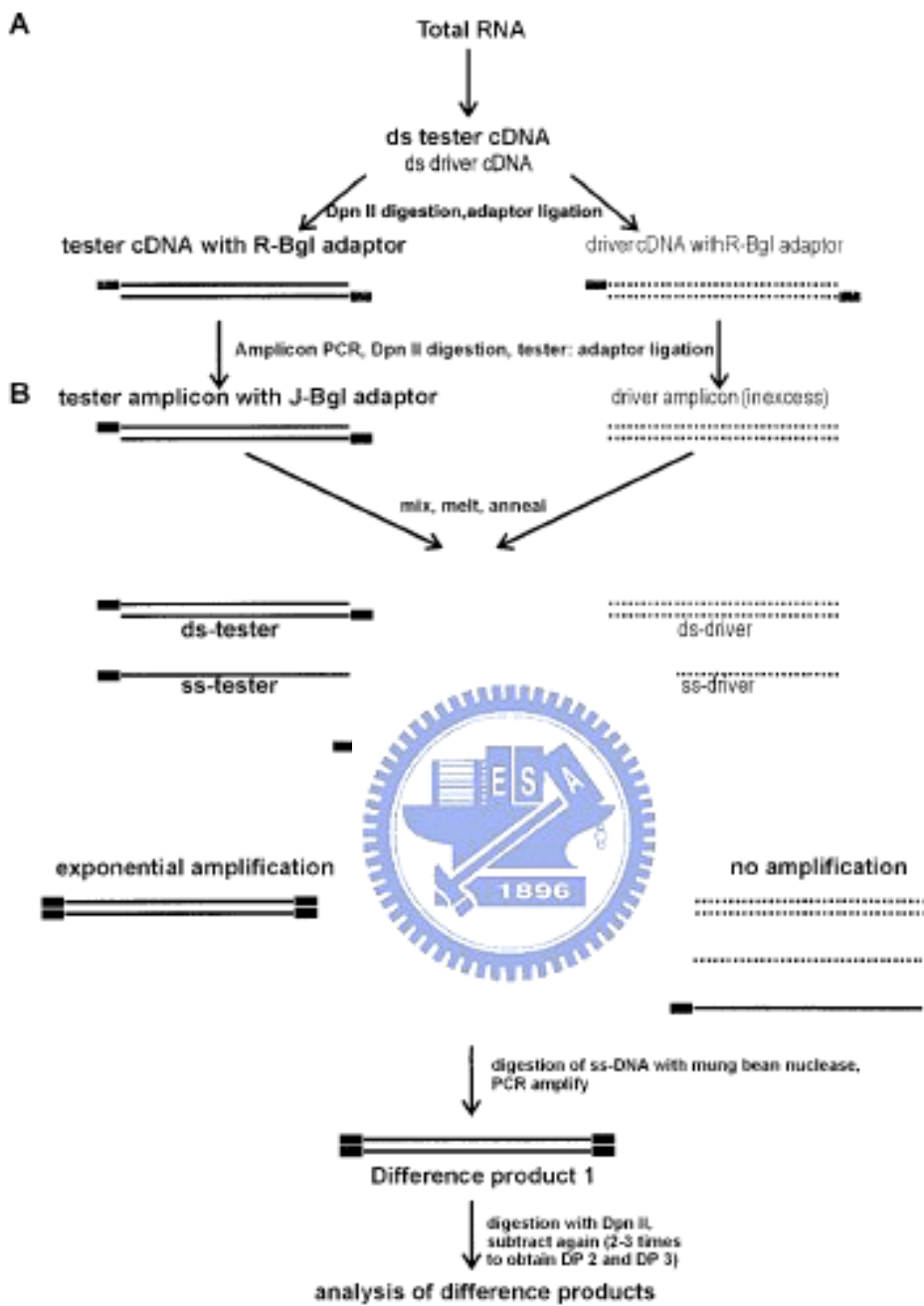


Appendix I



From: the protocol of Clontech PCR-Select™ Product

Appendix II

The sequences of target genes that are regulated by using a promoter trapping approach

krq01

1	TCGGTACCCG	GGGATCTAGG	CCTATCGACG	GATCAGATAA	TAGGTGGCGG
	AGCCATGGGC	CCCTAGATCC	GGATAGCTGC	CTAGTCTATT	ATCCACCCGCC
51	TGGGTAACAC	TTCACGGATG	GGGATCAAAA	ACGGCCGCAC	GAGGCGCTGA
	ACCCATTGTG	AAGTGCCTAC	CCCTAGTTTT	TGCCGGCGTG	CTCCGCGACT
101	AAGAGAGCCT	CAGCAAACCTG	AAGCTGGACT	ATGTGGATCT	GTCAACATTA
	TTCTCTCGGA	GTCGTTTGAC	TTCGACCTGA	TACACCTAGA	CAGTTGTAAT
151	AAATGGTTTA	ATGATTTTAA	CGCAGATTTT	GCAATAAGAT	TATTCTCGAG
	TTTACCAAAT	TACTAAAATT	GCGTCTAAAA	CGTTATTCTA	ATAAGAGCTC
201	CCAGAAAAGAT	TATTCTAAAC	GATTTCAAAA	TGTCAAAAGTC	GCTATCAAGA
	GGTCTTTCTA	ATAAGATTTG	CTAAAAGTTTT	ACAGTTTCAG	CGATAGTTCT
251	GTATTGCAAT	AATTGGCGCT	GGAATTAATA	ATTTACTGGA	TAGTTGTGCA
	CATAACGTTA	TAAACCGCGA	CCTTAATTTT	TAAATGACCT	ATCAACACGT
301	ATAAAAATTG	GTGGGGATGA	AAGAAATAGT	TCGCTCTTCA	CGATTGATTC
	TATTTTTAAC	CACCCCTACT	TTCTTTATCA	AGCGAGAAGT	GCTAACTAAG
351	AGTATCGATT	CAGGGGTGGC	GTACTACACT	AGCTTTTGAT	AATAATTCAT
	TCATAGCTAA	GTCCCCACCG	CATGATGTGA	TCGAAAATA	TTATTAAGTA
401	GGCGCATAAA	ATTTTGTGAT	TGCCATTTTT	TATGGGGAAA	TATTATTGCT
	CCGCGTATTT	TAAAACACTA	ACGGTAAAAA	ATACCCCTTT	ATAATAACGA
451	CCTCCCGGCA	ATAAGAATTC	TGGAGAGTGC	ATGGTATTTG	ATAATTGTAT
	GGAGGGCCGT	TATTC		ATAAAC	TATTAACATA
501	GTTTGCTGAT	AATCGC		TATGGT	GATTGGTTCT
	CAAACGACTA	TTAGC		ATACCA	CTAACCAAGA
551	TTTCAAAATG	TTCAT		ATTGTT	TGGTGATGCA
	AAAGTTTTAC	AAGTA		TAACAA	ACCACTACGT
601	AATGTCTTTA	TAAAC		CAGGGC	GAAAAACGAC
	TTACAGAAAT	ATTTG		GTCCCG	CTTTTGTCTG
651	AGATTTTACA	ATTGT		AGCTTT	GCTTCAGTTA
	TCTAAAATGT	TAACA		TCGAAA	CGAAGTCAAT
701	TTGACAGTTT	TATTT		AATAAT	AAATACGCCA
	AACTGTCAAA	ATAAA		TTATTA	TTTATGCGGT
751	TTGTTTTATG	TGATAAGTGA	CAATGAAAAAT	GGTTTATATG	TGCGTAATTT
	AACAAAATAC	ACTATTCACT	GTTACTTTTA	CCAAATATAC	ACGCATTAAA
801	ACGATTCCAG	GCAACTGAAA	ACTACAATCC	ATCAAAAGGC	ACGGAAAATG
	TGCTAAGGTC	CGTTGACTTT	TGATGTTAGG	TAGTTTCCG	TGCCTTTTAC
851	CTCTAGTACT	AGTAGGCGGT	GATGGGAAAGT	CTTATCTCGA	AAACGTTAGA
	GAGATCATGA	TCATCCGCCA	CTACCCTTCA	GAATAGAGCT	TTTGCAATCT
901	GTCTCATTAA	ATAATAAGTC	ATATTTAGCC	TTAAACAAAA	ATGATTCCTC
	CAGAGTAATT	TATTATTCAG	TATAAATCGG	AATTTGTTTT	TACTAAGAAG
951	TGTGTTGATG	AACTCA			
	ACACAACCTAC	TTGAGT			



krq02

1 GAATTCGAGC TCGGTACCGG GGATCTAGGC CTATCGACGG ATCGTCATTG
CTTAAGCTCG AGCCATGGCC CCTAGATCCG GATAGCTGCC TAGCAGTAAC
51 CCGCCGATAA CGACAGATTA GACGACAAGC CCAACACCGG CACAGAGCGT
GGCGGCTATT GCTGTCTAAT CTGCTGTTTCG GGTGTGGGCC GTGTCTCGCA
101 GCCGAGAAAG CGGCCTTATC CGTGGATGGT TATGTGTCCG TGCCGCCGAC
CGGCTCTTTC GCCGGAATAG GCACCTACCA ATACACAGGC ACGGCGGCTG
151 AGACTATAAG GCCGACTGGA ACGACTACCA CCAGCAGCAC GGGCTGGCAG
TCTGATATTC CGGCTGACCT TGCTGATGGT GGTTCGTCGTG CCCGACCGTC
201 CCGCTACAGC AGCCTTTAAC CATTTCGATGT ATCAACCGCA GGGGGGCAGC
GGCGATGTCG TCGGAAATTG GTAAGCTACA TAGTTGGCGT CCCCCGTCTG
251 GTGAAACCGC AGTTACAGGC CATTGAGGGC GGAAAGTCCG GCCAGCCAGA
CACTTTGGGC TCAATGTCCG GTAACCTCCG CCTTTCAGGC CGGTCCGTCT
301 GAAAGCCCGC CTTAAGCCGC ATGTTGAGAG CCGCGCAGAC GGCCTTTTCT
TCTTCTGGGC GAATTCGGCG TACAACCTCTC GCGCGCTCTG CCGCAAAAAG
351 GGGTAACACC GAAAGTGGAC AAGGACAGCG GGGAGGTTAT CAACCAGGAA
CCCATTGTGG CTTTTCACCTG TTCTGTGTCG CCTTCCAATA GTTGGTCTCT
401 GCGTGGCTGT GTTCGCCGCT GGAAGTGGTG GGCACCGGCC GGGATGATAA
CGCACCGACA CAAGCGGCGA CCTTACCAC CCCTGGCCCG CCTACTATT
451 AGACCAGTAC CTGATTATCC GCTGGCAGGC ATTCGGTGTG AGCGCGCTGA
TCTGGTCATG GACTAATAGG CGACCGTCCG TAAGCCACAG TCGCGGACT
501 CGACTGCCGC AATCCCCCTG GCTGATATTG GCGAGCGTGA AGGCTGGCGC
GCTGACGGCG TTAGGGGGAC CGACTATAAC CGCTCGCACT TCCGACCGCG
551 ACCCTGAAGG CGGGCGGGAT TAACGTCACC ACCAAAAGCA GCCTGCGCGC
TGGGACTTCC GCCCGCCCTA ATTGCAGTGG TGGTTTTCGT CGGACGCGCG
601 GATCCGGAGG TGCATC .CCGCCC GCGCCGGTGA
CTAGGCCTCC ACGTAC 'GGCGGG CGCGGCCACT
651 ACAAGGGCTG GCCATC .GAGGCG CAGCGCGCCA
TGTTCCCGAC CGGTAC 'CTCCGC GTCGCGCGGT
701 CTATTCTGGC CGACAT .CTGGCT GCCGGCGCCG
GATAAGACCG GCTGTG 'GACCGA CGGCCGCGGC
751 ACGGGCAACA CCATCC .TGGCGA CGCTGTGTAT
TGCCCGTTGT GGTAGC 'ACCGCT GCGACACATA
801 CGACGGTGGG AAAAAA .GATGTG CTGGGAGCCC
GCTGCCACCC TTTTTT 'CTACAC GACCCTCGGG
851 TGACCGGGGA TATGGC 'CGGCAA AATTACCGTC
ACTGGCCCT ATACCCGAAA CTACCACGGC TATAGCCGTT TTAATGGCAG
901 CACCCTGCC ACCTCTATGT GGCATCCCC GGAATTCCT GGCAGTTTAT
GTGGGACGGG TGCAGATACA CCGTAGGGG CCTTAAGGA CCGTCAAATA
951 GGCGGGCGTC CTGCCCCCA CCTCCGGGC CGTTGCTTCG CAACGTTCAA
CCGCCCCGAG GACGGGCGGT GGGAGCCCC GCAACGAAGC GTTGCAAGTT
1001 ATCCGCTN
TAGCGCAN



krq89

1 TGCCAGGAAT TCCCGGGGAT CGGCGAGGAA GCTTCTGGCC CAGCCGGGAA
ACGGTCCTTA AGGGCCCCTA GCCGCTCCTT CGAAGACCGG GTCGGCCCCTT
51 GATGGGCGAA CAGGCTGGCG ACCACCGTCG GAAACTGGGT GTCGTTATGC
CTACCCGCTT GTCCGACCGC TGGTGGCAGC CTTTGACCCA CAGCAATACG
101 TGCAGTTTGC TGTACACCAG GTTGAGCTCG TAGGCCAGCG AGGAGAGAAT
ACGTCAAACG ACATGTGGTC CAACTCGAGC ATCCGGTCGC TCCTCTCTTA
151 GATAGCCAGC GCGGTAAAAA CGATAAGACA GATAATGACG ATGGTCATCA
CTATCGGTTCG CCGCATTTTT GCTATTCTGT CTATTACTGC TACCAGTAGT
201 GCGCCGCGAG GCCATTGCGC TCGCCGAGGC GGGTACGGAT CGCCTGGCGC
CGCGGCGCTC CGGTAACGCG AGCGGCTCCG CCCATGCCTA GCGGACCGCG
251 AGTCGTCCGC GAGGGTCGGT GCCGGCAATG GTCTCGACGG TGATCGCGTT
TCAGCAGGCG CTCCCAGCCA CGGCCGTTAC CAGAGCTGCC ACTAGCGCAA
301 AACTCGGAT GACTGTCTGC GTCGCGACGT GATTAAAGCG TTGATCGCCC
TTGAGCGCTA CTGACAGACG CAGCGCTGCA CTAATTTTCG AACTAGCGGG
351 ACGGTCGCGA CAGCGACACG CCGGTGGCGG TAATTTCCCG TGGCACCCGT
TGCCAGCGCT GTCGCTGTGC GGCCACCGCC ATTAAAGGGC ACCGTGGGCA
401 GACGATCCGC TCTGCCGATT GCCAGTGTGC GAAGGTCGTT AATCGCCCCG
CTGCTAGGCG AGACGGCTAA CGGTCACACG CTTCCAGCAA TTAGCGGGCC
451 CGCAGCCCTT GCTCCAGATC ACGGGTGATC CGTCGATAGG CCTAGATCTG
GCGTCGGGGA CGAGGTCTAG TGCCCACTAG GCAGCTATCC GGATCTAGAC
501 ACTTACCGGG GTTTAATATG CAAATTAGCG ATACCGGCC GCAGCACACT
TGAATGGCCC CAAATTATAC GTTTAATCGC TATGGCCGGG CGTCGTGTGA
551 CCTGACTTTC ACGCCGTCTT CGCCCGTGAA GAACTGGNNT AACCCGAGCA
GGACTGAAAG TCGGCAGGA GCGGGCACTT CTTGACCNNA TTGGGCTCGT
601 TTAACCANCT TTACCC TTNCGC CAGCTGGCGN
AATTGGTNGA AATGGC AANGCG GTCGACCGCN
651 GANNACTTGC CNCCN CNCGGT NANTGGACGG
CTNNTGAACG GNNGN GNCGCA NTNACCTGCC
701 CGANGGCAGT TTTTTT
GCTNCCGTCA AAAAAA



krq32

1 ACCGGGGATC TAGGCCTATC GACGGATCTT AACGTCGGTG TACCCCTTCGC
TGGCCCTAG ATCCGGATAG CTGCCTAGAA TTGCAGCCAC ATGGGAAGCG
51 TTTTGGCGAT GGACAGCTGT ACGAGCTCCT CAGTGGCATC TGTGACGTTT
AAAACCGCTA CCTGTGACACA TGCTCGAGGA GTCACCGTAG AACTGCAAAA
101 TTACGGTTTTG CCTTACTCCT GGAGGTTGAT TTCAGGGTTG GAACGAAGAG
AATGCCAAAC GGAATGAGGA CCTCCAATA AAGTCCAAC CTTGCTTCTC
151 ACCGAGCCTC ATGAGCGCTA CCGGCTGCAC AGTACTGGTA CTGTTGACAT
TGGCTCGGAG TACTCGCGAT GGCCGACGTG TCATGACCAT GACAACTGTA
201 CCAGCTTAAC AACTTCGCCG GTTGTCTTGT CCACTTCCTG AAGATTGAGA
GGTCGAATTG TTGAAGCGGC CAACAGAACA GGTGAAGGAC TTCTAACTCT
251 GGTAAGCTGT TATTTTCGCT CGTCATCTTG ATCCCCGGA ATTCCTGGCA
CCATTCGACA ATAAAAGCGA GCAGTAGAAC TAGGGGCCCT TAAGGACCGT
301 GTTTATGGCG GCGTCCTGC CCGCACCTT CCGGGCCGTT GCTTCGCAAC
CAAATACCGC CCGCAGGACG GCGCGTGGA GGCCCGCAA CGAAGCTTG
351 GTTCAAATCC GCTCCCAGCG GATTTGTCTT ACTCAGGAGA GCGTTCACCG
CAAGTTTAGG CGAGGGCCGC CTAAACAGGA TGAGTCTCTT CGCAAGTGGC
401 ACAACAACA GATAAACGA AAGGCCAGT CTTTCGACTG AGCCTTTCGT
TGTTTGTGT CTATTTTGT TTCCGGGTCA GAAAGCTGAC TCGGAAAGCA
451 TTTATTTGAT GCCTGGAAA TCGATAAGCT TGGCGTAATC ATGGTCATAG
AAATAACTA CGGACCCTTT AGCTATTCGA ACCGCATTAG TACCAGTATC
501 CTGTTTCCTG TGTGAAATTG TTATCCGCTC ACAATTCAC ACAACATACG
GACAAAGGAC AACTTTAAC AATAGGCGAG TGTTAAGGTG TGTGTATGC
551 AGCCGGAAGC ATAAAGTGTA AAGCCTGGGG TGCCTAATGA GTGAGCTAAC
TCGGCCTTCG TATTTACACAT TTCGGACCCC ACGGATTACT CACTCGATTG
601 TCACATTAAT TGCGT' 'CCAGTC GGGAAACCTG
AGTGTAATTA ACGCA' .GGTCAG CCCTTTGGAC
651 TCGTGCCAGC TGCAT' 'CGGGGA GAGGCGGTTT
AGCACGGTCG ACGTA' 'GCCCT CTCCGCCAAA
701 GCGTATTGGG CGCTC' 'GACTCG CTGCGCTCGG
CGCATAACCC GCGAG' .CTGAGC GACGCGAGCC
751 TCGTTCGGCT GCGGC' 'AAAGGC GGTAAATACGG
AGCAAGCCGA CGCCG' 'TTTCCG CCATTATGCC
801 TTATCCACAG AATCA' 'ACATGT GAGCAAAAGG
AATAGGTGTC TTAGT' 'TGTACA CTCGTTTTCC
851 CCAGCAAAAG GCCAG' 'TTGCTG GCGTTTTTCC
GGTCGTTTTT CGGTCCTTGG CTTTTTTCCG GCGCAACGAC CGCAAAAAGG
901 ATAGGCTCCG CCCCCCTGAC GAGCATCACA AAAATCGACG CTCGAAGTCAG
TATCCGAGGC GGGGGGACTG CTCGTAGTGT TTTTAGCTGC GAGTTCAGTC
951 AGGTGGCGAA AC
TCCACCGCTT TG



Appendix III

The KvhA regulated sequences obtained by cDNA subtractive hybridization analysis

hag2-1

1	CAGGTCCGGG	AATTTGTAAT	ACGACTCACT	ATAGGGCGAG	CTCGGTACCC
	GTCCAGGCC	TTAAACATTA	TGCTGAGTGA	TATCCCGCTC	GAGCCATGGG
51	GGCGAATTC	CAAGCTTTCG	AGCGGCCGCC	CGGGCAGGTG	GCGGTGGTAT
	CCCCCTTAAG	GTTTCGAAAGC	TCGCCGGCGG	GCCCGTCCAC	CGCCACCATA
101	CGCTGGTGGC	GATGCTGGGA	AATTTTACCT	TTAACGGGAA	ACACCTGGAG
	GCGACCACCG	CTACGACCCT	TTAAAAATGGA	AATTGCCCTT	TGTGGACCTC
151	TTTCTCGATC	TGCTGGCGAT	AACCATTCCG	TCCACGCTGC	TGGGTATTTT
	AAAGAGCTAG	ACGACCCTTA	TTGGTAAGGC	AGGTGCGACG	ACCCATAAAA
201	AGCCATCGGT	ATCTTCAGCT	GGTTCCGCGG	CAAAGACCTG	GATCAGGACG
	TCGGTAGCCA	TAGAAGTCGA	CCAAGGCGCC	GTTTCTGGAC	CTAGTCCTGC
251	AAGCCTTCCA	GGCGTTTATT	GCGCTCCCGG	AGAACCGCCA	TTACGTCTAT
	TTCCGAAGGT	CCGCAAATAA	CGCGAGGGCC	TCTTGCGGGT	AATGCAGATA
301	GGCGATACGG	CGACCCTGCT	GGACAAAAAG	CTGCCGACCA	GCAACTGGAT
	CCGCTATGCC	GCTGGGACGA	CCTGTTTTTC	GACGGCTGGT	CGTTGACCTA
351	AGCAATGTGG	ATTTTTCTGG	CCTCTATCGC	CGTGGTGGCT	CTGCTTGGCG
	TCGTTACACC	TAAAAAGACC	GGAGATAGCG	GCACCACCGA	GACGAACCGC
401	CTTTTTCTGA	ACTGCGCCCG	GCCTTTGACG	GTAAACCACT	GTCGATGGTG
	GAAAAAGACT	TGACGCGGGC	CGGAAACTGC	CATTTGGTGA	CAGCTACCAC
451	CTGGTTATCC	AGATGTTTAT	GCTGCTCTCC	GGAGCGCTGA	TCATCATCAT
	GACCAATAGG	TCTACAAATA	CGACGAGAGG	CCTCGCGACT	AGTAGTAGTA
501	TACAAAAACT	AACCCCGCAT	CGATTTTCGA	AAATGAGGTC	TTTCGTCGGG
	ATGTTTTTGA	TTGGGC		CTCCAG	AAAGCAAGCC
551	GAATGATCGC	TATCGT		GTGGAT	GGCGGAGACC
	CTTACTAGCG	ATAGCT		CACCTA	CCGCCTCTGG
601	ATGTTTGGCG	CGCAC		TGCTTG	GTGAAGTGGT
	TACAAACCGC	GCGTGT		ACGAAC	CACTTCACCA
651	AAAAGAGTAT	CCGTGC		CTACTG	GTTTCGAAAT
	TTTTTCTCATA	GGCAC		GATGAC	CAAAGCTTTA
701	TTGTGAACTC	CCAGGC		TGTGCC	AGTAGCCCTG
	AACACTTGAG	GGTCCC		ACACGG	TCATCGGGAC
751	GCCATCGGGC	TCGACC		CTGCCG	CGGCCTGCTA
	CGGTAGCCGC	AGCTGGGCCG	GATATAGCAC	CGAAGACGCG	GCCGGACGAT
801	TGGCTACTAT	ATTCTGCCGA	CCTACCCAAG	CGATCTGGCG	GCGATCCAGT
	ACCGATGATA	TAAGACGGCT	GGATGGGTTT	GCTAGACCGC	CGCTAGGTCA
851	TTGATCGCTC	AGGGACCACC	CATATCGGCC	GCTTTGTGAT	TAACCACAGC
	AACTAGCGAG	TCCCTGGTGG	GTATAGCCGG	CGAAACACTA	ATTGGTGTCC
901	TTTATCCTGC	CCGGGCTCAT	CGGCGTCGGC	GTCTCTGCGT	GTTTGGCTGG
	AAATAGGACG	GGCCCAGATA	GCCGCAGCCG	CAGAGACGCA	CAAACCGACC
951	GTATCGCGCG	ATGTACCTCG	GCCGCGACAC	GCTGCCCTAT	AGATCTGGAT
	CATAGCGCGC	TACATGGAGC	CGGCGCTGTG	CGACGGGATA	TCTAGACCTA
1001	CCCCTCTAGA	GTCGACTGCA	GCATGCAGCT	GCGTATCATG	GTCATAGCTG
	GGGGAGATCT	CAGCTGACGT	CGTACGTCGA	CGCATAGTAC	CAGTATCGAC
1051	TTTCTGTGTT	TGAATTGTTA	TCCGCCTCCC	CATTCCAACC	AAACATACCG
	AAAGGACAAA	ACTTAACAAT	AGGCGGAGGG	GTAAGGTTGG	TTTGTATGGC
1101	CCGGGAAGCC	ATAAAGTGTG	TTAAG		
	GGCCCTTCGG	TATTTACACAC	AATTC		



hag3-4

1 CACCCCGGGC GGATTGTATA C GACTCACTA TAGGGCGAGC TCGGTACCCG
 GTGGGGCCCC CCTAACATAT GCTGAGTGAT ATCCCGCTCG AGCCATGGGC
 51 GGCGAATTCC AAGCTTTCGA GCGGCCCCCC GGGCAGGTAC AAATCTTTCT
 CCGCTTAAGG TTCGAAAGCT CGCCGGGGGG CCGTCCATG TTTAGAAAGA
 101 TACGCTTCCA GGTCCGCAAG ATCCTCGACG ATCTGTGTGG CAATCAGCTG
 ATGCGAAGGT CCAGCCGTTT TAGGAGCTGC TAGACACACC GTTAGTCGAC
 151 CAGCCGTTGC TGATCAAAAC ATTGCTCGAT CGCGCGGAAG GCGCGCTGCT
 GTCCGCAACG ACTAGTTTTG TAACGAGCTA GCGCGCCTTC GCGCGGACGA
 201 TATCAACGGC GAAGGGATCG ATCACGTCAG CCAGGCCGAG GAGATGGTCA
 ATAGTTGCCG CTTCCCTAGC TAGTGCAGTC GGTCCGGCTC CTCTACCAGT
 251 AGCTGGCCAC CGCCGTCGCC CACCTGATTG GTCGTTCCAA CTTTCGACGCC
 TCGACCGGTG GCGGCAGCGG GTGGACTAAC CAGCAAGGTT GAAGCTGCGG
 301 ATGAGCGGTC AGTATTACGC CCGCTTCGTG GTGAAAAATG TCGATAACTC
 TACTCGCCAG TCATAATGCG GGCGAAGCAC CACTTTTTAC AGTATTGAG
 351 AGACAGCTAC CTGCGCCAGC CGCACCGGGT GATGGAGCTG CACAACGACG
 TCTGTGATG GACGCGGTCG GCGTGGCCCA CTACCTCGAC GTGTTGCTGC
 401 GCACCTATGT CGAAGAGCAA ACCGACTACG TATTAATGAT GAAGATCGAC
 CGTGGATAACA GCTTCTCGTT TGGCTGATGC ATAATTACTA CTTCTAGCTG
 451 GAGCAGAACA TGCAGGGCGG CAATTCCCTG CTGCTGCATC TCGACGACTG
 CTCGTCTTGT ACGTCCCGCC GTTAAGGGAC GACGACGTAG AGCTGCTGAC
 501 GGAGCATCTC GACGAATTCT TCCGCGACC GCTCGCCCGC GCGCCGATGC
 CCTCGTAGAG CTGCTTAAGA AGGCGCTGGG CGAGCGGGCG GCGGGCTACG
 551 GCTGGGCCGC GCCGCCGAGC AAAAACGTCA GCAAGGATGT TTTCCACCCG
 CGACCCGGCG CGGCGGCTCG TTTTTCAGT CGTTCC TACA AAAGTGGGC
 601 GTGTTTCGACG TCGAT' .TGCCT ATATCGACCA
 CACAAGCTGC AGCTA' .ACGCGA TATAGCTGGT
 651 GTTCGTCCAG CCGAA' .TGGCTG AGCCGCCCTCT
 CAAGCAGGTC GGCTT' .ACCGAC TCGGCGGAGA
 701 CCGACGCGCA GGAGAC' .CATACC GGTGCCGGTG
 GGCTGCGCGT CCTCT' .GTATGG CCACGGCCAC
 751 GGCAAATTTT TGCTC' .TGCACG GCCCGACCCG
 CCGTTTAAAG ACGAG' .ACGTGC CGGCGCTGGC
 801 CTTTACGCCG CACCC' .GCTAGA TCTGGATCCC
 GAAATGCGGC GTGGG' .CGATCT AGACCTAGGG
 851 CTCTAGAGTC GACCT' .GTAATC ATGGTCATAG
 GAGATCTCAG CTGGACGTCC GTACGTTCTGA ACCGCATTAG TACCAGTATC
 901 CTGTTTCCTG TGTGAAATTG TTATCCGCTC ACATTCCACA CAACATACGA
 GACAAAGGAC AACTTTAAC AATAGGCGAG TGTAAGGTGT GTTGTATGCT
 951 GCCGGAAGCA TAAAGTGTA AGCCTGGGGG TGCTATGAG TGAGCTAACT
 CGGCCTTCGT ATTTACATT TCGGACCCCC ACGGATACTC ACTCGATTGA
 1001 CACATTATTG CGTTGCGCTC CCTGCCCGCT TTCCAGTCGG GAAACTTTTC
 GTGTAATAAC GCAACGCGAG GGACGGGCGA AAGGTCAGCC CTTTGAAAAG
 1051 GTGGCCAGCC TGCATTATTG AAATAGTGAA AGGG
 CACCGGTCGG ACGTAATAAC TTTATCACTT TCCC



hag4-1

1 GATGAAGGCG GATTGTATAC GACTCACTAT AGGGCGAGCT CGGTACCCGG
 CTACTTCCGC CTAACATATG CTGAGTGATA TCCCCTCGA GCCATGGGCC
 51 GCGAATTCCA AGCTTTTCGAG CGGCCGCCG GGCAGGTCGG TAAAGCGTTC
 CGCTTAAGGT TCGAAAGCTC GCCGGCGGGC CCGTCCAGCC ATTTTCGCAAG
 101 AAATAACGTC GAGCGAGACA GCGCGCAGTC GCGCGCCAGG GCGTTGACCG
 TTTATTGCAG CTCGCTCTGT CGCGCGTCAG CCGCCGGTCC CGCAACTGGC
 151 TCCAGCGTCT GCGGGGGGCC TGATGCAGCA GACGGATTGC CTGGGCGACC
 AGGTGCGAGA CCGCCCCCGG ACTACGTCGT CTGCCTAACG GACCCGCTGG
 201 CGCGGGTCGC TCAGCCCGCG CATCAGACCG GGAGTCGCGA GCGTGCCGGA
 GCGCCAGCG AGTCGGGCGC GTAGTCTGGC CCTCAGCGCT CGCACGGCCT
 251 GGAACGGATG GCCTCGATAA AAAGCAGCTC AACCATTCGC GCCAGAACGA
 CCTTGCCCTAC CGGAGCTATT TTTCGTCGAG TTGGTAAGCG CGGTCTTGCT
 301 AAGATTTGCC CCGCGGATCC GCCTGCGCTT CCTCTTTCAG CAGCCCGGCA
 TTCTAAACGG GCGCGTAGG CCGACGCGAA GGAGAAAGTC GTCGGGCCGT
 351 AAGGTGGCCA GCCGCGGTTG ATTGCGGACC ACCACAAAGG CTGGCAACAG
 TTCCACCGGT CCGCGCCAAC TAACGCCTGG TGGTGTTCCT GACCGTTGTC
 401 CGAGGCCAGC AGGCTGGCGT TCGTTGAACC GGAAGCGCAG TGGCCAATCA
 GCTCCGGTCG TCCGACCGCA AGCAACTTGG CCTTCGCGTC ACCGGTTAGT
 451 GCGCCTGCAT CTGCGCAGGC TTGTGAGGG CGCCAGCCG AAAGTGTCCC
 CGCGGACGTA GACGCGTCCG AACAGCTCCC GCGGGTCGGC TTTCACAGGG
 501 TCCGCCAGTT GCACCGGGAT GGTTCACAGC GCCGTTGAAG GGGGCGGGGT
 AGGCGGTCAA CGTGGCCCTA CCAAAGGTCG CGGCAACTTC CCCCGCCCA
 551 CAGGCTGGTT AGCGAGAAGC GGTACCTCGG CCGCGACCAC GCTAGATCTG
 GTCCGACCAA TCGCTCTTCG CCATGGAGCC GGCGCTGGTG CGATCTAGAC
 601 GATCCCTCTT AGAGT(:CTTGGC GTAATCATGG
 CTAGGGGAGA TCTCA(:GAACCG CATTAGTACC
 651 TCATAGCTGT TTCCT(:TCACAA TTCCACACAA
 AGTATCGACA AAGGA(:AGTGTT AAGGTGTGTT
 701 CATAAGAGCC GGAAG(:GGTGCC TAATGAGTGA
 GTATGCTCGG CCTTC(:CCACGG ATTACTCACT
 751 GCTAACTCAC ATTAAT :CGCTTT CCAGTCGGGA
 CGATTGAGTG TAATT? :GCGAAA GGTCAGCCCT
 801 AACCTGTCTG GCCAG(:AACGCG CGGGGAGAGG
 TTGGACAGCA CGGTC(:TTGCGC GCCCCTCTCC
 851 CGGTTTGCGT ATTGG(:CTCACT GACTCGCTGC
 GCCAAACGCA TAACCCGCGA GAAGGCGAAG GAGCGAGTGA CTGAGCGACC
 901 GCTCGGTCGT TCGCTGCGGC GAGCGGTATC AGCTCACTCA AAGCGGTAAT
 CGAGCCAGCA AGCGACGCCG CTCGCCATAG TCGAGTGAGT TTCGCCATTA
 951 ACGGTTATCC ACAGAATCAG GGGATAACGC AGAAAGAACA TGTGAGCAAA
 TGCCAATAGG TGTCTTAGTC CCCTATTGCG TCTTTCTTGT AACTCGTTT
 1001 GGCCAGCAAG GCCAGGAACC GTAAAAGGCC GCCGTTTCTT GGCGTTTTTC
 CCGGTCGTTT CCGTCCCTGG CATTTTCCGG CGGCAAAGAA CCGCAAAAAG
 1051 AATAGGCTCC GGCCCCCCT GGGACGAGAC TCACAAATTC CGACGCCCTC
 TTATCCGAGG CCGGGGGGGA CCCTGCTCTG AGTGTTTAAG GCTGCGGAGG
 1101 CAATCCTAAA ATTTT
 GTTAGGATTT TAAAG



hag5-1

1 AAAGCAAGC GTATGTATAC GACTCACTAT AGGGCGAGCT CGGTACCCGG
 TTTTCGTTTCG CATACATATG CTGAGTGATA TCCCCTCGA GCCATGGGCC
 51 GCGAATTCCA AGCTTTTCGAG CGGCCGCCC GGCAGGTGAG AGACACAACA
 CGCTTAAGGT TCGAAAGCTC GCCGGCGGGC CCGTCCACTC TCTGTGTTGT
 101 TATAGCGTCA CTGTCAGCTG TGAACACCGC TACATCTTGT GTTTTTTGGG
 ATATCGCAGT GACAGTCGAC ACTTGTGGCG ATGTAGAACA CAAAAAACCT
 151 CGCTGCGCGT CGGCGGGGCT GAAGCAGATC CGGAACTACC TTTTTTCGGA
 GCGACGCGCA GCCGCCCCGA CTTCGTCTAG GCCTTGATGG AAAAAAGCCT
 201 TCTCAGAAGT GGAGAAGAGG AATTATTGCA GCATGGACAC AAAAAACGAG
 AGAGTCTTCA CCTCTTCTCC TTAATAACGT CGTACCTGTG TTTTTTGCTC
 251 TAATTGTTGC GATGAGTGGA CTATTGCACC ATGAGCAAAT ATAAATGACT
 ATTAACAACG CTAATCACCT GATAACGTGG TACTCGTTTA TATTTACTGA
 301 TTATGCGTCC TCGGGTAAA TATATTGTTT ATGCCGTCCG GCTGAACCCC
 AATACGCAGG ACGCCAATTT ATATAACAAG TACGGCAGGC CGACTTGGGG
 351 GGGTCAGGGG ATCAGCCGGG AAGGAGAACG TCACGGTTCT GTTTTCCGCG
 CCCAGTCCCC TAGTCGGCCC TTCTCTTTC AGTGCCAAGA CAAAAGGCGC
 401 ACAGGGTCTT TCTTTTTTTC TTCAGGTACC TCGGCCGCGA CCACGCTAGA
 TGTCCAGAA AGAAAAAAG AAGTCCATGG AGCCGGCGCT GGTGCGATCT
 451 TCTGGATCCC CTCTAGAGTC GACCTGCAGG CATGCAAGCT TGGCGTAATC
 AGACCTAGGG GAGATCTCAG CTGGACGTCC GTACGTTCTG ACCGCATTAG
 501 ATGGTCATAG CTGTTTCTG TGTGAAATTG TTATCCGCTC ACAATTCAC
 TACCAGTATC GACAAAGGAC ACACTTTAAC AATAGGCGAG TGTTAAGGTG
 551 ACAACATACG AGCCGGAAGC ATAAAGTGTA AAGCCTGGGG TGCCTAATGA
 TGTTGTATGC TCGGCCTTCG TATTTACAT TTCCGGACCC ACGGATTACT
 601 GTGAGCTAAC TCACA? ?TGCCCG CTTTCCAGTC
 CACTCGATTG AGTGT? ?ACGGGC GAAAGGTCAG
 651 GGGAAACCTG TCGTGC? ?GGCCAA CGCGCGGGGA
 CCCTTTGGAC AGCAC? ?CCGGTT GCGCGCCCT
 701 GAGGCGGTTT GCGTA? ?CTCGCT CACTGACTCG
 CTCCGCCAAA CGCAT? ?GAGCGA GTGACTGAGC
 751 CTGCGCTCGG TCGTT? ?CAGCTC ACTCAAAGGC
 GACGCGAGCC AGCAA? ?GTCGAG TGAGTTTCCG
 801 GGTAATACGG TTATC? ?GCAGGA AAGAACATGT
 CCATTATGCC AATAG? ?CGTCCT TTCTTGTACA
 851 GAGCAAAAGG CCAGC? ?AAAGGC CGCGTTGCTG
 CTCGTTTTTC GGTCTTTTTT CCGTCCTTGG CATTTTTTCCG GCGCAACGAC
 901 GCGTTTTTTC ATAGGCTCCG CCCCCCTGAC GAGCATCACA AAAATCGACG
 CGCAAAAAGG TATCCGAGGC GGGGGGACTG CTCGTAGTGT TTTTAGCTGC
 951 CTCAAGTCAG AAGTGGGGGA AACCCGACAG GACTATAAAG ATACCAGGCG
 GAGTTCAGTC TTCACCCCTT TTGGGCTGTC CTGATATTTT TATGGTCCGC
 1001 TTTCCCCCTG GGAAGATCCC TCGTGGGCGC CTCCCGGTTT GGAACACTGG
 AAAGGGGGAC CCTTCTAGGG AGCACCCGCG GAGGGCCAAG CCTTGTGACC
 1051 CCGTTACCG AGATAACTAG TCCGCCTTTT TCCCTTCCCG GAAGAGGGGG
 GGCGAATGGC TCTATTGATC AGGCGGAAAA AGGGAAGGGC CTTCCTCCCC
 1101 GGGCTTTCTT CTATCTCAAC ACCGGAGGAT GCTCAACTTC CGTGGGTAGG
 CCCGAAAGAA GATAGAGTTG TGGCCTCCTA CGAGTTGAAG GCACCCATCC
 1151 GGGTGCCTC CCCACCCCGG GGTGTGTGGG CGAAAACCC AGCTAA
 CCCACGCGAG GGGTGGGGCC CCACACACCC GCTTTTGGGG TCGATT



hrg4-1

1 AGACAAAGAG TTTTTTTTTT GCGGGCGTGA TTGTATACGA CTCACTATAG
TCTGTTTCTC AAAAAAAAAA CGCCCGCACT AACATATGCT GAGTGATATC
51 GGCAGCTCG GTACCCGGGC GAATTCCAAG CTTTCGAGCG GCCGCCCGGG
CCGCTCGAGC CATGGGCCCG CTTAAGGTTT GAAAGCTCGC CGGCGGGCCC
101 CAGGTGCCGC CGCGATTGAT CATGTCGGAA ACGCTGTAAC GGTGAGCTG
GTCCACGGCG GCGCTAACTA GTACAGCCTT TGCGACATTG CCAGCTCGAC
151 AACCATACGG TCTTCATACA ATCCGAGGTA ACCGTCATAG ATTCCAAAAA
TTGGTATGCC AGAAGTATGT TAGGCTCCAT TGGCAGTATC TAAGGTTTTT
201 CTTCCAGACC TTCCGTTAAT GCCGCGCGCA CAACGCCACG AATTGCTGCG
GAAGGTCTGG AAGGCAATTA CGGCGCGCGT GTTGCGGTGC TTAACGACGC
251 TTCATGCCCC GCGCATCACC GCCACTTGTC AACACACCGA TTTTCTTAAT
AAGTACGGGC CGCGTAGTGG CGGTGAACAG TTGTGTGGCT AAAAGAAATTA
301 CATGACTACC TCTGAACTTA GGAAGCAAAA AATAATTCCG TTGCCTGAAG
GTACTGATGG AGACTTGAAT CCTTCGTTTT TTATTAAGGC AACGGACTTC
351 CGCTCCCTCA CCAGGAGTGC ACGACGCTAA TGACAATAGT ATATCAAACG
GCGAGGGAGT GGTCCCTCAG TGCTGCGATT ACTGTTATCA TATAGTTTGC
401 CCTCAAGCTG AATTGATTCA GGTGAGACCA AACGGCGGTA ATTTATACAA
GGAGTTCGAC TTAACTAAGT CCAGTCTGGT TTGCCGCCAT TAAATATGTT
451 AAAATGCCGA TCTGCCCCAC ATTTTTACAT CGAGTTTACA GACTATACCT
TTTTACGGCT AGACGGGGTG TAAAAATGTA GCTCAAATGT CTGATATGGA
501 CGGCCGCGAC CACGCTAGAT CTGGATCCC TCTAGAGTCG ACCTGCAGGC
GCCGGCGCTG GTGCGATCTA GACCTAGGGG AGATCTCAGC TGGACGTCCG
551 ATGCAAGCTT GCGTAATCA TGGTCATAGC TGTTCCTGT GTGAAATTGT
TACGTTTCGAA CCGCATTAGT ACCAGTATCG ACAAAGGACA CACTTTAACA
601 TATCCGCTCA CAATT(:GAAGCA TAAAGTGTA
ATAGGCGAGT GTTAA(:CTTCGT ATTTACATT
651 AGCCTGGGGT GCCTA(:TTAATT GCGTTGCGCT
TCGGACCCCA CGGAT(:AATTAA CGCAACGCGA
701 CACTGCCCCG TTTCC(:CCAGCT GCATTAATGA
GTGACGGGCG AAAGG(:GGTCGA CGTAATTACT
751 ATCGGCCAAC GCGCG(:TTGGGC GCTCTCCGC
TAGCCGGTTG CCGCG(:AACCCG CGAGAAGCG
801 TTCCTCGCTC ACTGA(:CGGCTG CGGCGAGCGG
AAGGAGCGAG TGACT(:GCCGAC GCCGCTCGCC
851 TATCAGCTCC CTCAA(:CACAGA ATCAGGGATA
ATAGTCGAGG GAGTTTCCGC CATTATGCCA ATAGGTGTCT TAGTCCCTAT
901 ACCGCAGGAA AGAACATGTT GGGCAAAGCA GCAAAGGCAG GACGTAAAAG
TGGCGTCCTT TCTTGTACAA CCCGTTTCGT CGTTTCCGTC CTGCATTTTC
951 GCCGGTTTGC TGGCGTTTCA TAGCTCCGCC CCCTGACAGC ATCCAAAATT
CGGCCCAACG ACCGCAAAGT ATCGAGGCGG GGGACTGTCT TAGGTTTTAA
1001 CGAGCTCAAG TCCAGTGCCA TCGACGACTT TAAGATTCAG GTCCTTGAGT
GCTCGAGTTC AGGTCACGGT AGCTGCTGAA ATTCTAAGTC CAGGAACTCA
1051 CCTTGGCCTT CGGTCGACTG GCTTACGGAT CGTCGCTTCC TGACGTGCTT
GGAACCGGAA GCCAGCTGAC CGAATGCCTA GCAGCGAAGG ACTGCACGAA
1101 CTAATCACTC AGATTATCG GTAGCTGCTC CAATGCGTGC ACCGTTCAAT
GATTAGTGAG TCTAAGTAGC CATCGACGAG GTTACGCACG TGGCAAGTTA
1151 GCTACGATCT GG
CGATGCTAGA CC



hrg1-6

1 GCCGATTTT TTTTTTTTTT TTGGGCGGAT TGTATACGAC TCACTATAGG
CGGCTAAAA AAAAAAAAAA AACCCGCCTA ACATATGCTG AGTGATATCC
51 GCGAGCTCGG TACCCGGGCG AATTCCAAGC TTAGCGTGGT CGCGGCCGAG
CGCTCGAGCC ATGGGCCCGC TTAAGGTTTCG AATCGCACCA GCGCCGCTC
101 GTACCGATTA TGACACGCTG GAAAAAGCTG GTATTGTGCC AAAAAGTACC
CATGGCTAAT ACTGTGCGAC CTTTTTCGAC CATAACACGG TTTTTGACGG
151 GCTGCAAAAT ACGCGTTAAC CCTGTTCCCG GCGCCCGCCG GGTAGAGTAT
CGACGTTTTA TGCGCAATTG GGACAAGGGC CGCGGGCGGC CCATCTCATA
201 TCAGGAGTAA GAACATGACC GCCATTCCCC AGGCGACGGC CATTTGACGGA
AGTCTCATT CTTGTACTGG CGGTAAGGGG TCCGCTGCCG GTAACAGCCT
251 TACCCCGGAT AAACATATCG TGCGCGCCAC CAACTCGGTG ATCTGGCGTA
ATGGGGCCTA TTTGTATAGC ACGCGCGGTG GTTGAGCCAC TAGACCGCAT
301 TCTCTATTTT TTATCTGTGC TCGATCTTTA TCGTCGTTCG CCTCATCCCA
AGAGATAAAA AATAGACACG AGCTAGAAAT AGCAGCAGCG GGAGTAGGGT
351 TGGAATATGC CGGGGCTGAA GAGCATCGGC TCCTACCGCT CGGTGCTCGA
ACCTTATACG GCCCCGACTT CTCGTAGCCG AGGATGGCGA GCCACGAGCT
401 GCTGCTGCAT ATTCCCTATG CCAAGCTGAT TATGGACGGC GTGATCCTCC
CGACGACGTA TAAGGGATAC GGTTCGACTA ATACCTGCCG CACTAGGAGG
451 TGTCGGTCAC CAGCTGTCTG AACTCGGCGC TGTATACCGC CTCGCGGATG
ACAGCCAGTG GTCGACAGAC TTGAGCCGCG ACATATGGCG GAGCGCCTAC
501 CTCTACTCCC TGAGCCGTCG CGGCGATGCG CCGGCCATTA TGGGCCGGAC
GAGATGAGGG ACTCGGCAGC GCCGCTACGC GGCCGGTAAT ACCCGGCCTG
551 CAACCGCAGC AAAACGCCAT ACGTGGCGGT GCTGCTCTCC ACCGCGCGCG
GTTGGCGTCG TTTTGGCGTA TGCACCGCCA CGACGAGAGG TGGCCGCGCC
601 CGTTCCCTCAC CGTAG? !CGCCAA GGTGTTTTAAA
GCAAGGAGTG GCATC? !GCGGTT CCACAAATTT
651 TTTACATCG ACAGC? !TGGTCT ATCTGGTGAT
AAAGTGTAGC TGTCG? !ACCAGA TAGACCACTA
701 CGCCGTTTCA CAACT? !CACGCG CAAGGCGGTG
CGGGCAAAGT GTTGA? !GTGCGC GTTCCGCCAC
751 AGATCCGTCT CAGAA? !GCCCGG GCGGCCGCTC
TCTAGGCAGA GTCTT? !CGGGCC CGCCGCGGAG
801 GAAGATCTGG ATCCC? !GCATGC AAGCTTGGCG
CTTCTAGACC TAGGG? !CGTACG TTCGAACCGC
851 TATCATGGTC ATAGC? !TATCCG CTCACATTC
ATAGTACCAG TATCGACAAA GGACACACTT TAACATAGGC GAGTGTAAGT
901 CACACATACG AGCCGAGCAT AAAGTGTAAG CTGGGTGCTA TGATGAGCTA
GTGTGTATGC TCGGCTCGTA TTTACATTC GACCCACGAT ACTACTCGAT
951 CTCACATATG CGTGCGCTCA CTGCCGCTTT CAGTCGAACT GTCTGCAGCT
GAGTGTATAC GCACGCGAGT GACGGCGAAA GTCAGCTTGA CAGACGTCGA
1001 GCATATGATC GCACGCCGGG AAGCGTGCGT ATGGCGCTCT CGCTCTCGCT
CGTATACTAG CGTGCGGCC TTCGCACGCA TACCGCGAGA GCGAGAGCGA
1051 CCTGACTCGC TGGTCGTCTG CGCTGCGCGA CGCATCAGCT CATCAAGCGT
GGACTGAGCG ACCAGCAGCA GCGACGCGCT GCGTAGTCGA GTAGTTTCGCA
1101 ATACCGTATC CAGAATCAGA AACGCAGAAG AACTTGTGAC ACGCCGAGCC
TATGGCATAG GTCTTAGTCT TTGCGTCTTC TTGAACACTG TGCGGCTCGG
1151 GAACGGTAGC GTGTGATCTA GTCGCTGATC AATACGCATC ATGCATCAGC
CTTGCCATCG CACACTAGAT CAGCGACTAG TTATGCGTAG TACGTAGTCCG
1201 TAGTCGGTCT GACCATAC
ATCAGCCAGA CTGGTATG



hrg2-6

1 AGGAAATCG GCCTTTTGGT GGAGGGGCAG GATTGTATAC GACTCACTAT
 TCCTTTAGC CGGAAAACCA CCTCCCCGTC CTAACATATG CTGAGTGATA
 51 AGGGCGAGCT CGGTACCCGG GCGAATTCCA AGCTTTCGAG CGGCCGCCCC
 TCCCGCTCGA GCCATGGGCC CGCTTAAGGT TCGAAAGCTC GCCGCGGGC
 101 GGCAGGTTAA TCAAATTCGT AACGCCTCGC CACGCCTCTA TAATAGCCAG
 CCGTCCAATT AGTTTAAGCA TTGCGGAGCG GTGCGGAGAT ATTATCGGTC
 151 AGCAACGTCT ATAACGCGCT GCGGGAGTGG CTGCGCGCCG GCGGCGACAC
 TCGTTGCAGA TATTGCGCGA CGCCCTCACC GACGCGCGGC CGCCGCTGTG
 201 CCGTACGCTG CGCCAGTTTG GCATCGACGC CTGGCAGATG CAGGGCGTCTG
 GGCATGCGAC GCGGTCAAAC CGTAGCTGCG GACCGTCTAC GTCCCGCAGC
 251 ACAACTATGG CAACGTCCAG TTTACCGGCT ACTACGCGCC GGTGGTGCAG
 TGTTGATAACC GTTGCAGGTC AAATGGCCGA TGATGCGCGG CCACCACGTC
 301 GCGCGCCACA CCCGTCAGGG CGAATTCCAG TACCCCATTT ACCGTATGCC
 CGCGCGGTGT GGGCAGTCCC GCTTAAGGTC ATGGGGTAAA TGGCATACGG
 351 GCCAAAGCGC GGTAAGCTGC CGTCCCGCGC CAGCATCTAT GCCGCGCGCG
 CGGTTTTCGCG CCATTTCGACG GCAGGGCGCG GTCGTAGATA CGGCCGCGCG
 401 TGAGCGATAA CTATGTGCTG GCGTATAGCA ACTCGTTGAT GGATAAATTTC
 ACTCGCTATT GATACACGAC CGCATATCGT TGAGCAACTA CCTATTGAAG
 451 ATCATGGACG TGCAGGGCAG CGGCTATATC GACTTTGGCG ATGGTTTCGCC
 TAGTACCTGC ACGTCCCGTC GCCGATATAG CTGAAACCGC TACCAAGCGG
 501 GCTCAACTTC TTTAGCTATG CCGGAAAAA CGGCTGGCCC TATCGCAGCA
 CGAGTTGAAG AAATCGATAC GGCCCTTTTT GCGACCGGG ATAGCGTCTG
 551 TCGGCAAAGT GCTAATCGAC CGCGGGGAAG TGAAGAAAGA AGATATGTCTG
 AGCCGTTTCA CGATTAGCTG GCGCCCTTTC ACTTCTTTCT TCTATACAGC
 601 ATGCAGGCGA TCCGCGC 'CGGCCG CGACCACGCT
 TACGTCCGCT AGGCGC 'GCCGGC GCTGGTGCGA
 651 AGATCTGGAT CCCCTC 'ATGCAA GCTTGGCGTA
 TCTAGACCTA GGGGAC 'TACGTT CGAACCGCAT
 701 ATCATGGTCA TAGCTC 'TATCCG CTCACAATTTC
 TAGTACCAGT ATCGAC 'ATAGGC GAGTGTAAAG
 751 CACACAACAT ACGAGC 'AGCCTG GGGTGCCTAA
 GTGTGTTGTA TGCTCC 'TCGGAC CCCACGGATT
 801 TGAGTGAGCT AACTC 'CACTGC CCGCTTTCCA
 ACTCACTCGA TTGAG 'GTGACG GGCGAAAGGT
 851 GTCGGGAAC CTGTCC 'ATCGGC CAACGCGCGG
 CAGCCCTTTG GACAGCACGG TCGACGTAAT TACTTAGCCG GTTGCGCGCC
 901 GGAGAGGCGG TTTGCGTATT GGCCTCTTC CGCTTCTCGC TCACTGACTC
 CCTCTCCGCC AAACGCATAA CCGCGAGAAG GCGAAGAGCG AGTGACTIONG
 951 GCTGCGCTCG TCGTCGCTGC GCGAGCGTAT CAGCTCACTC AAAGCGTATA
 CGACGCGAGC AGCAGCGACG CGCTCGCATA GTCGAGTGAG TTTTCGCATAT
 1001 CGTATCACAG ATCAGGGGATA CGCAGAAGAC ATGTGAGCAA AGCAGCAAAG
 GCATAGTGTC TAGTCCCTAT GCGTCTTCTG TACACTCGTT TCGTCTTTTC
 1051 CAGACGTAAA AGCGCGTTGC TGCGTTTCAT AGCTCGCCCC CTGACAGCAT
 GTCTGCATTT TCGCGCAACG ACGCAAAGTA TCGAGCGGGG GACTGTCTGTA
 1101 CAAAATCGA CCTCAGTCAG TGGCGATCGG ACGACTAAGA TCAGTCCCTG
 GGTTTTAGCT GGAGTCAGTC ACCGCTAGCC TGCTGATTCT AGTCCAGGAC
 1151 AGTCTCTGG CCTTCGTCGA CCTGCGTTAC GGAACGTGCT TCTCTGACGT
 TCAGGAGACC GGAAGCAGCT GGACGCAATG CCTTGCACGA AGAGACTGCA
 1201 GTTCATCACG TAGATCATCG TAGTGCTCAC TG
 CAAGTAGTGC ATCTAGTAGC ATCACGAGTG AC

