGCCC TCTCIAGCCC 300 CCGATTIAGG </td></td:></td:></td:>	TCTCIAGCCC TCTCIAGCCC 300 CCGATTIAGG
$ \begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$	300 CCCATTTAGG CCCATTTAGG CCCATTTAGG CCCATTTAGG GGGTAGGATT GGGTAGGATT GGGTAGGATT
AG exon AG intron * 220 * 240 * 260 * 280 * Ascdna : TGGTGGTCTCTATGAGTACGTCAACAAG Type1-1 : TGGTCGTCTCTATGAGTACGTAACAAG Type2-2 : TGGTCGGTCTTATGAGTACCTAACAACA Type2-2 : TGGTCGGTCTTATGAGTACCTAACAACA Type3 : TGGTCGGTCTTATGAGTACCTAACAACA * 320 * 340 * 360 * 380 * Ascdna : GTTTTATGTTATGATATGAACAACAAC Type1-1 : * Type1-2 : * Type1-2 : * Type2-2 : GTTTTATGTTATGATAAGGATAAAACTTTTACCAGGGGTTTCCGATCTTCTAGATTGGATTCTCTTCTTCTGAAGGCCTTATGGT Type2-2 : * * 320 * 340 * 360 * 380 * Ascdna : GTTTTATGTTATGTTAAAGCATAAACTTTTACCAGGGGTTTCCGATCTTCTAAGCTAGGATTGGATCTCTTCTTCTGAAGGCCTTATGGAT Type1-1 : * Type1-2 : * Type2-2 : * Type2-2 : * Type2-2 : * Type2-2 : * Type2-1 : * Type1-1 : * Type2-2 : * Type2-2 : * Type2-2 : * Type2-2 : * Type2-2 : * Type2-2 : * Type1-1 : * Type1-1 : * Type1-2 : * Type2-2 : * Type2	300 CCGATTTAGG
* 20 * 260 * 280 * AGtdna : TGGTCGTCTATCAGGIACTCTAACAACA GTTCTCTCTCTCTCTCGTGCTCTATCATTATAATATTAAATIATTTTTAAAGTC Type1-2 : TGGTCGTCTCTATCAGIACCTAACAACA Type2-1 : TGGTCGTCTCTATCAGIACTCTAACAACA Type2-2 : TGGTCGTCTCTATCAGIACTCTAACAACA Type3 : TGGTCGTCTCTATCAGIACTCTAACAACA Type2 : 340 * 360 * 380 * AGtdna : GTTTTATGTTAAGGCATAAATCTTTTACGAGGGGTTTCCGATCTTCTAAGGTAGATTTGATTCTTCTTCTCTGAAGGCCTCTATGG Type2 :	300 CCCATTTAGG
AGtdma : TGGTCGTCTATCAGIACTCTAACAACAG GTTTCTTCTTCTTCTCGTGCTCTATCATTATTATAATTAAT	CCGATTTAGG CCGATTTAGG CCGATTTAGG GGGTAGCATT GGGTAGCATT GGGTAGGATT
Type1-1: TGGTCGTCTTATGAGTACCTAACAACAG Type1-2: TGGTCGTCTTATGAGTACCTAACAACAG Type2-1: TGGTCGGTCTTATGAGTACCTAACAACAG Type3-2: TGGTCGGTCTTATGAGTACCTAACAACAG Type3-2: TGGTCGGTCTTATGAGTACCTAACAACAG Type3-2: TGGTCGGTCTTATGAGTACCTAACAACAG WT_AG TGGTCGTCTTATGAGTACCTAACAACAG WT_AG TGGTCGTCTTATGAGTACCTAACAACAG * 320 * 340 * 360 * 380 * AGstdna: GTTTTATGTTAAGTTAAAGCATAAACAG * 320 * 340 * 360 * 380 * Yppe1-1:	400 GGGTAGGATT GGGTAGGATT
Type1-2 : TGGTCGTCTCTATCAGTACCCTAACAACAC Type2-1 : TGGTCGTCTTATCAGTACCTAACAACAC Type3 : TGGTCGTCTTATCAGTACCTAACAACAC Type3 : TGGTCGTCTTATCAGTACCTAACAACAC TGGTCGTCTTATCAGTACCTAACAACAC TGGTCGTCTTATCAGTACCACAACAC Type3 : TGGTCGTCTTATGAGTACCTAACAACAC TGGTCGTCTTATGAGTACTCTAACAACAC TGGTCGTCTTTATGAGTACTCTAACAACAC * 320 * 340 * 360 * 380 * AGtdna : GTTTTATGTTATGTTAAGGCATAAACAC Type1-1 :	400 GGGTAGCATT GGGTAGCATT
Type2-1: TGGTCGTCTATCAGTACCTTAACAACA Type2-2: TGGTCGGTCTTATCAGTACCTTAACAACA Type3:: TGGTCGTCTTATCAGTACCTAACAACA WT_AG TGGTCGTCTTATCAGTACCTAACAACA WT_AG TGGTCGTCTTATGAGTACCTTAACAACA * 320 * 340 * 360 * 380 * AGtdna : GTTTTATGTTAAGGTACTTAACAACAC Type1-1:	400 GGGTAGGATT GGGTAGGATT
* 320 * 340 * 360 * 380 * AGtdna : GTTTTATGTTATGTTAAGAGAACATTTTTAAGAGAACATTTGATTATGTTAAGATAATTAAATTAAATTAAATTAAATTAAATTAAATTATT	CCCATTTAGG 400 GGGTAGGATT
* 320 * 340 * 360 * 380 * AGtdna : GTTTTATGTTATGTTAAGAACAG	400 GGGTAGCATT
ATTOC : SOURCECTATEGRACIACTECTATEGRACIACITICICIENTETETETETETETETETETETETETETETETETETET	400 GGGTAGGATT GGGTAGGATT
* 320 * 340 * 360 * 380 * AGtdna : GTTTTATGTTATGTTAAGCATAAATCTTTTACGAGGGTTTCGATCTTCTAAGCTAGATTTCATTCTCTTCTGAATGCTCTATGT Type1-1 : Type2-1 : Type2-2 : Type3 GTTTTATGTTATGTTAAGCATAAATCTTTTACGAGGGTTTCGATCTTCTAAGCTAGATTTCATTCTCTTCTGAATGCTCTATGG Type2-1 : Type3 GTTTTATGTTATGTTAAGCATAAATCTTTTACGAGGGTTTCGATCTTCTAAGCTAGATTTGATTCCTTCTTCTGAATGCTCTATGG WT_AG : * 420 * 440 * 460 * 480 * AGtdna : ATTTTTCACTTTTTCCTCATAATTCCACACACATATATATATAAACACACTAACATTAGTGGGAATATTTGTTTG	400 GGGTAGGATT
* 320 * 340 * 360 * 380 * AGtdma : GTTTTATGTTATGTTAAGGATAAAATCTTTTACGAGGGTTTTCGATCTTCTAAGGCTAGATTTCATTCTCTTCTTCTAAGGCTATAGGCTATGATTCGTTCTTCTAAGGCTATGGCTATGGCTATGGCTATGGTCTTATGTTAAGGCTAAAAGCATAAAATCTTTTACGAGGGTTTTCGATCTTCTAAGGCTAGATTTGATTCTCTTCTTCTAAGGCTACATTGATTCCTCTTCTGAATGGCTCTTATG Type2-2 :	400 GGGTAGGATT GGGTAGCATT
* 320 * 340 * 360 * 360 * 380 * Ascdna : GTTTTATGTTATGTTAAAGCATAAATCTTTTACCAGGGGTTTCGAATCTCTAAGCIACATTGATTCTCTTCTTCTGAATGCTCTAAG Type1-1 :	400 GGGTAGGATT
AGtdma : GTTTIATGTTIATGTTIAAGCATABATCTTTIACCAGGGTTTTCGATCTTCTAAGCTAGATTTCATTCTCTTCTTCTAAGGTCTTATAGGT Type1-1 : Type2-2 : Type3 : GTTTTATGTTIATGTTAAAGCATABATCTTTTACCAGGGTTTTCGATCTTCTAAGGTAGATTTGATTCTCTTCTTCTGAATGGTCTTATGGT Type3 : GTTTTATGTTIATGTTAAAGCATABATCTTTTACCAGGGGTTTTCGATCTTCTAAGGTAGATTTGATTCTTCTTCTGAATGGTCTTATGGT Type3 : GTTTTATGTTIATGTTAAAGCATABATCTTTTACCAGGGGTTTTCGATCTTCTAAGGTAGATTTGATTCTTTCT	GGGTAGGATT
Type1-1:	GGGTAGGATT
Type1-2 : Type2-1 : Type2-2 : Type3 : GTTTTATGTTIATGTTIATAGCATABATCTTTTACCAGGGTTTCCGATCTTCTAAGCTACATTTCATTCTTCTTCTAATGCTCTAATGCTCTATAGG MT_AG : * 420 * 440 * 460 * 480 * AGtdma : ATTTTTCACTTTTTTCCTCATATTCCACACACACATATATATATABACACACTAAGTGGGAATATTTGTTGATATGTTGATATGTTTATTTTAT Type1-1 :	GGGTAGCATT
Type2-1:	GGGTAGGATT
Type2-2 :	GGGTAGGATT
Type3 : GTTTTATGTTTATGTTAAAGCATAAATCTTTTACGAGGGGTTTCGATCTTCTAAGCTAGATTTGATTCCTTCTTCTGAATGCTCTTATGGT WT_AG :	GGGTAGGATT
WT_AG : 420 * 440 * 460 * 480 * AGtdna : ATTTTTCACTTTTTCCTCATATTCCACACACATATATAT	
* 420 * 440 * 460 * 480 * AGtdma : ATTTTTCACTTTTTCCTTCATATTCCACACACATATATATATAAACACACTAACATTAGTGGGAATATTTGTTGATATGTTGATATTTATT	
Type1-2 : Type2-1 : Type2-2 : Type3 : ATTTTTCACTTTTTCCTTCATATTCCACACACATATATATATATAACACACTAACATTAGTGGGAATATTTGTTTG	TTTACTTCGG
Type2-1 : Type2-2 : Type3 : ATTTTTCACTTTTTTCCTTCATATTCCACACACATATATATATAAACACACTAACATTAGTGGGGAATATTTGTTTG	
Type2-2 : Type3 : ATTTTTCACTTTTTTCCTTCATATTTCCACACACACATATATATATAAACACACTAACATTAGTGGGAATATTTGTTTG	
Type3 : ATTTTTCACTTTTTCCCTCATATTCCACACACATATATATATATAAACACCAC	
MT_AG :	TTTACTTCGG
-	
* 520 * 540 * 560 * 580 *	600
AGGUNA . GGGIIIIIGIAACAAFTTTGIAGATCIAATTTCTTGTGTTCTTCATGTGIAIATIAATTTTCCCTTAAGACTIAAAAAAGACAAGAAGAAGAAGAATTTG	ALALATATA
Type1-2 :	
Type2-1 :	
Type2-2 :	
TYDE3 : GGGTTTTTGIAACAATTTTGIAGATCIAATTTCTTGTTCTTGTTGIAIATTAATTTTCCCTIAAGACTIAAAAAAAAAA	
WT_AG :	AA
* 620 * 640 * 660 * 680 *	AA
AGGIDA : GATATATGAAGTGAGGGAAATGGTACAAAGTTAAAGGAGATCTGAGTGAG	AA 700
Typel-1 :	AA 700 FTTTTTCTAA
Type1-2 :	AA 700 TTTTTTCTAA
Type2-1 :	AA 700 TTTTTTCTAA
Type2-2 :	AA 700 TTTTTTCTAA
Type3 :	AA 700 TTTTTTCTAA
	AA 700 TTTTTTTCTAA
WT_AG :	700 TTTTTTCTAA
WT_AG :	AA 700 TTTTTTCTAA T-E
WT_AG :	AA 700 TTTTTTCTAA T-E 800
WT_AG :	700 TTTTTTCTAA
WT_AG :	700 TTTTTTCTAA T-E 800 AATCTTACTG
WT_AG :	700 TTTTTTCTAA
WT_AG :	700 TTTTTTCTAA
WT_AG * 720 * 740 * 760 * 780 * Actdna : TGTGGAGGTTTAGATTCAGTTTGTAGAACTAAGATTCACTTGTGGGGGTGTCTTTCTT	700 TTTTTTCTAA
WT_AG * 720 * 740 * 760 * 780 * AGtdna : TGTGGAGGTTTTAGCATTCAGTTTTGTAGAACTAAGATTCACTTTGTTGGGTGTTCTTTCT	700 TTTTTTCIAA

Figure S5 (continued)

	I-DNA			~						
Ctdpp	· American	820	*	840	*	860	*	880	*	900
pe1-1	:									:
e1-2	:						GA:	TATATTGTGGTG	TAAACAAAT	GAC :
pe2-1	:									:
pe2-2	:						GA	TATATTGTGGTG	TAAACAAAT	GAC :
be3	:									:
AG	:									:
	*	920	*	940	*	960	*	980	* 1	1000
Stdna	: GCTTAGACAACT	TAATAACACATTG	CGGACGTTT	TAATGTACTO	GGGTGGTTTT	TCTTTTCACC	AGTGAGACGG	GCAACAGCTGAT	TGCCCTTCA	CGC :
pe1-1	:						TGAGACGG	GCAACAGCTGAT	TGCCCTTCA	CGC :
pe1-2	: GCTTAGACAACT	TAATAACACATTG	CGGACGTTT	TAATGTACTO	GGGTGGTTTI	TCTTTTCACC	AGTGAGACGG	GCAACAGCTGAT	TGCCCTTCA	CGC :
ype2-1	:						TGAGACGG	GCAACAGCTGAT	TGCCCTTCA	CGC :
ype2-2	: GCTTAGACAACT	TAATAACACATTG	CGGACGTTT	TAATGTACTO	GGGTGGTTTI	TCTTTTCACC	AGTGAGACGG	GCAACAGCTGAT	TGCCCTTCA	CGC :
/pe3										:
1_A0										
Gtdna	* CTGGCCCTGAG	1020 AGAGTTGCAGCAAG	* CGGTCCACGO	1040 CTGGTTTGCCC	* CCAGCAGGCGA	1060 AAATCCTGTT	* TGATGGTGGT	1080 FCCGAAATCGGC	* AAAATCCCT	100 ATA :
ype1-1	: CTGGCCCTGAGA	AGAGTTGCAGCAAG	CGGTCCACG	CTGGTTTGCCC	CAGCAGGCGA	AAATCCTGTT	TGATGGTGGT	FCCGAAATCGGC	AAAATCCCTT	ATA :
pe1-2	: CTGGCCCTGAG	GAGTTGCAGCAAG	CGGTCCACG	TGGTTTGCCC	CAGCAGGCGA	AAATCCTGTT	TGATGGTGGT	ICCGAAATCGGC	AAAATCCCTT	ATA :
pe2-1	: CTGGCCCTGAG	AGAGTTGCAGCAAG	CGGTCCACG	CTGGTTTGCCC	CAGCAGGCGA	AAATCCTGTT	TGATGGTGGT	FCCGAAATCGGC	AAAATCCCTT	ATA :
pe2-2	CTGGCCCTGAG	GAGTTGCAGCAAG	CGGTCCACG	TGGTTTGCCC	CAGCAGGCGA	AAATCCTGTT	TGATGGTGGT	CCGAAATCGGC	AAAATCCCTT	ATA :
vpe3	:									:
AG	:									:
	*	1120	*	1140	*	1160	*	1180	* 1	200
Stdna	: AATCAAAAGAA	AGCCCGAGATAGG	GTTGAGTGT	GTTCCAGTT	GGAACAAGAG	TCCACTATIA	AAGAACGTGG	ACTCCAACGTCA	AAGGGCGAAA	AAC :
pe1-1	: АААААААААААА	AAAA								:
pe1-2	: ААААААААААА	AAAA								:
pe2-1	: AATCAAAAGAAT	(AGCCCGAGATAGG)	GTTGAGTGT	GTTCCAGTT	GGAACAAGAG	TCCACIATIA	AAGAACGTGG	ACTCCAACGTCA	AAGGGCGAAA	AAC :
rpe2-2	: AATCAAAAGAAT	DDATADADDDDDA	GTTGAGTGT	TGTTCCAGTT	GGAACAAGAG	TCCACTATTA	AAGAACGTGGA	ACTCCAACGTCA	AAGGGCGAAA	AAC :
vpe3	:									:
ype3 T_AG	:									:
ype3 T_AG Gtdna ype1-1	* : CGTCTATCAGGO	1220 GCGATGGCCCACTAC	* CGTGAACCAT	1240 FCACCCAPATO	* CAAGTTTTTTG	1260 GGGTCGAGGT	* GCCGTAAAGCI	1280 ACTAAATCGGAA	t * CCCTAAAGG	1300 EAGC :
ype3 I_AG Stdna ype1-1 ype1-2	* CGTCTATCAGGO	1220 SCGATGGCCCACTA(* CGTGAACCAT	1240 FCACCCAAATC	* .AAGTTTTTTG	1260 GGGGTCCAGGT	* GCCGTAAAGCI	1280 ACTAAATCGCAA	* 1 CCCTAAAGG	.300 AGC :
ype3 [_AG Gtdna ype1-1 ype1-2 ype2-1	CGTCTATCAGGO	1220 SCGATGGCCCACTAC SCGATGGCCCACTAC	* CGTGAACCAT CGTGAACCAT		* CAAGTTTTTTC	1260 GGGGTCCAGGT	* GCCGTAAAGCI	1280 ACTAAATCGGAA	* : CCCTAAAGGO CCCTAAAGGO	1300 AGC : : AGC :
Type3 TAG Stdna Mpe1-1 Mpe1-2 Mpe2-1 Mpe2-2 Type3	* CGTCTATCAGGG 	1220 SCGATGGCCCACTA(SCGATGGCCCACTA(SCGATGGCCCACTA(* CGTCAACCAT CGTCAACCAT CGTCAACCAT CGTCAACCAT	1240 ГСАСССАВАТС ГСАСССАВАТС ГСАСССАВАТС	* CAAGTTTTTTG CAAGTTTTTTG	1260 GGGTCCAGGT GGGTCCAGGT GGGGTCCAGGT	* GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI	1280 ACTAAATCGGAA ACTAAATCGGAA ACTAAATCGGAA	* 1 CCCTAAAGG CCCTAAAGG CCCTAAAGG	1300 EAGC : : EAGC : EAGC :
ype3 r_AG Gtdna ype1-1 ype1-2 ype2-1 ype2-2 ype3 r_AG	* CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC	1220 SCGATGGCCCACTA(SCGATGGCCCACTA(SCGATGGCCCACTA(* CGTGAACCAT CGTGAACCAT CGTGAACCAT	1240 ICACCCARATO	* :AAGTTTTTTC :AAGTTTTTTC :AAGTTTTTTC	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT	* GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI	1280 ACTAAATCGGAA ACTAAATCGGAA ACTAAATCGGAA	* 1 CCCTAPAGGO CCCTAPAGGO CCCTAPAGGO	.300 AGC : : AGC : : AGC : :
ype3 T_AG Stdna ype1-1 ype1-2 ype2-1 ype2-2 ype3 T_AG	* CGTCTATCAGGG CGTCTATCAGGG CGTCTATCAGGG CGTCTATCAGGG	1220 SCGATGGCCCACTAG SCGATGGCCCACTAG SCGATGGCCCACTAG SCGATGGCCCACTAG 1320	* CGTGAACCAT CGTGAACCAT CGTGAACCAT CGTGAACCAT	1240 PCACCCAPATO PCACCCAPATO PCACCCAPATO PCACCCAPATO PCACCCAPATO PCACCCAPATO PCACCCAPATO PCACCCAPATO PCACCCAPATO PCACAPATO PCACCAPATO PCACAPAT	* :AAGTTTTTTG :AAGTTTTTTG :AAGTTTTTTG :AAGTTTTTTG :	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360	* GCCGTAPAGC1 	1280 ACTARATCGGAR 	* 1 .ccctaaaggg .ccctaaaggg .ccctaaaggg .ccctaaaggg	: : AGC : : AGC : AGC : : : :
rpe3 r_AG Stdna rpe1-1 rpe1-2 rpe2-2 rpe3 r_AG Stdna	CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC	1220 SGGATGGCCCACTA(SGGATGGCCCACTA(SGGATGGCCCACTA(SGGATGGCCCACTA(1320 IGCTTCACGGGGAA)	* CGTGAACCAI CGTGAACCAI CGTGAACCAI CGTGAACCAI 	1240 FCACCCAPATO FCACCCAPATO FCACCCAPATO FCACCCAPATO FCACCCAPATO 1340 ACGTGGCCACA	* AAGTTTTTTG AAGTTTTTTG AAGTTTTTTG AAGTTTTTTG	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360 ACCAAAGCCAA	* GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI * AGCAGCGGCGGCG	1280 ACTARATCGGAA ACTARATCGGAA ACTARATCGGAA ACTARATCGGAA ACTARATCGGAA 1380 SCCATTCAGGCT	* 1 CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG	: : AGC : : AGC : AGC : : : : : : :
rpe3 r_AG Stdna ype1-1 ype2-1 ype2-2 ype23 r_AG Stdna ype1-1	* CGTCTATCAGGO CGTCTATCAGGO CGTCTATCAGGO CGTCTATCAGGO CGTCTATCAGGO CGTCTATCAGGO CCCCCGATTTAGA	1220 SCGATGGCCCACTA(SCGATGGCCCACTA(SCGATGGCCCACTA(1320 IGCTTCACGGGCAA)	* CGTGAACCAT CGTGAACCAT CGTGAACCAT CGTGAACCAT	1240 FCACCCAAATO FCACCCAAATO FCACCCAAATO FCACCCAAATO TCACCCAAATO T340 ACGTGGCCAGA	* :AAGTTTTTTG :AAGTTTTTTG :AAGTTTTTTG :	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360 IAGAAAGCCAA	* GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI AGCAGCGGGCC	1280 CCTARATCGGAR CCTARATCGGAR CCTARATCGGAR CCTARATCGGAR 1380 SCCATTCAGGCT	* 1 CCCTAPAGGC CCCTAPAGGC CCCTAPAGGC CCCTAPAGGC CCCTAPAGGC CCCTAPAGGC CCCTAPAGGC CCCTA	: : AGC : : AGC : : : : :
Ppe3 P_AG Stdna ppe1-1 ppe1-2 ppe2-1 ppe2-2 ppe3 P_AG Stdna ppe1-1 ppe1-2	* CGTCTATCAGGO CGTCTATCAGGO CGTCTATCAGGO CGTCTATCAGGO CGTCTATCAGGO CGTCTATCAGGO CGTCTATCAGGO CGTCTATCAGGO CCCCCGATTTAGA	1220 SCGATGGCCCACTA(SCGATGGCCCACTA(SCGATGGCCCACTA(SCGATGGCCCACTA(1320 IGCTTCACCGGGCAA)	* CGTGAACCAT CGTGAACCAT CGTGAACCAT CGTGAACCAT	1240 TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC 1340 ACGTGGCGAGA	* CAAGTTTTTTC CAAGTTTTTTC CAAGTTTTTTC CAAGTTTTTTC AAGTTTTTTC AAGTAGGAGGGA	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360 AGAAAGCCAA	* GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI AGCGGCGGGCC	1280 ACTAPATCGGAA ACTAPATCGGAA ACTAPATCGGAA ACTAPATCGGAA 1380 SCCATTCAGGCT	* 1 CCCTADAGG CCCTADAGG CCCTADAGG CCCTADAGG CCCTADAGG	1300 2AGC : : 2AGC : : : -AGC :
pe3 _AG pe1-1 pe2-2 pe2-2 pe3 _AG Stdna pe1-1 pe1-2 pe2-2 pe3 _AG	* CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CCCCCGATTTAGZ CCCCCGATTTAGZ	1220 SCGATGGCCCACTAG SCGATGGCCCACTAG SCGATGGCCCACTAG 1320 IGCTTCACGGGCAAA	* CGTCAACCA CGTCAACCA CGTCAACCA CGTCAACCA AGCCCGGCCA AGCCCGGCCA	1240 TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC 1340 ACGTGGCCAPA	* AAGTTTTTTC AAGTTTTTTC AAGTTTTTTC AAGTATTTTTC * AAGGAAGGGA	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360 AGAAAGCGAA	* GCCGTABAGC2 GCCGTABAGC2 GCCGTABAGC2 GCCGTABAGC3 AGCAGCGGGCC AGCAGCGGGCC	1280 ACTARATCGCAR ACTARATCGCAR ACTARATCGCAR 1380 SCCATTCAGGCT	* 1 CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG GCGCAPCTG GCGCAPCTG	1300 IAGC : IAGC
pe3 _AG Stdna pe1-1 pe2-2 pe3 _AG Stdna pe1-1 pe1-2 pe2-2 pe3 _AG Stdna pe1-1 pe2-2 pe2-2 pe2-2 pe2-2 pe2-2 pe2-2	* CGTCTATCAGGC GGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CCCCCGATTTAGZ CCCCCGATTTAGZ	1220 SCGATGGCCCACTA(SCGATGGCCCACTA) SCGATGGCCCACTA(1320 NGCTTCACGGGCAA) IGCTTCACGGGCAA)	* CGTGARCCA' CGTGARCCA' CGTGARCCA' CGTGARCCA' * AGCCGGCGGAI AGCCGGCGAI AGCCGGCGAI	1240 FCACCCAAATC FCACCCAAATC CACCCAAATC CACCCAAATC 1340 ACGTGGCGAGA	* AAGTTTTTTG AAGTTTTTTG AAGTTTTTTG AAGTATTTTTG AAGGAAGG	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360 IACAAAGCCAA IACAAAGCCAA	* GCCGTAAAGCA GCCGTAAAGCA GCCGTAAAGCA GCCGTAAAGCA GCCGTAAAGCA GCGGGGGGCG * AGGAGCGGGGCG AGGAGCGGGGCG	1280 ACTARATCGGAA ACTARATCGGAA ACTARATCGGAA ACTARATCGGAA ACTARATCGGAA ACTARATCGGAA SCCATTCAGGCT SCCATTCAGGCT	* 1 CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG GCGCAPCTG GCGCAPCTG	1300 AGC : AGC : AGC : AGC : AGC : AGC : TGG : TGG : TGG : TGG :
-pe3 -AG 	* CGTCTATCAGGG CGTCTATCAGGG CGTCTATCAGGG CGTCTATCAGGG CCCCCGATTTAGZ CCCCCGATTTAGZ	1220 SCGATGGCCCACTA SCGATGGCCCACTA SCGATGGCCCACTA 1320 AGCTTCACGGGCAAJ GCTTCACGGGCAAJ	* CGTCAACCA CGTCAACCA CGTCAACCA CGTCAACCA CGTCAACCA AGCCCGGCCA AGCCCGGCCA AGCCCGGCCA	1240 TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC 1340 ACGTGGCCACA ACGTGGCCACA	* CAAGTTTTTTC CAAGTTTTTTC CAAGTTTTTTC CAAGTTTTTTC AAGTATTTTTC AAGTAAGGGGG AAGGAAGGGG CAAGGGAAGGG	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360 ACAAAGCCAA	* GCCGTABAGC1 GCCGTABAGC2 GCCGTABAGC2 GCCGTABAGC3 GCCGTABAGC3 GCCGTABAGC3 AGCAGCGGGCC AGCAGCGGGCC	1280 ACTARATCGGAR ACTARATCGGAR ACTARATCGGAR 1380 SCCATTCAGGCT SCCATTCAGGCT	* 1 CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG GCGCAACTG GCGCAACTG GCGCAACTG	1300 130 13
AG Stdna Apel-1 Apel-2 Ape2-2 Ape2-2 AG Stdna Apel-2 AG Stdna Ape2-2 App2-2 App2-	* CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ	1220 SCGATGGCCCACTAI SCGATGGCCCACTAI SCGATGGCCCACTAI 1320 AGCTTCACGGGGCAAI IGCTTCACGGGGCAAI	* CGTCAACCA1 CGTCAACCA1 CGTCAACCA1 CGTCAACCA1 * AGCCCGGCCA1 AGCCCGGCCA1 AGCCCGGCCA1	1240 TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC	* AAGTTTTTTC AAGTTTTTTC AAGTTTTTTC AAGCAAGGGA AAGCAAGGGA AAGCAAGGGA AAGCAAGGGA AAGCAAGGGA AAGCAAGGGA AAGCAAGGGA AAGCAAGGGA AAGTTTTTCC *	1260 GGGTCCAGGT GGGTCCAGGT 1360 IAGAAAGCGAA AGAAAGCGAA IAGAAAGCGAA	* GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI AGCAGCGGGGC AGCAGCGGGGC AGCAGCGGGGC	1280 ACTAAATCGGAA ACTAAATCGGAA ACTAAATCGGAA SCCAATCAGGCT SCCATTCAGGCT SCCATTCAGGCT 1480	* 1 CCCTAPAGG CCCCAPC CCCTAPAGG CCCAPC CCCCAPC CCCC	
pe3 _AG _tdna pe1-1 pe2-2 pe2-1 pe2-2 pe3 _AG _tdna pe1-1 pe1-2 pe2-1 pe2-2 pe3 _AG _AG	<pre></pre>	1220 SCGATGGCCCACTAI SCGATGGCCCACTAI SCGATGGCCCACTAI 1320 AGCTTCACGGGGAAI AGCTTCACGGGGAAI AGCTTCACGGGGAAI SCTTCACGGGGAAI	* CGTCAACCA CGTCAACCA CGTCAACCA * AGCCGGCGA AGCCGGCGA AGCCGGCGA * CGCTATTACC	1240 TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC 1340 ACGTGGCGAGA ACGTGGCGAGA CGTGGCGCGAGA CGTGGCGCGAGA 1440 SCCAGCTGGCCGGCG	* CAAGTTTTTTC CAAGTTTTTTC CAAGTTTTTTC CAAGTTTTTTC AAAGCAAGGGA CAAGCAAGGGA CAAGCAAGGGA CAAGCAAG	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360 1360 AGAAAGCCAA JAGAAAGCCAA 1460 CGTGCTGCAAGC	* GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI AGCAGCGGGGCC * AGCAGCGGGGCC * GCCGTTAAGT	1280 ACTAPATCGGAA ACTAPATCGGAA ACTAPATCGGAA 1380 SCCATTCAGGCT SCCATTCAGGCT 1480 IGGGTAACGCCA	* 1 CCCTAPAGGC CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGGC CCCTAPAGC CCCTAPAGC CCCTAPAGC CCCTAPAGC CCCTAPAGC CCCTAPAGC CCCTAPAGC CCCC	1300 130 13
The second secon	* CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CCCCCATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ	1220 SCGATGGCCCACTAI SCGATGGCCCACTAI SCGATGGCCCACTAI 1320 16CTTCACGGGCAAA AGCTTCACGGGCAAA AGCTTCACGGGGAAA 1420 SGTGCGGGGCCTCTTC	* CGTCAACCAT CGTCAACCAT CGTCAACCAT CGTCAACCAT * AGCCCGGCCAJ AGCCCGGCCAJ AGCCGGCCAJ	1240 TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC 1340 ACGTGGCCACA ACGTGGCCACA ACGTGGCCACA ACGTGGCCACA 1440 SCCAGCTGGCC	* CAAGTTTTTTC CAAGTTTTTTC CAAGTTTTTTC CAAGTTTTTTC AAGCAAGGGCA CAAGCAAGGGCA CAAGCAAGGGCA CAAGCAAG	1260 GGGTCCAGGT GGGTCCAGGT 1360 IAGAAAGCGAA IAGAAAGCGAA IAGAAAGCGAA GGGCTGCAAG	* GCCGTAPAGC1 GCCGTAPAGC1 GCCGTAPAGC2 GCCGTAPAGC3 * AGCAGCGGGCC AGCAGCGGGCC AGCAGCGGGCC * GCCATTAPAGT	1280 ACTAPATCGGAA ACTAPATCGGAA ACTAPATCGGAA ACTAPATCGGAA I380 SCCATTCAGGCT SCCATTCAGGCT SCCATTCAGGCT I480 IGGGTAACGCCA	* 1 CCCTAPAGG CCCTACT CCCTAPAGG CCCTACT CCCTAPAGG CCCTACT CCCTACT CCCCACTG CCCCCCCCCC	1300 AGC : AGC
The second secon	* CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CCCCCGATTTACA CCCCCGATTTACA CCCCCGATTTACA CCCCCGATTTACA CCCCCGATTTACA CCCCCGATTTACA CCCCCGATTTACA	1220 SCGATGGCCCACTAI SCGATGGCCCACTAI SCGATGGCCCACTAI 1320 AGCTTCACGGGGCAAI AGCTTCACGGGGCAAI AGCTTCACGGGGCAAI 1420 3GTGCGGGGCCTCTTC	* CGTCAACCAT CGTCAACCAT CGTCAACCAT AGCCCGGCCAT AGCCCGGCCAT AGCCCGGCCAT	1240 ICACCCAPATC ICACCCAPATC ICACCCAPATC ICACCCAPATC ICACCGAGCGACA ACGTGGCGACA ACGTGGCGACA ACGTGGCCACCA ICAC	* AAGTTTTTTC AAGTTTTTTC AAGTTTTTTC AAGCAAGGGC AAGCAAGGGC AAGCAAGGGC AAGCAAGGGC AAGCAAGGGC AAGCAAGGGC AAGCAAGGGC AAGCAAGGGC AAGTTTTTCC AAGCAAGGCAAGGCC AAGCAAGGCC AAGCAAGGCAC AAGCAAGGCAAG	1260 GGGTCCAGGT GGGTCCAGGT 1360 IACAAAGCCAA IACAAAGCCAA IACAAAGCCAA IACAAAGCCAA IACAAAGCCAA IACAAAGCCAA	* GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI AGCAGCGGGCC AGCAGCGGGCC * GCCATTAAGT	1280 ACTAAATCGGAA ACTAAATCGGAA ACTAAATCGGAA ACTAAATCGGAA 3380 SCCATTCAGGCT SCCATTCAGGCT 3CCATTCAGGCT 1480 FGGGTAACGCCA	* 1 CCCTAPAGG CCCTACTG CCCTAPAGG CCCTACTG CCCTAPAGG CCCTACTG CCCTACTG CCCTACTG CCCTACTG CCCTACTG CCCTACTG CCCTACTG CCCTACTG CCCCACTG CCCTACTG CCCCCCCCCC	
-pe3 -AG -AG -AG -AG 	* CGTCTATCAGGO CGTCTATCAGGO CGTCTATCAGGO CGTCTATCAGGO CGTCTATCAGGO CCCCCGATTTAGA CCCCCGATTTAGA CCCCCGATTTAGA CCCCGATTTAGA CCCCGATTTAGA CCCCGATTTAGA CCCCGATTTAGA CCCCGATTTAGA CCCCGATTTAGA CCCCGATTTAGA CCCCGATTTAGA	1220 SCGATGGCCCACTAI SCGATGGCCCACTAI SCGATGGCCCACTAI 1320 AGCTTCACGGGGAAI AGCTTCACGGGGAAI SGCTCGACGGGCAAI 1420 SGTGCGGGCCTCTT	* CGTCAACCA CGTCAACCA CGTCAACCA * AGCCGGCGA AGCCGGCGA AGCCGGCGA CGCTATTACC	1240 TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC 1340 ACGTGGCGAGA ACGTGGCGAGA CGTGGCGAGA 1440 SCCAGCTGGCC	* CAAGTTTTTTC CAAGTTTTTTC CAAGTTTTTTC CAAGTTTTTTC AAAGCAAGGGA CAAAGCAAGGGA CAAAGCAAGGGA CAAAGGAGGGA CAAAGGGGGGA	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360 1360 AGAAAGCCAA GGAAGCCAA 1460 CGTGCTGCAAG	* GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI * AGCAGCGGGGCC AGCAGCGGGGCC * GCCATTAAGT	1280 ACTAPATCGGAP ACTAPATCGGAP ACTAPATCGGAP 1380 SCCATTCAGGCT SCCATTCAGGCT 1480 RGGGTAACGCCA RGGGTAACGCCA	* 1 CCCTAPAGG CCCTACG CCCTACCTG CCCCAPCTG CCCAPCTG CCCCAPCTG CCCAPCTG CCCAPCTG CCCAPCTG CCCAPCTG CCCAPCTG CCCAPCTG CCCAP	
The second secon	* CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTAGZ	1220 GGATGGCCCACTAI GGATGGCCCACTAI GCATGGCCCACTAI 1320 1350 1550	* CGTCAACCAT CGTCAACCAT CGTCAACCAT CGTCAACCAT * AGCCCGGCCAA AGCCGGCCAAT AGCCGGCCAAT CGCTATTACC CGCTATTACC	1240 TCACCCAAATC TCACCCAAATC TCACCCAAATC TCACCCAAATC 1340 ACGTGGCCACA ACGTGGCCACA ACGTGGCCACA 1440 SCCAGCTGGCC SCCAGCTGGCC	* AAGTTTTTTC AAGTTTTTTC AAGTTTTTTC AAGTTTTTTC * * AAGCAAGGGC AAGCAAGGGC * AAGCAAGGGC * AAGCAAGGGC * AAGCAAGGGC * AAGCAAGGGC *	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360 IAGAAAGCCAA IAGAAAGCCAA IAGAAAGCCAA GTGCTGCAAG GTGCTGCAAG GTGCTGCAAG	* GCCGTAPAGC1 GCCGTAPAGC1 GCCGTAPAGC2 GCCGTAPAGC3 * AGCAGCGGGCC AGCAGCGGGCC * GCCATTAPGT GCCATTAPGT	1280 ACTAPATCGGAA ACTAPATCGGAA ACTAPATCGGAA ACTAPATCGGAA ACTAPATCGGAA I380 GCATTCAGGCT GCCATTCAGGCT I480 IGGGTAACGCCA IGGGTAACGCCA IGGGTAACGCCA	* 1 CCCTAPAGG CCCCAPCT CCCTAPAGG CCCCAPCT CCCCAPC CCCCAPCT CCCCAPC CCCGCACTG CCCGCACTG CCCGCACTG CCCGCACTG CCCGCACTG CCCGCACTG CCCGCACTG CCCGCACTG CCCGCACTG CCCGCACTG CCCGCACTG CCCGCACTG CCCGCACTG CCCCACCG CCCCACTG CCCCACCG CCCCACCG CCCCACTG CCCCACCG CCCCACCG CCCCACCG CCCCACCG CCCCACCG CCCCACCG CCCCACCG CCCCACCG CCCCCCCC	
Tpe3 Tpe3 T_AG Stdna ype1-1 ype2-2 ype3 T_AG Stdna ype1-1 ype2-2 ype2-1 ype2-2 ype2-3 T_AG Stdna ype1-1 ype2-1 ype2-2 ype1-1 ype1-1 ype2-1 ype2-2 ype2-2 ype3 T_AG	* CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTAGZ	1220 GGATGGCCCACTAI GGATGGCCCACTAI GCATGGCCCACTAI 1320 1320 1320 GCTTCACGGGCAAAA GCTTCACGGGCAAAA GCTTCACGGGCAAAA 1420 GTGCGGGCCTCTTC GTGCGGGCCTCTTC	* CGTCAACCAT CGTCAACCAT CGTCAACCAT CGTCAACCAT AGCCCGGCCAA AGCCCGGCCAA AGCCCGGCCAAT CGCTATTACC CGCTATTACC	1240 TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC 1340 ACGTGGCCACA ACGTGGCCACA 1440 SCCAGCTGGCC SCCAGCTGGCC	* AAGTTTTTTC AAGTTTTTTC AAGTTTTTTC AAGTATTTTTC * * AAGCAAGGGC AAGCAAGGGC AAGCAAGGGC * * AAAGCAAGGGC * * AAAGCAAGGGC * *	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360 IAGAAAGCCAA IAGAAAGCCAA GTGCTGCAAG GTGCTGCAAG GTGCTGCAAG	* GCCGTAPAGC1 GCCGTAPAGC1 GCCGTAPAGC2 GCCGTAPAGC3 * AGCAGCGGGCC AGCAGCGGGCC * GCCATTAPGT GCCATTAPGT	1280 ACTAPATCGGAA ACTAPATCGGAA ACTAPATCGGAA ACTAPATCGGAA ACTAPATCGGAA I 380 GCATTCAGGCT GCCATTCAGGCT I 480 I	* 1 CCCTAPAGG CCCCACTG CCCGCACTG CCCGCACTG CCCGCACTG CCCCACCTG CCCGCACTG CCCCACCTG CCCGCACTG CCCCACCTG CCCCCCCCCC	
ype3 r_AG Stdna ype1-1 ype2-1 ype2-2 ype3 r_AG Stdna ype1-1 ype1-2 ype2-2 ype3 r_AG Stdna Stdna ype1-1 ype2-2 ype3 r_AG Stdna r_AG Stdna r_AG Stdna r_AG Stdna r_AG	* CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCGATTTAGA CCCCCGATTAGA CCCCCGATTAGA CCCCCGATTAGA CCCCCGATTAGA CCCCCGATTAGA CCCCCGATTAGA CCCCCGATTAGA CCCCCGATTAGA CCCCCGATTAGA	1220 SCGATGGCCCACTAI SCGATGGCCCACTAI SCGATGGCCCACTAI 1320 1320 1320 1320 1320 1320 1320 1320	* CGTCAACCAI CGTCAACCAI CGTCAACCAI AGCCGGCCAI AGCCGGCCAI AGCCGGCCAI CGCTATTACC CGCTATTACC	1240 TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC 1340 ACGTGGCGACZ CCGTGGCGACZ CCGTGGCGACZ 1440 SCCAGCTGGCC	* AAGTTTTTTC AAGTTTTTTC AAGTTTTTTC AAGCAAGGGC AAGCAAGGGC AAGCAAGGGCAA AAGCAAGGGCAA AAGCAGGGGCAA AAAGGGGGCAA AAAGGGGGCAA AAAGGGGGCAA	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360 IACAAAGCCAA IACAAAGCCAA IACAAAGCCAA CGTGCTGCAAG GTGCTGCAAG GTGCTGCAAG	* GCCGTAAAGCA GCCGTAAAGCA GCCGTAAAGCA GCCGTAAAGCA * AGCAGCGGGGCC AGCAGCGGGGCC AGCAGCGGGGCC CATTAAGT GCCATTAAGT GCCATTAAGT	1280 ACTARATCGGAR ACTARATCGGAR I380 SCCATTCAGGCT SCCATTCAGGCT I480 TGGGTRACGCCA TGGGTRACGCCA	* 1 CCCTAPAGG CCCTACTG CCCCAPCTG CCC	
pe3 _AG tdna pe1-1 pe2-2 pe2-1 pe2-2 pe3 _AG tdna pe1-1 pe1-2 pe2-2 pe3 _AG tdna pe1-1 pe1-2 pe2-2 pe3 _AG	* CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CCCCGATTTAGA CCCCGATTTAGA CCCCGATTTAGA CCCCGATTTAGA CCCCGATTTAGA CCCCGATTTAGA CCCCGATTTAGA CCCCGATTTAGA CCCCGATTAGA CCCCCGATCA CCCCGATCAC CCCCGAC CCCCGAC CCCCCGAC CCCCGAC CCCCCCCCCAC CCCCCCCCCC	1220 GCGATGGCCCACTAI GCGATGGCCCACTAI GCGATGGCCCACTAI 1320 1320 AGCTTCACGGGCAAI GCCTCGACGGGCAAI 1420 GGTGCGGGCCCTCTT GGTGCGGGCCCCTCTT GGTGCGGGCCCCTCTT GGTGCGGGCCCCCTTT GGTGCGGGCCCCCTTT	* CGTCAACCA CGTCAACCA CGTCAACCA * AGCCGGCGA AGCCGGCGA AGCCGGCGA CGCTATTACC CGCTATTACC	1240 TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC 1340 ACGTGGCGAGA ACGTGGCA ACGTGGCGAGA ACGTGGCGAGA ACGTGGCGAGA ACGTGGCGAGA ACGTGGCGAGA ACGTGGCGAGA ACGTGGCGAGA ACGTGGCGAGA ACGTGGCGAGA ACGTGGCGAGA ACGTGGCGAGA ACGTGGCGAGA ACGTGGCGAGA ACGTGGCA ACGTGGCGAGA ACGTGGCGAGA ACGTGGCGAGA ACGTGGCAGA ACGTGGGCAGA ACGTGGGCAGA ACGTGGGCAGA ACGTGGCAGA ACGTGGCAGA ACGTGGCAGA ACGTGGGGGAGA ACGTGGGGGAGA ACGTGGGGAGA ACGTGGGGAGA ACGTGGGGAGA ACGTGGGGAGA ACGTGGGGAGA ACGTGGGGAGA ACGTGGGGAGA ACGTGGGGAGA ACGTGGGGAGA ACGTGGGAGA ACGTGGGGAGA ACGTGGGGAGA ACGTGGGGGAGA ACGTGGGGGAGA ACGTGGGGGAGA AC	* AAGTTTTTTC AAGTTTTTTC AAGTTTTTTC AAGCAAGGGA AAGCAAGGGA AAGCAAGGGCA AAAGGAGGGGA * AAAGGAGGGGA AAAGGGGGGA AAAGGGGGGA AAAGGGGGGA AAAGGGGGGA AAAGGGGGGA	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360 1360 AGAAAGCCAA IAGAAAGCCAA GGTGCTGCAAG GGTGCTGCAAG GGTGCTGCAAG GGTGCTGCAAG GGTGCTGCAAG	* GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI * AGCAGCGGGGCC AGCAGCGGGGCC * GCCATTAAGT GCCATTAAGT GCCATTAAGT	1280 ACTAPATCGGAA ACTAPATCGGAA ACTAPATCGGAA ACTAPATCGGAA ACTAPATCGGAA I380 SCCATTCAGGCT SCCATTCAGGCT I480 FGGGTAACGCCA FGGGTAACGCCA FGGGTAACGCCA FGGGTAACGCCA	* 1 CCCTADAGG CCCTADAGG CCCTADAGG CCCTADAGG CCCTADAGG CCCTADAGG CCCTADAGG CCCTADAGG * 1 GGGCAACTG GGGCAACTG GGGCAACTG GGGTTTTCCC GGGTTTTCCC GGGTTTTCCC	
pe3 AG AG pe1-1 pe2-2 pe3 AG pe1-2 pe3 AG pe1-2 pe2-2 pe3 AG pe1-2 pe2-2 pe3 AG stdna pe1-2 pe2-2 pe3 AG stdna pe1-2 pe2-2 pe3 AG Stdna pe1-2 pe2-2 pe3 AG Stdna pe1-2 pe2-2 pe3 AG Stdna pe1-2 pe2-2 pe3 AG Stdna pe1-2 pe2-2 pe3 AG Stdna pe1-2 pe2-2 pe3 AG Stdna pe1-2 pe2-2 pe3 AG Stdna pe1-2 pe2-2 pe3 AG Stdna pe1-2 pe2-2 pe3 AG Stdna pe1-2 pe2-2 pe3 AG Stdna pe1-2 pe2-2 pe3 AG Stdna pe1-2 pe2-2 pe3 AG Stdna Pe1-2 pe2-2 pe3 AG Stdna Pe1-2 pe2-2 pe3 AG Stdna Pe1-2 pe2-2 pe3 AG Stdna Pe1-2 pe2-2 pe3 AG Stdna AG Stdna AG Stdna AG Stdna AG Stdna AG	* CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CACGGCGATCCC CAAGGGCCATCCC CAAGGGCCATCCC CAAGGGCCATCCC CAAGGGCCATCCC CAAGGGCCATCCC CAAGGGCCATCCC CAAGGGCCATCCC CCCCGACTTGZZ	1220 GCGATGGCCCACTAI GCGATGGCCCACTAI GCGATGGCCCACTAI 1320 130	* CGTCAACCAT CGTCAACCAT CGTCAACCAT CGTCAACCAT * AGCCCGGCCAA AGCCGGCCAA AGCCGGCCAATTACC CGCTATTACC CGCTATTACC	1240 TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCAPATC TCACCAPATC	* AAGTTTTTTC AAGTTTTTTC AAGTTTTTTC AAGCAAGGGA AAGCAAGGGA AAGCAAGGGCA AAAGGAGGGCA AAAGGGGCAT AAAGGGGCAT AAAGGGGCAT AAAGGGGCAT AAAGGGGCAT AAAGGGGCAT AAAGGGGCAT AAAGGGGCAT	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360 IAGAAAGCCAA IAGAAAGCCAA IAGAAAGCCAA GTGCTGCAAG GTGCTGCAAG GTGCTGCAAG GTGCTGCAAG GTGCTGCAAG GTGCTGCAAG	* GCCGTADAGCI GCCGTADAGCI GCCGTADAGCI GCCGTADAGCI AGCAGCGGGCC AGCAGCGGGCC AGCAGCGGGCC GCCATTAAGT GCCATTAAGT GCCATTAAGT	1280 ACTAPATCGGAA ACTAPATCGGAA ACTAPATCGGAA I 380 GCATTCAGGCT I 480 FGGGTAACGCCA FGGGTAACGCCA FGGGTAACGCCA FGGGTAACGCCA I 580 FAGTTTGCGCCCA	* 1 CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG GCCAPCTG GCGCAPCTG GCGCAPCTG GCGCAPCTG GCGCAPCTG GGGTTTCCC GGGTTTCCC GGGTTTCCC GGGTTTCCC	
pe3 AG 	* CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CCCCCGATTTAGA CCCCCGATTAGA CCCCCGATTAGA CCCCCGATTTAGA CCCCCGATTAGA CCCCCGATTTAGA CCCCCGATTTAGA CCCCCGATTAGA CCCCCCGATTAGA CCCCCCGATTAGA CCCCCCGATTAGA CCCCCCGATTAGA CCCCCCGATTAGA CCCCCCGATTAGA CCCCCCGATTAGA CCCCCCGATTAGA CCCCCCCGATTAGA CCCCCCGATTAGA CCCCCCGATTAGA CCCCCCCGATTAGA CCCCCCCCCCCCGATTAGA CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	1220 SCGATGGCCCACTAI SCGATGGCCCACTAI SCGATGGCCCACTAI 1320 1GCTTCACGGGGCAAA 4GCTTCACGGGGCAAA 4GCTTCACGGGGCAAA 1420 3GTGCGGGGCCTCTTC 3GTGCGGGGCCTCTTC 3GTGCGGGGCCTCTTC 1520 IAAACGACGGCCACG	* CGTCBACCAI CGTCBACCAI CGTCBACCAI CGTCBACCAI AGCCGGCCBAI AGCCGGCCBAI AGCCGGCCBAI CGCTBATTACC CGCTBATTACC CGCTATTACC	1240 TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC 1340 ACGTGGCGACZ ACGTGGCGACZ ACGTGGCGACZ ACGTGGCGACZ 1440 SCCAGCTGGCC SCCAGCTGGCC 1540 ACGTGGTAPAC	* CARAGTATTTTC CAAGTTTTTTC CAAGTTTTTTC A AAGCAAGGGC AAAGCAAGGGCA AAAAGGGGGCAA AAAAGGGGGCAA AAAAGGGGGCAA AAAAGGGGGCAA AAAAGGGGGCAA AAAAGGGGGCAA AAAAGGGGGCAA AAAAAGGGGGCAA AAAAAAAA	1260 GGGTCCAGGT GGGTCCAGGT 1360 1360 IACAAAGCCAAGC IACAAGCCAAGC IACAAGC IACAAGCCAAGC IACAAGC IACAAGCCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGC IACAAGCAAGCAAGCAAGC IACAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA	* GCCGTAPAGCA GCCGTAPAGCA GCCGTAPAGCA GCCGTAPAGCA 	1280 ACTAAATCGGAA ACTAAATCGGAA ACTAAATCGGAA ACTAAATCGGAA ACTAAATCGGAA I380 SCCATTCAGGCT SCCATTCAGGCT I480 IGGGTAACGCCA IGGGTAACGCCA I580 IAGTTGCGCGCG	* 1 CCCTAPAGG CCCTACTG CCCCAPCTG CCCG CACTG CCCCAPCTG CCCG CCCTACTG CCCCAPCTG CCCG CCCTACTG CCCCAPCTG CCCC	
<pre>pe3</pre>	<pre>* CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CCCCGATTTAGA CCCCCGATTTAGA CCCCGATTTAGA CCCCGATTGA CCCCGATTGA CCCCGATTGA CCCCGATTGA CCCCGATTGA CCCCGATTGA CCCCGATTGA CCCCGATTGA CCCCGATTAGA CCCCGATTGA CCCCGATTGA CCCCGATTGA CCCCGATTGA CCCCGATTAGA CCCCGATTAGA CCCCGATTGA CCCCGATTAGA CCCCGATTAGA CCCCGATTAGA CCCCGATTAGA CCCCGATTAGA CCCCGATTGA CCCCGATTGA CCCCGATTAGA CCCCGATTAGA CCCCGATTGA CCCCGATTGA CCCCGATTGA CCCCGATTGA CCCCCGATTGA CCCCCCGATTGA CCCCCCGATTGA CCCCCCGATTGA CCCCCCGATTGA CCCCCCGATTGA CCCCCCGATTGA CCCCCCCGATTGA CCCCCCCGATTGA CCCCCCCGATTGA CCCCCCGATTGA CCCCCCGATTGA CCCCCCGATTGA CCCCCCGATTGA CCCCCCCGATTGA CCCCCCCGATTGA CCCCCCGATTGA CCCCCCGATTGA CCCCCCGATTGA CCCCCCGATTGA CCCCCCGATTGA CCCCCCCGATTGA CCCCCCGATTGA CCCCCCGATTGA CCCCCCCGATTGA CCCCCCGATTGA CCCCCCGATTGA CCCCCGATTGA CCCCCCGATTGA</pre>	1220 SCGATGGCCCACTAI SCGATGGCCCACTAI SCGATGGCCCACTAI 1320 AGCTTCACGGGGCAAI AGCTTCACGGGGCAAI SGTGCGGGGCCTCTT SGTGCGGGGCCTCTTT SGTGCGGGGCCCTCTTT 1520 IADACCACGGCCACG	* CGTCAACCAT CGTCAACCAT CGTCAACCAT * AGCCGGCGAT AGCCGGCGAT AGCCGGCGAT CGCTATTACC CGCTATTACC CGCTATTACC	1240 TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCAPATC	* CAAGTTTTTTC CAAGTTTTTTTC CAAGTTTTTTTC CAAGTTTTTTTC CAAGTTTTTTTC CAAGTAAGGAAGGGA CAAGGGAAGGG	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360 1360 AGAAAGCCAA IAGAAAGCCAA GGGCGCGAAG GGGCGCGCAAG 1560 ICCGCGCGCCAAG	* GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI GCCGTAAAGCI * AGCAGCGGGGCC AGCAGCGGGGCC * GCCATTAAGT GCCATTAAGT GCCATTAAGT	1280 ACTAPATCGGAN ACTAPATCGGAN ACTAPATCGGAN ACTAPATCGGAN ACTAPATCGGAN I380 SCCATTCAGGCT SCCATTCAGGCT I380 IGGGTAACGCCA I580 ISGGTAACGCCA	* 1 CCCTAPAGG CCCAPCTG CCCGACTG CCCGACTG CCCGACTG CCCGCACTG CCCGCACTG CCCGCACTG CCCGCACTG CCCGCACTG CCCCCCCCCC	
Type3 F_AG Stdna ype1-1 ype2-2 ype3 F_AG Stdna ype1-2 ype2-2 ype3 F_AG Stdna ype1-2 ype2-1 ype2-2 ype3 F_AG Stdna ype1-1 ype2-2 ype3 F_AG Stdna ype1-1 ype2-2 ype3 F_AG Stdna ype1-1 ype2-2 ype3 F_AG	* CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CCCCGATTTAGZ CACGACGTGTGTZ	1220 GGATGGCCCACTAI GGATGGCCCACTAI GGATGGCCCACTAI 1320	* CGTCAACCAT CGTCAACCAT CGTCAACCAT CGTCAACCAT CGTCAACCAT * AGCCGGCCAA AGCCGGCCAA AGCCGGCCAAT CGCTATTACC CGCTATTACC CGCTATTACC CGCTATTACC	1240 TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCCAPATC TCACCAPATC	* CAAGGGGGGAT CAAGGGGGAAGGGGA CAAGGGGGGAAGGGGA CAAGGGGGGAAGGGGGA CAAGGGGGGAAGGGGGA CAAGGGGGGAAGGGGGA CAAAGGGGGGAA CAAAGGGGGAA CAAAGGGGGAA CAAAGGGGGAA CAAAGGGGGAA CAAAGGGGGAA CAAAGGGGGAA CAAAGGGGGAA CAAAGGGGGAA CAAAGGAGGGGAA CAAAGGGGGAA CAAAGGGGGAA CAAAGGGGGAA CAAAGGAGGGAA CAAAGGGGGAA CAAAGGGGGAA CAAAGGAGGGGAA CAAAGGAGGGGAA CAAAGGGGGAA CAAAGGGGGAA CAAAGGAGGGGAA CAAAGGGGGAA CAAAGGGGGAA CAAAGGAGGGAA CAAAGGAGGGAA CAAAGGAGGGAA CAAAGGAGGGAA CAAAGGAGGGAA CAAAGGAGGGAA CAAAGGAGGGAA CAAAGGAGGGGAA CAAAGGAGGGAA CAAAGGAGGGAA CAAAGGAGGGGAA CAAAGGAGGGAA CAAAGGAGGGGAA CAAAGGAGGGGAA CAAAGGAGGGGAA CAAAGGAGGGAA CAAAGGAGGGAA CAAAGGAGGAA CAAAGGAGGAA CAAAGGAGAAGAACAAA CAAAGGAGAAGAA CAAAGGAGAA CAAAGGAGAAGAA CAAAGGAGAAGAACAAAAAAAA	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360 IAGAAAGCCAA IAGAAAGCCAA IAGAAAGCCAA GTGCTGCAAG GTGCTGCACG GTGCTGCAAG GTGCTGCACG GTGCTGCACG GTGCTGCACG GTGCTGCACG GTGCTGCCACG GTGCTGCACG GTGCTGCACG GTGCTGCACG GTGCTGCACG GTGCTGCACG GTGCTGCAAG GTGCTGCACG GTGCACG GTGCA	* GCCGTADAGCI GCCGTADAGCI GCCGTADAGCI GCCGTADAGCI * AGCAGCGGGCC AGCAGCGGGCC * GCCATTAAGT GCCATTAAGT GCCATTAAGT GCCATTAAGT GCCATTAAGT	1280 ACTAPATCGGAA ACTAPATCGGAA ACTAPATCGGAA I 380 GCATTCAGGCT GCCATTCAGGCT I 480 FGGGTAACGCCA FGGGTAACGCCA FGGGTAACGCCA I 580 I 580 I AGTTTGCGCCC	* 1 CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG CCCTAPAGG GCGCAPCTG GCGCAPCTG GCGCAPCTG GCGCAPCTG GCGCAPCTG GCGCAPCTG GCGCTTTCCC GGGTTTTCCC GGGTTTTCCC GGGTTTTCCC GGGTTTTCCC GGGTTTTCCC GGGTTTTCCC GGGTTTTCCC GGGTTTTCCC GGGTTTTCCC GGGTTTTCCC GGGTTTTCCC GGGTTTTCCC GGGTTTTCCC GGGTTTTCCC	
ype3 r_AG Stdna ype1-1 ype2-1 ype2-2 ype3 r_AG Stdna ype1-2 ype2-1 ype3 r_AG Stdna ype1-2 ype2-1 ype2-2 ype3 r_AG Stdna ype1-1 ype2-2 ype3 r_AG Stdna ype1-2 ype2-2 ype3 r_AG Stdna ype1-2 ype2-2 ype3 r_AG Stdna ype1-2 ype2-1 ype2-2 ype2-2 ype2-1 ype2-2 ype2-1 ype2-2 ype2-1 ype2-1 ype2-2 ype2-1 ype2-1	<pre>* CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCTATCAGGC CGTCATTACA CCCCCGATTTACA CCCCCGATTTACA CCCCCGATTTACA CCCCCGATTTACA CCCCCGATTTACA CCCCGATTTACA CCCCCGATTTACA CCCCGATTTACA CCCCGATTTACA CCCCGATTTACA CCCCGATTTACA CCCCGATTACA CCCCGATTTACA CCCCCGATTTACA CCCCCGATT</pre>	1220 SCGATGGCCCACTAI SCGATGGCCCACTAI SCGATGGCCCACTAI 1320 AGCTTGACGGGGCAAI AGCTTGACGGGGCAAI AGCTTGACGGGGCAAI 3GTGCGGGGCCTCTTG 3GTGCGGGGCCTCTTG 1520 IAAACGACGGCCACG IAAACCACNGCCACG	* CGTCAACCAI CGTCAACCAI CGTCAACCAI * AGCCGGCCAI AGCCGGCCAI AGCCGGCCAI CGCTATTACC CGCTATTACC CGCTATTACC CGCTATTACC CGCTATTACC CGCTATTACC CGCTATTACC CGCTATTACC	1240 TCACCCAPATC TCACCACCACCACACA TCACCACCACCACACA TCACCACCACCACACACA	* CAAGTTTTTTC CAAGTTTTTTC CAAGTTTTTTC CAAGTTTTTTC AAGCAAGGGC AAAGCAAGGGCA AAAGGAGGGGAA AAAGGGGGGAA AAAGGGGGGAA AAAGGGGGG	1260 GGGTCCAGGT GGGTCCAGGT GGGTCCAGGT 1360 IACAAAGCCAA IACAAAGCCAA IACAAAGCCAA GGGCTGCAAG GGGCTGCAAG GGGCTGCAAG I560 ICCGCCCCCCA	* GCCGTAPAGCI GCCGTAPAGCI GCCGTAPAGCI GCCGTAPAGCI * AGCAGCGGGGCC AGCAGCGGGGCC AGCAGCGGGGCC GCCATTAAGT GCCATTAAGT GCCATTAAGT GCCATTAAGT GCCATTAAGT GCCATTAAGT GCCATTACT	1280 ACTARATCGGAR ACTARATCGGAR ACTARATCGGAR I380 SCCATTCAGGCT SCCATTCAGGCT I480 IGGGTRACGCCA I580 IGGGTRACGCCA I580 IAGGTTACGCCGC IAGTTGCGCGCC	* 1 CCCTAPAGG CCCTACTG CCCCAPCTG CCC	

Figure S5 (continued)

		T-DNA												
		\rightarrow	*	1620	*	1640	*	1660	*	1680	*	1700		
AGtdna	:	CTATCGCGT	TATTAA	IGTATAATI	GCGGGGACTCT	AATCATAAA	AACCCATCTC	ATAAATAACGT	CATGCATTAC	ATGTTAATTA	TTACATGCTI	ATDOAA	:	1700
Type1-1	:												:	-
Type1-2	:												:	-
Type2-1	:	CTATCGCGI	TATTAA	IGTATAATI	GCGGGGACTCT	ААТСАТААА	AACCCATCTC	ATAAAAAAAAA	ААААААААА	A			:	938
Type2-2	:	CTATCGCGI	IATTAAA'	IGTATAATI	GCGGGGACTCT	ААТСАТААА	AACCCATCTC	ATAAATAAAAA	ААААААААА	AAA			:	1034
Туре3	:												:	-
WT AG	:												:	-

в



Supplementary Figure 5. (A) Transcript isoforms of AGAMOUS (AG) containing introns or T-DNA sequences identified by sequencing of RACE products from ag, ag^* , ag^{*ibm2} , and ag^{*edm2} . Sequences are aligned against the WT AG genomic DNA sequence from the start codons. Black shading represents the exon sequence of AG. Types 1-1 and 1-2 share similar polyadenylation sites, but have different acceptor sites. Types 2-1 and 2-2 share similar polyadenylation sites and share the same splice acceptor site as Types1-1 and 1-2, respectively. (B) AG expression level measured by qRT-PCR. ag; parental ag mutant, ag^* ; suppressed ag, WT; non-transgenic Columbia. Bars are mean +/- SEM (n=3).



Supplementary Figure 6. McrBC-qPCR of *ag*, *ag**, *ag***ibm2*, and *ag***edm2* to measure the level of DNA methylation of T-DNA regions. Error bars represent the standard errors of means from four biological replicates of each line. *ASA1* and LTR regions have been selected as unmethylated and methylated controls for the McrBC digest, respectively. The *ASA1* region was not digested and the LTR shows almost complete digestion by McrBC, as expected. Both the T-DNA splice acceptor site (approximately 100 bp flanking the acceptor site) and the 3' border region are unmethylated in *ag*, but heavily methylated in *ag**, *ag***ibm2*, and *ag***edm2*.

Figure S7



Supplementary Figure 7. BS-seq of the 5' border of the T-DNA insertion site and the flanking AG intron. (A) A graphical representation of the DNA methylation status (CG, CHG, and CHH) of representative samples with indicated genotypes. Numbers in parentheses indicate the number of independent clones sequenced for each genotype. The dotted line represents the border between the AG intron and T-DNA insertion. (B) DNA methylation context-dependent summary of samples analysed in (A).



Supplementary Figure 8. BS-seq of the 35S promoter sequence of T-DNA inserted in the *AGAMOUS* intron. (A) A graphical representation of the DNA methylation status (CG, CHG, and CHH) of representative samples with indicated genotypes. Numbers in parentheses indicate the number of independent clones sequenced for each genotype. Note that at least two 35S promoter sequences are found in the T-DNA sequence inserted into the *AG* locus; therefore, results are from mixed PCR products for these sequences. (B) DNA methylation context-dependent summary of samples analysed in (A).



Supplementary Figure 9. BS-seq analyses of the 5' and 3' borders of the T-DNA insertion site and flanking intron sequences, as well as T-DNA regions in *gl2* single mutants. (A) A graphical representation of DNA methylation status (CG, CHG, and CHH) of representative samples with the indicated genotypes/epigenotypes. *gl2* represents DNA methylation in the T-DNA regions of T4 homozygous plants before crossing. For *gl2* (F2), genomic DNA from *gl2* single mutants after crossing to *ag* plants was used, in which the absence of T-DNA in the *ag* locus was confirmed by PCR. Gray lines represent the number of cytosines covered by BS-seq reads. Cytosines that covered fewer than 4 reads were excluded from the analysis. (B) A summary of DNA methylation analysed in (A).

gl2 gl2(F2)

3'

20 0

gl2 gl2(F2)

5'

gl2 gl2(F2)

T-DNA



Supplementary Figure 10. BS-PCR of the 5' border sequence of T-DNA inserted in the *AGAMOUS*, *LEAFY* and *GLABRA2* loci. (A) A graphical representation of DNA methylation status (CG, CHG, and CHH) of representative samples with the indicated genotypes. Ten to sixteen clones were sequenced for each genotype. 5'-flanking regions analysed by BS-seq were indicated in Supplementary Fig. S1-S3. (B) DNA methylation context-dependent summary of samples analysed in (A).



Supplementary Figure S11. BS-seq of the 5' border of the T-DNA insertion site and the flanking AG intron. A graphical representation of DNA methylation status (CG, CHG, and CHH) of representative samples with indicated genotypes. Numbers in parentheses indicate the number of independent clones sequenced for each genotype. The dotted line represents the border between the AG intron and T-DNA insertion. For visualisation, DNA methylation status in the reverse strand is shown. A summary graph is shown in Fig. 5D.

Supplementary Table S1. Primers used in this study.

Target	Primers	Sequence (5' to 3')	Application
AGAMOUS	AG F1	TATGAAGTGAGGGAAATGGTAC	Genotyping
	AG R3	ACAATGGAGGATGGATGATCAC	Genotyping and McrBC qPCR of T-DNA 3'border
	LBa1	TGGTTCACGTAGTGGGCCATCG	Genotyping and McrBC qPCR of T-DNA 3'border
	KO293	TTYYYTTAAGAYTTAAATAAAAAGAGAGAATTTGT	BS-PCR (5' region of AG T-DNA: Fig. S7, S10)
	KO294	TRRTRAAAAAAAAAAAACCACCCCARTAC	BS-PCR (5' region of AG T-DNA: Fig. S7, S10)
	358 BS F1	GTATGTTGTGTGGGAATTGTGAGYGGATAA	BS-PCR (AG T-DNA: Fig. S8)
	35S BS R1	CCTTTTTARARACTCCAATCTCTATTACTT	BS-PCR (AG T-DNA: Fig. S7, S10)
	T-DNA LB Bs R2	TTGATTTGGGTGATGGTTYAYGTAGTGGG	BS-PCR (AG T-DNA: Fig. 5, S11)
	Ag int BS F2	TTCTTTCTTCACTCATTTCTRTTATTATAA	BS-PCR (AG T-DNA: Fig. 5, S11)
	KO085	ACGGCGTACCAATCG	3' RACE region 1 outer primer
	AGRTF1	ATCGGAGCTAGGAGGAGATTCC	3' RACE region 1 inner primer
	AGF1	TATGAAGTGAGGGAAATGGTAC	3' RACE region 2 outer primer
	KO104	ATGGTACAAAGTTAAAGGAGATC	3' RACE region 2 inner primer
	KO098	CGTCATCACTCAGATATTATTTC	3' RACE region 3 outer primer
	KO105	AGATATTATTTCTTTTTTTTTTTCACTTG	3' RACE region 3 inner primer
	KO099	ACAAGAATCAGCCAAATTG	3' RACE region 4 outer primer
	KO106	TGCGTCAACAAATAATCAG	3' RACE region 4 inner primer
	KO101	GTCGCCAAGACCAAAC	3' RACE region 5 outer primer
	KO108	AAACCGCTCTCCAGTTAG	3' RACE region 5 inner primer
	KO158	GCCGTGGTCGTCTCTATGAG	qRT-PCR of AG exon1-T-DNA
	Lbc1	TAAGGGATTTTGCCGATTTCGGA	qRT-PCR of AG exon1-T-DNA
	KO123	CCAAACCGCTCTCCAGTTAG	qRT-PCR of AG 3' CDS
	KO124	GGCCATTTCCTTCAGCCTAT	qRT-PCR of AG 3' CDS
	KO156	GGGGAGAGAGTAAGGAAGGACT	qRT-PCR of AG intron
	KO157	ACTCTCACTTACCATCACATGTGT	qRT-PCR of AG intron
	KO160	TCAAAATGCAGATTTAAGCGTAGA	qRT-PCR of AG intron acceptor junction
	KO161	TCCGGTGTTAGAATTGTCCGA	qRT-PCR of AG intron acceptor junction
	KO259	AGGACTAGCCCAACCTTCAC	McrBC qPCR of ASA1
	KO260	GATCCCGACGGTGGTGAATT	McrBC qPCR of ASA1
	KO255	TTTGCTCTCAAACTCTCAATTGAAGTTT	McrBC qPCR of LTR
	KO256	TAGGGTTCTTAGTTGATCTTGTATTGAGCTC	McrBC qPCR of LTR
	KO262	TTGCTGCAACTCTCTCAGGG	McrBC qPCR of T-DNA acceptor site, qRT-PCR
	KO263	AACACATTGCGGACGTTTTT	McrBC qPCR of T-DNA acceptor site, qRT-PCR
LEAFY	LFY F3	CTATAGCTATAATCATGGACAG	Genotyping
	LFY R2	TCTGTACTATCATCACTAGAGG	Genotyping
	LBa1	TGGTTCACGTAGTGGGCCATCG	Genotyping
	LFY BS F1	AAAATTTAGGTTYTAATTTATTAAATTTT	BS-PCR (5' region of T-DNA insertion)
	358 BS C3 F2	CACCCCARRCTTTACACTTTATRCTTCC	BS-PCR (5' region of T-DNA insertion)
	LFY BS R1	CAAAGAAACAACTATGTCCTTTCCCTAACTC	BS-PCR (3' region of T-DNA insertion)
	LB BS G5 F2	TTGATTTGGGTGATGGTTYAYGTAGTGGG	BS-PCR (3' region of T-DNA insertion)
	KO125	ATTGGTTCAAGCACCACCTC	qRT-PCR of LFY upstream
	KO126	CAAGAAGCTCCCAACGAAAG	qRT-PCR of LFY upstream
	KO127	GGTACGCGAAGAAATCAGGA	qRT-PCR of LFY downstream

	KO128	ATGACGACAAGCGATGTTCA	qRT-PCR of LFY downstream
GRABLA2	GL2 F1	GTCACCACCGATCAGATCAG	Genotyping
	GL2 R1	CTTGCTCAGCTGCTGTCTTTGC	Genotyping
	LBa1	TGGTTCACGTAGTGGGCCATCG	Genotyping
	GL2 BS F1	AGTTAGGGTTYAGTTGYATGTAAAGATTTT	BS-PCR (5' region of T-DNA insertion)
	LB BS C3 F1	CTCCTTTCRCTTTCTTCCCTTCCTTTCTC	BS-PCR (5' region of T-DNA insertion)
edm2-9	edm2-9 XbaI P1	GTAGTATCTGACTGTTTATATTTTTGGAATTATGTC	Genotyping (Xbal digestion after amplification)
edm2-9	edm2-9 Xbal P1 edm2-9 P2	GTAGTATCTGACTGTTTATATTTTTGGAATTATGTC CCATATAAGCACATATGATGAC	Genotyping (Xbal digestion after amplification) Genotyping (Xbal digestion after amplification)
edm2-9 Universal	edm2-9 XbaI P1 edm2-9 P2 oligodT T7 2-3	GTAGTATCTGACTGTTTATATTTTTGGAATTATGTC CCATATAAGCACATATGATGAC CAGTGAATTGTAATACGACTCACTATAGGNVTTTTT	Genotyping (Xbal digestion after amplification) Genotyping (Xbal digestion after amplification) 3'-RACE
edm2-9 Universal Outer adaptor	edm2-9 Xbal P1 edm2-9 P2 oligodT T7 2-3 T7 Primer3	GTAGTATCTGACTGTTTATATTTTTGGAATTATGTC CCATATAAGCACATATGATGAC CAGTGAATTGTAATACGACTCACTATAGGNVTTTTT CAGTGAATTGTAATACGACTC	Genotyping (Xbal digestion after amplification) Genotyping (Xbal digestion after amplification) 3'-RACE 3'-RACE
edm2-9 Universal Outer adaptor Inner adaptor	edm2-9 XbaI P1 edm2-9 P2 oligodT T7 2-3 T7 Primer3 T7	GTAGTATCTGACTGTTTATATTTTTGGAATTATGTC CCATATAAGCACATATGATGAC CAGTGAATTGTAATACGACTCACTATAGGNVTTTTT CAGTGAATTGTAATACGACTC TAATACGACTCACTATAGGG	Genotyping (Xbal digestion after amplification) Genotyping (Xbal digestion after amplification) 3'-RACE 3'-RACE 3'-RACE