

**Table 4. Repeat type and abundance in *GRP* genes**

Repeat*	Species	Sequence <sup>†</sup>	Copy no.	% Identity	Overlapping repeats <sup>‡</sup>
<b><i>GRP14</i> (7510)</b>					
I	At	GGACGTAGGAGATTTGGG G R R R F G	3.2	90	
	Aa1	GGAGGTAGGAGATTTGGG G G R R F G	3.2	94	
	Aa2	GGAGGTAGGAGATTTGGG G G R R F G	3.2	94	
II	At	GGTGGAGGTTTACCTGGAGGACTTGGAGGATTAGGA G G G L P G G L G G L G	2.4	94	15, 5.4, 70 9, 9.7, 67
	Aa1	GGGGGAGGTTTACCTGGAGGACTTGGAGGCCCTTGGGA G G G L P G G L G G L G	3.2	94	9, 10.7, 69
	Aa2	GGGGGAGGTTTACCTGGAGGACTTGGAGGCCCTTGGGA G G G L P G G L G G L G	3.1	94	9, 10.7, 69
	Op	GGAGGTCTACCTGGAGCCGCAGGT G G L P G A A G	6.0	91	12, 5.3, 64
III	At	GAAACTGCACCTGCCGCT E T A P A A	2.8	87	
	Aa1	GGAGCTGCACCAGCCGCT G A A P A A	6.2	90	6, 17.2, 58
	Aa2	GGAGCTGCACCTGCCGCT G A A P A A	4.2	84	
	Op	CCACCTGCTGCTGGAGCTGCTACACCACCTGCGCCTGGAGC TAGT P P A A G A A T P P A P G A  S	2.9	82	21, 6.0, 71
	Si	GGAGGAGCTTACCCGCAGGAGGAGCTGCACCAGCG G G A S P A G G A A P A	3.2	92	18, 6.3, 80
	Bo	GCACCTGCACCCGCG A P A P A	2.3	100	
<b><i>GRP18</i> (7520)</b>					
I	Aa1	GGTGGAGGTGCACCTGGAGGGGGCGGAAATCCATTTGCAAA CATCTCAAAATGGTTCGGACCAGGGGCCGCA G G G A P G G G G N P F A N  I S K W F G P G A A	2.2	93	
	Aa2	GGTGGAGGTGCACCAGGAGGGGGCGGAAATCCATTTGCAAA CATCTCAAAATGGTTCGGACCAGGGGCCGCA G G G A P G G G G N P F A N	2.2	90	

Repeat*	Species	Sequence <sup>†</sup>	Copy no.	% Identity	Overlapping repeats <sup>‡</sup>
		I S K W F G P G A A			
II	At	GCCGCTGGAGCTGCGCCT A A G A A P	9.2	89	First 2 copies interrupted with 63 bp
	Aa1	GCCGCTGGAGCTGCACCA A A G A A P	10.0	87	First 3 copies interrupted with the same 51-bp sequence
	Aa2	GCCGCTGGAGCTGCACCT A A G A A P	10.0	90	First 3 copies interrupted with the same 51-bp sequence
	Op	CCTGCGGCTGGAGCTGCTACACCACCTGCGGCTGGAGCTAG TGCA P A A G A A T P P A A G A S  A	3.8	88	21, 11.8, 68
	Cr	GCGGCCGGAGCTGCTAACCT A A G A A K P	2.6	82	
	Si	GCGGGTGGAGCTGAACCA A G G A E P	6.4	84	
	Bo	GCAGCACCAGCAGCTGAA A A P A A E	1.91	88	18, 4.1, 79 overlaps
<b>GRP17 (7530)</b>					
ExonI	Op	CAAAAACCATCATCGGCTCAGTCTCTGTCA Q K P S S A Q S L S	3.1	95	Before oleosin domain
	Cr	CAAAAACCATCATCAGCTCAGTCTCTGTCAATGAGAGAGGG CAGAAAGTTTCCTTTTCTGTCTCTCCA Q N P S S A Q S L S M R E G  R K F P F L S L P	2.4	93	Before oleosin domain
I	At	GGAGGTAAAAAAGTAAATTC G G K K S K F	3.1	67	
	Aa1	GGAGGTAAATCCAAAAGTCGA G G K S K S R	2.9	79	
	Aa2	GGTAAATTCGGAGGTAAAAAGGCAAGTCCGAAGGTGAAGA A G K F G G K K G K S E G E E	2.1	79	
	Cr	TCTCTGCCACAAAAACAATCATCAGCTCAG S L P Q K Q S S A Q	1.9	85	
	Bo	ATACCGGAGAGTATTAACCAAGTAATATT I P E S I K P S N I	4.0	97	15, 7.2, 66

Repeat*	Species	Sequence <sup>†</sup>	Copy no.	% Identity	Overlapping repeats <sup>‡</sup>
	Bo	AAACACGGAAGTGGAGGT K H G S G G	17.5	68	21, 2, 81 21, 4.7, 83
II	At (a)	GAAGGTATGTCTGGAAGTGAAGGAGGTATGTCCGGAGGTGG AGGAAGTAAATCCAAAAGTAAAAAAGTAAACTCAAAGCTA AATTGGGAAAGAAAAAGGTATGTCCGGAGGCATGTCAGGA AGTGAA E G M S G S E G G M S G G G  G S K S K S K K S K L K A K L G K K K G M S G G M S G  S E	4.3	93	21, 3.1, 67 21, 4.8, 72 42, 2.9, 67 21, 2, 95 21, 2.4, 90
	At (b)	GGCATGTCAGGAAGTGAAGGAGGTATGTCTCGAAGTGAAGG AAGTATATCCGGAGGTGGTATGTCTGGGGCAGCGGAAGTA AACACAAAATTGGAGGAGGTAAACACGGAGGTCTTAGAGGT AAATTCGGAAGAAAAGAGGCATGTCAGGAAGTGAAGGA G M S G S E G G M S R S E G  S I S G G G M S G G S G S K H K I G G G K H G G L R G  K F G K K R G M S G S E G	2.2	93	Overlaps with previous repeat by 60 bp 21, 23.9, 60 (spans a and b) 141, 2.1, 80
	Aa1	GGAGGTAAATCCAAAAGAGAGGAGGTAAACTCGGAGGTCT TCGAGATAAATACGGAAAGAAAGGAGGTAAAGTCCGGAGGTG GAGGAGGCATGTCAGGAAGTGAAGGAGCTATGTCTGGAAGT GAAGGAGGTATGTCC G G K S K K R G G K L G G L  R D K Y G K K G G K S G G G G G M S G S E G A M S G S  E G G M S	2.0	90	21, 9, 65 42, 4.7, 66 15, 4.2, 93 21, 13, 63 (spans II and III)
	Aa2 (a)	TCCGAAGGTGAAGAAGGTAGTCGTCTGGGGAAGAAGGTATG TCCGGAGGTGAAGGAGGTAAATCCAAAAGGGAGGGAGTAA ATCTAAAAGTCGACGAGATAAATCAAAAAGTAAAAAGGCA TGTCCGGAGGG S E G E E G M S S G E E G M  S G G E G G K S K K G G S K S K S R R D K S K S K K G  M S G G E	2	76	Partial gene, exon 1 is interrupted 21, 2.9, 74 21, 9.3, 63 (spans a and b)

Repeat*	Species	Sequence <sup>†</sup>	Copy no.	% Identity	Overlapping repeats <sup>‡</sup>
	Aa2 (b)	GGAGGTAAATCCAAAAAGAGAGGAGGTAAACTCGGAGGTCT TTCGAGATAAATACGGAAAGAAAGGAGGCAAGTCCGGAGGTG GAGGAGGCATGTCAAGAAGTGAAGGAGCTATGTCTGGAAGT GAAGGAGGTATGTCC G G K S K K R G G K L G G L  R D K Y G K K G G K S G G G G G M S R S E G A M S G S  E G G M S	2.0	88	
	Op	GGAGGTATGTCTGGAAGGAAGGAGGTATGTCTCAGGAGGTGGA GGAGGTAAATCCAAAGGTGGACATAAATCCAGAGGTCTAA GAGGTAAGTCCAAAAGTAAAAAAGGTATGTCCGGAAGCGAA  G G M S G R K E V C Q E V E  E V N P K V G H K S R G L R G K S K S K K G M S G S E	2.4	85	21, 34.8, 67 15, 3.2, 84
	Op	GGTATGTCTGGAAGCGAAGGAGGTATGTCTGGAAGCGAAGG AGGTAAACACGGAGGCGGAGGAGGTAAACACGGAGGTAAAC ATGGAGGTGGA G M S G S E G G M S G S E G  G K H G G G G G K H G G K H G G G	2.1	89	Overlaps with a by 69 bp 12, 2.5, 94 21, 7.4, 76 12, 2.5, 100 33, 2.9, 65
III	Aa1	GGTGGTATGTCCGGA G G M S G	2.4	85	
	Aa1	GGAGGTATGTCCGAGGTGGA G G M S G G G	5.8	78	15, 1.9, 92
	Aa2	GGAGGTATGTCCGGAAGTGAA G G M S G S A	13	66	15, 3.3, 97
	Aa2	GGTGGTATGTCCGGA G G M S G	2.4	85	
	Aa2	GGAGGTATGTCCGAGGTGGA G G M S G G G	5.8	78	15, 1.9, 92
	Cr	GGAGGTATGTCCGGAAGTGAA G G M S G S A	21.2	67	12, 3.1, 100 15, 4.7, 92 24, 2.3, 96 15, 6.7, 79
IV	At	GGTAAACACAAAATCGGAGGA G K H K I G G	2.7	91	
	Aa1	GGTAAACACAAAATCGGAGGA G K H K I G G	2.7	94	

Repeat*	Species	Sequence <sup>†</sup>	Copy no.	% Identity	Overlapping repeats <sup>‡</sup>
	Aa2	GGTAAACACAAAATCGGAGGA G K H K I G G	2.7	94	
	Si	GGAAAAAATCCAAAAGT G K K S K S	2.2	86	
	Si	TCTATGGGTGGAGGTAAACATGGAAGTGAAGTTCA S M G G G K H G S G S S	7.2	70	18, 8.7, 66 18, 12.9, 60
	Bo	GGTGGAGGTAAACACGGAAGCGGAGGTAAACATGGAAGTGG AGGTTCTGCTATG G G G K H G S G G K H G S G  G S S M	4.9	79	
V	Bo	AGTTCAGATGGA S S D G	5.3	86	27, 2.7, 82
<b>GRP16 (7540)</b>					
I	At	GGAGGAGCGTCG G G A S	15.5	81	42, 2.1, 82 24, 11.8, 75 36, 7.1, 75
	Aa1	GAAGGAGCGTCTGGAGATAAACCGGGAGGAATGTCAGGAGC GGGAGGAGCGTCAGGAGATAAACCA E G A S G D K P G G M S G A  G G A S G D K P	2.6	92	24, 10.8, 71 18, 2, 83 18, 2, 88
	Aa2	GGAGGAGCGTCTGGAGATAAACCGGGAGGAATGTCAGGAGC GGGAGGAGCGTCGGGAGATAAACCA G G A S G D K P G G M S G A  G G A S G D K P	2.6	95	24, 9.8, 72 12, 12.5, 64
	Op	GATAAACCGGGAGGAGCGTCGGGA D K P G G A S G	3.0	78	48, 6, 76 (spans I and II)
	Cr	GGAGCATCGGGAGGAGCGTCAAGTGGAGCGTCATCAGGAAC ATCGGGTGAAGCGTCATCA G A S G G A S S G A S S G T  S G E A S S	4.8	78	12, 23.8, 64 36, 9.2, 71
	Si	GATAAACCGACAAAAATGCCAGCAGATAAACCGACTGGAGG GCCAAAA D K P T K M P A D K P T G G  P K	4.1	85	12, 3.8, 72 24, 3.5, 75
	Bo		3.1	77	36, 6.3, 76

Repeat*	Species	Sequence†	Copy no.	% Identity	Overlapping repeats‡
		GGAGGAGCAGCAGATAAACCCAAAAGATAAATCCGCTGGAGG AGCAACAGATAAGCCAGGAAGTAAACCAGCT G G A A D K P K D K S A G G  A T D K P G S K P A			24, 3.5, 70
II	At	GGGGGAGCATCGGGAGATAAACCG G G A S G D K P	5	66	
	Aa1	GGAGGAGCGTCGGGAGGAGCGCCT G G A S G G A P	4.2	81	12, 9.7, 59 36, 2.8, 74 36, 4.2, 67 (spans II and III)
	Op	GGAGCGCCGGGAGGAGCGTCAGGAGGAGCAGGAGCACCGGG G G A P G G A S G G A G A P G	4.7	78	12, 23.9, 61 42, 5, 82 (extends beyond) 24, 12.2, 62
	Cr	AAATCTGGAGGAACGAAGGGTGATAAATCCGGAGGAGCAAA GGGAGAT K S G G T K G D K S G G A K  G D	7.1	73	24, 14.9, 69 72, 4.6, 68
	Si	GGAGATAAACCGGTTGGAGGGTCAACAGATAAGCTA G D K P V G G S T D K	2.9	86	
	Bo	GATAAACCCGCTGGTGGACCAACAGATAAGCCAGAAAAGTAA ACCAGCTGGAGAGGCATCAAATAAACCAAAA D K P A G G P T D K P E S K  P A G E A S N K P K	5.4	88	36, 11.1, 70 12, 4.7, 68
III	Aa1	GGAGCGCGAGGAGATAAACCCGAGGAAACACCAGGACATAA ACCGGCG G A R G D K P E E T P G H K  P A	2.4	74	24, 4.9, 76
	Aa2	GGAGCCCGAGGAGATAAACTCGGAGGAACACCAGGACATAA ACACGGG G A R G D K L G G T P G H K  H G	2.5	74	24, 4.9, 77 36, 4.2, 69
	Bo	AAACCGACCGGAGGGACAGAAAAT K P T G G T E N	3.6	66	
<b>GRP19 (7550)</b>					
		No conserved repeats detected			

Repeat*	Species	Sequence†	Copy no.	% Identity	Overlapping repeats‡
<b>GRP20 (7560)</b>					
I	At	CCAGCAGCTAAA P A A K	2.7	100	
	Aa1	GCTGCCCCA A A P	4.3	76	
	Aa1	CCAGCAGCTAAA P A A K	4.2	89	
	Aa2	CCAGCAGCTAAA P A A K	3.2	92	
	Op	GGAGGTGGAGCTGCCCCAGCA G G G A A P A	2.8	80	
	Cr	GGAGTGCATCAGGT G A A S G	2.4	100	
	Cr	CCGGCAGCTAAAAA P A A K K	4.6	87	
	Si	GCAAAACCAGCAGCCAAACAAGCAGCCAAACCTGCA A K P A A K Q A A K P A	1.9	76	12, 4.8, 86
<b>GRP21 (7565)</b>					
I	At	GATGCACCTTCACCA D A P S P	8.3	62	
	At	GGACCTCCATCAGATGCACCTACATCAGGT G P P S D A P T S G	6.1	62	
	At	TCAGGTGGACCTCCA S G G P P	15.9	64	
	At	GGACTTATACCGAGACTTCCTGCACCAAGTGGGCCTGCATC GAGTGGACCTTCACCAGGTGGAGATGCAATGACTGGACCAA CACCGGAGGACCAACACCGAGTGGACTTACACCAGATGGA CATTACCTACTGGAGATATATCAAGTGGAAATTTCCAGG TGGATCTCCATCGAATGCACCTTCACCAGGTGGATCTATGA TTGATGGACTTATACCAAATGGGCCTACATCGGGTGGACCT CCATCAGATGCACCTTCATCAAGTGGACCTCAATAGATGT ACCTGCAACGGGCAATCAATACTGGATGGACTTATACCAA AAGGACTTATACCGAGACGACGAATATTGAGA G L I P R L P A P S G P A S  S G P S P G G D A M T G P T P G G P T P S G L T P D G  H L P T G D I S S G N F P G  G S P S N A P S P G G S M I D G L I P N G P T S G G P  P S D A P S S S G P P I D V	3.1	95	5, 7.2, 64 30, 3.5, 66 15, 3.3, 72 15, 4, 73 15, 3, 83 15, 3.3, 77 30, 2.1, 85 15, 4, 73 15, 6.2, 64 15, 3.3, 77 30, 2.4, 83 15, 4, 70 15, 8, 64

Repeat*	Species	Sequence†	Copy no.	% Identity	Overlapping repeats‡
		P A T G K S I L D G L I P K G L I P R R R I L R			
	At	CCGAAAGGACTTATA P K G L I	4.6	78	
	At	CCAAGTGGACCTTCACCTGGTGGAGATAAATCAGGTGGTAC TTCA P S G P S P G G D K S G G T  S	2.4	71	
	At	GGACTTATACCGAGA G L I P R	5.4	71	
	Aa1	CCAGGTGAAGATGCATCAAATGGAGCTTCCCAAGTGAAGA TACAACAAGTGGAGCTACA P G E D A S N G A F P S E D  T T S G A T	2.3	75	15, 3.6, 70 30, 3.3, 70
	Aa1	GCATCGGGTGGACCTCCATCAGATGCATCT A S G G P P S D A S	10.4	68	15, 2.3, 85 15, 4.3, 65 45, 7, 62
	Aa1	GGTGGACCCGCACCAGATGCACCTGCACCT G G P A P D A P A P	2.2	73	15, 32.5, 57 (spans this and previous repeats)
	Aa1	GCACCAGGTGGACCTCCATCGGATACACCT A P G G P P S D T P	2.5	86	15, 3.5, 67 15, 5.3, 67
	Aa1	CTTATACCGAGACGA L I P R R	4.9	90	
	Aa1	CCTACACCGAGTGGG P T P S G	2.9	80	
	Aa1	CCTACACCAGATGGACCTCCATCGGATGCA P T P D G P P S D A	2.5	96	15, 16.8, 72
	Aa1	GGACTTATACCGAGAAGACTTATACCGAAA G L I P R R L I P K	2.5	84	15, 4.9, 81
	Aa1	GCACCAGATGGACCTTACCGGGTGGAGAT A P D G P L P G G D	3.4	63	
	Aa1	CCTACACCAGGTGGACCTCCATCAGATGCA P T P G G P P S D A	2.5	80	15, 6.4, 68
	Aa1	GGACTTATACCGAGA G L I P R	5	86	
	Aa1	AGTGGACCCGCACCGGATGGACCTTACCTGGTGGAGATGC AATG S G P A P D G P S P G G D A	1.9	81	15, 5.2, 68



Repeat*	Species	Sequence <sup>†</sup>	Copy no.	% Identity	Overlapping repeats <sup>‡</sup>
		M			
	Aa1	CTTATACCGAGACGA L I P R R	5	81	
	Aa1	GCACGAGTGGACCTTCACCTGGTGGACAT A Q S G P S P G G H	2.8	72	15, 6.6, 74
	Aa2	CCAGGTGAAGATGCATCAAATGGAGCTTTACCAAGTGAAGA TACAACGAGTGCAGCTACA P G E D A S N G A L P S E D  T T S A A T	2.3	80	30, 4.9, 69 15, 10.4, 61
	Aa2	CCACCGGATGCATCTGCACCGGATAGATCT P P D A S A P D R S	4.5	76	15, 9, 75 45, 2.9, 73
	Aa2	CCTGCACCGGGTGGACCTCCATCAGATGCA P A P G G P P S D A	4.5	83	15, 47.3, 56 (spans this and the next repeat)
	Aa2	CCAGGTGGACCTATA P G G P I	32.7	61	15, 6, 78 15, 6.3, 76
	Aa2	GGTGGACTTACATTAGGTGGACCTAAAACA G G L T L G G P K T	3.8	73	15, 6.3, 72
	Aa2	GGACTTATACCAAGA G L I P R	4	84	
	Aa2	CCTGCACCGGGTGGGA P A P G G	17	62	
	Aa2	GGACTTATACCGAGA G L I P R	8	92	
	Aa2	GCACCGAGTGGACCC A P S G P	2.7	100	
	Aa2	AGTGGACCCGCACCGAGTGGGACTTCACTGGATGAAGATGC AATG S G P A P S G T S L D E D A  M	1.9	75	
	Aa2	GGACTTATACCGAGA G L I P R	4	91	
	Op	GGAGATGCATCAGGGGGAGGTGCACCAGTCGGAGATGCATC AGGGGGAGGTGCACCGTTGGAGATGCATCAGGGGGAGGTTT GCCGTTT G D A S G G G A P V G D A S G G G A P L E M H Q G E V C R	21.3	71	30, 58.1, 77 15, 11.6, 63 15, 14.7, 64 15, 13.3, 71 60, 32.4, 72 30, 7.3, 85
	Op	GGAGATGCATCGGGTGGAGCTGCACCGATT G D A S G G A A P I	3.4	70	15, 3.1, 67

Repeat*	Species	Sequence†	Copy no.	% Identity	Overlapping repeats‡
	Cr	CCGATTGGAGACACACCGGTTGGAGACACATCGGGTGGAGC CGCACCA P I G D T P V G D T S G G A  A P	2	82	15, 6.8, 67 30, 2.1, 84
II	At	CCATCATCTAAACCGACATCTAAA P S S K P T S K	19.1	77	12, 37.4, 68
	Aa1	CCATCATCAAAAACCGACAGCAAAA P S S K P T A K	48	65	12, 43.4, 65 48, 9.4, 73 36, 19.0, 66
	Aa2	CCGACATCAAAAACCGACAACAAAA P T S K P T T K	48.7	69	12, 38, 70 24, 24.3, 71 36, 31.5, 67
	Op	AAACCAACAAC K P T T	33.8	65	
	Op	AAACGACAGCAAAAACCGACAGCAAAAACCGACAGTAAACCG GCAGCTA K T T A K P T A K P T V K P  A A	18.1	70	24, 36.3, 71 12, 63.3, 69 12, 86.9, 60
	Op	CCGGCTAAACCGTAACCTAAA P A K P V T K	2	85	
	Cr	AACCGACATCAAAAACCAACATCTAAACCGAAAACCTA P T S K P T S K P K T K	2.9	69	12, 8.5, 70
<b>GRP22 (7600)</b>					
I	At	GCTGGT A G	20.2	77	3, 41.7, 66
	Aa1	GCAGGTGCAAGTACACCT A G A M T P	4.1	67	
	Aa2	GCTGGT A G	16.2	82	
	Cr	GCACCCGACGCAAGTACACCC A P A A S T P	3.2	75	15, 5.4, 68

As in Table 3, repeated sequences, numbered (\*) were categorized by length and sequence content and are grouped in order of occurrence in each gene; species are abbreviated as in Fig. 2. Consensus repeat sequences (†), adjusted to begin at the first position of a codon and to maintain periodicity between species; nucleotide sequence (upper row); protein sequence (lower row). Sequence length, copy number, and % identity are shown, respectively, for overlapping, internal repeats (‡).