

(19) United States

(12) Patent Application Publication (10) Pub. No.: US 2006/0127988 A1 Wood et al.

Jun. 15, 2006 (43) Pub. Date:

(54) SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF **PREPARATION**

(75) Inventors: Keith V. Wood, Madison, WI (US); Monika G. Gruber, Madison, WI (US); Yao Zhuang, Madison, WI (US); Aileen Paguio, Madison, WI (US)

> Correspondence Address: Schwegman, Lundberg, Woessner & Kluth, P.A. P.O. Box 2938 Minneapolis, MN 55402 (US)

(73) Assignee: Promega Corporation

(21) Appl. No.: 11/316,042

(22) Filed: Dec. 22, 2005

Related U.S. Application Data

(62) Division of application No. 09/645,706, filed on Aug. 24, 2000.

Publication Classification

(51) Int. Cl. C12P 19/34 C07H 21/04 (2006.01)(2006.01)C12N 15/63 (2006.01)

(52) **U.S. Cl.** 435/91.1; 536/23.2; 435/320.1

(57)**ABSTRACT**

A method to prepare synthetic nucleic acid molecules having reduced inappropriate or unintended transcriptional characteristics when expressed in a particular host cell.

Figure 1
The Genetic Code

First Position (5' end)	Secoi	nd posi	tion		Third position (3' end)
Ĺ	U	С	Α	G	
ļ	Phe	Ser	Tyr	Cys	ี บ
	Phe	Ser	Tyr	Cys	C
U	Leu	Ser	Stop	Stop	A
	Leu	Ser	Stop	Tro	G
	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
C	Leu	Pro	Gln	Arg	Α
	Leu	Pro	Gln	Arg	G
	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
A	Ile	Thr	Lys	Arg	Α
	<u>Met</u>	Thr	Lvs	Arg	G
	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
G	Val	Ala	Glu	Gly	Α
	Val	Ala	Glu	Gly	G

Figure 2

	ATGATGA	AACGCGAAAA	GAACGTGATCTACGGCCCAGAAC 40
GR6.SEQ	ATGATGA	A A C G C G A A A A	GAACGTGATCTACGGCCCAGAAC 40
GRVER5.SEQ			GAACGTGATCTACGGCCCAGAAC 40
GRVER4.SEQ	ATGATGA	A A A C G C G A A A A A	GAACGTGATCTACGGCCCAGAAC 40
GRVER3.SEQ			GAACGTGATCTACGGCCCAGAAC 40
GRVER2.SEQ	ATGATGA	AAC GCG AAA AC	GAACGTCATCTACGGCCCAGAGC 40
RVER1.SEQ	ATGATGA	AlAC GICIG ALA ALC	GAACGTCATCTACGGCCCAGAGC 40
G81-6G1.SEC	ATGATGA	AGCGAGAGAA	AAATGTTATATATGGACCCGAAC 40
RDVER1.SEQ	ATGATGA	AGCGTGAGAA	AAATGTGATTTATGGTCCTGAAC 40
RDVER2.SEQ	ATGATGA	AGCGTGAGAA	AAATGTGATTTATGGTCCTGAAC 40
DVER3.SEQ	ATGATGA	AGCGTGAGAA	AAATGTCATCTATGGCCCTGAGC 40
DVER4.SEQ	ATGATGA	AGCGTGAGAA	AAATGTCATCTATGGCCCTGAGC 40
DVER5.SEQ	ATCATCA	AGCGTGAGAAA	AAATGTCATCTATGGCCCTGAGC 40
D7.SEQ	ATCATCA	AGCGIIGAGAAA	A A A T G T C A T C T A T G G C C C T G A G C 40
	AIGAIGA	AGCGTGAGAAA	AAATGTCATCTATGGCCCTGAGC 40
OURDES ORG	ATGATGA	AGCGTGAGAAA	AAATGTCATCTATGGCCCTGAGC 40
UVERSZ.SEQ	ATGATGA	AGCGTGAGAAA	AAATGTCATCTATGGCCCTGAGC 40
DIS61H9.SEQ	A T G A T A A	AGCGTGAGAAA	AAATGTCATCTATGGCCCTGAGC 40
		0 ==	
	CACTGCA	TCCACTGGAAG	GACCTCACCGCTGGTGAGATGCT 80
R6.SEQ	CAC TIGIC A	TICCIACITGGAAG	GACCTCACCGCTGGTGAGATGCT 80
RVER5.SEQ	CACTGC A	TCCACTGGAAG	GACCTCACCGCTGGTGAGATGCT 80
RVER4.SEQ	CACTGCA	TCCACTGGAAG	GACCTCACCGCTGGTGAGATGCT 80
RVER3.SEQ	CACTGCA	TCCA <u>C</u> TGGAAG	GAC C T C AC C GCTGG T GA G ATGCT 80
RVER2.SEQ	CTCTGCA	CCCATTGGAAG	GACCTGACCGCTGGTGAGATGTT 80
RVER1.SEQ	CICTGCA	CCCATTGGAAG	GACCTGACCGCCGGTGAGATGTT 80
G81-6G1.SEQ	CCCTACA	CCCCTTGGAAG	GACTTAACAGCTGGAGAAATGCT 80
DVER1.SEQ	CATTGCA	TCCTCTGGAGG	GATTTGACTGCTGGCGAAATGCT 80
DVER2.SEQ	CATITIGICA	TCCTCTGGAGG	GATTTGACTGCCGGCGAAATGCT 80
DVER3.SEQ	CTTTGCA	CCCTTTGGAGG	GATTTGACTGCCGGCGAAATGCT 80
DVER4.SEQ	CTTTGCA	TIC CITIT T G G A G G	GATTTGACTGCCGGCGAAATGCT 80
DVER5.SEQ	CITIC TICK	TICCITITEGALG	GATTTGACTGCCGGCGAAATGCT 80
D7.SEQ	CITIC TICIC A	TIC CITIT T G G A G G	GATTTGACTGCCGGCGAAATGCT 80
	CITIC TICICA	TC CTT T C C N C C	
DVER52 SEC		TICCITTTCCA	GA T TT G AC T GC C GG C GAAATGCT 80 GA T TT G AC T GC C GG C GAAATGCT 80
D1561H9 SEA	CITIC TICK	TIC CITE TO GALGE	GAITT TIGACITIC CICIG GICIG A A A T G C T 80
D1501115.5EQ	CELC LEIC K	TICCHILLOGNE	GATTTGACTGCCGGCGAAATGCT 80
RVER51.SEO	Сттсска	е състес села	A A C A TAGT C ACCTCC CTC AAG CA 120
R6.SEQ	CTTCCG	G CACTEC GTA	A A C A T A G T C A C C T C C C T C A A G C A 120
RVER5.SEQ	CTTCCGI	G CIAIC TIGIC GITTA A	A A C A T A G T C A C C T C C C T C A A G C A 120
EVER4.SEQ		G CACTICO CIMA	A A C A T A GT C A C C T C C C T C A A G C T 120
RVER3.SEQ	GTTCCCT	C C C C B C C C B A A	A A C A TIA GIT C AIC CITICIS CITIC A AIG C T 120
EVER2.SEQ	CTTCCGT	C C T C T G C G T A A	AACATAGCCACCTGCCTCAAGCT 120
WED1 COA		COMO DICIO CITIA A	AACATTCTCACTTGCCTCAAGCC 120
RVER1.SEQ	orrecet	G CITIC TIGIC GITA A	AACATTCTCACTTGCCTCAAGCC 120
or-odr.SEO		GUCCTTCGAAA	AACATTCTCATTTACCGCAGGCT 120
OVER1.SEQ	G T T T C G C	G C CITITIGIC GICIA A	AGC ACAGCCATCTGC CACAGGCT 120
OVER2.SEQ	ᄗᅚᅚᆙᅚᆝᅂᅝᆙᇊ	CCLLLCCCCV V	AGCACAGCATCTGCCACAAGCT 120
OVER3.SEQ		G C T T T G C G T A A	AGCACTCTCATTTGCCTCAAGCC 120
DVER4.SEQ	G T T T C G T	G CIT TITIGIC GITIA A	A A C A C T C T C A T T T G C C T C A A G C C 120
DVER5.SEQ	G т т т с G т	G CITIC TICIC GICIA A	AGC ACT CT CATTTGC CTC AAG CC 120
D7.SEQ	GTTTCGT	G CITIC TICIC GICIA A	AIGIC AICIT C TITIA T T TIGIC CITIC AIAIG CICI 120
DVER51.SEQ	G Т Т Т С G Т	G CITIC TICIC GICIA A	AGCACTCTCATTTGCCTCAAGCC 120
DVER52.SEQ	G T T T C G T	G CITIC TICIC GICIA A	A G C A C T C A T T T G C C T C A A G C C 120
1561H9.SEQ	G т т т с с т	G CITIC TICIC GICIA A	AGCACTCTCATTTGCCTCAAGCC 120

Figure 2 (cont.)

- 1 Iguro 2 (
GRVER51.SEQ CTCGTGGAC	G T C G T G G G A G A C G A G A G C C T C T C C T A C A A A G 160 G T C G T G G G A G A C G A G A A C C T C T C C T A C A A A G 160
GRVERS. SEQ CIT CIG TIGIG A C	GTCGTGGGAGACGAGAGCCTCTCCTACAAAG 160
GRVER4. SEQ CITICIG TIGIG A C	GTCGTGGGAGACGAGAGCCTCTCTTACAAAG 160
GRVER3. SED CTCGTGGAC	GTCGTGGGTGACGAGAGCCTGTCTTACAAAG 160
GRVER2. SEQ CTGGTCGAT	GTCGTGGGCGACGAGAGCTTGTCTTTATAAGG 160
GRVER1.SEQ CTGGTGGAT	GTCGTGGGCGACGAAAGCTTGTCTTATAAGG 160
YG81-6G1. SEQT TAGTAGAT	GTGGTTGGCGACGAATCGCTTTCCTATAAAG 160
RDVER1.SEQ TTGGTCGAC	GTGGTCGGTGATGAGTCTCTGAGCTACAAAG 160
RDVER2.SEQ TTGGTGGAC	GTGGTCGGTGATGAATCTCTGAGCTACAAAG 160
RDVER3.SEQ TTGGTCGAT	GTGGTCGGCGATGAATCTTTGAGCTATAAGG 160
RDVER4.SEQ T T G G T C G A T	GTGGTCIGGCGAITIGAATCITTITIGAGICTACIA AICIG 160
KDVERS. SEQ TTGGTCGAT	GTGGTCGGCGATGAATCTTTTGAGCTACAAGC 160
KDI. SEQ TTIGIGTICIGAT	G T G G T C G G C G A T G A A T C T T T T G A G C T A C A A C C C C C C C C C C C C C C
RDVERSI. SEQ T TIGIG TICIG A T	G T G G TICIG G C G AITIG A A T CIT TITIG A CIC T AICIA AICIG 160
RDVER52.SEQ T T G G T C G A T	G T G G T C G G G G A T G A A T C T T T T G A G C T A C A A C C A C C
RD1561H9.SEQTTGGTCGAT	GTGGTCGGCGATGAATCTTTGAGCTACAAGG 160
GRVER51.SEQ A A T T T T C G	AAGCTACTGTGCTGTTGGCCCCAAAGCCCTCCA 200
GK6.SEQ AATTTTCG	A A G CITIA CITIG TIGIC TIGITITISIS CICIC A A A GIOC T C C A 200
GRVERS. SEQ A A T T T T T C G	A A G CITIA CITIG TIGIC TIGIG CICIC A A A GICIC T C C A 200
GRYENA. SEQ A A T T T T T C G	A A G C[T]A C T G T G C T G T T G G C C C A A A G C C T C C A 200
GRVERS. SEQ A A T T T T T C G	A A G CITIA CITIG TIGIC TIGIG CICIC A A GICIC TIGIC A 200
GRVER2.SEQ AATTTTTCG.	AAGCTACTGTCCTGTTGGCCCCAATCTCTGCA 200
GRVER1.SEQ AGTTTTTCG.	AAGCTACTGTCCTGTTGGCCCAGTCTCTGCA 200
RDVER1.SEQ AATTCTTTT	AAGCGACAGTCCTCCTAGCGCAAAGTCTCCA 200
RDVER2.SEQ AGTTCTTTG	AGG CCA CCG TG TTGC TGG CTC A A A GC TTGC A 200
RDVER3.SEQ AGTTTTTG	A G G C A A C C G T G T T G C T G G C T C A G A G C T T G C A 200 A G G C A A C C G T C T T G C T G G C T C A G T C T T T G C A 200
RDVER4.SEQ AGTTTTTG	AGGCAACCGTCTTGCTGGCTCAGTCTTTGCA 200
RDVER5.SEQ AGTTTTTG	AGG CAACCGTCTTGCTGGCTCAGTCCCTCCA 200
RD/.SEQ AGTTTTTG	AIGIG CIAIA CICIG T CITITIGIC TIGIG CITIC AIG T C CIC T C C A 200
KDYEKSI.SEQ AGTTTTTG	A IGIG CIAIA CICIG T CITITIGIC TIGIG CITIC A IGIT C CICIT C C A 2000
ROVERSZ.SEQ AGTTTTTG	AIGIG CIAIA CICIG T CITITIGIC TIGIG CITIC AIG T C CIC T C C A 200
RD1561H9.SEQAGTTTTTG	AGGCAACCGTCTTGCTGGCTCAGTCCCTCCA 200
_	
GR6.SEQ TAATTGTGGG	GTACAAAATGAACGATGTGGTGAGCATTTGT 240
GRVER5.SEQ TAATTGTGG	GTACAAAATGAACGATGTGGTGAGCATTTGT 240
GRVER4.SEQ TAATTGTGG	GTACAAAATGAACGATGTGGTGAGCATTTGT 240
GRVER3.SEQ TAATTGTGG	ATACAAAATGAACGATGTGGTGAGCATTTGT 240
GRVER2.SEQ TAATTGCGGG	TTACAAAATGAACGATGTGGTGAGCATCTGT 240 TTACAAAATGAACGATGTGGTCAGCATTTGT 240
GRVER1. SEQ TAATTGCGGT	
YG81-6G1. SEQC A A T T G T G G A	TTACAABATGAACGATGTGGT <u>CAGC</u> ATTTGT 240 ATACAAGATGAATGA <u>T</u> GT <u>A</u> GTGTC <u>G</u> ATCTGC 240
RDVER1.SEQ CAACTGTGG	CTATAAGATGAATGACGTCGTGTCTATCTGC 240
RDVER2.SEQ CAACTGTGG	CTATAAGATGAATGACGTCGTGTCTATCTGC 240
KUVERS. SEQ T AATTG C GG C	CITACAAGATGAAICIGAICIGTICITCITATITGT 240
RDVER4. SEQ TAATTGTGG	CIT A C A A G A T G A AICIG AICIG TICIG TICIT CICIA TITIT GITI 240
RDVERS. SEQ CAATTGTGG	C TACAAGATGAA C GA C GT C GT TAGT ATCTG T 240
KD1.SEQ CAATTGTGG	CITACAAGATGAA CIGA CIGT CIGT TAGT ATCTG T 240
RDVER51.SEQ CAATTGTGG	CITACAAGATGAAICIGAICIGTITAGTIATCTGITI 240
PD156149 SECON TO TO TO G	CTACAAGATGAACGACGTCGTTAGTATCTGT 240
	CTACAAGATGAACGACGTCGTTAGTATCTGT 240

Figure 2 (cont.)

GRVER51.SEQ G C T G A G A A T A A C A C T C G C T T C T T T A T T C C T G T A A T C G C T G 280
GR6.SEQ GC T GAGAATAA C AC TC G C TT C TTTATTCC T GT A AT C GC T G 280
GRVERS.SEQ GC T GAGAATAA C AC TC G C TT C TTTATTCC T GT AAT C GC T G 280
GRVER4. SEQ G C T G A G A A T A A C A C T C G C T T T T T T T C C C T G T T A T C G C T G 280
GRVER3.SEQ G C T G A G A A T A A C A C T C G C T T T T T T A T C C C T G T G A T C G C T G 280
GRVER2.SEQ G C T G A G A A T A A C A C C C G C T T T T T C A T C C C A G T G A T T G C C G 280
GRVER1.SEQ G CTG A G A A T A A C A C C C C C T T T T T C A T C C C A G T G A T T G C C G 280
YGB1-6G1.SEQG C C G A G A A T A A T A C A A G A T T T T T T A T T C C C G T T A T T G C A G 280
RDVER1.SEQ GCCGAAAAACAATACTCGTTTTCTTTATTCCTGTCATCGCTG 280
RDVER2.SEQ GCCGA A AA C AATAC TC G T TTC TTTATTCC T GT C AT C GC T G 280
RDVER3.SEQ GCCGA A AA C AATAC CC G T TT C TT C ATTCC A GT C AT C GC C G 280
RDVER4.SEQ G C A A A A A A C A A T A C C C G T T T T C T T C A T T C C A G T C A T C G C C G 280
RDVER5.SEQ G C T G A A A C A A T A C C C G T T T C A T T C C A G T C A T C G C C G 280
RD7. SEQ G C T G A A A C A T A C C C G T T T C A T T C C A G T C A T C G C C G 280
RDVER51. SEQ G C T G A A A A C A A T A C C C G T T T C T T C A T T C C A G T C A T C G C C G 280
RDVER52. SEQ G C T G A A A C A A T A C C C G T T C T T C A T T C C A G T C A T C G C C G 280
RD1561H9.SEQG CTG ALAA CAATA CC CGTTTTCTTCLATT C CLAGTCATCG CCG 280
GRVER51.SEQ CTTGGTACATCGGCATGATTGTCGCCCCTGTGAATGAATC 320
GR6. SEQ CTTGGTA C AT C GG C ATGATTGT C GC CCTGT G AATGAA TC 320
GRVERS.SEQ CTTGGTA C AT C GG C ATGATTGT C GC CCTGT G AATGAA TC 320
GRVER4.SEQ CTTGGTA C AT C GG C ATGATTGT C GC CCTGT G AATGAA TC 320
GRVER3.SEQ CTTGGTA C AT C GG C ATGATTGT C GC CCTGT G AATGAA TC 320
GRVER2.SEQ CTTGGTA C AT C GG C ATGATTGT C GC CCTGT G AATGAA TC 320
GRVER1.SEQ CTTGGTACATCGGCATGATTGTCGCCCTGTGAATGAA <u>TC</u> 320
YGBI-6GI.SEQCITGGTATATTGGTATGATTGTAGCACCIGTTAATGAAAG 320
ROVER1.SEQ CCTGGTATATTGGTATGATCGTGGCTCCAGTCAACGAGAG 320
ROVERS SEQ CICT GGT AT AT T GGT AT GATICIG T GG C T C C A G T C C A G G A G A G 320
ROVER3. SEQ CICIT G G T A T A T C G G T A T G A T C G T G G C T C C A G T C A A C G A G A G 320
RDVER4. SEQ CATGGTATATCGGTATGATCGTGGCTCCAGTCAACGAGAG 320
RDVERS.SEQ CATGGTATATCGGTATGATCGTGGCTCCAGTCAACGAGAG 320
RD7.SEQ CATGGTATATCGGTATGATCGTGGCTCCAACGAGAGAGAG
RDVER51. SEQ CATGGTATATICGGTATGATICGTIGGCTCCAACTCAACGAGAG 320
ROVER52.SEQ CATGGTATATCGGTATGATCGTGGCTCCAGTCAACGAGAG 320
RD1561H9.SEQCATGGTATATCGGGTATGATCGTGGCTCCAGTCAACGAGAG 320
GRVER51.SEQ TTACATCCCAGATGAGCTGTGTAAGGTTATGGGTATTAGC 360
GR6.SEQ TTACATCCCAGATGAGCTGTGTAAGGTTATGGGTATTAGC 360
GRVERS. SEQ TTACATCCCAGATGAGCTGTGTAAGGTTATGGGTATTAGC 360
GRVER4. SEQ TTACATCCCAGATGAGCTGTGTAAGGTTATGGGTATTAGC 360
GRVER3. SEQ TTACATCCCAGATGAGTTGTGTAAGGTGATGGGTATTAGC 360
GRVER2. SEQ TTATATCCCAGACGAGTTGTGCAAGGTCATGGGTATTAGC 360
GRVER1.SEQ TTATATCCCAGAGGGTTGTGCGAAGGTCATGGGTATTAGC 360
YG81-6G1.SEQTTACATCCCAGATGAACTCTGTAAGGTGATGGGTATATCG 360
RDVER1. SEQ CTACATTCCTGATGAACTGTGTAAAAGTGATGGGCATCTTCT 360
RDVER2. SEQ CTACATTCCTGATGAACTGTGTAAAGTGATGGGCATCTCT 360
ROVERS. SEQ CTACATTCCTGACGAACTGTGTAAAGTCATGGGTATCTCTT 360
ROVER4. SEQ CTACATTCCCCGACGAACTGTGTAAAGTCATGGGTATCTCTT 360
RDVERS. SEQ CTACATTCCCCGACGAACTGTGTAAAGTCATGGGTATCTCTT 360
RD7.SEQ CTACATTCCCGACGAACTGTGTAAAGTCATGGGTATCTCTT 360
RDVER51.SEQ CTACATTCCCGACGAACTGTGTAAAGTCATGGGTATCTCTT 360
RDVERS2.SEQ CTACATTCCCGGACGAACTGTGTAAAGTCATGGGTATCTCTTT 360
RD1561H9.SEQCTACATTCCCGACGAACTGTGTAAAGTCATGGGTATCTCTT 360

Figure 2 (cont.)

GRVERS1. SEQ A A A C C T C A A A T C G T C T T T A C T A C C A A A A A C A T C T T G A A T A 400
GR6.SEQ AAACC T CAAAT C GT C TTTAC T AC C AA A AACAT C TT G AATA 400
GRVERS.SEQ AAACCITCAAATCIGTTCITTTACITACICAAAAAACATCITTIGAATA 400
GRVER4. SEQ AAACCITCAAATCGTICTTTACITACCAAAAATATCCITGAATA 400
GRVER3.SEQ AAACCTCAAATCGTCTTTACTACCAAAAACATCCTGAAAAAAAA
GRVERZ.SEQ AAACCICAAATCGTGTTTACITACCAAGAACATTCITGAATA 400
GRVER1.SEQ AAACCTCAAATCGTGTTTACTACCAAGAACATTCTGAATA 400
YG81-6G1. SEQAAACCACAAATAGTTTTTACGACAAAGAACATTTTAAATA 400
RDVER1.SEQ A AGC C A CAGA TTG TCT TCA CCACTA AAA ATA TCT TGA ACA 400
RDVER2.SEQ A A G C C A C A G A T T G T C T T C A C C A C T A A A A T A T C T T G A A C A 400
RDVER3.SEQ A A G C C A C A G A T T G T G T T C A C C A C T A A G A A T A T T T T G A A C A 400
RDVER4.SEQ AAGCCACAGATTGTCTTCACCACTAAGAATATTCTGAACA 400
RDVERS.SEQ AAGCCACAGATTGTCTTCACCACTAAGAATATTCTGAACA 400
RD7. SEQ A A G C C A C A G A T T G T C T T C A C C A C T A A G A A T A T T C T G A A C A 400
RDVERS1. SEQ A A G C C A C A G A T T G T C T T C A C C A C T A A G A A T A T T C T G A A C A 400
RDVER52.SEQ A A G C C A C A G A T T G T C T T C A C C A C T A A G A A T A T T C T G A A C A 400
RD1561H9. SEQA AGC C A C AGA T TG TCT T CA CCA CTA A G A ATA T TCTGA ACA 400
ANTIONIO 1924 A GOLO A O A GOLA I EL A EL A CELA CELA CELA CELA CELA CELA
GRVER51. SEQ AGGTCTTGGAAAGTCCAGTCTCGTACTAACTTCATCAAACG 440
GRE. SEQ AGGTCTTGGAAGTCCAGTCTCGTACTAACTTCATCAAACG440
GRVERS. SEQ AGGTCTTGGAAGTCCAGTCTCGTACTAACTTCATCAAACG440
GRVER4.SEQ AGGT CTTGGAAGTCCAGTCTCGTACTAACTTCATCAAACG 440 GRVER3.SEQ AGGT CTTGGAAGTCCAGTCTCGTACTAATTTCATCAAACG 440
GRVERZ. SEQ AGGTCTTGGAAGTGCAGTCTCGTACTAACTTCATCAAGCG 440
GRVER1. SEQ AAG TCTTGGAAGCG 440
YG81-6G1. SEQAGGTATTGGAGGTACAGAGCAGAACTAATTTCATAAAAAG 440
RDVERL.SEQ AGGTGCTGGAGGTCCAAAGCCGCACCAATTTTAATAAACG 440
RDVER2. SEQ A A G T G C T G G A G G T C C A A A G C C G C A C C A A T T T T A T T A A A C G 440
RDVER3.SEQ A A G T G C T G G A A G T C C A A A G C C G C A C C C A A C T T T A T T A A G C G 440
ROVER4. SEQ A A G T C C T G G A A G T C C A A A G C C G C A C C A A C T T T A T T A A G C G 440
RDVERS.SEQ A A G T C C T G G A A G T C C A A A G C C G C A C C A A C T T T A T T A A G C G 440
RD7.SEQ A A G T C C T G G A A G T C C A A A G C C G C A A C T T T A T T A A G C G 440
RDVER51.SEQ A A G T C C T G G A A G T C C A A A G C C G C A A C T T T A T T A A G C G 440
RDVER52.SEQ A A G T C C T G G A A G T C C A A A G C C G C A C C A A C T T T A T T A A G C G 440
RD1561H9. SEQA AG T C C T G G A AG T C C A AA G C C G C A C C T T T A T T A A G C G 440
GRVERS1. SEQ CATCATTATTCTGGATACCGTCGAAAACATCCACGGCTGT 480
GR6.SEQ CATCATTATTCTGGATACCGTCGAAAACATCCACGGCTGT 480
GRVERS. SEQ CATCATITATITCTGGATACCGTCGAAAACATCCACGGCTGT 480
GRVER4.SEQ CATCATTATTCTGGATACCGTCGAAAACATCCATGGCTGT 480
GRVER3. SEQ CATTATTATTCTGGATACCGTCGAAAACATCCACGGCTGT 480
GRVER2. SEQ CATTATCATTCTGGATACCGTCGAGAATATCCACGGCTGT 480
GRVERL. SEQ CATTATCATTCTGGATACCGTCGAGAATATCCACGGCTGT 480
YG91-6G1.SEQGATCATACTTGATACTGTAGAAAACATACACGGTTGT 480
RDVER1. SEQ TATCATTATCTTGGACACTGTGGAAAACATTCATGGTTGC 480
RDVER2.SEQ TATCATTATCTTGGACACTGTGGAAAACATTCATGGTTGC 480
RDVER3.SEQ TATCATCATCTTGGACACTGTGGAATATTCATGGTTGC 480
RDVER4.SEQ TATCATCATCTTGGACACTGTGGAATATTCACGGTTGC 480
RDVERS. SEQ TATCATCATCTTGGACACTGTGGAGAATATTCACGGTTGC 480
RD7. SEQ TATCATCATCTTGGACACTGTGGAGAATATTCACGGTTGC 480
RDVER51. SEQ TATCATCATCTTGGACACTGTGGAGAATATTCACGGTTGC 480
RDVERS2.SEQ TATCATCATCTT GGACACTGT GGAGAATATT CACGGTT GC 480
RD1561H9. SEQTATCATCATCTTGGACACTGTGGAGAATATTCACGGTTGC 480

Figure 2 (cont.)

GRVERS1. SEQ G A G A G C C T C C C T A A C T T C A T C T C T C	
GRO. SEQ GAGGA GICIC TICIC CITIA AICIT TICIA TICIT C T C G T T AIC A C CIC A T C CITIA 520	
GRVERS. SEQ GAIGIA GICIC TICIC CITIA AICIT TICIA TICIT C T C G T T AIC A G CIG A T G CITIA 520	
GRADAL SEC GRIGH GICLE TIGIC CITIA AICIT TICIA TICIT C T C G T T AIC A G C G A T C C T A C A	
GRVERS. SEQ GAIGIA GICTITIGIC CITIA AICIT T TA TICIT C T C G T T AIC A G C G A T G C ITIA 500	
GREEKZ. SEQ G A GG A G C T T T G C C A A A C C T T A T T T C T C G T T A T A G C G A C C C C T A SOC	
GRADEL SEQ GARAGIC TITIGIC CIAIA AICIT T T A T T T C T C G T T A T I A C C C A C C C MIN 500	
1G01-6G1. SEQGAAAGTCTTCCCAATTTTATTCTCGCTTATTCCGCATCCAA 520	
ADVERTIBLE GAGT CIT C TIGIC CITIA A T T TICIA TIC A G CIC GICIT AICHT CITIC A T C COL SOC	
RUVERZ. SEQ GAAIT CIT CITIGIC CITIA AIT TITICIA TIC AGCIC GICIT AICHT CITIGIAIT COLOR FOR	
REVERSURE G A ALL CIT C TIGIC CITIA A T T TICIA T TIA G CIC GICIT A T T CITIC A CIC CICIA FOR	
REVERSIBLE OF A AIT CITITITIGIC CITIA A T T T T A T TIA G CIC GICIT A T T CIAIC AIGIG A B COO	
REVERSISED GAAT CITITIFICIC CITIA AT TITICIA TICIT CITICITA TITICIA CIDICIA CON SOCIEDA	
A A I ULITINGS CITIA A T T TICIA TICIT C T C GICIT A T T CIAIG A ICIC CICIA 600	
AND LINE SECURE A AIT CITITITITITICS CITIA A T T TICIA TICIT C T C GICIT A T T CINIC AIGIC GIGIN FOR	
RUVERSZ. SEQ G A AIT CITITITIGIC CITIA A T T TICIA TICIT C T C GICIT A T T CIAIC AIGIC CIGIT SCO	
RD1561H9. SEQG A AT CTTTGC CTAATTTCATCTCGCTATTCAGGCG CA 520	
GRVER51.SEQ A T A T C G C T A A T T T C A A G C C C T T G C A T T T T G A T C C A G T C G A 560	
ATATICIS CITIA AITIT T CA AIGIC CICIT TIGICAT TITIGAT C CIAIG TICICA	
ONVERSURE A TATICIS CITIA AITIT T C A AIGIC CICIT TIGIC A T T TITIC A T C CINIC TICIC A SCO	
GRYERS, SEC. A TATICIS CITIA AITIT T C A A A C CIA CITIGIC A T T TITIG A T C CIAIG TICIC A TCO	
ON BROADE A TATION CITIE ALT T C A AIGIC CIA CITIGIC A T T TITIG A T C CIAIC TICIO A 500	
GRYERZ, SEV A LA LIUIG CITIA A C T T C A AIGIC C TICITIGIC A T T TITIG A T C CIAIC TICIC A CCA	
GRYENI. SEQ. A T A T C G C T T A A C T T C A A G C C T C T G C A T T T T T G A T C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C C A C	
1GOI GOI GEORA CARCATTO A A A CONTRACATO TO CARO CONTRACA ECO	
ADVERTISED A CONTROL TO CONTROL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL	
ADVERS. SEQ A COM T T G C C A AITIT TITIA A A C CIAIT TIGIC AICIT T C C A CIC C T C T C T C T C T C T C T C T C T	
ACTUAL TOTAL CONTROL OF THE CONTROL	
ADVERTISED A CA TICIG C CAACT TITIA A GIC C TICITICICAT T T C G A CC C T G TICICA SECO	
AND TAKES OF A POINT POINT A A C T TITIA A A C CIA CITICIC AICIT T C G AICIC C T G TIGIG A 560	
ADI-350 AICIA TICIG CIAIA A C. T. TITIA A A C. CIA CITICIC AICIT TO C. C. AICIC O T. O. T. C. C. AICIC O T	
RDVER51. SEQ A CAT CG CAAACTTTAAACCACTCCACTTCGACCCTGTGGA 560	
TO SECOND A POINT POINT A PROPERTY AND A PROPERTY AND A POINT A POINT AND A PO	
RD1561H9.SEQACATCGCAAACTTTAAACCACTCCACTTCGACCCTGTGGA 560	
GRYERS1, SEC. G. C. A. A. G. C.	
GRYERS1.SEQ G C A A G T G G C C G C T A T T T T G T G C T C C T C C G G C A C C A C T G G T G G G G G G G G G G G G G G G	
GRVER2. SEO. G.C.A.A.G.T.G.C.G.C.G.C.G.A.R.T.T.T.T.G.T.G.G.G.C.A.C.T.G.G.T. 600	
GRVER1.SEQ G C A A G T C G C C G C T A T T T T G T G C T C T A G C G G C A C C A C C G G T 600	
YG81-6G1.SEQG CAAGTGG CAGCTATCTTATGTTCGTCAGGCACTACTGGA 600	
RDVER1. SEQ A C A G G T G G C T G C C A T C C T G T G T A G C T C T G G T A C T G G C 600	
RDVER2.SEQ ACAGGTGGCTGCCATCCTGTGTAGCTCTGGTACCTACTGGC 600	
RDVER3. SEQ ACAAGTGGCTGCTATCCTGTGTAGCAGCGGTACTACTGGC 600	
RDVER4.SEQ A C A A G T T G C T G C A A T C C T G T A G C A G C G G T A C T A C T G G A 600	
RDVER5. SEQ A C A A G T T G C A G C C A T T C T G T A G C A G C G G T A C T A C T G G A 600	
RD7. SEQ A C A A G T T G C A G C C A T T C T G T A G C A G C G G T A C T A C T G G A 600	
ROVER51. SEQ A C A A G T T G C A G C C A T T C T G T A G C A G C G G T A C T A C T G G A 600	
ADVENDED BY A A G TITIG C A G CICIA TIT CITIGIT G TIA G C A G CIG GITIA C T A C T G C A SOO	
RD1561H9.SEQACAAGTTGCAGCCATTCTGTGTAGCAGCGGTACTACTGGA 600	

Figure 2 (cont.)

GRVER51.SEQ TTGCCTAAAGGTGTCATGCAGGACTCACCAGAATATCTGTG 640
GRE. SEQ TTGCCTAAAGGTGTCATGCAGACTCACCAGAATATCTGTG 640
GRVERS.SEQ TTGCCTAAAGGTGTCATGCAGACTCACCAGAATATCTGTG 640
GRVER4.SEQ TTIGICCITAAAGGTGT C ATGCA G ACTCACCA G AATAT C TGTG 640
GRVER3.SEQ TTIGICCITIAAAGGTGTICIATGCAIGIACTCACCAIGIAATATICITGTG 640
GRVER2.SEQ CTGCCTAAAAGGCGTGATGCAGACTCACCAAAATATCTGTG 640
GRVER1. SEQ CTGCCTAAAGGCGTGATGCAGACTCACCAAAATATCTGTG 640
YG81-6G1.SEQT TACCGAAAGGTGTAATGCAAACTCACCAAAATATTTGTG 640
RDVER1. SEQ TTGCCAAAGGGTGTCATGCAAACCCATCAGAACATTTGCG 640
RDVER2. SEQ TTGCCAAAGGGTGTCATGCAAACCCATCAGAACATTTGCG 640
RDVER3. SEQ CTCCCAAAGGGCCGTCATGCAGACCCATCAAAACATTTGCG 640
RDVER4. SEQ CT CC CA A A G G G A G T C A T G C A G A C C C A T C A A A A C A T T T G C G 640
RDVERS. SEQ CT CC CA A A G G G A G T C A T G C A G A C C C A T C A A A A C A T T T G C G 640
RD7. SEQ CT CC CA A A G G G A G T CA T G C A G A C C C A T C A A A A C A T T T G C G 640
RDVER51.SEQ CT CC CA A A G G G A G T C A T G C A G A C C C A T C A A A A C A T T T G C G 640
RDVER52.SEQ CTCCCAAAAGGGGAGTCATGCAGGACCCATCAAAAACATTTGCG 640
RD1561H9. SEQUITOCCAA AGGGGAGTOATGCAGACOCATCAAAAOATTTGOG 640
GRVER51.SEQ TIGIC GIT TITIGIA TICIC AICIG C TICITICIGIA C C CIT CIGITIGIT GIG GITIA CITICIA 680
GRE-SEQ TIGIC GIT TITIGIA TICICA CIGICITICIGA CCCIT CIGITIGIT GIG GITIA CITICA 680
GRVERS. SEQ TIGIC GITTITIGIA TICIC A CIGICITICIGIA C C CIT CIGITIGIT GG GITIA CITICIA 680
GRVER4. SEQ TIGIC GITTITIGIA TICIC AICIG C TICITICIGIA C C CIT CIGITIGIT GIG GITIA CITIC A 680
GRVER3. SEQ TEC CETTE GATE CALCECE CETE CEACECTE CETE GET GE GET A CETE A 680
GRVER2.SEQ TCCG CT T G AT CATGC CC T G GACCC AC G T G TG GG T AC T CA 680
GRVER1.SEQ TCCGCTTGATTCATGCCCTGGACCCACGTGTGGGGTACCCA 680
YG81-6G1.5EQT C C G A C T T A T A C A T G C T T T A G A C C C C A G G G C A G G A A C G C A 680
RDVER1. SEQ TGCGTCTGATCCACGCTCTCGTCGATCCTCGCTACGGCACTCA 680
RDVER2. SEQ T G C C G A T C C A C G C T C G A T C C T C G C T A C G G C A C C A 680
RDVER3.SEQ TGCGTCTGATCCATGCTCTCGATCCACGCTACGGCACTCA 680
RDVER4.SEQ TIGIC GITIC TIGIA TICIC A TIGIC TICIG AITIC CA CIGIC TA CIGICIA CITIC A 680
RDVER5. SEQ TIGIC GITIC TIGIA TICIC AT G C TICITICIGAITIC CIA CIGIC T A CIG G ICIA CITIC A 680
RD7. SEQ TGCGTCTGATCCATGCTCTCGATCCACGCTACGGCACTCA 680
RDVER51.SEQ TGCGTCTGATGCTCTCGATCCACGCTACGGCACTCA 680
RDVER52.SEQ TGCGTCTGATCCATGCTCTCGATCCACGCTACGGCACTCA 680
RD1561H9. SEQTEC GTC TEATECT CTCE ATC CACECTACE GCACTCA 680
GRVER51. SEQ A T T G A T C C C T G G C G T G A C T G T G C T G G T G T A T C T G C C T T T C 720
GR6.SEQ ANTIGATICTCTGGCGTGACTGTGCTGGTGTATCTGCCTTTC 720
GRVER5. SEQ A T T G AT C CCTGG C GTGAC T GT GC T G GT G TATCTGCCTTT C 720
GRVER4.SEQ A T T G AT C CCTGG C GTGAC T GT GC T G GT G TAT <u>C</u> TGCCTTT C 720
GRVER3. SEQ ATTGATIC CTGG CGTGACTGTGTGCTGGTATTTGCCCTTTC 720
GRVER2. SEQ GTTGATCCTGGCGTGACTGTCCTGGTGTACTTGCCATTC 720
GRVER1. SEQ GTTGATCCCTGGCGTGACTGTCCTGGTGTACTTGCCATTC 720
YG81-6G1. SEQACTTATTCCTGGTGTGACAGTCTTAGTATATCTGCCTTTT 720
RDVER1. SEQ: ACTGATTCCAGGTGTCACCGTGTTGGTCTATCTGCCTTTT 720
RDVER2. SEQ ACTGATTCCTGGTGTCACCGTGTTGGTTCTATCTGCCTTTT 720
RDVER3.SEQ GCTGATTCCTGGTGTCACCGTCTTGGTCTACCTGCCTTTC 720
RDVER4.SEQ GCTGATTCCTGGTGTCACCGTCTTGGTTCTACTTGCCTTTC 720
RDVER5. SEQ G CT G ATTCCTGGTGT C AC C GTCTT G GT C TA CTTGCCTTT C 720
RD7. SEQ GCTGATTCCTGGTGTCACCGTCTTGGTCTACTTGCCTTTC 720
RDVER51. SEQ G C T G A T T C C T G G T G T C A C C G T C T T G G T C T T G C C T T T C 720
RDVER52. SEQ G C T G A T T C C T G G T G T C A C C G T C T T G G T C T A C T T G C C T T T C 720
RD1561H9. SEQGCTGATTCCTGGTGTCACCGTCTTGGTCTACTTGCCTTTC 720

Figure 2 (cont.)

CPUEDS1 SEQ T TRIC - Clare Color	
GRVER51.SEQ T T T C A C G C C T T T G G T T T C T C T A T T A C C C T G G G C T A T T T C A 760	_
The state of the s	
The state of the s	
The state of the court in the court in the court in the court of the court in the c	
The second of th	
The state of the s	
RD1561H9. SEQT T C C A T G C T T T C G G C T T T C A T A T T A C T T T G G G T T A C T T T A 760	
GRVER51. SEQ T G G T C G G C T T G C G T G T C A T C A T G T T T C G T C G C T T C G A C C A 800	
TO THE SECOND CONTRACTOR OF THE CONTRACTOR OF TH	
The state of the s	
GRVER1.SEQ TGGTCGGTTTGCGCGCGTGATCATGTTTCGTCGCTTCGATCA 800 YG81-6G1.SEQTGGTGGGTCTTCGTTCGTTCGTTCGTTCGATCA 800	
TO THE PROPERTY OF THE PROPERT	
TO THE SECOND SE	
RD7. SEQ T G G T C T C C G C G T G A T T A T G T T C C G C C G T T T T G A T C A 800 RDVER51. SEQ T G G T C T C C G C G T G A T T A T G T T C C G C C G T T T T G A T C A 800	
RDVER51.SEQ T G G T C G G T C T C C G C G T G A T T A T G T T C C G C C G T T T T G A T C A 800 RDVER52.SEQ T G G T C G G T C T C C G C G T G A T T A T G T T C C G C C G T T T T G A T C A 800 RD1561H9.SEO T G G T C G G T C T C C G C G T G A T T A T G T T C C G C C G T T T T G A T C A 800	
RD1561H9.SEQT G G T C G C G C G T G A T T A T G T T C C G C C G T T T T G A T C A 800	
TO SEE THE TENT CE SECRET TT GAT CA 800	
GRVER51. SEQ AGAAGCCTTCTTGAAGGCTATTCAAGGACTACGAGGTGCGT 840	
TO A A GUICLE FICIO TIGIA RIGIG C T A T T C A A C R C M A C C A C A C A C A C A C A C A C A	
TO THE TOTAL OF CHOICE TO THE ANGLE CICIA TOTAL AND A MORE AND A M	
YG81-6G1.SEQAGAAGCATTTCTAAAAGCTATTCAGGATTATGAAGTTCGA 840 RDVER1.SEO GGAGGCCTTCTTTCTAAAAGCTATTCAGGATTATGAAGTTCGA 840	
RDVER1. SEQ GG AGG CCT TC TTGA A A G C T A TCC AAG A T T A T G A A G T G C GC 840	
RDVER2. SEQ GG A GG C T T T C T T G A A A G C T A T C C A A G A T T A T G A A G T G C G C 840	
IN A PORT TO TAKE THE TOTAL A A G CICIA TICIC ALAIG A T T A T C A A C TICIC CICI ALO	
THE TOTAL TO THE TITIES AND CICIA TICIC AIRIG A TITIES AND THE TIT	
TO THE TIME	
The state of the s	
TO LINE SUB BIG AIGIG CITIT TIC TITIGIA A A C CICIA TICIC AIAIC A M T A T A T A T A T A T A T A T A T A	
The state of the s	
RO1561H9. SEQGG AGG CTT TCTTGA A A G CCATCC AAGATTATGA AGTCC GC 840	

Figure 2 (cont.)

GRESSED T C C G T G A T C A C G T C C T T C A G T C C T T T C A G T C A T T T T G T C C T G A G C A 880 GRESSED T C C G T G A T C A A C G T C C T T C A G T C A T T T T G T T C C T G A G C A 880 GRESSES T C C G T G A T C A A C G T C C T T C A G T C A T T T T G T T C C T G A G C A 880 GRESSES T C C G T G A T C A A C G T C C T T C A G T C A T T T T G T T C C T G A C C A 880 GREENS SED T C T G T G A T C A A T G T C C T T T C A G T C A T T T T G T T C C T G A C C A 880 GREENS SED A G C G T G A T C A A T G T C A T T C T T G T G A T T T T T G T T C C T G A C C A 880 GREENS SED A G C G T G A T C A A T G T C A T T C T T G T G A T T T T T G T T C C T G A C C A 880 GREENS SED A G C G T G A T C A A T G T C A T T C T T G T G A T T T T T G T T C C T G A C C A 880 GREENS SED A G C G T G A T C A A T T A C G T C C T T C T G T G A T T T T T G T T C C T G A C C A 880 GREENS SED A T C T A A T T A A C G T T C C A T C A G T A A T T T T G T T C T T T C C T G A C C A 880 GREENS SED A T C T A A T T A A T T A C G T T C C A T C A G T A A T T T T G T T C T T A T C G A S 880 GREENS SED A T C T A A T T A A T T A C G T T C C A T C A G T A A T T T G T T C T T A T C G A C C A 880 GREENS SED A T C T A A T T A A T T A C T T G C A A G C G T C A T C C T T T T T T G T T C T T T A T C G T C A T C A A G C T C A T C C T G T T T T T T T G T T C T T C T T C A G T C A T C T C	Ŭ., .	•	
GRVERAS.SEQ	GRVER51.SEQ	CCGTGATCAACGTCCCTTCAGTCATTTTGTTCCTGAGCA 880	
GRVERA: SEQ T C T T C T T C A A T T G T C C T T C A G T C A T T T T G T T C C T A G C A 880 GRVERA: SEQ A G C T G A T C A A T G T C C T T C C T T T G G T T C C T A G C A 880 GRVERA: SEQ A G C T G A T C A A C G T C C T T C T C T G T G A T T T T T G T T C C T A G C A 880 GRVERA: SEQ A G C T G A T C A A C G T C C C T T C T G G T T T T T G T T C C T A G C A 880 GRVERA: SEQ A G C G T G A T C A T T A A C G T C C C T T C T G T G A T T T T T G T T C C T A G C A 880 GRVERA: SEQ A G C T C A T T A A T T A C G C C A A G C A T T T T T T T T T T T T T T T T T T	GR6.SEQ	CCGTGATCAACGTCCCTTCAGTCATTTTTTCTCTTGAGCA 880	
GRVERA: SEQ	GRVERS.SEQ	CEGTIGIATICIA A C GTICIC CITIT C A GTICIA TITIT T G T T C C ITIG A G C IA 880	
GRVERA: SEQ A GLG TIGATICA A C G TIC CITT CITG TIGATIT T G TT CITTG A G C A 880 (ROVERA: SEQ) A GLG TIGATICA TO THE G TIC CAT TO	GRVER4.SEQ	CIT G T C A T C A A T G T C C C T T C A G T C A T T T T G T T C C T G A G C A 880	
GRVERA: SEQ A GLG TIGATICA A C G TIC CITT CITG TIGATIT T G TT CITTG A G C A 880 (ROVERA: SEQ) A GLG TIGATICA TO THE G TIC CAT TO	GRVER3.SEQ	<u>CIT</u> G TIGIA TICIA A TIG TICIC CAT CITIG TICIA TITIT TIGIT TICITIG A GICIA 880	
GRVER1.SED A GEG TEGATIGA A C G TECC CITT CIT G TEGATIGAT TOTT T COTTE A G A 880 ROUVER1.SED AGT GT A A ATA TA T G T T C C T T C T T A T C G A 880 ROUVER1.SED TC T G T G A T A A T A T G T G C C A A G C G T G A T C A T T A A T G A 880 ROUVER1.SED TC T G T G T G A T A A T G T G C C A A G C G T G A T C C T G T T T T T T G T C T A 880 ROUVER1.SED TC T G T G A T A A T G T G C C A G C G T G A T C C T G T T T T T T G T C T A 880 ROUVER3.SED A G C G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROUVERS.SED A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROUVERS.SED A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROUVERS.SED A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROUVERS.SED A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROUVERS.SED A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROUVERS.SED A G T G T C A T C A C A G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROUVERS.SED A G T G T C A T C A C A G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROUVERS.SED A G T G T C A T C A C C G T G C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROUVERS.SED A G T G T C A T C A C C A G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROUVERS.SED A G T G T C A T C A C C A G T G C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROUVERS.SED A A T C C C T T T T T G T T G A C A A G T A G C A G C T G C T T T T T T T G T T C T A G C G T G C T A G C G T G C T A G C A G C T T G C T G T T T T T T T G T T C T A G C G T G C T A G C A G C A G C T T G C T G T G C C T A G C G T G C T C T T T T T T T G T T C C T T T G T T G T C C T T T G T T T T	GRVER2.SEQ	. G C G T G A T C A A C G T C C C T T C T G T G A T T T T G T T C C T G A G C A 880	
ROVERAL SED T C T C T T A A T A A T A T G C C A A G C G T C A T C C T T T T T T C T A 880 ROVERAL SED T C T C T C A T T A A T G T G C C A A G C G T C A T C C T G T T T T T G T C T A 880 ROVERAL SED A G G G T C A T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T G T C T A 880 ROVERAL SED A G G G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROVERAL SED A G T G T C A T C A A C G T G C C T A G C G T G A T C T G T T T T T T T C T C T A 880 ROVERAL SED A G T G T C A T C A A C G T G C C T A G C G T G A T C T G T T T T T T T T C T C T A 880 ROVERAL SED A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROVERAL SED A G T G T C A T C A A C G T G C C T A G C G T G A T C C T T T T T T T G T C T A 880 ROVERAL SED A G T G T C A T C A A C G T G C T A G C G T G A T C C T T T T T T T G T C T A 880 ROVERAL SED A G T G T C A T C A A C G T G C T A G C G T G A T C T T T T T T T T T T T T T T T T T	GRVER1. SEQ	GCGTGATCAACGTCCCTTCTGTGATTTTTTGTTCCTGAGCA 880	
ROVERAL SED T C T C T T A A T A A T A T G C C A A G C G T C A T C C T T T T T T C T A 880 ROVERAL SED T C T C T C A T T A A T G T G C C A A G C G T C A T C C T G T T T T T G T C T A 880 ROVERAL SED A G G G T C A T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T G T C T A 880 ROVERAL SED A G G G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROVERAL SED A G T G T C A T C A A C G T G C C T A G C G T G A T C T G T T T T T T T C T C T A 880 ROVERAL SED A G T G T C A T C A A C G T G C C T A G C G T G A T C T G T T T T T T T T C T C T A 880 ROVERAL SED A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROVERAL SED A G T G T C A T C A A C G T G C C T A G C G T G A T C C T T T T T T T G T C T A 880 ROVERAL SED A G T G T C A T C A A C G T G C T A G C G T G A T C C T T T T T T T G T C T A 880 ROVERAL SED A G T G T C A T C A A C G T G C T A G C G T G A T C T T T T T T T T T T T T T T T T T	YG81-6G1.SEQ	GTGTAATTAACGTTCCATCAGTAATATTGTTCTTATCGA 880	
ROVERS. SEQ A GE G T C A T T A A T G T G C C A A G C G T C A T C C T G T T T T G T C T A 880 ROVERS. SEQ A GE G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T G T C T A 880 ROVERS. SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T G T C T A 880 ROVERS. SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T G T C T A 880 ROVERS. SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROVERS. SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROVERS. SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROVERS. SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROVERS. SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROVERS. SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROVERS. SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROVERS. SEQ A T C C C T T T G G T T G A C A A G T A T G A T C T G A C A G C T T G C G 920 GRVERS. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERA. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C G A G C T T G C G 920 GRVERA. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C G A G C T T G C G 920 GRVERA. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C G A G C T T G C G 920 GRVERA. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C G A C A G C T T G C G 920 GRVERA. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C G A C A G C T T G C G 920 GRVERA. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A C T G A C A G C T T G C G 920 GRVERA. SEQ A A T C T C C T T T G G T T G A C A A G T A T T G A C C A G C T T G C G 920 GRVERA. SEQ A A T C T C C T T T G G T T G A C A A G T A T T G A C C T T T T T G C G 920 GRVERA. SEQ A T C T C C T T T G G T T G A C A A G T A T T G	RDVER1.SEQ	CTGTCATTAATGTGCCAAGCGTCATCCTGTTTTTTGTCTA 880	
ROVERS. SEQ A GIGST CATTA ACGTIGC CTAGCGTGATC CTGTTTTT GTCTT B 880 ROVERS. SEQ A GTGT CATCA ACGTIGC CTAGCGTGATC CTGTTTTTTGTCTA 880 ROVERS. SEQ A GTGTCATCA TCAACGTGCCTAGCGTGATCCTGTTTTTTGTCTA 880 ROVERS. SEQ A GTGTCATCATCACGTGCCTAGCGTGATCCTGTTTTTTGTCTA 880 ROVERS. SEQ AGTGTCATCATCACGTGCCTAGCGTGATCCTGTTTTTTGTCTA 880 ROVERS. SEQ AGTGTCATCATCACGTGCCTAGCGTGATCCTGTTTTTTGTCTTA 880 ROVERS. SEQ AGTGTCCTTTTGGTTGACAAGTATTGATCTGAGCAGCTTTGCGG 920 ROVERS. SEQ AATCCTCCTTTGGTTGACAAGTATTGATCTGAGCAGCTTGCGG 920 ROVERS. SEQ AATCCTCCTTTTGGTTGACAAGTATTGATCTGAGCAGCTTTGCGG 920 ROVERS. SEQ AATCCTCCTTTGGTTGACAAGTATTGATCTGAGCAGCTTTGCGG 920 ROVERS. SEQ AATCCTCCTTTGGTTGACAAGTATTGATCTGAGCAGCTTTGCGG 920 ROVERS. SEQ AATCCTCCTTTGGTTGACAAGTATTGATCTGAGCAGCTTTGCGG 920 ROVERS. SEQ AATCCTCCTTTGGTTGACAAGTATTGATCTTGAGCAGCTTTGCGG 920 ROVERS. SEQ AATCCTCCTTTGGTTGACAAGTATTGATCTTGAGCTAGCGGTTTGCGG 920 ROVERS. SEQ AATCCTCCATTTGGTTGACAAAGTATTGATCTTGAGCTAGCGTTTGCGG 920 ROVERS. SEQ AATCCTCCATTTGGTTGACAAAGTATTGATCTTTATCAAGTTTTAAG 920 ROVERS. SEQ AGGGCTCCACTCTGTGGTGGACAGCTTTGCGG 920 ROVERS. SEQ AGGGCCCCTCCTGTGGTGACAAGTATTGACCTTTTATCAAGTTTAAG 920 ROVERS. SEQ AGGGCCCCTCCTGTGGGACAAGTACGACTTTGTCTTTTATCAAGTTTAAG 920 ROVERS. SEQ AGGGCCCCACTCTGTGGGACAAGTACGACTTTGTCTTCTCTTCTGCGG 920 ROVERS. SEQ AGGGCCCCACTCCTGTGGACAAGTACGACTTTGTCTTTCACTGCG 920 ROVERS. SEQ AGGGCCCCACTCCTGTGGACAAGTACGACTTTGTCTTTCACTGCG 920 ROVERS. SEQ AGGGCCCACTCCTTCTGTGGACAAGTACGACTTTGTCTTTCACTGCG 920 ROVERS. SEQ AGGGCCCACATGCCCTTTTGGACCAAGTACGACTTTGTCTTTCACTGCG 920 ROVERS. SEQ AGGGCCCACATGCCCCCCCCCCCCCCCCCCCCCCCCCCTTTTGGCCGAAAGAAGTACTG 960 ROVERS. SEQ TGAACTGTGTGGCGCGCCCCCCCCCCCCCCCCCCCCCCC	RDVER2.SEQ	CIT G TICIA T T A AITIG TIGIC C AIA G CIG TICIA TIC CIT G T TITIT TIGIT CITIA 880	
ROVERS.SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T G T C T A 880 ROY-SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROY-SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROY-SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROY-SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROY-SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T T G T C T A 880 ROY-SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T T T T T T T T T T	RDVER3.SEQ .	GEGTICATTAACGTIGCCITAGCIGTIGATICCITGTTITTIGTCITA 880	
ROY-RSS. SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROY-RSS. SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROY-RSS. SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROY-RSS. SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROY-RSS. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G T T T T T T T G T C T T A 880 ROY-RSS. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 ROY-RSS. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 ROY-RSS. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 ROY-RSS. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 ROY-RSS. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 ROY-RSS. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 ROY-RSS. SEQ A A T C T C C C A T T G G T C G A C A A G T A T G A T C T G A G C A G C T T G C G 920 ROY-RSS. SEQ A A T C T C C C A T T G G T C G A C A A G T A T G A T C T G A G C A G C T T G C G 920 ROY-RSS. SEQ A A T C T C C C A T T G G T C G A C A A G T A T G A C C T G A G C A G C T T T G C 920 ROY-RSS. SEQ A G G C C T C C C A C T C G T G G G C A A A T A C G A T T T T T T C C T G C 920 ROY-RSS. SEQ A G G C C C C C C C C T T T G G T G G A C A A A T A C G A T T T T T C C T G C G 920 ROY-RSS. SEQ A G G G C C C C C C C T C T G T G G A C A A A T A C G A C T T G T C T C G C 920 ROY-RSS. SEQ A G G G C C C C C C C T C T G T G G A C A A G T A C G A C T T G T C T C T C G C 920 ROY-RSS. SEQ A G G G C C C A C T C C T G T G G G C A A G T A C G A C T T G T C T T C C T G C G 920 ROY-RSS. SEQ A G G G C C C A C T C C T G T G G C C A A G T A C G A C T T G T T T T C C T T C G C 920 ROY-RSS. SEQ A G G G C C C A C T C C T G T G G C C C T	RDVER4.SEQ .	GTGT C AT C AACGT G CC TAGC GT G AT CC TGTT TT G TC T A 880	
ROY-SEQ A GT GT C A T C A A C GT G C C T A G C G T G A T C C T T G T C T T T T T G T C T A 880 ROVERSI. SEQ A GT GT C A T C C A C G T G C C T A G C G T G A T C C T G T T T T T G T C T A 880 ROVERSI. SEQ A GT GT C A T C A A C GT G C C T A G C G T G A T C C T G T T T T T T G T C T A 880 ROVERSI. SEQ A GT G T C A T C A T C A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T T A 880 ROVERSI. SEQ A GT G T C A T C A T C A C G T G C C T A G C G T G A T C C T G T T T T T T G T C T T A 880 GRVERSI. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 VOSI - 601. SEQ A A G T C C C T T T G G T T G A C A A G T A T G A T T T A T C A A G T T T A A G 920 ROVERS. SEQ A G T G C C T C T G G T G G A C A A A T A C G A T T T A T C A A G T T T A A G 920 ROVERS. SEQ A G G C C C T C T G G T G G A C A A A T A C G A T T T G T C C T T G C G 920 ROVERS. SEQ A G A G C C C T C T G G T G G A C A A A T A C G A T T T T T T C C T T C C T G C 920 ROVERS. SEQ A G A G C C C C T C T G G T G G A C A A G T A C G A C T T G T C T T C C C T G C G 920 ROVERS. SEQ A G A G C C C C T C T G T G G A C A A G T A C G A C T T G T T T T T C A C T G C G 920 ROVERS. SEQ A G A G C C C C C C T C T G T G G C C T G C T G C C C C	RDVERS.SEQ .	GTGTCATCAACGTGCCTAGCGTGATCCTGTTTTTTTCTA880	
ROVERS1. SEQ A GT GT C A T C A A C GT G C C T A G C G T G A T C C T T T T T T GT C T A 880 ROVERS2. SEQ A GT GT C A T C A A C GT G C C T A G C G T G A T C C T G T T T T T GT C T A 880 ROTS61H9. SEQ A GT GT C A T C A A C GT G C C T A G C G T G A T C C T G T T T T T GT C T A 880 ROTS61H9. SEQ A GT GT C A T C A A C GT G C C T A G C G T G A T C C T G T T T T T GT C T A 880 GRVERS1. SEQ A A T C T C C T T T G G T T G A C A A GT A T G A T C T G A G A G C T T G C G 920 GRVERS4. SEQ A A T C T C C T T T G G T T G A C A A GT A T G A T C T G A G C A G C T T G C G 920 GRVERS4. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS4. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS4. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS4. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS4. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS4. SEQ A A T C T C C A T T G G T C G A G A A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS4. SEQ A A T C T C C A T T G G T C G A G A A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS4. SEQ A A T C T C C A T T G G T C G A T A A G A T A T T A T C G A G C A G C T T G C G 920 GRVERS4. SEQ A A T C T C C T T T G G T T G A C A A G T A T A G A T C G A T T T T T T C T G C G 920 GRVERS5. SEQ A G G C C C T C T G G T G G A C A A A T A C G A T T T A T C A A G T T T A A G 920 RDVERS1. SEQ A G G G C C T C T G G T G G A C A A A T A C G A T T T T T T C A C T G C G 920 RDVERS5. SEQ A G G G C C C T C T G T G G A C A A G T A C G A C T T T T T T T T T C T C T G C G 920 RDVERS5. SEQ A G G G C C C C T C G T G G A C A A G T A C G A C T T T T T T T C A C T G C G 920 RDVERS5. SEQ A G G G C C C C C T C G T G G A C A A G T A C G A C T T T T T T T C A C T G C G 920 RDVERS5. SEQ A G G G C C C C C T C G T G G C C C T C C T T T T	RD7.SEQ	GTGT C AT C AACGT G CC TAGC GT G AT CC TGTT TTT G TC T A 880	
ROVERS1.SEQ A GT GT C A T C A A C GT G C C T A G C GT G A T C C T G T T T T T T G T C T A B80 ROTS61H9.SEQA GT GT C A T C A C GT G C C T A G C GT G A T C C T G T T T T T T G T C T A B80 GRVERS1.SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS.SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS.SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS.SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS.SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS.SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS.SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS.SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS.SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERS.SEQ A A T C T C C T T T G G T T G A C A A G T A T G A C C T G A G C A G C T T G C G 920 GRVERS.SEQ A G T T C C C T T T G G T T G A C A A G T A T T G A C C T G A G C C T G T G G G G C C T G T G G T G A C A A A T A C G A T T T A T C A A G T T T A C C A G 920 YG81-6G1.SEQA A A G T C C T T T G G T T G A C A A A T A C G A T T T T A T C A A G T T T A C G G 920 ROVERS.SEQ A G G A G C C C C T C T G T G G G A C A A A T A C G A T T T T G T C T T T T T T A C G G 920 ROVERS.SEQ A G A G C C C C A C T C G T G G A C A A A G T A C G A C T T T T T T C T C C T G C G 920 ROVERS.SEQ A G A G C C C C A C T C G T G G A C A A G T A C G A C T T G T T T T C A C T G C G 920 ROVERS.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T T T T C A C T G C G 920 ROVERS.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T T G T T C A C T G C G 920 ROVERS.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T T T T C A C T G C G 920 ROVERS.SEQ A G A C T G T G C T G T G G	RDVER51.SEQ	GTGTCATCAACGTGCCTAGCGTGATCCTGTTTTTTTTTT	
ROTSCHIPS. SEQ A G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T G T C T A 880 GRVER51. SEQ A A T C T C C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVER5. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVER5. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVER5. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVER5. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVER5. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVER5. SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVER1. SEQ A A T C T C C A T T G G T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVER1. SEQ A A T C T C C A T T G G T G A T A A G T A T G A T C T G A G C A G C T T G C G 920 GRVER1. SEQ A A T C C T T T G G T T G A C A A A T A C G A T T T A T C A A G T T T G C G 920 ROVER1. SEQ A A G T C C T T T G G T T G A C A A A T A C G A T T T A T C A A G T T T G C G 920 ROVER1. SEQ A G G C C C T C T G G T G G A C A A A T A C G A T T T G T C T T G C G 920 ROVER1. SEQ A G G C C C T C T G G T G G A C A A A T A C G A T T T G T C T T C T G C G 920 ROVER1. SEQ A G A G C C C T C T G G T G G A C A A G T A C G A C T T T T T T C C C T G C G 920 ROVERS. SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T T G T C T T C A C T G C G 920 ROVERS. SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T T T C T T C A C T G C G 920 ROVERS. SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T T C T T C A C T G C G 920 ROVERS. SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T T C T T C A C T G C G 920 ROVERS. SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T T C T T C A C T G C G 920 ROVERS. SEQ A G A G C C C A C T C G T G G C C C T C C T T T G G C C A A A G A A G A A G A G T G G C C C A C T T C G	RDVER52.SEQ .	GTGT C AT C AACGT G CC TAGC GT G AT CC TGTT TT G TC T A 880	
GRVER51.SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVER5.SEQ A A T C T C C T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVER5.SEQ A A T C T C C T T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVER1.SEQ A A T C T C C T T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVER1.SEQ A A T C T C C T T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVER1.SEQ A A T C T C C A T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVER1.SEQ A A T C T C C A T T G G T T G A C A A G T A T G A T C T C A G C A G C T T G C G 920 GRVER1.SEQ A A T C T C C A T T G G T G A C A A G T A T G A T C T C A G C A G C T T G C G 920 GRVER1.SEQ A A T C T C C A T T G G T G A C A A A T A C G A T T T A T C A A G T T T A G G 920 ROVER3.SEQ A A T C T C C A T T G G T G G A C A A A T A C G A T T T A T C A A G T T T A G G 920 ROVER3.SEQ A G G C C C T C T G G T G G A C A A A T A C G A T T T A T C A A G T T T A G G 920 ROVER5.SEQ A G A G C C C T C T G G T G G A C A A A T A C G A T T T T G T C T T C T G C G 920 ROVER5.SEQ A G A G C C C A C T C G T G G A C A A A T A C G A T T T T G T C T T C C T G C G 920 ROVER5.SEQ A G A G C C C A C T C G T G G A C A A A T A C G A T T T T G T C T T C C T G C G 920 ROVER5.SEQ A G A G C C C A C T C G T G G A C A A A T A C G A C T T G T T T C A C T G C G 920 ROVER5.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T T T C A C T G C G 920 ROVER5.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T T C T T C A C T G C G 920 ROVER5.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T T C T T C A C T G C G 920 ROVER5.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T T C T T C A C T G C G 920 ROVER5.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T T C T T C A C T G C G 920 ROVER5.SEQ A G A C T G T G C T G T G C T G T G C T G C T G C T G C T C C T T T G G C C A A A G A A G A G T G G C C A C T T	RD1561H9.SEQ	GTGTCATCAACGTGCCTAGCGTGATCCTGTTTTTTGTCTTA 880	
GRVERS. SEQ			
GRVERS. SEQ	GRVER51.SEQ	ATCTCCTTTGGTTGACAAGTATGATCTGAGCAGCTTGCC 920	
GRVERAS. SEQ A AT CIT C C TIT T G G T T G A C A A G T A T G A C T T G A G C A A G C T T G C G S S S C REVERAS. SEQ A A T CIT C C T T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G S S S C R G C T T G C G S S S C R G C T T G C G S S S C R G C T T G C G S S S C R G C T T G C G S S S C R G C T T G C G S S C R G C T T G C G S S C R G C T T G C G S S C R G C T T G C G S S C R G C T T G C G S S C R G C T T G C G S S C R G C T T G C G S S C R G C T T G C G S S C R G C T T G C G S S C R G C T T G C G S S C R G C T T G C G S S C R G C T T T G G T C G A T A A G T A T G A C C T G A G C T T T T G C G S S C R G C C T G C T G C T G C G C T G C G C T G C G C	GR6.SEQ	AIT CIT C C T T T G G T T G A C A AIGIT AITIG A TICITIG A G CIA GICIT TIG CIG 920	
GRVERAL SEQ A AT C T C C T T T G G T T G A C A A G T A T G A T C T G A C C T G C G 920 GRVERAL SEQ A A T C T C C T T T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERAL SEQ A A T C T C C A T T G G T T G A C A A G T A T G A T C T G A G C A G C T T G C G 920 GRVERAL SEQ A A A T C T C C A T T G G T C G A T A A G T A T T A T C A A G T T T T G C G 920 GRVERAL SEQ A A A T C T C C A T T G G T C G A T A A G T A T T A T C A A G T T T T G C G 920 ROVERAL SEQ A A A G T C C T T T G G T T G A C A A A T A C G A T T T A T C A A G T T T A G 920 ROVERAL SEQ A G A G C C C T C T G G T G G A C A A A T A C G A T T T A T C A A G T T T A G 920 ROVERAL SEQ A G A G C C C T C T G G T G G A C A A A T A C G A T T T A T C A A G T T C T G C G 920 ROVERAL SEQ A G A G C C C A C T C T G G T G G A C A A A T A C G A T T T G T C T T G C G 920 ROVERAL SEQ A G A G C C C A C T C T G G T G G A C A A G T A C G A C T T G T C T T C T C C T G C G 920 ROVERAL SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 ROVERAL SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 ROVERAL SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 ROVERAL SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 ROVERAL SEQ A G A G C C C A C T C C T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 ROVERAL SEQ A G A G C C C A C T C C T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 ROVERAL SEQ A G A G C C C A C T C C T G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 ROVERAL SEQ A G A G C C C A C T C C T G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 ROVERAL SEQ A G A G C C C A C T C C T G T G G C C T G C T C C T T T G G C C A A A G A A G T G G G C G C T G C T C C T T T G G C C A A A G A A G T G G G C G C T G C T C C T T T G G C C A A A G A A G T G G G G C G C T G C T C C T T T G G C C A A A G A A G T G G G G C G C T G C T C C C T T T T	GRVER5.SEQ	ATCITCCTTTGGTTGACAAGTATGATCTGAGCAGCTTGCG 920	
GRVER2.SEQ A A T C T C C T T T G G T T G A C A A G T A T G A C C T G A G C T T G C G 920 GRVER1.SEQ A A T C T C C A T T G G T C G A T A A G T A T G A C C T G A G C T T T G C G 920 GRVER1.SEQ A A T C T C C A T T G G T C G A T A A G T A T G A C C T G A G C T T T T G C G 920 GRVER1.SEQ A G G C C C T C T G G T G G A C A A A T A C G A T T T A T C A A G T T T A A G 920 RDVER3.SEQ A G G C C C T C T G G T G G A C A A A T A C G A T T T G T C T T T C C C G 920 RDVER3.SEQ A G G C C C C C T C T G G G G A C A A A T A C G A T T T G T C T T C T C T G G 920 RDVER3.SEQ A G G C C C C A C T C G T G G A C A A G T A C G A T T T G T C T T C A C T G C G 920 RDVER3.SEQ A G G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER3.SEQ A G G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G G C C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ T G A G A G C C C A C T C G T G G C C T C C T T T G G C C A A A G A G T G G C C A C T C G T G G C C A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ T G A G C C C A C T C G T G G C C C C T C C T T T G G C C A A A G A A G T G 960 GRVER3.SEQ T G A A C T G T G C T G T G C C C C T T T G G C C A A A G A A G T G 960 GRVER3.SEQ T G A A C T G T G C T G T G C C C C T T T G G C C A A A G A A G T G 960 GRVER3.SEQ T G A A C T G T G C T G T G C C C C T C C T T T G G C C A A G G A G T G 960 RDVER3.SEQ T G A A C T G T G C T G T G	GRVER4.SEQ	AIT CIT C C T T T G G T T G A C A AIGIT AITIG A TICITIG A G CIA GICIT TIC CIC 920	
GRVERI.SEQ A AIT C T C C AIT T G G T C G A T A A G T A T G A C C T G A G C A G C T T G C G 920 YG81-6G1.SEQ A A T C T C C AT T G G T C G A T A A G T A T G A C C T G A G C T C T T T G G G 920 RDVERI.SEQ A A G T C C T T T G G T C G A C A A A T A C G A T T T A T C A A G T T T A G C G 920 RDVERZ.SEQ A G A G C C C T C T G G T G G A C A A A T A C G A T T T G T C T T T C T C T G G G 920 RDVERZ.SEQ A G A G C C C T C T G G T G G A C A A A T A C G A T T T G T C T T C T C T G G 920 RDVERZ.SEQ A G A G C C C T C T G G T G G A C A A A T A C G A T T T G T C T T C T C T G C G 920 RDVERS.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C T C T G C G 920 RDVERS.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVERS.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVERS.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVERS.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVERS.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVERS.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVERS.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVERS.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVERS.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T T G T C T T C A C T G C G 920 GRVERS.SEQ A G A G C C C A C T C G T G G G C C G C T C C T T T G G C C A A A G T A G G A C T T G T C T T C A C T G C G 920 GRVERS.SEQ T G A A C T G T G T G C T G T G G C C T C C T T T G G C C A A A G A G T G 960 GRVERS.SEQ T G A A C T G T G C T G T G T G C C G C T C C T T T G G C C A A A G A G T G 960 GRVERS.SEQ T G A A C T G T G C T G T G G C G C T C C T C T	GRVER3.SEQ	AIT CIT C C T T T G G T T G A C A AIGIT AITIG A TICITIG A G CIA GICIT TIG CIG 920	
GRVERS1.SEQ A A GT C C C A C T C G G T G G A C A A A T A C G A T T T A T C A A G T T T A A G G 920 RDVER1.SEQ A A A GT C C T T T G G T G G A C A A A T A C G A T T T A T C A A G T T T A A G 920 RDVER3.SEQ A G A G C C C T C T G G T G G A C A A A T A C G A T T T G T C T T C C T G C 920 RDVER3.SEQ A G A G C C C A C T C G T G G A C A A A T A C G A T T T G T C T T C C C T G C 920 RDVER3.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C 920 RDVER5.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C 920 RDVER5.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C 920 RDVER5.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C 920 RDVER51.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G A G C C C A C T C G T G G G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER51.SEQ T G A G C T G T G C T G T G G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER3.SEQ T G A A C T G T G C T G T G G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER3.SEQ T G A A C T G T G C T G T G G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER3.SEQ T G A A C T G T G C T G T G G C C C C C T T T G G C C A A A G A A G T G 960 GRVER3.SEQ T G A A C T G T G C T G T G G C C C C C T T T G G C C A A A G A A G T G 960 GRVER3.SEQ T G A A C T G T G C T G T G C G C T C C C T T T G G C C A A A G A A G T C 960 RDVER3.SEQ T G A A T T G T G T T G C G G T G C C C C C C	GRVER2.SEQ	AT CIT C CAT T G G T C G ATTA A G T A T G A C C T G A G C A G C T T G C G 920	
YGG1-6G1.SEQ A A G T C C T T T G G T T G A C A A A T A C G A T T T A T C A A G T T T A A G 920 RDVER1.SEQ A G A G C C C T C T G G T G G A C A A A T A C G A T T T G T C T G C T G C G 920 RDVER2.SEQ A G A G C C C T C T G G T G G A C A A G T A C G A T T T G T C T T C C C T G C G 920 RDVER4.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER4.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER5.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER5.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51.SEQ A G A G C C C A C T C G T G G C A C A G T A C G A C T T G T C T T C A C T G C G 920 GRVER51.SEQ T G A G C C C A C T C G T G G C C T G C T C T T T G T C T T C A C T G C G 920 GRVER51.SEQ T G A G C C T G T G C T G T G G C C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER51.SEQ T G A G C T G T G C T G T G G C C T G C T C C T T T T	GRVER1.SEQ	AT CIT C CALT T G G T C G A T A A G T A T G A C C T G A G C T CIT T T G C G 920	
RDVER2.SEQ	YG81-6G1.SEQ	AAGTCCTTTGGTTGACAAATACGATTTATCAAGTTTAAG 920	
ROVER3.SEQ	RDVER1.SEQ	GAGCCCTCTGGTGGACAAATACGATTTGTCTAGCCTGCGG20	
RDVER3. SEQ	RDVER2.SEQ	GAGCCCTCTGGTGGACAAATACGATTTGTCTTCTCTGTGCGG920	
RDVER4. SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER5. SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51. SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51. SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER52. SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER52. SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER51. SEQ T G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 GRVER51. SEQ T G A G C T G T G C T G T G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER5. SEQ T G A G C T G T G C T G T G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER3. SEQ T G A G C T G T G C T G T G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER3. SEQ T G A G C T G T G C T G T G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER3. SEQ T G A A C T G T G C T G T G G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER1. SEQ C G A A C T G T G C T G T G G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER1. SEQ C G A A C T G T G C T G T G G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER1. SEQ C G A A C T G T G C T G T G G C G C T G C T C C T T T G G C C A A A G A A G T G 960 RDVER1. SEQ T G A G T T G T G C G G T G C C C C T T T G G C C A A A G A A G T T 960 RDVER2. SEQ T G A G T T G T G T G C G G T G C C C C T T T G G C C A A G G A G T C 960 RDVER3. SEQ T G A G T T G T G T G C G G T G C C C C T T T G G C C A A G G A G T C 960 RDVER3. SEQ T G A A T T G T G T T G C G G T G C C C C T C C A C T G G C T A A G G A G T C 960 RDVER3. SEQ T G A A T T G T G T T G C G G T G C C C C C T C C A C T G G C T A A G G A G T C 960 RDVER5. SEQ T G A A T T G T G T T G C G G T G C C C C C C T C C A C T G G C T A A G G A G T C 960 RDVER5. SEQ T G A A T T G T G T T G C G G T G C C C C C C C	RDVER3.SEQ 2	GA GCC CACTEG TGG A C A AGT A C G ACT TGT CTT C C CTG CG 920	
RDVERS. SEQ A GAGCCCACTCGTGGACAAGTACGACTTGTCTTCACTGCG920 RD7.SEQ A GAGCCCACTCGTGGACAAGTACGACTTGTCTTCACTGCG920 RDVERS1.SEQ A GAGCCCACTCGTGGACAAGTACGACTTGTCTTCACTGCG920 RDVERS2.SEQ A GAGCCCACTCGTGGACAAGTACGACTTGTCTTCACTGCG920 RDVERS2.SEQ A GAGCCCACTCGTGGACAAGTACGACTTGTCTTCACTGCG920 RD1561H9.SEQAGAGCCCACTCGTGGACAAGTACGACTTTGTCTTCACTGCG920 RD1561H9.SEQAGAGCCCACTCGTGGCCGCTGCTTTGTCTTCACTGCG920 RRVERS1.SEQ TGAGCTGTGTGGCGGCGCTGCTTTTGTCTTTAACTGCG920 RRVERS1.SEQ TGAGCTGTGTGGCGGCTGCTCCTTTGGCCAAAGAAGTAGG960 RRVERS.SEQ TGAGCTGTGTGGCGGCTGCTCCTTTTGGCCAAAGAAGTG960 RRVERS.SEQ TGAACTGTGTGGCGGCTGCTCCTTTTGGCCAAAGAAGTG960 RRVERS.SEQ TGAACTGTGTGGCGGCTGCTCCTTTTGGCCAAAAGAAGTG960 RRVERS.SEQ TGAACTGTGTGCTGTTGGCGCCCTTTTGGCCAAAAGAAGTG960 RRVERS.SEQ TGAACTGTGTGCTGTTGGCGCCCTTTTGGCCAAAAGAAGTG960 RRVERS.SEQ TGAACTGTGTGCTGTTGGCGCCCTTTTGGCCAAAAGAAGTG960 RRVERS.SEQ TGAACTGTGTGCTGTTGGCGCCCTTTTGGCCAAAAGAAGTG960 RRVERS.SEQ TGAACTGTGTGCTGTTGGCGCCCTTTTGGCCAAAAGAAGTG960 RRVERS.SEQ TGAACTGTGTGCTGTTGCCCCTTTTGGCCAAAAGAAGTG960 RRVERS.SEQ TGAACTGTGTGCGGCCCCCTTTTGGCCAAAAGAAGTC960 RRVERS.SEQ TGAACTGTGTGCGGCGCTCCACTGGCTAAAGGAAGTC960 RRVERS.SEQ TGAACTTGTGTGTGCGGCGTTCCACTTGGCCAAAGGAAGTC960 RRVERS.SEQ TGAACTTGTGTGTGCGGCGCTCCACTGGCTAAAGGAAGTC960 RRVERS.SEQ TGAACTTGTGTGTGCGGTGCCCCCTTCCACTGGCTAAAGGAAGTC960 RRVERS.SEQ TGAATTGTGTGTGCGGTGCCGCTCCACTGGCTAAAGGAAGTC960 RRVERS.SEQ TGAATTGTGTGTGCGGTGCCCGCTCCACTGGCTAAAGGAAGTC960 RRVERS.SEQ TGAATTGTGTGTTGCGGTTGCCCGCTCCACTGGCTAAAGGAAGTC960 RRVERS.SEQ TGAATTGTGTGTGCGGTGCCCGCTCCACTGGCTAAAGGAAGTC960 RRVERS.SEQ TGAATTGTGTGTGCGGTGCCCGCTCCACTGGCTAAAGGAAGTC960 RRVERS.SEQ TGAATTGTGTGTGTGCGGTGCCCCCCTTCCACTGGCTAAAGGAAGTC960 RRVERS.SEQ TGAATTGTGTGTGCGGTGCCCCCCTCCACTGGCTAAAGGAAGTC960	RDVER4.SEQ	GA GCC CA CTICG TIGGA CAAGTA CGAICTTIGT CITT CA CTIG CG 920	
ROY-SEQ A GAGCCCACTCGTGGACAAGTACGACTTGTCTTCACTGCG 920 ROVER51.SEQ A GAGCCCACTCGTGGACAAGTACGACTTGTCTTCACTGCG 920 ROVER52.SEQAGAGCCCACTCGTGGACAAGTACGACTTGTCTTCACTGCG 920 ROVER51.SEQAGAGCCCACTCGTGGACAAGTACGACTTGTCTTCACTGCG 920 ROVER51.SEQTGAGCCCACTCGTGGACAAGTACGACTTGTCTTCACTTGCG 920 ROVER51.SEQTGAGCCCACTCGTGGACAAGTACGACTTTGTCTTCACTTGCG 920 GRVER51.SEQTGAGCCCACTCGTGGCGCTGCTCCTTTTGGCCAAAGAAGTG 960 GRVER51.SEQTGAGCCTGTGCTGTGGCGCTGCTCCTTTTGGCCAAAGAAGTG 960 GRVER3.SEQTGAACTGTGCTGTGGCGCTGCTCCTTTTGGCCAAAGAAGTG 960 GRVER3.SEQTGAAACTGTGCTGTGGCGCTGCTCCTTTTGGCCAAAGAAGTG 960 GRVER3.SEQCGAAACTGTGCTGTGGCGCTGCTCCTTTTGGCCAAAGAAGTG 960 GRVER3.SEQCGAAACTGTGCTGTGGCGCTGCTCCTTTTGGCCAAAGAAGTG 960 GRVER3.SEQCGAAACTGTGCTGTGGCCGCTGCCCCTTTTGGCCAAAGAAGTG 960 ROVER3.SEQCTGAAACTGTGCTGTGGCCGCTGCCCCCTTTTGGCCAAAGAAGTG 960 ROVER3.SEQCTGAAACTGTGTGTGCGGCGCTGCCCCCTTTTGGCCAAAGAAGTG 960 ROVER3.SEQCTGAAACTGTGTGTTGCCGGTGCCCCCTTTTGGCCAAAGAAGTG 960 ROVER3.SEQTTGAAACTGTGTGTTGCCGGTGCCCCCCTTTTGGCCAAAGAAGTC 960 ROVER3.SEQTTGAAATTGTGTTTGCGGTGCCGCTCCACTTGGCCAAAGGAAGTC 960 ROVER3.SEQTTGAAATTGTGTTTGCGGTGCCGCTCCACTTGGCTAAAGGAGGTC 960 ROVER3.SEQTTGAAATTGTGTTTGCGGTGCCGCTCCACTTGGCTAAAGGAGGTC 960 ROVER3.SEQTTGAAATTGTGTTTGCGGTGCCGCTCCACTTGGCTAAAGGAAGTC 960 ROVER3.SEQTTGAAATTGTGTTTGCGGTGCCGCTCCACTTGGCTAAAGGAAGTC 960 ROVER3.SEQTTGAAATTGTGTTTGCGGTGCCGCTCCACTTGGCTAAAGGAAGTC 960 ROVER51.SEQTTGAAATTGTGTTTGCGGTGCCGCTCCACTTGGCTAAAGGAAGTC 960 ROVER51.SEQTTGAAATTGTGTTTGCGGTGCCGCTCCACTTGGCTTAAGGAAGTC 960 ROVER51.SEQTTGAAATTGTGTTTGCGGTGCCGCTCCACTTGGCTTAAGGAAGTC 960 ROVER51.SEQTTGAAATTGTGTTTGCGGTGCCGCTCCACTTGGCTTAAGGAAGTC 960 ROVER51.SEQTTGAAATTGTGTTTGCGGTGCCGCTCCACTTGGCTTAAGGAAGTC 960 ROVER51.SEQTTGAAATTGTGTTTGCGGTGCCGCTCCACTTGGCTTAAGGAAGTC 960	RDVER5.SEQ 2	GAGCCCACTCGTGGACAAGTACGACTTGTCTTCACTGCG 920	
RDVER51. SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RDVER52. SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 RD1561H9. SEQA G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 GRVER51. SEQ T G A G C T G T G C T G T G G C G C T G C T C C T T T T	RD7.SEQ	GAGICIC CACITICS TIGGACAAGTACGAICITTIGT CITTICACTICICS 220	
RDVER52.SEQ A G A G C C C A C T C G T G G A C A A G T A C G A C T T G T C T T C A C T G C G 920 GRVER51.SEQ T G A G C T G T G C T G T G G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GR6.SEQ T G A G C T G T G C T G T G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER5.SEQ T G A G C T G T G C T G T G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER5.SEQ T G A G C T G T G C T G T G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER4.SEQ T G A G C T G T G C T G T G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER3.SEQ T G A A C T G T G C T G T G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER2.SEQ C G A A C T G T G C T G T G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVER1.SEQ C G A A C T G T G C T G T G C G C T G C C C T T T T	RDVER51.SEQ	GA GICC CIA CITICG TIGIGA CA AIGITA CGA CITTGIT CITT CA CITIGIG 920	
GRVERS1.SEQ TG A G C T G T G C T G T G G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GR6.SEQ TG A G C T G T G C T G T G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVERS1.SEQ TG A G C T G T G C T G T G G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVERS.SEQ TG A G C T G T G C T G T G G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVERS.SEQ TG A G C T G T G C T G T G G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVERS.SEQ TG A A C T G T G C T G T G G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVERS.SEQ C G A A C T G T G C T G T G G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVERS.SEQ C G A A C T G T G C T G T G G C G C T G C T C C T T T G G C C A A A G A A G T G 960 GRVERS.SEQ C G A A C T G T G C T G T G C G C T G C C C T T T G G C T A A A G A G T G 960 GRVERS.SEQ T G A G T T G T G T G C G G T G C C C C T T T G G C T A A A G A G T G 960 ROVERS.SEQ T G A G T T G T G T G C G G T G C C C C T T T G G C C A A A G A A G T C 960 ROVERS.SEQ T G A G T T G T G T T G C G G T G C C C C C A C T G G C C A A A G G A G T C 960 ROVERS.SEQ T G A A T T G T G T T G C G G T G C C C C C A C T G G C T A A G G A G G T C 960 ROVERS.SEQ T G A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 ROVERS.SEQ T G A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 ROVERS.SEQ T G A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 ROVERS.SEQ T G A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 ROVERS.SEQ T G A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 ROVERS.SEQ T G A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 ROVERS.SEQ T G A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 ROVERS.SEQ T G A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 ROVERS.SEQ T G A A T T G T G T T G C G G T G C C G C T C C A C T G G C	RDVER52.SEQ	GAGCCCACTCGTGGACAAGTACGACTTGTCTTCACTGCG	
GRVERS1.SEQ TG AG C T G T G C T G T G C G C T G C T C C T T T T	RD1561H9.SEQ1	GAGICIC CIA CITICIG TIGIG A CAAIGIT A CGAICIT TIGIT CIT T CACITIC CIC 920	
GR6. SEQ			
GR6. SEQ	GRVER51.SEQ	GAG CIT G T GCIT GTIG GCIG CTIG CTIC CTIT TIGG CCIA A A G A G TIG 960	
GRVER5.SEQ TG AG CT GT GCT GT GCT GCT GCT CCTTTT GGCCAAAGAAGT G960 GRVER4.SEQ TG AG CT GT GCT GT GGCT GCT CCTTTT GGCCAAAGAAGT G960 GRVER2.SEQ CGAACT GT GCT GT GGCT GCT CCTTTT GGCCAAAGAAGT G960 GRVER1.SEQ CGAACT GT GCT GT GGCT GCT CCTTTT GGCTAAAGAGT G960 YG81-6G1.SEQGGAACT GT GCT GT GCGCT GCCCTT TT GGCTAAAGAGT G960 RDVER1.SEQ TGAGT TGT GT TGCGGT GCCCATT TT GGCCAAAAAAGAT TGT GT GCGCT GCCCTT TT GGCCAAAGGT G960 RDVER1.SEQ TGAGT TGT GT GCGGT GCCCACT GGCCAAAGGAAGT T960 RDVER3.SEQ TGAGT TGT GT GT GCGGT GCCCCT CCACT GGCCAAAGGAAGT C960 RDVER4.SEQ TGAATT GT GT TGCGGT GCCCCT CCACT GGCTAAAGGAAGT C960 RDVER4.SEQ TGAATT GT GT TGCGGT GCCCCT CCACT GGCTAAAGGAAGT C960 RDVER5.SEQ TGAATT GT GT TGCGGT GCCCCT CCACT GGCTAAAGGAAGT C960 RDVER5.SEQ TGAATT GT GT TGCGGT GCCGCT CCACT GGCTAAAGGAAGT C960 RDVER5.SEQ TGAATT GT GT TGCGGT GCCGCT CCACT GGCTAAAGGAAGT C960 RDVER5.SEQ TGAATT GT GT TGCGGT GCCGCT CCACT GGCT AAGGAAGT C960 RDVER5.SEQ TGAATT GT GT TGCGGT GCCGCT CCACT GGCT AAGGAAGT C960 RDVER5.SEQ TGAATT GT GT TGCGGT GCCGCT CCACT GGCT AAGGAAGT C960 RDVER5.SEQ TGAATT GT GT TGCGGT GCCGCT CCACT GGCT AAGGAAGT C960 RDVER5.SEQ TGAATT GT GT TGCGGT GCCGCT CCACT GGCT AAGGAAGT C960 RDVER5.SEQ TGAATT GT GT TGCGGT GCCGCT CCACT GGCT TAAGGAAGGT C960	GR6.SEQ	GAIG CIT GT GICT GITIG GICIG CITIG CITIT TIGIG CICIA A G A A G TICL SEN	
GRVER4.SEQ TG A A C T G T G C T G T G C G C T G C T C C T T T T	GRVER5.SEQ	GAIG CIT GT GICIT GITIG GICIG CITIC CITIT TIGIG CICIA A C A A C TICI SEN	
GRVER3.SEQ TIGAACT GTGCTGTGGCGCTGCTTTTGGCCAAAGAAGTGGGGGGGCGCTGCCCTTTTGGCCTAAAGAGGTGGGGGCGCTGCCCTTTTGGCCTAAAGAGGTGGGGGCGCTGCCCTTTTGGCCTAAAGAGGTGGGCGGCTGGCCCCTTTTGGCCTAAAGAGGTGGGCGGCTGCCCCTTTTGGCCTAAAGAGGTGGGCGGCTGCCCCTTTTGGCCTAAAGAGGTGGGCCGCTCCACTGGCCTAAAGAGGTCGGCCAAGGAAGTTGGCCGCTCCACTGGCCTAAAGGAAGTCCGGCCAAGGAAGTCCGGCCTCCACTGGCCTAAAGGAAGTCCGGCCAAGGAAGTCCGGCCTCCACTGGCCTAAAGGAAGTCCGGCTCCACTGGCCTAAAGGAAGTCCGGCTCCACTGGCCTAAAGGAAGTCCGGCTCCACTGGCCTAAAGGAAGTCCGGCTCCACTGGCCTAAAGGAAGTCCGGCTCCACTGGCCTAAAGGAAGTCCGGCTCCACTGGCCTAAAGGAAGTCCGGCTCCACTGGCCTAAAGGAAGTCCGGCTCCACTGGCCTAAAGGAAGTCCGGCTCCACTGGCCTAAAGGAAGG	GRVER4.SEQ	GAIG CIT GT GICIT GITIG GICIG CITIC CITIT TIGIG CICIA A A G A A G TIGI GAN	
GRVER1.SEQ CG A A CT G T G CT GT G C CG CT G C C C CTT T T G G C T A A A G A G G T G 960 YG81-6G1.SEQ G A A CT G T G C G G T G C G C C C T T T T G G C T A A A G A G G T G 960 YG81-6G1.SEQ G A A T T G T G T G C G G T G C G C A C T T A G C A A A G A A G T T 960 RDVER1.SEQ T G A G T T G T G T G C G G T G C C G C T C C A C T G G C C A A G G A A G T C 960 RDVER2.SEQ T G A G T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER3.SEQ T G A G T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER4.SEQ T G A A T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER5.SEQ T G A A T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER51.SEQ T G A A T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER51.SEQ T G A A T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER51.SEQ T G A A T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER51.SEQ T G A A T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER51.SEQ T G A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER52.SEQ T G A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960	GRVER3.SEQ	GAACT GT GCT GTG GCG CTG CTT CTTTTGG CCAAAGAAGTG SOO	
GRVER1.SEQ CG A A CT G T G CT G T G C G C T G C C C T T T G G C T A A A G A G G T G 960 YG81-6G1.SEQG G A A T T G T G T G C G G T G C G C T C C A C T G G C C A A A G A A G T C 960 RDVER1.SEQ T G A G T T G T G C G G T G C C G C T C C A C T G G C C A A G G A A G T C 960 RDVER3.SEQ T G A G T T G T G T G C G G T G C C G C C C C	GRVER2.SEQ	GAAICIT GT GICIT GITIG GICIG CITIG CICIC CITIT TIGIG CITIA A A GAIGE TIGI SEO	
YG81-6G1. SEOG G A A T T G T G T T G C G G T G C G G C A C C A T T A G C A A A A G A A G T T 960 RDVER1. SEQ T G A G T T G T G T G C G G T G C C G C T C C A C T G G C C A A G G A A G T C 960 RDVER2. SEQ T G A G T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G T C 960 RDVER4. SEQ T G A A T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER5. SEQ T G A A T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER5. SEQ T G A A T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER51. SEQ T G A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER52. SEQ T G A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER52. SEQ T G A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960	GRVER1.SEQ	GAACIT GT GICIT GITIG GICIG CITIG CICIC CITIT TIGIG CITIA A A GAIGIG TIGIG COM	
RDVER1.SEQ TG A G T T G T G T G C G G T G C C G C T C C A C T G G C C A A G G A A G T C 960 RDVER3.SEQ TG A G T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A A G T C 960 RDVER4.SEQ TG A A T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER5.SEQ TG A A T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER5.SEQ TG A A T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER51.SEQ TG A A T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER52.SEQ TG A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER51.SEQ TG A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER52.SEQ TG A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960	YG81-6G1.SEO	GAATIGIGITICCGGTGCGGCACCATTACCAAAACAACTT	
RDVER3.SEQ TG A G T T G T G T G T G C G G T G C C G C T C C A C T G G C C A A G G A A G T C 960 RDVER4.SEQ TG A A T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER5.SEQ TG A A T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER5.SEQ TG A A T T G T G T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER51.SEQ TG A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER52.SEQ TG A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER52.SEQ TG A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960	ROVER1.SEO	GAGTTGTGTTTGCGGTGCCGCTTCCACTTGGCCAACGAAGGAAG	
RDVER3.SEQ TG AGTTGTGTTGCGGTTGCCGCCCCACTGGCTAAGGATGCTC 960 RDVER4.SEQ TG AATTGTGTTTGCGGTTGCCGCTCCACTGGCTAAGGATGCTC 960 RDVER5.SEQ TG AATTGTGTTTGCGGTTGCCGCTCCACTGGCTAAGGATGCTC 960 RDVER51.SEQ TG AATTGTGTTTGCGGTTGCCGCTCCACTGGCTTAAGGAGGTC 960 RDVER51.SEQ TG AATTGTGTTTGCGGTTGCCGCTCCACTTGGCTTAAGGAGGTC 960 RDVER52.SEQ TG AATTGTGTTTGCGGTTGCCGCTCCACTTGGCTTAAGGAGGTC 960	RDVER2.SEO	GAGTTGTGTTGCGGTGCCGGCTCCALCTCGCCCAACGAACGAACTC	
RDVER4.SEQ TGAATTGTGTTTGCGGTTGCCGCTTCACTGGCTAAGGAGGTC960 RDVER5.SEQ TGAATTGTGTTTGCGGTTGCCGCTTCACTGGCTAAGGAGGTC960 RO7.SEQ TGAATTGTGTTTGCGGTTGCCGCTTCACTGGCTAAGGAGGTC960 RDVER51.SEQ TGAATTGTGTTTGCGGTTGCCGCTTCACTGGCTAAGGAGGTC960 RDVER52.SEQ TGAATTGTGTTTGCGGTTGCCGCTCCACTGGCTTAAGGAGGTC960			
RDVERS.SEQ TGAATTGTGTTTGCGGTTGCCGCTCCACTGGCTAAGGAGGTC 960 RO7.SEQ TGAATTGTGTTTGCGGTTGCCGCTCCACTGGCTAAGGAGGTC 960 RDVER51.SEQ TGAATTGTGTTTGCGGTTGCCGCTCCACTGGCTAAGGAGGTC 960 RDVER52.SEQ TGAATTGTGTTTGCGGTTGCCGCTCCACTGGCTTAAGGAGGTC 960	- 1		
RD7.SEQ TGAATTGTGTTGCGGTGCCGCTCCACTGGCTAAGGAGGTC 960 RDVER51.SEQ TGAATTGTGTTGCGGTGCCGCTCCACTGGCTAAGGAGGTC 960 RDVER52.SEQ TGAATTGTGTTGCGGTGCCGCTCCACTGGCTAAGGAGGTC 960	ROVERS SEO	GAATTGTGTTTGCGGTGCCCTCCCTCCACCTCCCCTAAACCCACCTAAACCCACCAACAAC	
RDVER51. SEQ TG A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960 RDVER52. SEQ T G A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960	RD7.SEO	GAATTGTGTTTGCGGTCCCCTCCCTCCACTCCACTAAAGGAACCAACA	
ROVERS2. SEQ TG A A T T G T G T T G C G G T G C C G C T C C A C T G G C T A A G G A G G T C 960	ROVERS1 SEO	GAATTGTGTTGCGGTCCCCCCCCCCCCCCCCCCCCCCCC	
RO1561H9.SEQTG A A T T G T G T G C G G T G C G C T C C A C T G G C T A A G G A G G T C 960		GARTIGIGITIGE GIGGETTE CALCITIGE CITIA AIGIGAIGIG TE 960	
	RD156149 9FA	GAATTGTGTTGCGGTGCCCCCCCCCTCCACTGGCCTAAAGGAAGG	
		ORALIGICA CONTROL OF CONTROL AND THE CONTROL AND A SECOND CONTROL OF CONTROL	

Figure 2 (cont.)

GRVER51.SEQ G C C G A G G T C G C T G C T A A G C G T C T G A A C C T C C C T G G T A T	C C 1000
GRESEQ GCCGAGGTCGCTGCTTAAGCCGTCTGAACCTCCCTGGTAT	c c 1000
GRVER5. SEQ G CICIG A G G TICIG CITIG CITIA A GC GT CITIGIA A CICITICIC CITIG GITIA T	C C 1000
GRVER4. SEQ GCCGAGGTCGCTGCTAAGCCGTCTGAACCTCCCTGGTAT	C C 1000
GRVER3. SEQ G C C G A G G T C G C T G C T A A G C G T C T G A A C C T C C C T G G T A T	C C 1000
GRVER2. SEQ G C C G A A G T C G C T G C C A A G C G T C T G A A T T T G C C A G G T A T	C C 1000
GRVER1. SEQ G CCG A A G T C G CTG C CA A G C G T C T G A A T T T G C C A G G T A T	C C 1000
YG81-6G1. SEQG C T G A G G T T G C A G C A A A A C G A T T A A A C T T G C C A G G A A T	T.C. 1000
RDVER1. SEQ GCTGAGGTGGCCGCTAAACGCTTGAACCTGCCTGGCAT	T C 1000
RDVER2.SEQ GCTGAGGTGGCCGCTAAACGCTTGAACCTTGCCTTGGCAT	T C 1000
RDVER3. SEQ GCTGAAGTGGCCGCCAAACGCTTGAATCTGCCAGGCAT	T C 1000
RDVER4. SEQ GCTGAAGTGGCCGCCAAACGCTTGAATCTGCCCGGCAT	T C 1000
RDVER5. SEQ GCTGAAGTGGCCGCCAAACGCTTGAATCTTCCAAGGGAT	T C 1000
RD7. SEQ GCTGAAGTGGCCGCCAAACGCTTGAATCTTCCAAGGAT	T C 1000
RDVER51.SEQ G C T G A A G T G G C C G C C A A A C G C T T G A A T C T T C C A G G A T	T C 1000
RDVER52.SEQ G C T G A A G T G G C C G C C A A A C G C T T G A A T C T T C C A G G A T	1 C 1000
RD1561H9.SEQG C T G A A G T G G C C C C A A A C G C T T G A A T C T T C C A G G G A T	T C 1000
ADITOTION OF THE CONTROL OF THE CONT	T C 1000
COURTS STORE OF THE COMPANY OF THE C	
GRVERS1.SEQ GCTGCGGTTTTTGGTTTGACTGAGAGCACTTCTTGCTAAC	AT 1040
GR6. SEQ GCTGCGGTTTTGGTTTGACTGAGAGCACTTCTGCTAAC	A T 1040
GRVERS. SEQ GCTGCGGTTTTGGTTTGACTGAGAGCACTTCTGCTAAC	A T 1040
GRVER4.SEQ GCTGCGGTTTTGGTTTGACTGAGAGCACTTCTGCTAAC	A T 1040
GRVERS. SEQ GCTGCGGTTTTGGTTTGACTGAGAGCACTTCTGCCAAC	A T 1040
GRVER2. SEQ GCTGCGGCTTTGGTCTGACTGAGAGCACCTCTGCTAAC	A T 1040
GRVER1.SEQ GCTGCGGCTTTGGTCTGACTGAGAGCACCTCTGCTAAC	A T 1040
YG81-6G1.SEQGCTGTGGATTTGGTTTGACAGAATCTACTTCAGCTAAT	A T 1040
RDVER1.SEQ GTTGTGGTTTTCGGCTTGACCGAATCTACTAGCGCCATT	A T 1040
ROVER2. SEQ GITGGGGTTTCGGCTTGACCGAATCTACTAGCGCCATT	A T 1040
ROVERS. SEQ GTTGTGGGCTTCGGCCTCACCGAATCTACCAGCGCTATT.	A T 1040
RDVER4.SEQ GITTGTGGGCTTCGGCCTCACCGAATCTACCAGCGCTATT	A T 1040
RDVERS. SEQ GITGGGCTTCGGCCTCACCGAATCTACCAGCGCTATT	A T 1040
RD7. SEQ GTTGTGGCTTCGGCCTTCACCGAATCTACCAGCGCTATT	A T 1040
RDVER51.SEQ GITTGTGGCTTCGGCCTTCACCGAATCTACCAGCGCTATT	AT 1040
RDVER52.SEQ GITTGTGGGCTTCGGCCTCACCGAATCTACCAGCGCTATT	A T 1040
RD1561H9.SEQGTTGTGGGCTTCGGCCTCACCGAATCTAC <u>CAGT</u> GCGATT	A T 1040
GRVER51.SEQ CC ATAGCTTG CGAGACGAGTTTAAGTCTGGTAGCCTGG	3 T 1080
GR6.SEQ CCATAGCTTGCGAGAGAGTTTAAGTCTGGTAGCCTGGC	3 T 1080
GRVERS.SEQ CCATAGCTTGCGAGAGAGTTTAAGTCTGGTAGCCTGG	3 T 1080
GRVER4. SEQ CCATAGCTTGCGAGAGACGAGTTTAAGTCTGGTAGCCTGGC	G T 1080
GRVER3. SEQ CCATAGCTTGCGTGACGAGTTTAAATCTGGTAGCCTGGC	G T 1080
GRVER2. SEQ TICATAGCTTGCGTGATGAGTTCAAATCTGGCAGCCTGGC	F T 1080
GRVER1. SEQ TC ATAGCTTG CGTG ATG AATTCAAATCTG GC AG CC TGG	FT 1080
YG81-6G1. SEQA CACAGTCTTAGGGATGAATTTAAATCAGGATCACTTG	A 1080
RDVERL. SEQ CCAATCTCTGCGCGACGAGTTTAAGAGCGGTTCTTTGGC	
RDVER2. SEQ CCAATCTCTGCGCGACGAATTTAAGAGCGGTTCTTTGGC	S C 1080
RDVER3. SEQ TC AATCTCTCCGCGATGAGTTTAAGAGCGGCTCTTTGGG	G C 1080
RDVER4. SEQ T C A G T C T C C G C G A T G A G T T T A A G A G C G G C T C T T T G G C	3 C 1080
RDVERS.SEQ TCAGTCTCTCCGCGATGAGTTTAAGAGCGGCTCTTTGGG	G C 1080
RD7. SEQ TCAGTCTCCGCGATGAGTTTAAGAGCGGCTCTTTGGG	G C 1080
RDVERS1. SEQ T C A G T C T C C G C G A T G A G T T T A A G A G C G G C T T T T G G G	[C] 1080
RDVER52.SEQ T C A G T C T C T C G G G G A T G A G T T T A A G A G C G G C T C T T T G G G	C 1080
RD1561H9.SEQCCAGACTCTCGGGGATGAGTTTAAGAGCGGCTCTTTGGG	(C) 1080

Figure 2 (Cont.)

GRVERS1. SEQ CGCG TGACT CCTCTTATGGCTGCAAAGATCGCCGACCGTG 1120	_
GRE. SEQ CGCGTGACT CCTCTTAT GGCTGCAAAAGATCGGCCGACCGTG 1120	
GRVERS. SEQ CGCGCGTGACTCCTCTTATGCCTGCAAAGATCGCCGGCGACCGTG 1120	
GRVER4.SEQ CGCGTGACTCCTCTTATGCCTGCAAAGATCGCCGACCGTG 1120	
GRVER3.SEQ CCCGTGACCCCTTTGATGGCTGCAAAGATCGCCGACCGTG 1120	
GRVER2. SEQ C G C G T G A C T C C T T T G A T G C C G C T A A G A T C G C G A C C G T G 1120	
GRVER1.SEQ CGCGTGACTCCTTTGATGGCCGCTAAGATCGCCGACCGTG 1120	
YG81-6G1.SEQAGAGTTACTCCTTTAATGGCAGCTAAAATAGCAAGATAGGG 1120	
RDVER1. SEQ CGTGTCACCCCACTGATGGCCGAAAATTGCCTGATCGCCG 1120	
RDVER2. SEQ CGTGTGTCACCCCCACTGATGGTTGCCAAAATTGCTTGATCGCG	
RDVER3. SEQ CONTIGET CACT COACT CAT GOCT GCTAAAATCGCT GAT CGC G 1120	
RDVER4. SEQ CIGITIG TICIA CITICIA TIGIGICITIG CITIA AGA TICIGICIG A TICIGICIG 1120	
RDVERS. SEQ CGTGTGTCACTCCACTCATGGCTGCTAAGATCGCTGATCGCG1120	
RD7. SEQ CGTGTCACTCCACTCATGCTTGCTAAGATCGCTGATCGCGCG 1120	
RDVER51. SEQ C G T G T C A C T C C A C T C A T G G C T G C T A A G A T C G C T G A T C G C G 1120	
RDVER52. SEQ CGTGTCACTCCACTCATGGCTGCTAAGATCGCTGATCGCG 1120	
RD1561H9.SEQCGTGTGACTCACTCATGGCTGCTAAGATCGCTGATCGCG 1120	
GRVER51. SEQ AGA CCIG GCA A A G C ACT G G GCIC C A A A T C A A G T C G G T G A A T T 1160	
GRE. SEQ AGACCGGCAAAGCACTGGGCCCAAATCAAGTCGGTGAATT 1160	
GRVERS. SEQ AGACCCGCCAAAGCACCCAAATCAAGTCGGTGAATT 1160	
GRVER4.SEQ AGACCGGCAAAGCACTGGGCCCAAATCAAGTCGGTGAATT 1160	
GRVER3. SEQ A G A C C G G C C A A A G C C C T G G G C C C A A A T C A G G T C G G T G A A T T 1160	
GRVER2. SEQ AGACOGGCAAAGCTCTGGGTCCAAATCAAGTCGGCGAATT 1160	
GRVERL SEQ AGACOGGOAAAGCTCTGGGTCCAAATCAAGTCGGCGAATT 1160	
YG91-6G1. SEQAAACTGGTAAAGCATTGGGACCAAATCAAGTTGGTGAATT 1160	
RDVERL. SEQ A A A C T G G T A A G G C C T T G G G C C T A A C C A G G T G A G C T 1160	
RDVER2. SEQ. A A A C T G G T A A G G C C T T G G G C C C T A A C C A G G T G A G C T 1160	
RDVER3.SEQ AAACTGGTAAGGGCTTTTGGGGCCCTAAGCCAAGTGGGGCGAGCT 1160	
RDVER4 SEQ AAACTGGTAAGGCTTTGGGGCCTTAACCAAGTGGGCGAGCT 1160	
RDVERS. SEQ A A A C T G G T A A G G C T T T G G G C C T A A C C A A G T G G G C G A G C T 1160	
RD7. SEQ AAACTGGTAA GGC TTTTGGGC CCGAACTCGAAGTGGGCGA GC T 1160	
RDVER51. SEQ A A A C T G G T A A G G C T T T G G G C C C G A A C T G G G C G A G C T 1160	
RDVER52. SEQ A A A C T G G T A A G G C T T T G G G C C G A A C T G G G C G A G C T 1160	
RD1561H9.SEQA A A C T G G T A A G G C T T T G G G C C C G A A C C A A G T G G G C G A G C T 1160	
GRVER51.SEQ GT GTATTAAGG GCC CTATGGTCTAAAGGCTACGTGAAC 1200	
GR6. SEQ GTGTATTAAGGGGCCCTATGGTCTCTAAAGGCCTACGTGAAC 1200	
GRVER5. SEQ GTGTATTAAGGGGCCCTATGGTCTCTAAAGGCCTACGTGAAC 1200	
GRVER4. SEQ GTGTATTAAGGGGCCCTATGGTCTCTAAAGGCCTACGTGAAC 1200	
GRVER3.SEQ GTGCATTAAGGGGCCCTATGGTCTCTAAAGGCTACGTGAAC 1200	
GRVER2. SEQ GT GTATTAAGG GT C CTATG GT GT CTAAAG G CTAC GT CAAC 1200	
GRVER1. SEQ GTGTATTAAGGGTCCTATGGTGTCTAAAGGCTACGTCAAC 1200	
YG81-6G1. SEQAT G CAT TAAAG G T C C CAT G G T AT C G A A A G G T T A C G T G A A C 1200	
RDVER1.SEQ GTGCATCAAAGGCCCAATGGTCAGCAAGGGTTATGTGAAT 1200	
RDVER2. SEQ GTGCATCAAAGGCCCAATGGTCAGCAAGGGTTATGTGAAT 1200	
RDVER3. SEQ GT GT A T C A A A G G C C C T A T G G T G A G C A A G G G T T A T G T C A A T 1200	
RDVER4. SEQ GT GT A T C A A A G G C C CT A T G G T G A G C A A G G G T T A T G T C A A T 1200	
RDVERS. SEQ GT GT A T C A A A G G C C C T A T G G T G A G C A A G G G T T A T G T C A A T 1200	
RD7. SEQ GT GTAT CAAAGGCCCTATGGTGAGCAAGGGTTATGTCAAT 1200	
RDVERS1.SEQ GT GT A T CAA A G G C C CT A T G G T G A G C A A G G G T T A T G T CAA T 1200	
RDVER52.SEQ GT GTATCAAAGGCCCTATGGTGAGCAAGGGGTTATGTCAAT 1200	
RD1561H9. SEQGT GTATCAAAGGCCCTATGGTGAGCAAGGGTTATGTCAAT 1200	

Figure 2 (cont.)

GRVER51. SEQ AATGTGGGAGGCCACTAAAGAAGCCATTGATGATGATGGCT 1240	
GR6.SEQ AATGTGGGAGGCCACTAAAGAAGCCATTGATGATGATGGCT 1240	
GRVER5. SEQ AATGT G G A G G C C A C T A A A G A A G C C A T T G A T G A T G G C T 1240	
GRVER4. SEQ AATGTGGGAGGCCCACTAAAGAAGCCCATTGATGATGATGCCT 1240	
GRVER3.SEQ AATGTGGGAGGCCACTAAAGAAGCTATTGATGATGGTT 1240	
GRVER2.SEQ AATGTGGAGGCCACTAAGGAAGCTATTGATGACGATGGTT 1240	
GRVER1. SEQ AATGTGGAGGCCCACTAAGGAAGCTATCGATGACGATGGTT 1240	
YG81-6G1. SEQAATGTAGAAGCTACCAAAGAAGCTATTGATGATGATGATGATGATGATGATGATGATGATGAT	
RDVER1. SEQ A A C G T C G A A G C T A C C A A A G A G G C C A T T G A C G G A T G A C G G C T 1240	
RDVER2.SEQ AACGTCGAAGCTACCAAAGAGGCCATCGAAGACGATGACGGCT 1240	
RDVER3. SEQ A A C G T C G A A G C T A C C A A G G A G G C C A T C G A C G A C G G C T 1240	
RDVER4.SEQ A A C G T C G A A G C T A C C A A G G G A G G C T A T C G A C G A C G A C G G C T 1240	
RDVER5. SEQ A A C G T C G A A G C T A C C A A G G A G G C A T C G A C G A C G G C T 1240	
RD7. SEQ A A C G T T G A A G C T A C C A A G G A G G A C C C C C C C C C	
RDVER51. SEQ A A C G T T G A A G C T A C C A A G G A G G C C A T C G A C G A C G A C G G C T 1240	
RDVER52. SEQ A A C G T T G A A G C T A C C A A G G A G G C C A T C G A C G A C G G C T 1240	
RD1561H9. SEQA ACIG TITG A A G C T A C C A AGIG AGIG CCA TCG ACIG ACIG ACIG GCT 1240	
GRVER51. SEQ GGCTCCATAGCGGCGACTTCGGGTTACTATGATGAGGACGA 1280	
GR6. SEQ GGCTC CA TAGC GG C GACTT C GGT TACTATGATGAGGA C GA 1280	
GRVER5. SEQ GGCT C CA TAGC GG C GACTT C GG T TACTATGATGAGGA C GA 1280	
GRVER4.SEQ GGCT C CA TAGC GG C GACTT C GG T TACTATGATGAGGA C GA 1280	
GRVER3.SEQ G G [T] T G C A T A G C G G C C G A C T T C C G G T T A T T A T G A T G A G G A C G A 1280	
GRVER2.SEQ GGCTGCACAGCGGCGACTTTGGTTTATTACGATGAGGACGA 1280	
GRVER1. SEQ GGCT GCACAGCGCCGACTTTGGTTTAITTAICGATGAGGAICGA 1280	
YG81-6G1. SEQGGCTTCACTCTGGAGACTTTGGATACTATGATGAGGATGA 1280	
RDVER1.SEQ GG T T G CA T TCTGG T GA T TT C GG C TACTATGA C GA A GATGA 1280	
RDVER2.SEQ GGTTGCATTCTGGTGATTTCGGGCTACTATGACGAAGATGA 1280	
RDVER3.SEQ GGCTGCATTCTGGTGATTTTGGCTACTACGACGAAGATGA 1280	
RDVER4. SEQ GGTTGCATTCTGGTGATTTTGGATATTACGACGACAAGATGA 1280	
RDVERS. SEQ GGTTGCATTCTGGTGATTTTGGATATTACGACGACGAAGATGA 1280	
RD7. SEQ GGTTGCATTCTGGTGATTTTGGATATTACGACGAAGATGA 1280	
RDVER51. SEQ G G T T G C A T T C T G G T G A T T T T G G A T A T T A C G A C G A A G A T G A 1280	
RDVER52.SEQ G G T T G C A T T C T G G T G A T T T T G G A T A T T A C G A C G A A G A T G A 1280	
RD1561H9.SEQGGTTGCATTCTGGTGATTTTGGATATTACGACGAAGATGA 1280	
CDUEDS 1 STO NO SOR TO CONTROL OF SOR TO SOR	
GRVER51. SEQ ACACTTCTATGTGGTCGATCGCTACAAAGAATTGATTAAG 1320 GR6. SEQ ACACTTCTATGTGGTCGATCGCTACAAAGAATTGATTAAG 1320	
YG81-6G1. SEQG CATTTCTATGTGGTGGACCGTTACAAGGAATTGATTAAA 1320 RDVER1. SEQ GCACTTTTTACGTGGTCGACCGTTATTAAGGAACTGATCAAA 1320	
RDVER3. SEQ GCATTTTTTACGTCGTGGATCGTTACAAGGAGCTGATCAAA 1320 RDVER4. SEQ GCATTTTTTACGTTCGTGGATCGTTACAAGGAGCTGATCAAA 1320	
a series a series and a series and a series of the series	
a a a lala salala a lala a a a salala a a a	
RDVER51. SEQ G C A T T T T T A C G T C G T G G A T C G T T A C A A G G A G C T G A T C A A A 1320	
RDVER52.SEQ G C A T T T T T A C G T C G T G G A T C G T T A C A A G G A G C T G A T C A A A 1320	
RD1561H9.SEQG C A T T TTT A C G T C G T G G A T C G T T A C A A G G A G C T G A T C A A A 1320	

Figure 2 (cont.)

rigure 2 (cont.)	
GRVER51.SEQ TACAAAGCTCTCAAGTCGCACCAGCCGAACTGGAAGAAA 1360	
GR6. SEQ TAICIA AIAIG G C T C T C AIAIG TICIG C A C CIAIG CICIG A A C TIGIG A A G AIA 1360	
GRVERS. SEQ TAICIA AIAIG G C T C T CAIAIG TICIG C A C CIAIG CICIG A A C TIGIG A A G AIAIA 1350	
GRVER4. SEQ TAICIA AIAIG G C T C T C AIAIG TICIG CICIC CIAIG CICIG A A C TIGIG A A G AIAI A 1360	
GRVER3. SEQ TAICIA A A A G G C T C T C A A A G T C I C I C I C I C I C I C I C I C I C	
GRVER2.SEQ TATAAAGGCTCTCAAGTCGCCCCAGCTGAGCTGGAAGAAA 1360	
GRVER1. SEQ TATAAAGGCTCTCAGGTCGCCCCAGCTGAGCTGGAAGAGAGAG	
YG81-6G1. SEQT A T A A G G G C T C T C A G G T A G C A C C T G C A G A A C T A G A A G A G A 1360	
ROVERL. SEQ TACAAGGGTAGCCAAGTGGCTCCTGCCGAATTGGAGGAAA 1360	
RDVER2. SEQ TACAAGGGTAGCCAAGTGGCTCCTGCCGAATTGGAGGAGA 1360	
ROVERS. SEQ TACAAGGGTAGCCAGGTGGCTCCAGCCGAGTTTGGAGGAGA 1360	
RDVER4.SEQ TACAAGGGTAGCCAGGTTGCTCCAGCTTGAGTTGGAGGAGA 1360	
ROVERS. SEQ TACAAGGGTAGCCAGGTTGCTCCAGCTGAGTTGGAGTTGGAGAA1360	
RD7. SEQ TACAAGGGTAGCCAGGTTGCTCCAGCTGAGTTGGAGGAGA 1360	
RDVER51.SEQ TACAAGGTAGCCAGGTTGCTCCAGCTGAGTTGGAGGAGA 1360	
RDVER52.SEQ T A C A A G G G T A G C C A G G T T G C T C C A G C T G A G T T G G A G G A G A 1360	
RD1561H9. SEQT A C A A G G G T A G C C A G G T T G C T C C A G C T G A G T T G G A G G A G A 1360	
1300	
GRVER51.SEQ TTTTGCTGAAGAACCCTTTGTATCCGCGACGTGGCCGTCGT 1400	
GR6. SEQ T T T T G C T G A A G A A C C C T T G T A T C C G C G A C G T G G C C G T C G T 1400	
GRVERS. SEQ TTTTGCTGAAGAACCCTTGTATCCGGCGACGTGGCCGTCGT 1400	
GRVER4. SEQ TTTTGCTGAAGAACCCTTGTATCCGGCGACGTGGCCGTCGT 1400	
GRVER3. SEQ TTTTGCTGAAGAACCCTTGTATTCGCGACGTGGCCGTCGT 1400	
COVERS OF COMMISSION OF CAMBRIDA AND AND AND AND AND AND AND AND AND A	
YG81-6G1. SEQT T T T A T T G A A A A A T C C A T G T A T C A G A G A T G T T G C T G G T 1400	
RDVER1. SEQ TTCTGTTGAAAATCCATGTATCCGCGATGTCGCTGTGT 1400	
RDVER2. SEQ TTCTGTTGAAAATCCATGTATCCGCGATGTCGCTTGGT 1400	
RDVER3. SEQ TTCTGTTGAAAATCCATGCATCCGTTGATGTCGCTGTGGT 1400	
ROVER4. SEQ TTCTGTTGAAAATCCATGCATTCGCGATGTCGCTGTGGT 1400	
RDVERS. SEQ TTCTGTTGAAAATCCATGCATTCGCGATGTCGCTGTGGT 1400	
RD7. SEQ TTCTGTTGAAAATCCATGCATTCGCGATGTCGCTGTGGT 1400	
ROVER51. SEQ T T C T G T T G A A A A A T C C A T G C A T T C G C G A T G T C G C T G T G G T 1400	
RDVER52.SEQ TT CT GTT GAAAAAT CCAT GCAT T CGC GAT GT CG CT GT G GT 1400	
RD1561H9.SEQT TCTGT TGAAAAAT CCATGCATTCGCGATGTCGCTGTGGT 1400	
The state of the s	
GRVER51. SEQ GGGTATCCCAGACTTGGAAGCTGGCGAGTTGCCTAGCGCC 1440	
CDMPDE CRO ICIO O PINICIPIA AIGINA AIGIN	
GRVER4. SEQ GGGTATCCCAGGACTTGGGAAGCTGGCGAGTTGCCTAGCGCC1440	
GRVER3. SEQ GGGTATCCCAGACTTGGAAGCTGGCGAGTTGCCTAGCGCC1440	
GRVER1. SEQ GG G T A T C C C A G A T T T T GG G A G C T G G C G T G C C T A G C G C C 1440	
YG81-6G1.SEQT G G T A T T C C T G A T C T A G A A G C T G G A A A C T G C C A T C T G C G 1440	
RDVERL SEQ CIG GCIATTCCTGACCTGGAGGCCCGGTGAATTCCCATCTGCT	
=	
a lala alalia a a a a a a a a a a a a a	
RDVER52.SEQ CG G CATTCCTGATCTGGAGGCCGGCGAACTGCCTTCTGCT 1440	
RD1561H9.SEQCGGCATTCCTGATCTGGAGGCCGGCGAACTGCCTTCTGCT 1440	

Figure 2 (cont.)

GRVERS1. SEQ T T T G T G G T G A A A C A A C C C G G C A A G G A G A
GR6.SEQ TTTGTGGT G AAACA A CCCGG C AAGGAGAT C AC T GCTAA G G 1480
GRVERS. SEQ TTTGTGGT[G]AAACA A CCCGG C AAGGAGAT C AC T GCTAA G G 1480
GRVER4. SEQ TTTGTGGTGAAACAACCTGGAAAGGAGATCACTGCTAAGG 1480
GRVER3. SEQ TITGTGGTGAAACAACCTGGCAAGGAGATTACTGCTAAGG 1480
GRVERZ. SEQ TITGICGIGAAACAACCAGGCAAGGAAATTACCGCTAAAG 1880
GRVER1.SEQ TTTGTCGTGAAACAACCAGGTAAGGAAATTACCGCTAAAG 1480
YG81-6G1. SEQT T T G T G G T T A A A C A G C C C G G A A A G G A G A T T A C A G C T A A A G 1480
RDVERL SEQ T TEGT G G TEA AGE A G C CTG G CA A AG A G A TEA CTG CEA AGG 1480
RDVERZ. SEQ TTCGTGGTCAAGCAGCCTGGTAAAGAGATCACTGCCAAGG 1480
RDVER3. SEQ TTCGTCGTCAAGCAGCCTGGTAAAGAAATCACCGCCAAAG 1480
RDVER4. SEQ TTCGTTGTCAAGCAGCCTGGTAAAGAAATTACCGCCAAAG 1480
RDVER5. SEQ T T C G T T G T C A A G C C T G G T A A A G A A T T A C C G C C A A G 1480
and the first of the first of the state of t
RD7. SEQ TTCGTTGTCAAGCAGCCTGGTAAAGAATTACCGCCAAAG 1480
RDVER51. SEQ TTCGTTGTCAAGCAGCCTGGTAAAGAATTACCGCCAAAG 1480
RDVER52. SEQ TTCGTTGTCAAGCAGCCTGGTAAAGAATTACCGCCAAAG 1480
RD1561H9.SEQTTCGTTGTCAAGCAGCCTGGTACAGAAATTACCGCCAAAG 1480
GRVER51. SEQ AGG TOT A CG ACT A TITTGG C CG A GOGGG TGT CTC ACA COA A 1520
GR6. SEQ A G G T C T A C G A C T A T T T G G C C G A G C G T G T C T C A C A C C A A 1520
GRVER5.SEQ A G G T C T A C G A C T A T T T G G C C G A G C G T G T C T C A C A C C A A 1520
GRVER4.SEQ AGGTCTACGACTATTTGGCCGAGCGCGTGTCTCACACCAA 1520
GRVER3.SEQ A G G T C T A C G A C T A T T T G G C C G A G C G C G T C T C A C A C T A A 1520
GRVER2.SEQ A G G T C T A C G A C T A T T T G G C C G A C G C G C T G T C T C A C A C T A A 1520
GRVER1. SEQ AGG T CT A C G A C T A T T T G G C C G A A C G C G T G T C T C A C A C T A A 1520
YG81-6G1.SEQAAGTGTACGATTATCTTGCCGAGAGGGTC <u>TC</u> CCATAC <u>A</u> AA 1520
RDVER1. SEQ AAGTGTATGATTACCTGGCTGAGCGTGTCAGCCATACCAA 1520
RDVER2.SEQ AAGTGTATGATTACCTCGCCTGAACGTGTCAGCCATACCAA 1520
RDVER3.SEQ AAGTGTA T GATTA C CT G GC T GA AC G T GT GAG CCATAC C AA 1520
RDVER4.SEQ AAGTGTATGATTACCTGGCTGAACGTGTGAGCCATACTAA 1520
RDVER5.SEQ AAGTGTA T GATTA C CT G GC T GA AC G T GAG CCATAC T AA 1520
RD7. SEQ AAGTGTA T GATTA C CT G GC T GA AC G T GT GAG CCATAC T AA 1520
RDVER51. SEQ AAGTGTATGATTACCTGGCTGAACGTGTGTGAGCCATACTAA 1520
RDVER52. SEQ AAGTGTATGATTACCTGGCTGALACGTGTGTGAGCCATACTAA 1520
RD1561H9.SEQAAGTGTATGATTACCCTGGCTGAACGTGTGTGAGCCATACTAA 1520
GRVERS1. SEQ ATATCT GCGTGGCGGCGTCCGCTTCGTCGATTCTATTCCA 1560
GR6.SEQ ATATICITGCGTGGICIGGICIGTICICGTTCGATITCTATITCCA 1560
GRVERS. SEQ ATATICIT G C G T G G C G G C G T C C G C T T C G T C G A T T C T A T T C C A 1560
GRVER4. SEQ ATATICIT G C G T G G C G G C C G T C C G C T T C G T C G A T T C C A T C C C A 1560
GRVER3. SEQ. AT A TICT G C G T G G C G G C G T C C G C T T C G T C G A T T C T A T C C C T 1560
GRVER2.SEQ GTACCTGCGTGGCGGTGTCCCGCTTCGTCGATAGCATCCCT1 1560
GRVER1.SEQ GTACCTGGGTGGCGGTGTCCGGCTTCGTGGATAGCATCCCT 1560
YG81-6G1.SEQGTATTTGCGTGGAGGGGTTCGATTGATAGCATACCA 1560
RDVER1. SEQ ATATTTGCGCGGTGGCGTGCGTTTTTGTCGACTCTATTCCA 1560
RDVER2. SEQ ATATTTGCGCGCGGTGGGGTTTTTGTGGACTCTATTCCA 1560
RDVER3.SEQ GTACTTGCGTGGCGTGCGTTTTTGTGGGACAGCATTCCA 1560
RDVER4.SEQ GTACTTGCGTGGCGGCGTGCGTTTTTGTGGATAGCATTCCT 1560
RDVERS. SEQ GTACTTGCGTGGCGGCGTGCGTTTTTGTTGACTCCATCCCT 1560
RD7. SEQ GTACTTGCGTGGCGGCGTGCGTTTTTGTTGACTCCATCCCTT 1560
RDVERSI. SEQ GTACTTGCGTGGCGGCGTGCGTTTTTGTTGACTCCATCCCTT 1560
RDVER52. SEQ GTACTTGCGTGGCGGCGTGCGTTTTGTTGACTCCATCCCTT 1560
RD1561H9. SEQG T A C T T G C G T G G C G T G C G T T T T G T T G A C T C C A T C C C T 1560

Figure 2 (cont.)

GRVER51. SEQ C G C A A C G T T A C C G G T A A G A T C A C T C G T A A A A G A G T T G C T G A 1600 GRVER5. SEQ C G C A A C G T T A C C C G G T A A G A T C A C T C G T A A A A G A G T T G C T G A 1600 GRVER5. SEQ C G C A A C G T T A C C C G G T A A G A T C A C T C G T A A A A G A G T T G C T G A 1600 GRVER4. SEQ C G C C A A C G T G A C C G G T A A G A T C A C T C G T A A A G A G T T G C T G A 1600 GRVER4. SEQ C G C C A A C G T C A C C G G C A A G A T C A C T C G T A A A G A G T T G C T G A 1600 GRVER4. SEQ C G C C A A C G T C A C C C G G C A A G A T C A C T C G T A A A A G A G T T G C T G A 1600 GRVER4. SEQ C G C C A A T G T C A C C C G G C A A A A A T T A C T C G T A A A G A G T T G C T G A 1600 GRVER4. SEQ C G C A A T G T C A C C G G C A A A A T T A C T C G T A A G A G T T G C T G A 1600 GRVER4. SEQ C G C A A T G T C A C C G G C A A A A T T A C T C G T A A G A G G T T G C T G A 1600 GRVER1. SEQ C G C A A T G T C A C C G G C A A A A T T A C T C G T A A G A C T G T G A 1600 ROVER1. SEQ C C G T A A C G T G G T A A G A T C A C C C G C A A A G A A C T G T T G A 1600 ROVER3. SEQ C C G T A A C G T G G T A A G A T C A C C C G C A A A G A A C T G T T G A 1600 ROVER3. SEQ C C G T A A T G T G A C T G G T A A G A T T A C C C C G C A A A G A A C T G T T G A 1600 ROVER3. SEQ C C G T A A T G T G A C T G G T A A A A T T A C C C C G C A A A G A A C T G T T G A 1600 ROVER3. SEQ C C G T A A T G T G A C T G G C A A A A T T A C C C C G C A A G G A C T G T T G A 1600 ROVER3. SEQ C C G T A C G T G A C T G G C A A A A T T A C C C C G C A A G G A C T G T T G A 1600 ROVER3. SEQ C C G T A C G C T G T T G A C T G G C A A A A T T A C C C C G C A A G G A G C T G T T G A 1600 ROVER3. SEQ C C G T A C G C T G T T G A 1600 ROVER3. SEQ C C G T A C G C T G T T G A 1600 ROVER3. SEQ C C G T A C G C T A A C G C C A A A A T T A C C C C G C A A G G A G C T G T T G A 1600 ROVER3. SEQ C C G T A C G C T A C G C C A A A T T A C C C C G C A A G G A G C T G T T G A 1600 ROVER3. SEQ C C G T A C C T A C G C C A A A A T T A C C C C G C A A G G C C T T T G A		`	′											
GRVER4. SEQ	GRVER51.SEQ	CGC	AAC	GTT	ACC	GGTA	AGA	TCA	CTC	GTAA	AGAG	TGCT	G A 1600	
GRVERA1. SEQ	_	cec	AAC	GTT	ACC	GGTA	AGA	TCA	CTC	G T A A	A G A G 1	TGCT	G A 1600	
GRVERAS.SEQ C G C A C C G C A C C G C A A G A T T C G T G C T G A C G C A A A A T T C G T A A G A G T T G C T G A 1600 GRVERAI.SEQ C G C A A T G T C A C C G G C A A A A T T C G T A A G A G T T G C T G A 1600 ROVERAI.SEQ G G A A T G T G A C G G C A A A A T T C C G G A A A G A C T G T G A 1600 ROVERAI.SEQ C G T A A G T G G T A A A A T T A C C G C A A A A A T T C T G A A G A C T G T T G A 1600 ROVERAI.SEQ C G T A A G T G A C T G T A A A A T T A C C G C A A G A C T T T G A 1600 ROVERAI.SEQ C G T A A G T G A C T G T A A A A T T A C C G C A A G G T T G A 1600 ROVERAI.SEQ C G T A A C G T A A A A T T A C C G C A A G G T G T T G A 1600 ROVERAI.SEQ C G T A A C G T A A A A T T A C C G C A A G G G T G T T G A 1600 ROVERAI.SEQ C G T A A C G T A A A A T T A C C G C A A G G G G T T G A 1600 ROVERSI.SEQ C G T A A C G T A A A A T T A C C G C A A G G G G G G G G	_	CGC	A A C	GTI	ACC	GGTA	AGA	TCA	CTC	GTAA.	A G A G 1	TGCT	G A 1600	
GRVER1.SEQ C G C A A T G T C A C C G G C A A A A A T T A C T C G T A A G G T T G G C G A A A A T T A C T C G T A A G G T T G A 1600 ROVER1.SEQ C G T A A C G T A A A A A A A A A	_	CGC	A A C	G T G	ACC	GGTA	AGA	TCA	CITC	GTAA	A G A A 1	г д Б С т	G A 1600	
GRVER1.SEQ C G C A A T G T C A C C G G C A A A A A T T A C T C G T A A G G T T G G C G A A A A T T A C T C G T A A G G T T G A 1600 ROVER1.SEQ C G T A A C G T A A A A A A A A A	-	CGC	A A C]GT C	ACC	GGCA	AGA	TCA	CTC	G T A A L	A G A G 1	г т Б С т	G A 1600	
YGS1-6G1. SEQA G G A A T G T T A C A G G T A A A A T T A C A A G A A A G G A A C T T C T G A 1600 RDVER1. SEQ	_	CGC	AAT	GTC	ACC	GGCA	AAA	TTA	CITC	GITAA	GGAG1	r Tigle 🕆 .	G A 1600	
ROVER1. SEQ C G TTA A C G T G A C T G G T A A G A T C A C C G G C A A A G A A C T G T T G A 1600 ROVER3. SEQ C G T A A C G T G A C T G G T A A A A A T T A C C C G C A A G G A C T G T T G A 1600 ROVER3. SEQ C G T A A C G T G A C T G G T A A A A T T A C C C G C A A G G A C T G T T G A 1600 ROVER3. SEQ C G T A A C G T A A C A G G C A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVER3. SEQ C G T A A C G T A A C A G G C A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVER3. SEQ C G T A A C G T A A C A G G C A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVER3. SEQ C G T A A C G T A A C A G G C A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVER3. SEQ C G T A A C G T A A C A G G C A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVER3. SEQ C G T A A C G T A A C A G G C A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVER3. SEQ C G T A A C G T A A C A G G C A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVER3. SEQ C G T A A C T C C T C G A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVER3. SEQ A G C T A A C T C C T C G A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVER3. SEQ A G C A A C T C C T C G A A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVER3. SEQ A G C A A C T C C T C G A A A A A A A G C T G G C G C G C G C G C G C G C G C	GRVER1.SEQ	CGC	AAT	GIC	ACC	G G C A	AAA	TTA	CTC	GTAA (GGAG 1	гтыст.	G A 1600	
ROVER3. SEQ C G T A A C G T G A C T G G T A A G A T T A C C C G C A A A G A A C T G T T G A 1600 ROVERS. SEQ C G T A A A G T G T G G C A A A A A T T A C C C G C A A G G A A C T G T T G A 1600 ROVERS. SEQ C G T A A C G T A A A A T T A C C C G C A A G G A G T G T T G A 1600 ROVERS. SEQ C G T A A C G T A A C A G G C A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVERS. SEQ C G T A A C G T A A C A G G C A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVERS. SEQ C G T A A C G T A A C A G G C A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVERS. SEQ C G T A A C G T A A C A G G C A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVERS. SEQ C G T A A C G T A A C A G G C A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVERS. SEQ C G T A A C G T A A C A G G C A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVERS. SEQ C G T A A C T C C T C G A A A A A T T A C C C G C A A G G A G C T G T T G A 1600 ROVERS. SEQ A G C A A C T C C T C G A A A A A A A A C C C G C A A G G A G C T G T T G A 1600 ROVERS. SEQ A G C A A C T C C T C G A A A A A A A A A C C C G C A A G G A G C T G T T G A 1600 ROVERS. SEQ A G C A A C T C C T C G A A A A A A A A G C T G G C G		QAGG	AAT	GTT	ACA	GGTA	AAA	TTA	CAA	GAAA 9	GGAAC	сттст	G A 1600	
ROVER4. SEQ	_	CGT	AAC	GTG	ACT	GGTA	A G A	TCA	CCC	GCAA	AGAAC	TGTT	G A 1600	
ROVER4. SEQ	_	CGT	AALC	GTG	ACT	GGTA	AGA	TCA	CCC	G C A A	AGAAC	т G т т	G A 1600	
ROVERS. SEQ	_	CGT	ААТ	GTG	A C T	GGTA	AAA	TTA	cic cic	GICIAA	GGAAC	ствтт	G A 1600	
RD7.SEQ	_	CGC	A A <u>T</u>	GTG	ACT	GGCA	AAA	T T A	c c c c	GCAA	GGAGC	ст в т т	G A 1600	
RDVER51.SEQ	_	CGT	AAC	GTA	ACA	GGCA	AAA	TTA	c c c c	GCAA	GGAGG	т стт	G A 1600	
ROVER52.SEQ C G T A A C G T A A C A G G C A A A T T A C C G C A A G G G T T G A 1600 GRUER51.SEQ A G C A A C T C C T C G A A A A G C T G G G G C GRVER51.SEQ A G C A A C T C C T C G A A A A G C T G G G G C GRVER51.SEQ A G C A C T C C T C G A A A A G C T G G G G C GRVER51.SEQ A G C A C T C C T C G A A A A G C T G G G G C GRVER51.SEQ A G C A C T C C T C G A A A A G C T G G G G C GRVER51.SEQ A G C A C T C C T C G A A A A G C T G G G G C GRVER51.SEQ A A C A C T C C T C G A A A A G C T G G G G C GRVER51.SEQ A A C A C T G G A A A A G C T G G G G C GRVER51.SEQ A A C A C T G G A A A A G C T G G G G G C GRVER51.SEQ A A C A C T G T T G G G G G G		CGT	A A C	GTA	ACA	GGCA	AAA	T T A	c c c c	G C A A (GGAGG	т втт	G A 1600	
ROVER52.SEQ C G T A A C G T A A C A G G C A A A T T A C C G C A A G G G T T G A 1600 GRUER51.SEQ A G C A A C T C C T C G A A A A G C T G G G G C GRVER51.SEQ A G C A A C T C C T C G A A A A G C T G G G G C GRVER51.SEQ A G C A C T C C T C G A A A A G C T G G G G C GRVER51.SEQ A G C A C T C C T C G A A A A G C T G G G G C GRVER51.SEQ A G C A C T C C T C G A A A A G C T G G G G C GRVER51.SEQ A G C A C T C C T C G A A A A G C T G G G G C GRVER51.SEQ A A C A C T C C T C G A A A A G C T G G G G C GRVER51.SEQ A A C A C T G G A A A A G C T G G G G C GRVER51.SEQ A A C A C T G G A A A A G C T G G G G G C GRVER51.SEQ A A C A C T G T T G G G G G G	RDVER51.SEQ	CGT	A A C	GTA	ACA	GGCA	AAA	T T A	clc clc	CAAC	GGAGG	т стт	G A 1600	
RD1561H9.SEQCGTA A C G T A A C A G G C A A A A T T A C C C G C A A G G C T G T G A 1600 GRVER51.SEQ A G C A A C T C C T C G A A A A A A G C T G G C G G C 1626 GRVER5.SEQ A G C A A C T C C T C G A A A A A A G C T G G C G G C 1626 GRVER5.SEQ A G C A A C T C C T C G A A A A A A G C T G G C G G C 1626 GRVER4.SEQ A G C A A C T C C T C G A A A A A A G C T G G C G G C 1626 GRVER3.SEQ A A C A A C T C C T C G A A A A A A G C T G G C G G C 1626 GRVER3.SEQ A A C A A T T G C T C G A A A A A G C T G G C G G C 1626 GRVER1.SEQ A A C A A T T G C T G G A A A A G G C T G G T G G C G G C 1626 ROVER1.SEQ A G C A G T T G C T G G A A A A A G G C T G G T G G C	RDVER52.SEQ	CGT	A A C	GTA	ACA	G G C A	AAA	тта	cle ele	SICIAA	g g Algic	т в т т	G A 1600	
GROUPERS. SEQ A G C A A C T C C C T C G A A A A A G C T G G C G G C GROVERS. SEQ A G C A A C T C C T C G A A A A A G C T G G C G G C GROVERS. SEQ A G C A A C T C C T C G A A A A A G C T G G C G G C GROVERS. SEQ A A C A A T T G C T G G A A A A A G C T G G C G	RD1561H9.SE(OCCIT.	A A C	G T A	ACA	G G C A	A A A	T T A	c <u>c c</u> c	GCAA (GGAGC	тстт	G A 1600	
GROUPERS. SEQ A G C A A C T C C C T C G A A A A A G C T G G C G G C GROVERS. SEQ A G C A A C T C C T C G A A A A A G C T G G C G G C GROVERS. SEQ A G C A A C T C C T C G A A A A A G C T G G C G G C GROVERS. SEQ A A C A A T T G C T G G A A A A A G C T G G C G	CRIMBEL CRO			م اما م	-6-		I	□I		_				
GRVER5.SEQ A G C A A C T C C T C G A A A A A G C T G G C G G C GRVER4.SEQ A G C A A C T C C T C G A A A A A G C T G G C G G C GRVER3.SEQ A A C A A C T C C T C G A A A A A G C T G G C G G C GRVER3.SEQ A A C A A T T G C T C G A A A A A G C T G G C G G C GRVER4.SEQ A A C A G T T G C T G G A A A A G G C T G G T G G C GRVER1.SEQ A A C A G T T G C T G G A A A A G G C T G G T G G C ROVER1.SEQ A G C A A C T G T T G G A G A A A G G C C G G G G G T ROVER2.SEQ A G C A A C T G T T G G A G A A A G C C G G G G G T 1626 ROVER3.SEQ A G C A A T T G T T G G A G A A A G G C C G G C G G T 1626 ROVER5.SEQ A G C A A T T G T T G G A G A A G G C C G G C G T 1626 ROVER5.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 ROVER5.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 ROVER5.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 ROVER5.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 ROVER5.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 ROVER5.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 ROVER51.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 ROVER51.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 ROVER52.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 ROVER52.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626														
GRVER4.SEQ	-													
GRVER3.SEQ	_													
SRVER2.SEQ A A C A G T T G C T G G A A A A G G C T G G T G G C 1626 SRVER1.SEQ A A C A G T T G C T G G A A A A G G C T G G T G G C 1626 RDVER1.SEQ A G C A A C T G T T G G A G A A A G C C G G C G G T 1626 RDVER3.SEQ A G C A A T T G T T G G A G A A A G C C G G C G G T 1626 RDVER3.SEQ A G C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER4.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER5.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER5.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER5.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER5.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER5.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER5.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER5.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER55.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER55.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER55.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER55.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626	_													
A A C A A T T G T T G G A A A A A G C C G G G A G T 1626 RDVER1.SEQ A G C A A C T G T T G G A G A A A G C C G G G G G T 1626 RDVER2.SEQ A G C A A C T G T T G G A G A A A G C C G G C G G T 1626 RDVER3.SEQ A G C A A T T G T T G G A G A A G C C G G C G G T 1626 RDVER4.SEQ A A C A A T T G T T G G A G A A G C C G G C G G T 1626 RDVER5.SEQ A A C A A T T G T T G G A G A A G C C G G C G G T 1626 RDVER5.SEQ A A C A A T T G T T G G A G A A G C C G G C G G T 1626 RDVER5.SEQ A A C A A T T G T T G G A G A A G C C G G C G G T 1626 RDVER5.SEQ A A C A A T T G T T G G A G A A G C C G G C G G T 1626 RDVER51.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER52.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER52.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER52.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER52.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626	_	AAAC	AAT	TGC	TCG	AAA	AJGC	TIGGI	CIG GIG					
1626 1626	-	AAA	AGT	TGC	TGG	AAAA	GGC	TIG G	TIE GIC	7				
REVERT. SEQ A G C A A C T G T T G G A G A A A G C C G G C G G T 1626 REVERZ. SEQ A G C A A C T G T T G G A G A A A G C C G G C G G T 1626 REVERZ. SEQ A G C A A T T G T T G G A G A A G G C C G G C G G T 1626 REVERS. SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 REVERS. SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 REVERS. SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 REVERS. SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 REVERS. SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 REVERS. SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 REVERS. SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 REVERS. SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 REVERS. SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 REVERS. SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626	-	A A A C	AGT	TGC	TGG	AAAA	GGC	<u>T</u> JG G[TIGGE	3				
RDVER2.SEQ A G C A A C T G T T G G A G A A A G C C G G C G G T 1626 RDVER3.SEQ A G C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER4.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER5.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER51.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER51.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER52.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626 RDVER52.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626														
RDVER3.SEQ														
RDVER4.SEQ AACAATTGTTGGAGAAGGCCGGCCGGT 1626 RDVER5.SEQ AACAATTGTTGGAGAAGGCCGGCCGGT 1626 RD7.SEQ AACAATTGTTTGGAGAAGGCCGGCCGGT 1626 ROVER51.SEQ AACAATTGTTTGGAGAAGGCCGGCCGGT 1626 RDVER52.SEQ AACAATTGTTTGGAGAAGGCCGGCCGGT 1626	_	AGC	AC	I G T	TGG	a G A A	MG C		cie e 1					
RDVER5.SEQ AACAATTGTTGGAGAAGGCCGGCGGT 1626 RD7.SEQ AACAATTGTTGGAGAAGGCCGGCGGT 1626 ROVER51.SEQ AACAATTGTTGGAGAAGGCCGGCGGT 1626 RDVER52.SEQ AACAATTGTTGGAGAAGGCCGGCGGT 1626	_	A G C A	AAT	TGT	TGG	AGAA	GGC	cle el	CGGT	:				
RD7.SEQ AACAATTGTTGGAGAAGGCCGGCGGT 1626 RDVER51.SEQ AACAATTGTTGGAGAAGGCCGGCGGT 1626 RDVER52.SEQ AACAATTGTTGGAGAAGGCCGGCGGT 1626	-													
RDVER51.SEQ AACAATTGTTGGAGAAGGCCGGGCGGT 1626 RDVER52.SEQ AACAATTGTTGGAGAAGGCCGGCGGT 1626	_													
RDVER52.SEQ A A C A A T T G T T G G A G A A G G C C G G C G G T 1626	_													
1626														
	VDIDGINA. SEC	AWENCI	AAT	TGT	TGG	IGA A	GGC	Cle el	CIGGT	ı			1626	

Figure 3

<u>-</u>	<u> </u>	_																																	
GRVER51.SEQ	М	M	ĸ	R	E I	K b	1 A	1	Y	G	P	5 P	Ļ	н	P	L	Ē	D	L 1	r A	G	Ë	M :	L F	R	A I	L P	K	Н	S	H L	P	Q A	118)
GR6.SEQ	M	M	K	R	E 1	K	1 A	I	Y	G	P	P	L	н	P	L	Ε	D :	L 7	ГΑ	G	Ε	M :	L F	R '	A I	LF	K	н	S	H L	P	QA	118	j
GRVER5.SEQ	М	M	K	R	E	K 1	1 A	I	Y	G :	PI	P	L	Н	₽	L	Ε	D :	L 1	ΓА	G	E	М	L E	R	A 1	LP	K	H	5	H L	P	QA	118)
GRVER4.SEQ	М	M	K	R	E I	K N	1 V	I	Y	G	P 1	. P	L	н	P	L	E	D :	LI	ΓА	G	Ε	M	L F	R	A I	LP	K	н	S	H L	P	QA	118	ì
GRVER3.SEQ	М	M	ĸ	R	E	K 1	1 A	I	Y	G	P	: P	L	н	P	L	E	D :	L 1	r A	G	Ε	M	L F	R	A I	LR	K	н	S	H L	₽	QA	118	J
GRVER2.SEQ	М	M	K	R	£	۲.	J V	I	Y	G :	P	P	L	н	P	L	Ė	D :	LI	ГΑ	G	E	M	LE	R	A I	LR	K	н	S	H L	P	QA	118	J
GRVER1.SEQ	М	M	K	R	E	()	ı v	I	Y	G	P	P	L	н	P	L	E	D :	L 1	· A	G	E	M I	LF	R	A I	LP	ĸ	н	S	нL	P	QA	118	j
YG81-6G1.SEC	M	M	K	R	ΕΙ	()	1 V	I	Y	G :	P	P	L	н	P	L	Ε	D.	L	' A	G	Ε	M I	L F	R	A I	LP	K	Н	S	H L	P	O A	118	ł
ROVER1.SEQ	М	М	K	R	E I	()	ıν	I	Y	G :	P	P	L	н	P	L	Ė	0	L 3	A	G	E	M I	LF	R	A I	LB	ĸ	Н	S !	нL	P	O A	118	ı
ROVER2.SEQ																																		118	
RDVER3.SEQ																																		118	
RDVER4.SEQ																																		118	
RDVER5.SEQ																																		118	
RD7.SEQ																																		118	
RDVER51.SEQ																																			
RDVER52.SEQ																																			
RD1561H9.SEQ																																			
KDIJOINJI DBQ	•••	اث.	••	•				•	•	٠,			-	•••	•	. بد	_				G				•	<i>n</i> 1			п	3 1	7 L		Ų A	118	
GRVER51.SEQ	r	v	D	v	v c	: -		_	1		, ,		F	F		х -	т .	., ,			_		, ,		_	٠.				ъ,				020	
GR6.SEQ																																			
GRVER5.SEQ																																	ΙC	238	
GRVER4.SEQ																																	IC	238	
																																	IC	238	
GRVER3.SEQ																																	ıç	238	
GRVER2.SEQ																																	ΙC	238	
GRVER1.SEQ																																	I C	238	
YG81-6G1.SEQ																																		238	
RDVER1.SEQ																																		238	
RDVER2.SEQ																					_										V	-	-	238	
RDVER3.SEQ																																	I C	238	
RDVER4.SEQ																																	I C	238	
RDVERS.SEQ																																	I C	238	
RD7.SEQ																																		238	
RDVER51.SEQ																																		238	
RDVER52.SEQ																																		238	
RD1561H9.SEQ	L	V	D '	V.	V G	ם	Ε	S	L	S	<i>.</i>	E	F	F	E	A '	ľ	V	L	A	Q	S	LH	N	С	G Y	K	M	N	D /	1 V	S:	I C	238	
GRVER51.SEQ																																		358	
GR6.SEQ																															/ M			358	
																															/ M			358	
GRVER4.SEQ	A :	E	N	N	T F	F	F	I	P	V I	P	. A	W	Y	Ι	G	M :	Ľ۱	/ A	P	٧	N I	E S	Y	1	P D	E	L	С	K V	7 M	G I	(S	358	
GRVER3.SEQ	A :	E	N	N	T F	F	F	I	P	V]	. A	. A	W	Y	I	G I	4	I \	A	P	V	N I	E S	Y	1	P D	E	L	C.	ΚV	/ M	G I	[S	358	
GRVER2.SEQ	A :	E	N	N	T F	l F	F	I	P	V]	. 4	. A	W	Y	I	G I	M	, I	7 A	P	V	N	E S	Y	I	P D	E	L	С	K V	/ M	G I	C S	358	
GRVER1.SEQ	A :	E	N	N	T F	ŀ	F	I	P	V I		A	W	Y	I	G I	4	. I	7 A	P	v	N I	E S	Y	I	P D	E	L	Ç	KΥ	/ M	G 1	[S	358	
YG81-6G1.SEQ	A	E I	N	H	T P	ŀ	F	I	P	V J	I A	. A	W	Y	I	G I	4	I \	7 A	P	٧	N I	E \$	Y	I	P D	E	L	С	K V	/ M	G I	I S	358	
RDVER1.SEQ	A :	Ė	N I	N	T F	F	F	I	P	V I		A	w	Y	I	G I	4	ı ۱	/ A	P	V	N I	E S	Y	I	P D	E	L	С	K V	/ M	G J	I S	358	
RDVER2.SEQ	A :	E I	N I	N	T F	F	F	I	P	V I		A	W	Y	I	G I	4	ľ	/ A	P	V	N I	E S	Y	1	P D	E	L	С	K V	7 M	G :	I S	358	
RDVER3.SEQ	A :	E :	N I	N '	T F	F	F	1	P	v i		A	W	Y	I	G I	4	ι,	7 A	P	V	N I	E S	Y	I	P D	E	L	С	K V	/ M	G :	ı s	358	
ROVER4.SEQ	A :	E I	N I	N '	T F	F	F	ľ	ъ.	v i	. A	A	W	Y	I	G I	4		7 A	P	V	N I	E S	Y	I	P D	E	L	С	K V	7 M	G :	r s	358	
																															M			358	
RD7.SEQ																															/ M			358	
RDVER51.SEQ																																		358	
RDVER52.SEQ																																		358	
RD1561H9.SEQ																																			
						-	-									-				-				-	-	-	_	-					-		

Figure 3 (cont.)

			_	_																																		
GRVER51.SEQ																																						
GR6.SEQ																																						478
GRVER5.SEQ																																						478
GRVER4.SEQ																																						478
GRVER3.SEQ																																						478
GRVER2.SEQ																																					G C	
GRVER1.SEQ	K	Б	Q	1	٧	F	T	T	K	N	I	L	N	K	٧	L	Ε	٧	Q	S	R :	r N	F	I	κ	R	I	1	L	D	T	٧	E	N	I	н	G C	478
YG81-6G1.SEC																																						
RDVER1.SEQ	K	P	Q	I	V	F	Т	T	K	N	I	L	N	K	V	L	E	٧	Q	S	R '	ΓN	F	I	ĸ	R	I :	I	L	D	T	v	Е	N	I	н (G C	478
RDVER2.SEQ	к	P	Q	I	٧	F	T	T	ĸ	N	I	L	N	K	٧	L	E	٧	Q	S	R :	ΓN	E	I	K	R	Į.	I	L	D	T	v	E	N	I	н	G C	478
RDVER3.SEQ	K	P	Q	I	٧	F	T	T	K	N	I	L	N	K	V	L	Ε	٧	Q	S	R :	ľN	F	I	K	R	I	I	L	Đ	T	v	Ε	N	I	н	G C	478
RDVER4.SEQ																																					G C	
RDVER5.SEQ	К	P	Q	I	V	F	Т	T	K	N	I	L	N	ĸ	v	L	E	٧	Q	S	R 1	ΓΝ	F	1	ĸ	R	I)	I	L	D	T	v	Е	N	I	н	G C	478
RD7.SEQ																																					G Ç	
RDVER51.SEQ	K	P	Q	Ι	٧	F	T	T	K	N	I	L	N	K	٧	L	Ε	٧	Q.	S	R 1	r N	E	I	K	R	I 1	I	L	D	т	٧	Е	N	Ι	н	G C	478
RDVER52.SEQ	к	P	Q	I	٧	F	T	T	ĸ	N	I	L	N	ĸ	v	L	Е	٧	ō.	s:	R 1	ı N	F	I	к	R	1 1	ı	L	D	т	ν	E	N	T	н	G C	478
RD1561H9.SEQ	K	₽	Q	1	٧	F	T	T	к	N	I	L	N	к	v	L	E	v .	ō:	s :	R 1	'N	F	т	к	R	T 3	7	Τ.	ח	т	v	E	N	7	н с		478
												-		-		_	_		-			-	-	-	•	•	•	_	_	-	•	•	-	••	•		, .	470
GRVER51.SEQ	Ε	s	L	P	N	F	1	s	R	Y	s	D	G	N	I	А	N	F	ĸ :	P:	î. F	F	D	P	v	F. (, A	Δ	7	τ.	c	ç	ç	c	т.	r c	598
GR6.SEQ	Ξ	s	L	Р	N	F	I	s	R	Ÿ	5	D	G	N	_ T	A	N	F	ĸ	p ·	 1. k	F	D	P	v	F	- ·	,		T	ī	Č				T 1		598
GRVER5.SEQ																																						598
GRVER4.SEQ	E	s	L	P	N	F	I	s	R	Y	S	D	G	N	- T	A	N	F	K '	 P 1	 I. 1-	F	ם	P	v	E (יי אר	, A	_ A	Ť	1	~			٠.	T 1		598
GRVER3.SEQ																																					rg	
GRVER2.SEQ	E	s	L	P	N	F	T	S	R	Ÿ	s	D	G	N	7	Δ	N	F	K	P 1			n	D	7	E (2 4		Δ Δ	+	7	c	S	3 '	G .	r 1		598 598
GRVER1.SEQ																																						598 598
YG81-6G1.SEQ	F	s	7.	P	N	F	Ť	ç	P	Ŷ	9	ח	c	N	Ť	^	LV NJ	E 1		6 . D 1			ם		V :	e (2 4		~			_	S .	2	G :	1.1	. G	598
RDVER1.SEQ																																						
RDVER2.SEQ	E	٥		, D	N		+	۰	D	v	٠	n	c	27	+	~	I.I	r :		r 1				F	V .	E (2 V		A.		-	٠	5	٥ '	: تا	II	فا:	598 598
RDVER3.SEQ	F	٥		D	M	-	+		D	v	0	מ	0	P.I	·	^	IN N	F I					ט	P	V .	- 4	2 V	A	A	1	_	-	5 .	5	G :	r 1) G	598
RDVER4.SEQ	F	۰		r D	м		T	٥	D .	v	٥	ם	6	EN N	7	^	N	E 1		r 1			ע		V 1	E (, v	Α.	A	1	<u>.</u>	Ü.	5	5 (G 1	r I	G	598
RDVER5.SEQ	-		ī		M	E.	+	٥		v	c	ם	G	N N	+		IA IA	E 1		rı			ט	Ρ.	v ,	E (} v	Α.	Α.	_ T	r	Ċ.	S :	S (G 7	r I	G	598
RD7.SEQ	-		t	£.	N	r	<u>+</u>	٥	7	ı	2	ט	<u>د</u>	N	-		N	F 1			. H		ע		v ,	E (2 V	Α.	Α -	1	L	c :	3 :	S (G 1	r I	G	598
	-			P.	М		+	٥		ĭ	5	יי	9	F.	-	^	N	F 1			L H		0	-	v ,	E (} v	A	Α.	_	Tr -	Ç :	S :	5 (G 1	r I	G	598
RDVER51.SEQ	E		7	r r	N		+	٥	ĸ	I	2		6	N	1	A	N	t 1	K 1		ь н	F	0	Ρ.	V 1	EÇ	2 V	A	A	I	r.	C :	S :	S	3 1	ГТ	: G	598
RDVER52.SEQ	-	2						5	-	ľ	5	-	G	N	Τ.	A	D.	- I	к і	P 1	LH		D	P	V 1	E (2 V	A	A	I	L	C	S	S	3 7	ΓŢ	G	598
RD1561H9.SEQ	£	3	L	۲	N	F	1	5	ĸ	Y	S	D	G	N	I	A	N	F	KI	PI	ЬН	F	D	P	V I	Ε (V	Α	A	Ι	L	Ç:	s :	S	3 7	гт	G	598
GRVER51.SEQ	T.	Þ	ĸ	G	v	м	^	т	н	0	N	т	_	17	D	ŧ.	т :		ח ה		۰.	-	G)	<u>.</u>	T /	٠,			_	**		.,		., ,				210
GR6.SEQ	т.	P	ĸ	Č	v	м	۸	·		~	ม	Ť	_	v			1 :	11		, ,	, ,],	~ '	<u></u>	2 .	, ,	á	~	٧	T.	v ,	L 1	v :				718
GRVER5.SEQ	т.	P	ĸ	6	v	M	×	·		<u>۷</u>	LT NI	Ť	~	T			a i				, ,		<u> `</u> ,	G C	- 1 -	2 1		ē	<u>د</u>	٧	T .	v ,	L 1	v :		. P		718
GRVER4.SEQ	T.	D	Ľ	6	v	17 M	×	·	n u	ν.	EV N	+	<u>ر</u>	٧	r.	·	1 1 7 1	n 4			, ,		ĽI	G	T	2 1			6	٧	T .	v ,	ь \ 	v :	r 1			718
																																						718
GRVER2.SEQ	T	E D	r	9	11		ž		п.	•	N	-	ر م	٠.	ĸ		1 1		A 1		, ,	K	ĽI	G	т (3 T	. 1	2	G	v 	T '	v 1	L \	V :	<i>r</i> 1	. P	F	718
_		F		9	٠,		ă	T	n	v ·	N	Ţ		٧	K	L			A L	ייי) P	R	ĽI	G	T	5 T	. I	-	G	v 	T '	V 1	ь ,	V 1	rı	. 2	F	718
GRVER1.SEQ		-		-	٧		ů	T -	H	Q.	N	Ţ	C .	٧	K	L	I i	н А	A I	. [P	R	W	G	т (2 I	. I	P	G	V	T	V 1	r /	V)	ľ	. P	F	718
YG81-6G1.SEQ																																						
RDVER1.SEQ																																						718
																																						718
																																						718
																																						718
																																						718
RD7.SEQ	L	P	K	G	V	М	Q	T	Н	Q I	N	I	C '	V :	R	L	I	1 7	A L	. 0	P	R	Y	G '	T	Į	I	P	G	V	T '	V J	L \	٧)	(I	. P	F	718
RDVER51.SEQ	L	P	K	G	٧	M	Q	T	н	QΙ	N :	I	C 1	V :	R	L	I	1 4	A L		9	R	Y	G:	rç	2 L	I	P	G	v ·	T 1	v .I	LV	V Y	ľ	. Р	F	718
RDVER52.SEQ	L	₽	K	G	٧	M	Q	T	н	QI	N :	I	י ם	V I	R	L	I	1 7	A L	. 0	2	R	Y	G 1	T () L	I	P	G	v ·	r v	V I	LV	VY	/ L	. Р	F	718
RD1561H9.SEQ	L	P	K	G	v	M	Q	T	н	QI	N :	Ι (C 1	V	R	L	I	1 4	۱ L	. 0	P	R	Y	G :	r (L	I	P	G	v	r v	V I	L	7 1	<i>.</i> I	P	F	718
																							_															

Figure 3 (cont.)

				_																														
GRVER51.SEQ	FF	A	F	G	F	S	T	L	G	Y	M	ν	G	L	R١	/ I	M	F	R F	F	D	QE	A	F	LK	Α	I	Q	D Y	Ε	V F	₹ 8	38	_
GR6.SEQ	FF	I A	F	G	F	SI	ΙŢ	L	G '	Y	M	٧	G	L	R	/ I	M	F	RR	F	D	QΕ	Α	F	ĹΚ	Α	I	Q	Y	Ε	V F	₹ 8	38	
GRVER5.SEQ	FH	ΙA	F	G	F	S I	T	L	G	Y	M	٧	G	L	R \	/ I	M	F	RP	F	D	QΕ	Α	F	LK	Α	I	Ō 1	Y	E	V F	₹ 8	38	
GRVER4.SEQ	FH	IA	F	G	F	S 1	Т	L	G :	Y	M	٧	G	L	R	/ I	M	F	R P	F	D	QΕ	Α	F	LK	A	1	Q I	Y	E	V F	8 9	38	
GRVER3.SEQ	FH	A	F	G	F	S I	T	L	G :	YE	M	٧	G	L	RY	/ I	M	F	RR	F	D	QΕ	Α	F :	Ŀĸ	A	I	QI	Y	E	V F	8	38	
GRVER2.SEQ	FH	A	F	G	E	\$ 1	T	L	G :	YE	М	٧	G	L	RV	/ I	M	F	RR	F	D	QΕ	Α	F	LK	A	1	QI	Y	E	V F	8 3	38	
GRVER1.SEQ	FH	A	F	G	F	S I	T	L	G :	YE	M	٧	G	L	RV	/ I	M	F	R R	F	D (QΕ	Α	F	LK	А	1 (Q I	Y	E	V P	8 3	38	
YG81-6G1.SEC)FH	A	F	G	F	<u>s</u> 1	T	L	G :	YE	M	٧	G	L	RV	/ I	M	F	RR	F	D	QE	Α	F	Ŀĸ	A	I	Qf	Y	E	V P	t 8	38	
RDVER1.SEQ	FH	A	F	G	F	H 1	T	L	G 1	Y E	M	٧	G	L .	RV	<i>'</i> I	M	F	RR	F	D (QΕ	Α	F	LK	A	1	Q	Y	E	V R	8	38	
RDVER2.SEQ	FH	A	F	G	F	H I	T	L	G :	YE	M	V	G	L	R V	ľ	M	F	R R	F	D	QΕ	A	F	LK	А	1 (Q	Y	Ε	V R	8	38	
RDVER3.SEQ	F H	Α	F	G	F	H 1	T	L	G Y	Y F	, W	٧	G	L	RV	'I	M	F	R R	F	D (QΕ	Α	F	K	А	I	0 [Y	Ε	V R	8 3	38	
RDVER4.SEQ	FH	A	F	Ģ	F	H I	τ	L	G Y	ľ	M ·	٧	G	L	R V	' I	М	F	R R	F	D (QΕ	Α	F	K	Α	Ι (QI	Y	E	V R	8	38	
RDVER5.SEQ	FH	A	F	G	F	H I	T	L	G Y	/ E	. М	٧	Ģ	L	R V	ľ	М	F	R R	F	D (QΕ	Α	F I	K	Α	Ι (Q [Y	E	V R	8	38	
RD7.SEQ	FΗ	A	F	G	F	4 I	T	L	G Y	ľE	. М	٧	G	L	RV	ľ	М	F I	R R	F	D (QΕ	Α	F I	K	Α	I	2 0	Y	Ē	V R	. 8	38	
RDVER51.SEQ	FH	A	F	G	F	1 1	T	L	G Y	(F	. М	٧	G	LI	R V	ľ	М	FI	R R	F	D (QΕ	Α	FI	K	Α	Ι (2 [Y	Е	V R	. 8	38	
RDVER52.SEQ	FH	Α	F	G	F	ΙI	T	L	GY	ľ E	. М	V	G	LI	Rν	I	М	F 1	R R	F	D (QΕ	A	F I	K	Α	I (2 0	Y	Ε	V R	. 8	38	
RD1561H9.SEC	F H	A	F	G	F	į I	Т	L	G Y	′ F	M	٧	G	LI	Rν	I	M	F	R R	F	D (2 E	A	FI	. к	A	I	2 0	Y	E	V R	. 8	38	
GRVER51.SEQ	s v	I	N	V	P S	S V	I	L	FI	S	K	S	P	L	V D	K	Y	D I	LS	s	L	R E	L	c	: G	A	A	? L	. A	к	ΕV	9:	58	
GR6.SEQ	s v	I	N	٧	P S	v	I	L	FI	s	K	\$	P	L	V D	K	Y	D I	LS	s	LE	RE	L	c c	G	А	A I	PL	A	K	ΕV	9:	58	
GRVER5.SEQ	s v	I	N	V	P S	SV	Ι	L	FI	s	K	s	P	L	V D	K	Y	D I	LS	s	LE	RE	L	c c	G	Α	A I	? L	A	K	Eν	9	58	
GRVER4.SEQ	s v	1	N	٧	P S	5 V	Ι	L	F I	. \$	K	s	P	L	V D	K	Y	D I	LS	s	LF	RE	L	c c	G	A	A E	P L	A	ĸ	ΕV	9:	58	
GRVER3.SEQ	s v	Ι	N	V	PS	V	Ι	L	FI	S	ĸ	S	P	L	ם ע	K	Y	D I	LS	S.	LE	RE	L	¢ c	G	A	A E	P L	A	ĸ	ΕV	99	58	
GRVER2.SEQ	s v	Ι	N	٧	P 5	v	Ι	L	FI	S	K	S	₽ :	L	V D	K	Y	D I	. \$	S	LE	E	L	c c	G	Α.	A E	L	Α	ĸ	EΥ	9	58	
GRVER1.SEQ	s v	Ι	N	V	P 5	V	Ι	L	FI	S	ĸ	s	P :	L١	D V	K	Y	DI	S	S	LE	RΕ	L	c c	G	Α.	A E	L	Α	ĸ	Eν	9:	58	
YG81-6G1.SEQ	s v	I	N	٧	P 5	V	Ι	L	FI	S	K	s	₽ :	L	/ D	K	Y	D I	S	S	L F	E	L	c c	G	Α.	A E	L	А	ĸ	EΥ	9:	58	
RDVER1.SEQ	s v	I	N	V	P S	v	Ι	L	FL	S	K	s	P	LI	/ D	K	Y	O I	S	S	LF	E	L	cc	G	Α.	A E	L	Α	K	ΕV	9:	58	
RDVER2.SEQ	sν	Ι	N	٧	P S	3 V	Ι	L	F L	s	К	s	P :	L	/ D	K	Y	D I	S	S :	LF	E	L	c c	G	Α.	A É	L	A	ĸ.	ΕV	95	58	
RDVER3.SEQ	s v	I	N	V	P S	V	I	L	F L	. \$	К	S	₽ 1	L١	/ D	K	Y	D I	S	S	LF	E	L	c c	G	Α.	A E	L	A	K	E V	95	58	
RDVER4.SEQ	s v	I	N	٧	P S	v	Ι	L	F L	S	K	S	P :	LV	/ D	K	Y	D I	S	S	LF	E	L	c c	G	A.	A P	L	Α	K	E V	95	8	
RDVER5.SEQ	s v	Ι	N	V	P S	V	I	L	F L	, s	K	S	P:	LV	, D	ĸ	Y	D I	s	\$	LF	E	L	cc	G	A.	A F	L	A	K	E V	95	5 8	
RD7.SEQ	s v	Ι	N	V	P S	V	I	L	F L	S	K	S	P	LV	7 D	ĸ	Y	D L	s	S	LF	E	L	cc	G	A i	A P	, r	A	K	E V	95	5 8	
RDVER51.SEQ	s v	Ι	N	٧	P S	v	I	L	F L	s	K	S	P 1	LV	, D	ĸ	Y	D I	s	S	LA	E	L	cc	G	A	A P	L	A	K I	E V	95	8	
RDVER52.SEQ	s v	I	N	V	P S	v	I	L	F L	S	K	S	P 1	LV	D	ĸ	Y	D I	. \$	S	LF	E	L (cc	G	A	A P	, r	A	K	E V	95	8	
RD1561H9.SEQ	s v	I	N	V	P S	v	I	L	F L	. \$	ĸ	\$	P 1	LV	D	ĸ	Y.	D I	s	S	J	E	L	C	G	A	A P	r	A	ĸ	e v	95	88	
GRVER51.SEQ	AΕ	٧	Α.	A.	K P	L	N	L	ΡG	I	R	С	G 1	E G	L	T	E	s t	S	A	ı	H	SI	L R	D	E I	F K	S	G	S	L G	10	78	
GR6.SEQ	ΑE	٧	A.	A :	KR	L	И	L	ΡĢ	I	R	C	G 1	FG	L	T	E	ST	S	A 1	1	H	S I	L R	D	E I	FK	S	G	S	G	10	178	
GRVER5.SEQ	ΑE	٧	A.	A :	K R	L	N	L.	PG	I	R	C ·	G I	F G	L	T	E	SI	S	A i	1	Н	SI	L R	D	E I	FK	S	G	S	LG	10	78	
GRVER4.SEQ	ΑE	٧	A.	A I	K P	L	N	L :	PG	I	R	C	G I	FG	L	Ť	E :	s T	s	A i	1	H	S I	L R	D	E I	F K	S	G	S	L G	10	78	
GRVER3.SEQ	ΑE	٧	A.	A	K R	L	N	L	PG	Ι	R	C	GI	FG	L	T	E :	ST	S	A i	Į	Н	SI	. R	D	E I	FK	S	G	S	LG	10	78	
GRVER2.SEQ	ΑE	٧	A.	A I	K R	L	N	L :	PG	I	R	C	G I	FG	L	T	E :	ST	\$	A I	I	H	S I	. R	D	E	FK	S	G	S 1	LG	10	78	
GRVER1.SEQ	ΑE	٧	A	A I	KR	L	N	L	ΡG	I	R	C	G I	FG	L	T	E :	S T	s	A i	ı,	н	S I	R	D	E 1	FK	S	G	S I	G	10	78	
YG81-6G1.SEQ	ΑE	V	A A	A I	K R	L	N	L	PG	I	R	C	G I	FG	L	T	E :	S T	s	A M	I	H	S I	R	D	E	FK	S	G	S 1	L G	10	78	
RDVER1.SEQ	ΑE	٧	A.	A I	KR	L	N	L	PG	I	R	C	G I	FG	L	T	E :	S T	S	Αĺ	Ī	ଢ	SI	R	D	E I	e K	S	G	s I	G	10	78	
RDVER2.SEQ	ΑE	V	A i	A I	K R	L	N	L	PG	I	R	C	G I	r G	L	T	E :	s T	s	A I	I	0	s I	R	D	E I	FK	S	G	S 1	G	10	78	
RDVER3.SEQ	AΕ	٧	A i	A I	KR	L	N	L	ΡG	I	R ·	C	GI	G	L	T	E :	ST	S	Alı	ılı	lol	SI	R	D	E I	· K	S	G	S 1	G	10	78	
RDVER4.SEQ	ΑE	V	A A	A I	K R	L	N	L	PG	I	R	¢	G E	G	L	T	E :	S T	S	A I	ijΙ		S I	R	D	E 1	FK	S	G	S 1	LG	10	78	
RDVER5.SEQ	AΕ	V	A	ΑI	ΚR	L	N	L	ΡG	I	R	C	GE	G	L	Ţ	E S	5 T	S	Alı	ıΊΙ	lol	SI	R	D	E I	K	S	G	SI	G	10	78	
RD7.SEQ	ΑE	٧	A A	A I	K R	L	N	L	PG	I	R	C	GE	G	L	T	E S	S T	S	A	I		SI	R	D	E I	K	s	G	S I	G	10	78	
RDVER51.SEQ	AΕ	٧	A A	A I	< R	L	N	L	PG	1	R I	C	G E	G	L	Т	E 5	S T	S	AI	I	lol	SI	R	D	E I	K	S	G	S I	G	10	78	
RDVER52.SEQ	ΑE	٧	A A	A I	C R	L	N	L	PG	I	R	c	GE	G	L	T	E 5	S T	S	A	ı	lal	S I	٠G	D.	E	K	s	G	s 1	G	10	78	
RD1561H9.SEQ	ΑE	V	A A	A 1	C R	L	N	LI	PG	I	R	c (G E	G	L	T	E S	з т	s	A	1	۵	TI	٠LG	D:	E	K	s	G	s I	G	10	78	
																				_	_				•									

Figure 3 (cont.)

<u>-</u>	_			_																					_													
GRVER51.SEQ	R	٧	T	F	, 1		1 F	A A	K	ï	A	D	R	E	T	G	К	A	L	G	P	N	Q V	7 G	Ε	L	C		G	P	М	v	s	K	3 Y	v	N	1198
GR6.SEQ	R	٧	T		, 1		1 7	A A	K	I	Α	D	R	E	T	G	ĸ	А	L	G	P	N (QV	7 G	E	L	C I	C #	G	P	М	v	s	K (G Y	v	N	1198
GRVER5.SEQ	R	٧	T	F	, 1		1 7	A A	K	Ĺ	Α	D	R	E	Ţ	G	K	A	L	G	P	N (Q V	/ G	E	L	C I		G	P	М	٧	s	K (3 Y	v	N	1198
GRVER4.SEQ	R	٧	Т	, 6) I		1	A A	K	I	Α	D	R	Ε	T	G	K	A	L	G	P	N (Q V	G	E	L	C I		G	P	M	V	s	ĸ	3 Y	v	N	1198
GRVER3.SEQ	R	٧	Ţ	. 6	ľ	. M	1	A	K	Ι	A	D	R	ε	T	G	К	A	L	G	P	N (Q V	G	E	L	C I		G	P	М	٧	S	ĸ	3 Y	v	N	1198
GRVER2.SEQ	R	٧	T	P	1	. M	A	A A	K	: I	A	D	R	E	Т	G	ĸ	A	L	G	P	N (Q V	G	E	L	C I	I N	G	₽	М	v	s	ĸ (3 Y	v :	N	1198
GRVER1.SEQ	R	v	T	P	ı	. 14	P	A A	К	: 1	A	٥	R	E	T	Ģ	ĸ	A	L	G	P	N (2 V	G	E	L	C I	K	G	₽	M	v	s	ĸ	3 Y	v	N	1198
YG81-6G1.SEC	R	٧	T	P	L	M	A	A	K	I	A	D	R	E	Т	G	ĸ	A	L	G	P	N (Q V	G	E	L	C I	K	G	P	М	v	S	ĸ	3 Y	v	N	1198
RDVER1.SEQ	R	٧	T	P	L	. M	A	۱ A	K	I	A	D	R	Ε	т	G	ĸ	A	L	G	P	N (Q V	G	E	L	C I	K	G	₽	М	ν	s :	ĸ	3 Y	v	N	1198
RDVER2.SEQ	R	٧	T	P	L	M	A	A	K	Ι	A	D	R	Ē	т	G	ĸ	A	L	G	P	N (o v	G	E	L	C I	K	G	P	М	ν	s	KC	3 Y	v	N	1198
RDVER3.SEQ	R	٧	T	P	L	. м	A	A	K	I	Α	Þ	R	E	т	G	K	Α	L	G	P	N (2 V	G	E	L	C I	K	G	P	М	v	s	ĸ	Y	V	N	1198
RDVER4.SEQ	R	V	T	P	L	, м	Α	A	K	I	Α	D	R	Е	т	G	K	Α	L	G	Р	N (o v	G	Ε	L	CI	K	G	P	М	v	S	ĸc		V	N	1198
RDVER5.SEQ	R	٧	Т	P	L	. М	Α	A	ĸ	I	Α	D	R	E	т	G	ĸ	A	L	G	Р	N C	o v	G	E	t.	СІ	K	G	P	м	v	S	ĸ c	. v	v	N	1198
RD7.SEQ	R	v	т	p	Ĺ	. м	A	A	K	I	А	D	R	E	т	G	ĸ	A	L	G	Р	N C	- > v	G	E	t.	C I	K	G	P	м	v	SI	ĸ c		17 1	NJ	1198
RDVER51.SEQ	R	v	т	P	L	. м	Α	A	ĸ	I	Α	D	R	E	т	G	ĸ	А	L	G	P	N C	יעכ	G	E	т.	СТ	ĸ	G	P	м	v	5 1	KG	. v	v	N	1100
RDVER52.SEQ	R	ν	T	P	L	. м	А	A	ĸ	I	A	D	R	E.	т	G	ĸ	A	Τ.	G	P	N C	יי	·G	E	1.	C T	ĸ	6	- -	м	v	S 1			17 1	NI.	1100
RD1561H9.SEQ	R	v	Т	P	Ĺ	M	А	A	K	I	A	D	R	E.	Ţ	G	ĸ	Α.	t.	G	P		יי	٦	F	L	c r	K	6	D	м	v	5 i			17 1	LV NT	1100
				Ī		•				-	••	_	••	_	•	Ü		•	_	٠	•		. •	٠	_	_	٠.		G	•	М	٧	3 1	٠. ٥	, 1	٠.	N	1196
GRVER51.SEQ	N	v	E	Α	Т	ĸ	E	A	r	D	D	D	G	W	L	н	s	G	ם	F	G ·	ΥY	(D	Е	D	E	H F		v	ν	D	₽ .	y i	K E	: т	T s	ĸ.	1310
GR6.SEQ	N	v	E	A	т	K	E	A	I	D	D	D	G	W	L	н	Ś	G	D	F	G ·	ΥV	. D	F	n	E	н г	v	v	v	D.	ъ.	. : V 1	, E		7 -		1318
GRVER5.SEQ	N	v	Е	A	T	ĸ	E	A	T	D	ם	ם	G :	w	τ.	н	s	c	n	F	٠.		, ,	F	ח			·	٧	v v	ח	ь,		· E		T .	λ υ	1318
GRVER4.SEQ	N	v	E	A	T	ĸ	F	Α	т	- D	n	n	G I	w	~ τ.	ч	ç	c	n				. D	-	5	-			٧,	.,	2	n ,			. 1	1.5		1318
GRVER3.SEQ	N	v	E	A	т	ĸ	E	A	7	D	n	n	G I	w	τ.	н	0	c	n	E 4	σ,			-	7	E 1			v	٧,	ם	n ,				1 5		1318
	N	v	E	A	т	ĸ	F.	A	Ţ	D	n	D	G I	 U	7.	H	~	G.	ח		g,		, ,	-	2	E 1		v	v	.,	ט	K :		(E	. т		Λ.	1318
GRVER1.SEQ	N	v	E	Δ	т	ĸ	F	Δ.	T	n	ח	ņ	G 1	· ω				٥.	_	E ,				5	D	E 1		1	٠,			r .				_ F	Α	1318
YG81-6G1.SEQ	N	v	E	Δ.	T	ĸ	F	Δ.	7	n	n	ח	G I	w	7		٠.	G ·	D	E 1			ם י	E		E 1		1	.,	.,	ט ה	R :			. L	1 5	۸.	1318
RDVER1.SEQ	N	v	E	A	т	ĸ	E	A	T	n	ח	n	G 1	w	τ.	и	9	G :	ח	E (י ה	-	ח	5 i			.,	V 17	ים מ	R :				1 5	,	1318
RDVER2.SEQ	N	v	E	Α	т	ĸ	E	A	T	D	n	D	ر د د	ω	T.	н	9	G .	n .	E (ם.	-	ח				v	11	ים מ	R ;	, ,			1 5		1318
	N	v	E	A	т	ĸ	E	А	T	ם	n	ם	G 1	w	ī.	H	9	د	D .	E /				E-	ח	E 1		,	.,	17	יט הי					1 5		1318
RDVER4.SEQ	N	v	E.	A	T	ĸ	E	Α	T	ח	D	D	6 1	w	ī.	н	9 1	G 1	n .	 	- 1	• •		-	<u>_</u>	E 1		÷	.,	·	ים הים		. r			1 5		1318
-	N	v	F.	A	т	ĸ	E	A	T	ם	b	n	د د	w	- T.	н	9 1	G 1	ה ח	F (. ה	D	D.	E 1	, ,	v	17	v	ים נים		; n			IN	,	1318
RD7.SEQ	N	v	E	Δ.	т	ĸ	E	ь	T	'n	n	ח	G 1	ω. ω	T.	u		G 1	ה ה	E (- ,			-	2	E 1			٠,	٠,	י מי					1 5		1318
RDVER51. SEQ	N	v	E	Α	т	ĸ	E	A	Ţ	ח	D	ח	ים הו	ω .	1	n		G I	י רו	E (ם	E	ע	E 1	1 F	ı	.,	٧	U .	K 1	;	. E		IK	-	1318
RDVER52. SEQ	N	v	Ē	Δ	т	ĸ	-	Δ.	Ť	ם	ח	<u></u>	C 1					C 1	, ,	E (, .		ַ	5	D .		1 5	ı	٧,	٧.,	ים					I K	٠.	1318
RD1561H9.SEQ	N	v	E	Δ	т	K	F	Δ	Ť	D	ח	n	ים הי		 T	n .		C 1	N 1	E (, ,		-	ים מ		1 E	I	٧	٠.	יט	K 1		. =	ъ.	IK	٠.	1318
	•••	•	_	•	•	•	_	^	•	Ŭ				•	_	n			,		,		U	E.	υ.	E 1		1	V	ν.	וט	K 1		. E	٠	1 6	٠.	1318
GRVER51.SEQ	Y	ĸ	G	s	o	v	A	P	A	E	Ţ.	F.	E '	т.	T.	r. :	ĸ ı	N 1	P (וח		ם פ	v	Δ	v	v	. T	ъ	n	T . 1	.				D	e 2		1430
GR6.SEQ	Y	ĸ	G	s	ā	v	A	P	A	E	L	E	E :	T :	T.	r. i	 К I	N I	P (. 1		ת פ	v	2	v	0 0		Þ	0	L. I	6 J				E D	9 2		1438
GRVERS.SEQ	Y	к	G	s	ō	v	A	P	A	E	1.	E	E i	7	T.	L 1	K 1	N I	P (0 1		ת מו	v	Δ.	v	v 6	. T		ח	Σ. 1			: F		F D	9 7		1438
	Y	ĸ	G	s	ō	v	A	P	A	E	ī.	E	E 1	- ·	T.	T. 1	K 1	NT I	- ·	2 1		חו	v	Δ.	v	v 6	. T	D	ח	7. 1	E 1		. E		D	9 2		1438
GRVER3.SEQ	Y	ĸ	G	s	ō	v		P	Δ	E	7.	E	6 . E 1		 T.	L 1	K 1	N 1		~ 1		חי	v	A	v v v	v (, T	r D	ב	14 J					F	<i>o</i> n		1438
GRVER2.SEQ																																						1438
GRVER1.SEQ	Ÿ	ĸ	G	s	0	v			_	E	7.	r F	6 1 6 1	 	_	 T		17 I		- 1 - 1		, כ	17	~	¥ '			<i>T</i>	ם	11 J	6 / F 1		, E		r			1438
YG81-6G1,SEQ	Ÿ	ĸ	~	~	ă	v	î	D	~		T		D 1					N .		. I			٧.	^	v '	v (יט			1 6	, 6		2	3 A		1438
RDVER1.SEQ	÷	ĸ	6	•	~	7	~	E D	~			e. E			,			N .					٧	^	· ·			r	יט	ь <i>і</i>		4 6	, E	ı		3 A	۱.	1438
																																						1438
																																						1438
-																																						1438
-																																						1438
																																						1438
_	·	v	<u>ح</u>		ž	v	~	r	~	E	ا ط ت ت	ا ت د	5 <i>1</i>					1 E	. (נט	٧	A	v \	v G		2	ָ ע	. I	. F	G	. E	į.	F :	۲A	. 1	1438
RDVER51.SEQ	'n	r.	<u>-</u>	2	2	٧	A.	r	A .	E	ا با ت	د ا د	5 I		: مسا	i	•	۱ E			. R		٧	A '	v 1	v G	. I	۲	ם כ	ا با -		۱ G	E.	L	P :	s A	.]	1438
RDVER52.SEQ	ī		<u>ر</u>	2	Š	٧	A.	۲,	A.	E.	. i	c 1	5 J	. 1	ا ما	ן נו		۷ E		; [. R	ם י	٧	Α '	۷ ۱	/ G	I	4	0) -	i l	5 7	G	Ε	L	P :	s A	, 1	1438
RD1561H9.SEQ	Ľ	K	G	3	Q	٧	A	٢	A	Ľ	L, I	. 1	5 I	. 1		L	. 1	4 E	• (: I	. R	D	٧	A '	V	/ G	I	P	ם ו	LI	. <i>F</i>	l G	E	L	Ρ:	S A	. 1	1438

Figure 3 (cont.)

GREENSEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVERS. SEQ		<u> </u>			=	GRVER51.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558																																					
GROVERS.SEQ FV V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERS.SEQ F V V K Q P G K E I	GRVER51.SEQ	F	٧	٧	K	() F	, 6		E	I	T	A	K	E	V	Y	, D	Y	L	A	E	R	٧	5	Н	T	ĸ	Y	L	R	G	G	v	R	F	٧	D	s	ī	P	1550	A
GRVERAS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERAS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERAS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERAS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERAS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERAS.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD V R S I F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD V R S I F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD V R S I F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD V R S I F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD V R S I F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD V R S I F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD I F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD I F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD I F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD I F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD I F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD I F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD I F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD I F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD I F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD I F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D	GR6.SEQ	F	V	V	K	() F	. 6	, K	E	1	T	Α	K	Ε	٧	Y	. 0	Y	L	Α	E	R	٧	S	н	Т	К	Y	L	R	G	G	v	R	F	v	n	9	Ť	P	1550	
GRVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 ROVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 ROVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 ROVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 ROVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 ROVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 ROVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 ROVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 ROVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 ROVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 ROVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 ROVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 ROVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 ROVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 ROVERAI.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 ROVERAI.SEQ R N V T G K I T R K E L L K Q L L E K A G G 1624 GRVERAI.SEQ R N V T G K I T R K E L L K Q L L E K A G G 1624 GRVERAI.SEQ R N V T G K I T R K E L L K Q L L E K A G G 1624 GRVERAI.SEQ R N V T G K I T R K E L L K Q L L E K A G G 1624 GRVERAI.SEQ R N V T G K I T R K E L L K Q L L E K A G G	GRVER5. SEQ	F	٧	٧	K	ς ς) F	, G	K	E	I	T	Α	K	E	V	Y	D	Y	Ļ	A	ε	R	v	s	Н	Т	к	Y	L	R	G	G	v	R	F	ν	n	s	T	0	1550	
GRVER1.SEQ	GRVER4.SEQ	F	V	٧	K	(() P	Ġ	; K	E	I	Ţ	Α	K	Ε	٧	Y	ם	Y	L	A	E	R	٧	S	н	T	к	Y	L	R	G	G	v	Ŕ	F	υ	n	c	Ŧ	ъ	1660	
GRVERI.SEQ	GRVER3.SEQ	F	v	V	K	(() P	, G	K	E	Ι	T	А	K	Ε	V	Y	D	Y	L	A	E	R	٧	s	н	T	ĸ	Y	L	R	G	G	v	R	F	v	n	s	т	D	1550	5
GRVERT.SEQ	GRVER2.SEQ	F	٧	٧	K	C	2 2	G	K	E	1	T	A	K	Ε	V	Y	D	Y	L	Α	E	R	V	s	н	Т	ĸ	Y	L	R	G	G	v	R	F	v	<u>_</u>	~	Ť	r D	1550	,
TABLE-8GI.SEQ F V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVER1.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVER3.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVER3.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVER4.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVER5.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVER5.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVER51.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVER52.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVER52.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVER52.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVER52.SEQ F V V K Q P G T E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G G G G G G G G G G G G G G G G G	GRVER1.SEQ	F	V	٧	K	C	₽ (Ģ	K	E	I	T	Α	К	Ε	٧	Y	D	Y	L	Α	Ε	R	V	s	Н	T	K	Y	L	R	G	G	v	R	F	v	n	c	т	D	1550	
ROVER2.SEQ	YG81-6G1.SE	QF	٧	٧	K	Ç	P	G	K	ε	I	Т	A	к	E	٧	Y	D	Y	L	Α	E	R	v	s	н	т	ĸ	Ÿ	1.	R	G	6	v	D	F	v	<u>_</u>	Č	Ť	D.	1550	,
RDVER3.SEQ	RDVER1.SEQ	F	V	٧	K	Q	P	G	к	E	I	Т	A	ĸ	Ε	v	Y	D	Y	L	А	Е	R	v	s	н	T	ĸ	Ÿ	Ţ.	R		6	v	D.	·	1/	5		+	r D	1550	,
REVERS. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 R R V S R S S S S S S S S S S S S S S S	RDVER2.SEQ	F	٧	٧	K	Q	9	G	K	E	I	Т	А	ĸ	Ε	٧	Y	D	Y	L	А	Е	R	v	s	н	т	ĸ	Ÿ	Ť.	R	6	c	v	D	ŗ	v	ŗ		±	r n	1550	,
ROVER5.SEQ	RDVER3.SEQ	F	٧	٧	K	Q	P	G	к	E	I	Т	A	K	E	v	Y	D	Y	L	A	E	R	v	S	н	T	ĸ	Ÿ	ī.	R	6	c	v		F	77	Ď		+	E D	1550	,
ROVERS. SEQ	RDVER4.SEQ	F	٧	٧	ĸ	Q	P	G	ĸ	E	1	Т	A	ĸ	Ε	٧	Y	D	Y	L	A	E	R	v	s	н	T	ĸ	v	r.	B	c	6	v	D .	F	v	5	0	+	r	1008	,
RDV.RSQ	RDVER5.SEQ	F	٧	٧	K	Q	P	G	K	Ε	1	Т	Α	ĸ	Ε	ν	Y	D	Y	L	A	E	R	v	s	н	T	ĸ	Ÿ	T.	R	c	٥	v	D.	F	v	ח		T		1550	
REDVER51.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 REDVER52.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RED1561H9.SEQF V V K Q P G T E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RED1561H9.SEQF V V K Q P G T E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RED1561H9.SEQF V V K Q P G T E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RED1561H9.SEQF V V K Q P G T E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RED1561H9.SEQF V V K Q P G T E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RED1561H9.SEQF V V K Q P G T E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RED1561H9.SEQF V V K Q P G T E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RED1561H9.SEQF V V K Q P G T E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RED1561H9.SEQ R N V T G K I T R K E L L K Q L L E K A G G RED1561H9.SEQF V V K Q P G T E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RED1561H9.SEQ F V V K Q P G T E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RED1561H9.SEQ F V V K Q P G T E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RED1561H9.SEQ F V V K Q P G T E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RED1561H9.SEQ F V V K Q P G T E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RED1561H9.SEQ F V V K Q P G T E I T A K E V Y D Y L A E R V S H T K Y L R G G V R T T R K E L L K Q L L E K A G G V R T T R K E L L K Q L L E K A G G V R T T R K E L L K Q L L E K A G G V R T T R K E L L K Q L L E K A G G V R T T R K E L L K Q L L E K A G G V R T T T T T T T T T T T T T T T T T T	RD7.SEQ	F	٧	٧	ĸ	Q	P	G	K	Ε	I	Т	А	к	Ε	٧	Y	D	Y	L	A	E	R	v	S	н	7	ĸ	v	r.	R	c	c	v	D		17			± T		1550	
REDUSABLY. SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 RD1561H9. SEQ F V V K Q P G T E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P 1558 GRVER\$1. SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER\$1. SEQ R N V T	RDVER51.SEQ	F	v	v	К	Q	P	G	ĸ	Ε	Ι	т	А	ĸ	Ε	v	Y	D	Y	L	A	E	R	v	S	н	т Т	ĸ	Ÿ	T.	D.	c	c	17			17	ם		± •	r D	1228	
GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER1.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER1.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER1.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER1.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER1.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER1.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER2.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER5.	RDVER52.SEQ	2	V	٧	K	Q	P	G	К	Ε	Ι	Т	Α	K	Ε	V	Υ	D	Y	L	Α	E	R	v	S	н	т	ĸ	Y	r.	R	c	c	v	b	F .	37	n		τ.	n	1660	
GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GR6.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER4.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER1.SEQ R N V T G K I T R K E L L K Q L L E K A G G GROVER1.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER1.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER2.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G ROVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G	RD1561H9.SE(g F	v	ν	ĸ	Q	₽	G	T	E	I	т	А	ĸ	E	v	Y	D	Y	L	A	E	R	v	s	H	T	ĸ	Ÿ	- T.	D.	6	٠ د	v.	r D	F	٧		0	<u>.</u>	F	1550	
GR6.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVERS.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER4.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER4.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER1.SEQ R N									Т	•												_	•	•	_	••	•	••	٠	_	•	•	•	•		E	٧	υ.	٥	1	P	1338	
GR6.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVERS.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER4.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER4.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER1.SEQ R N	GRVER51.SEQ	R	N	v	T	G	K	I	т	R	ĸ	E	L	L	ĸ	0	L	L	E	к	A	G	G																			1.004	
GRVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER4.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER1.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER1.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER1.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER1.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G	GR6.SEQ																																•										
1624 1624	GRVERS.SEQ																																										
GRVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G 1624																																											
1624 1624																																											
1624 1624 1624 1624 1624 1624 1624 1624																																											
1624 1624	GRVER1.SEQ																																										
1624 1624	YG81-6G1.SEC	R	N	v	T	G	ĸ	I	Т	R	ĸ	Ε	L	L	ĸ	ō	L	L	E	ĸ	Α.	G	G																				
RDVER2.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER4.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G RD7.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER55.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER55.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER55.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER55.SEQ R N V T G K I T R K E L L K Q L L E K A G G																																											
RDVER3.SEQ R N V T G K I T R K E L L K Q L L E K A G G 1624 RDVER4.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER55.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER55.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER55.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER55.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER55.SEQ R N V T G K I T R K E L L K Q L L E K A G G																																											
RDVER4.SEQ R N V T G K I T R K E L L K Q L L E K A G G 1624 RDVER5.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G RDVER54.SEQ R N V T G K I T R K E L L K Q L L E K A G G	RDVER3.SEQ	R	N	v	T	G	K	Ī	T	R	ĸ	E	t.	T.	ĸ	ŏ	T.	r.	E	K	΄.	c i	2																				
RDVER5.SEQ RNVTGKITRKELLKQLLEKAGG 1624 RD7.SEQ RNVTGKITRKELLKQLLEKAGG 1624 RDVER51.SEQ RNVTGKITRKELLKQLLEKAGG 1624 RDVER52.SEQ RNVTGKITRKELLKQLLEKAGG 1624 RD156149 SFORNVTGKITRKELLKQLLEKAGG 1624	RDVER4.SEQ	R	N	v	т	Ġ	к	Ī	т	R	ĸ	F	T.	T.	ĸ	ň	7.	T.	-	r.		c /	-																				
RD7.SEQ RNVTGKITRKELLKQLLEKAGG 1624 RDVER51.SEQ RNVTGKITRKELLKQLLEKAGG 1624 RDVER52.SEQ RNVTGKITRKELLKQLLEKAGG 1624	RDVER5.SEQ																																										
RDVERS1.SEQ RNVTGKITRKELLKQLLEKAGG 1624 RDVERS2.SEQ RNVTGKITRKELLKQLLEKAGG 1624																																											
RDVER52.SEQ R N V T G K I T R K E L L K Q L L E K A G G	RDVER51.SEO	R	N	v	T	G	ĸ	- T	т	Ř	ĸ	E	- Τ.	T.	ĸ	~	T.	L			n .	۱ بود د م	-																				
ID1561H9 SEOP N V m C V T m D V m T + V m	RDVER52.SEO	R	N	v	T	G	ĸ	T	Т	R	ĸ	E.	ت T.	T.	ĸ	~	1.	L.	F		π. '	יטי כי	3																				
1624	RD1561H9.SEO	R	N	v	T	G	ĸ	Ť	Ť	R	ĸ	F	т.	~ T.	ĸ	~		ř			n. '	٠,	-																				
					-	•		-	•	•	••	_	_	_	•	~	_	- L	<u>*</u>)		n '	9 (,																			1624	

Figure 4 Codon Usage Analysis

per 542 fet	roders YGFE1-6G	wer! GR	veri RD	ward GR	vers RD	HUM
CGA		0	0	2	0	3
COC		13	Ö	ú	12	6
coc		ő	0	ö	0	6
		13		ŭ	14	ì
COT		ö	13	ö	Ö	š
AGA		ŏ		å		
Are AGG			0		<u> </u>	_ •
CTA		0	0	0	Q.	3
CTC		0	1	12	E1	11
100		28	27	19	18	23
cm		0	0	1.	t	6
ITΛ	17	q	0	o	0	3
Leu TTG	13	27	27	23	25	6
TCA	6	0	0	1	2	5
1 700	2	0	٥	4	2	10
TCG	7	0	0	0	Ó	2
TCT	7	16	15	EI	12	7
AGC	2	15	15	14	12	10
Ser ACT	7	o	ő	î	2	5
ACA	10	-0	0	- ;	1 -	
1		-		_		
ACC	2	11	11	8	II.	12
ACG	2	0	0	0	Q.	4
The ACT		- 11		14	10	7
CCV	9	14	14	9	12	8
1,000	8	0	0	2	1	11
cca	2	Ú	0	a	0	4
Pro CCT	9	14	14	17	15	8
GCA	14	0	0	5	4	8
GCC	4	19	18	14	42	16
GCG	5	o	0	0	ō	4
AL GCT	15	18	19	18	21	11
GGA	18	0	0	-;-	3	- ; -
GGC	3	20	19	21	, 21	14
GGG	2	0	0		21 I	,
	16			!		
Gly GGT		19	20	16	14	6
GTA	13	0	0	1	1	3
GTC	4	25	24	21	26	9
GTG	12	25	25	25	17	17
ValGTT	20	. 0	0	3	5	6
***	23	17	18	19	13	12
Lysaag	12	18	17	16	22	19
AAC	6	Ή	11	13	12	12
TAA RIA	16	11	10	9	9	9
CAA		7	1			6
Gln CAG	6	7	i	ij	i	18
CAC	- 6 -	-		- ; -	- i	
His CAT	7	6	7	6	9	5
GAA	26	19	19	19	18	- 13
Glu GAG	12					
		19	19	19	20	22
GAC	6	13	13	14	12	16
ASP CAT	20	- 13	13	12		12
TAC	8	10	10	12	13	10
Tyr TAT		9	10	7	7	7
TGC	3	6	5)	4	8
Cys TGT		5	6	8	7	5
πc	11	13	12	15	12	12
Phe TTT	14	12	13	10	(3	9
ATA	12	0	- 13	- 6	0	- 7
ATC	7	19	19	23	20	13
IK ATT	19	19	20		19	
Met ATG	- 11			- 15		
			11	- ti	- 11	12
To TGG	2	2	?	2	2	7

relative codo	a usuge	for each as (vers GR	*(00) ver5 R.D	HUM
CGA		**************************************	443,400	10
	27		46	
CCC	4	42		11
സ	0	a	0	19
CGT	19	50	54	9
AGA	23	o	٥	19
Are AGG	27	0	0	21
CTA	9	0	0	6
CTC	7	22	20	21
CTG	i	35	33	44
CIT	22	1	2	- ii - l
		-	ô	6
ПΛ	31	0		
Lev TTG	24	42	45	!
TCA	18	J	7	13
TCC	6	13	7	25
TCG	23	0	0	6
101	23	35	40	18
AGC	6	45	40	26
Ser AGT	23	j	7	13
			- 5	
ACA	45	0		25
ACC	9	36	50	40
ACG	9	0	0	12 {
The ACT	36	64	45	22
CCA	32	32	43	26
ccc	29	7	4	35
ccc	,	Ü	ò	12
		61	54	27
Pro CCT	32			
GÇA	37	13	11	19
GCC	11	37	32	40
GCG	13	0	0	10
Ala GCT	39	47	55	27
GGA	46			24
GGC	1	54	54	36
GGG	š	3	3	25
		41	36	16
Gty GGT	41		- 20 -	9
GTA	27	2	_	
GTC	Į.	42	53	25
GTG	24	50	35	48
AN CLL	41	6	10	16
AAA	66	54	37	39
	34	46	63	61
Lys AAG	27	59	57	58
AAC				6
TAA meA	73	41	43	
CAA	57	79	47	25
Gla CAG	43	21	53	76
CAC	46	54	31	59
His CAT	54	46	69	39
GAA	68	50	47	39
GH GAG	32	50	53	61
GAC	23	54	46	- 56
			*	42
AspGAT	77	46		- 60
TAC	42	63	65	
Tyr TAT	54	37	35	40
TGC	27	27	36	8
Cys TGT	73	73	64	41
110	44	60	48	58
	56	40	52	41
Phc TTT			- 6 -	13
ATA	32			55
ATC	18	61	51	
lic ATT	50	39	49	34
Maatg	100	100	100	100
Trp TGG	100	100	100	(00)

Patent Application Publication Jun. 15, 2006 Sheet 22 of 65 US 2006/0127988 A1

Figure 5A

Code	on Usage	YG#8	1-6G01	(yello	w-green	1)					
TTT	Phe	14	TCT	Ser	7	TAT	Tyr	11	TGT	Cys	8
TTC	Phe	11	TCC	Ser	2	TAC	Tyr	8	TGC	Cys	3
TTA	Leu	17	TCA	Ser	6	TAA	***	ō	TGA	***	0
TTG	Leu	13	TCG	Ser	7	TAG	***	ō	TGG	Trp	2
CTT	Leu	12	CCT	Pro	9	CAT	His	7	CGT	Arg	5
CTC	Leu	4	CCC	Pro	8	CAC	His	6	CGC	Arg	1
CTA	Leu	5	CCA	Pro	9	CAA	Gln	8	CGA	Arg	7
CTG	Leu	4	CCG	Pro	2	CAG	Gln	6	CGG	Arg	ó
ATT	Ile	19	ACT	Thr	8	AAT	Asn	16	AGT	Ser	7
ATC	Ile	7	ACC	Thr	2	AAC	Asn	6	AGC	Ser	2
ATA	Ile	12	ACA	Thr	10	AAA	Lys	23	AGA	Arg	6
ATG	Met	11	ACG	Thr	2	AAG	Lys	12	AGG	Arg	7
GTT	Val	20	GCT	Ala	15	GAT	Asp	20	GGT	Gly	16
GTC	Val	4	GCC	Ala	4	GAC	Asp	6	GGC	Gly	3
GTA	Val	13	GCA	Ala	14	GAA	Glu	26	GGA	Gly	18
GTG	Val	12	GCG	Ala	5	GAG	Glu	12	GGG	Gly	2

Figure 5B

(Codor	usage:	GRver	1								
•	TTT	Phe	12	TCT	Ser	16	TAT	Tyr	9	TGT	Сув	5
•	TTC	Phe	13	TCC	Ser	0	TAC	Tyr	10	TGC	Сув	6
1	ΓTA	Leu	0	TCA	Ser	0	TAA	***	0	TGA	***	0
-	PTG	Leu	27	TCG	Ser	0	TAG	***	0	TGG	Trp	2
(CTT	Leu	0	CCT	Pro	14	CAT	His	6	CGT	Arg	13
(CTC	Leu	0	CCC	Pro	0	CAC	His	7	CGC	Arg	13
(CTA	Leu	0	CCA	Pro	14	CAA	Gln	7	CGA	Arg	0
(CTG	Leu	28	CCG	Pro	0	CAG	Gln	7	CGG	Arg	0
2	ATT	Ile	19	ACT	Thr	11	AAT	Asn	11	AGT	Ser	0
7	ATC	Ile	19	ACC	Thr	11	AAC	Asn	11	AGC	Ser	15
2	ATA	Ile	0	ACA	Thr	0	AAA	Lys	17	AGA	Arq	0
I	ATG	Met	11	ACG	Thr	0	AAG	Lys	18	AGG	Arg	0
C	STT	Val	0	GCT	Ala	18	GAT	Asp	13	GGT	Gly	19
C	STC	Val	25	GCC	Ala	19	GAC	Asp	13	GGC	Gly	20
C	STA	Val	0	GCA	Ala	0	GAA	Glu	19	GGA	Gly	0
(3TG	Val	25	GCG	Ala	0	GAG	Glu	19	GGG	Gly	0

Figure 5C

Codo	on Usage:	RDver	1								
TTT	Phe	13	TCT	Ser	15	TAT	Tyr	10	TGT	Cys	6
TTC	Phe	12	TCC	Ser	0	TAC	Tyr	10	TGC	Cys	5
TTA	Leu	0	TCA	Ser	0	TAA	***	0	TGA	***	0
TTG	Leu	27	TCG	Ser	0	TAG	***	0	TGG	Trp	2
CTT	Leu	0	CCT	Pro	14	CAT	His	7	CGT	Arq	13
CTC	Leu	1	CCC	Pro	0	CAC	His	6	CGC	Arg	13
CTA	Leu	0	CCA	Pro	14	CAA	Gln	8	CGA	Arq	0
CTG	Leu	27	CCG	Pro	0	CAG	Gln	7	CGG	Arg	o
ATT	Ile	20	ACT	Thr	11	AAT	Asn	10	AGT	Ser	0
ATC	Ile	19	ACC	Thr	11	AAC	Asn	11	AGC	Ser	15
ATA	Ile	0	ACA	Thr	0	AAA	Lys	18	AGA	Arg	0
ATG	Met	11	ACG	Thr	0	AAG	Lys	17	AGG	Arg	0
GTT	Val	0	GCT	Ala	19	GAT	Asp	13	GGT	Gly	20
GTC	Val	24	GCC	Ala	18	GAC	Asp	13	GGC	Gly	19
GTA	Val	0	GCA	Ala	0	GAA	Glu	19	GGA	Gly	0
GTG	Val	25	GCG	Ala	0	GAG	Glu	19	GGG	Gly	0

Figure 5D

Codo	on Usage:	Grver	2								
TTT	Phe	12	TCT	Ser	15	TAT	Tyr	9	TGT	Cys	5
TTC	Phe	13	TCC	Ser	0	TAC	Tyr	10	TGC	Cys	6
TTA	Leu	0	TCA	Ser	0	AAT	***	0	TGA	***	0
TTG	Leu	27	TCG	Ser	0	TAG	***	0	TGG	Trp	2
CTT	Leu	0	CCT	Pro	14	CAT	His	6	CGT	Arg	13
CTC	Leu	0	CCC	Pro	0	CAC	His	7	CGC	Arq	13
CTA	Leu	0	CCA	Pro	14	CAA	Gln	10	CGA	Arg	0
CTG	Leu	28	CCG	Pro	0	CAG	Gln	4	CGG	Arg	0
ATT	Ile	20	ACT	Thr	11	AAT	Asn	11	AGT	Ser	0
ATC	Ile	18	ACC	Thr	11	AAC	Asn	11	AGC	Ser	16
ATA	Ile	0	ACA	Thr	0	AAA	Lys	16	AGA	Arg	0
ATG	Met	11	ACG	Thr	0	AAG	Lys	19	AGG	Arg	ō
GTT	Val	0	GCT	Ala	18	GAT	Asp	13	GGT	Gly	18
GTC	Val	28	GCC	Ala	19	GAC	Asp	13	GGC	Gly	21
GTA	Val	0	GCA	Ala	0	GAA	Glu	17	GGA	Gly	0
GTG	Val	22	GCG	Ala	0	GAG	Glu	21	aaa	Gly	Ŏ

Figure 5E

Codon	Usage	:Rdver2
-------	-------	---------

TTT	Phe	13	TCT	Ser	16	TAT	Tyr	10	TGT	Cys	6
TTC	Phe	12	TCC	Ser	0	TAC	Tyr	10	TGC	Сув	5
TTA	Leu	0	TCA	Ser	0	TAA	***	0		***	
TTG	Leu	27	TCG	Ser	-			_	TGA		0
	Dea	21	100	ser	. 0	TAG	***	0	TGG	Trp	2
CTT	Leu	0	CCT	Pro	15	CAT	His	7	CGT	Arg	13
CTC	Leu	1	CCC	Pro	0	CAC	His	6	CGC	Arg	13
CTA	Leu	0	CCA	Pro	13	CAA	Gln	8		_	
CTG	Leu	27	CCG	Pro					CGA	Arg	0
	200	21	CCG	PIO	0	CAG	Gln	7	CGG	Arg	0
ATT	Ile	19	ACT	Thr	11	AAT	Asn	10	AGT	Ser	0
ATC	Ile	20	ACC	Thr	11	AAC	Asn	11	AGC	Ser	14
ATA	Ile	0	ACA	Thr	0	AAA	Lys	19	AGA	Arq	
ATG	Met	11	ACG	Thr	Ö		-			_	0
_			1100	****	V	AAG	ГÀв	16	AGG	Arg	0
GTT	Val	0	GCT	Ala	19	GAT	qaA	13	GGT	Gly	21
GTC	Val	21	GCC	Ala	17	GAC	Asp	13	GGC	-	
GTA	Val	0	GCA	Ala	1	GAA	_			Gly	18
GTG	Val	_					Glu	21	GGA	Gly	0
310	AGI	28	GCG	Ala	0	GAG	Glu	17	GGG	Glv	0

Figure 5F

Code	on Usage:	GRver	.3								
TTT	Phe	13	TCT	Ser	16	TAT	Tyr	9	TGT	Сув	7
TTC	Phe	12	TCC	Ser	0	TAC	Tyr	10	TGC	Сув	4
TTA	Leu	0	TCA	Ser	0	TAA	***	0	TGA	***	0
TTG	Leu	26	TCG	Ser	0	TAG	***	0	TGG	Trp	2
CTT	Leu	0	CCT	Pro	18	CAT	His	6	CGT	Arg	14
CTC	Leu	5	CCC	Pro	0	CAC	His	7	CGC	Arg	12
CTA	Leu	0	CCA	Pro	10	CAA	Gln	9	CGA	Arg	0
CTG	Leu	24	CCG	Pro	0	CAG	Gln	5	CGG	Arg	ő
ATT	Ile	14	ACT	Thr	14	AAT	Asn	11	AGT	Ser	0
ATC	Ile	24	ACC	Thr	8	AAC	Asn	11	AGC	Ser	15
ATA	Ile	0	ACA	Thr	0	AAA	Lys	21	AGA	Arg	0
ATG	Met	11	ACG	Thr	0	AAG	Lys	14	AGG	Arg	ŏ
GTT	Val	1	GCT	Ala	18	GAT	Asp	12	GGT	Gly	18
GTC	Val	22	GCC	Ala	18	GAC	Asp	14	GGC	Gly	21
GTA	Val	0	GCA	Ala	ı	GAA	Glu	20	GGA	Gly	0
GTG	Val	27	GCG	Ala	0	GAG	Glu	18	GGG	Gly	Ö

Figure 5G

Code	n Usage:	RDver	.3								
TTT	Phe	13	TCT	Ser	14	TAT	Tyr	7	TGT	Cys	6
TTC	Phe	12	TCC	Ser	1	TAC	Tyr	13	TGC	Cys	5
TTA	Leu	0	TCA	Ser	0	TAA	***	0	TGA	***	0
TTG	Leu	27	TCG	Ser	0	TAG	***	0	TGG	Trp	2
CTT	Leu	0	CCT	Pro	16	CAT	His	10	CGT	Arq	16
CTC	Leu	6	CCC	Pro	0	CAC	His	3	CGC	Arg	10
CTA	Leu	0	CCA	Pro	12	CAA	Gln	8	CGA	Arg	0
CTG	Leu	22	CCG	Pro	0	CAG	Gln	7	CGG	Arg	ō
ATT	Ile	20	ACT	Thr	10	AAT	Asn	10	AGT	Ser	0
ATC	Ile	19	ACC	Thr	12	AAC	Asn	11	AGC	Ser	15
ATA	Ile	0	ACA	Thr	0	AAA	Lys	13	AGA	Arg	0
ATG	Met	11	ACG	Thr	0	AAG	Lys	22	AGG	Arg	Ŏ
GTT	Val	0	GCT	Ala	20	GAT	Asp	14	GGT	Gly	16
GTC	Val	27	GCC	Ala	16	GAC	Авр	12	GGC	Gly	23
GTA	Val	0	GCA	Ala	1	GAA	Glu	18	GGA	Gly	0
GTG	Val	22	GCG	Ala	0	GAG	Glu	20	GGG	Gly	ō

Figure 5H

Code	Codon Usage: GRver4											
TTT	Phe	11	TCT	Ser	13	TAT	Tyr	7	TGT	Сув	8	
TTC	Phe	14	TCC	Ser	2	TAC	Tyr	12	TGC	Cys	3	
TTA	Leu	0	TCA	Ser	1	AAT	***	0	TGA	***	0	
TTG	Leu	21	TCG	Ser	0	TAG	* * *	0	TGG	Trp	2	
CTT	Leu	1	CCT	Pro	18	CAT	His	7	CGT	Arg	14	
CTC	Leu	11	CCC	Pro	0	CAC	His	6	CGC	Arg	11	
CTA	Leu	Q	CCA	Pro	10	CAA	Gln	11	CGA	Arg	1	
CTG	Leu	22	CCG	Pro	0	CAG	Gln	3	CGG	Arg	ō	
ATT	Ile	13	ACT	Thr	14	AAT	Asn	11	AGT	Ser	1	
ATC	Ile	25	ACC	Thr	8	AAC	Asn	11	AGC	Ser	14	
ATA	Ile	0	ACA	Thr	0	AAA	Lys	20	AGA	Arg	0	
ATG	Met	11	ACG	Thr	0	AAG	Lys	15	AGG	Arg	ō	
GTT	Val	3	GCT	Ala	19	GAT	Asp	12	GGT	Gly	17	
GTC	Val	22	GCC	Ala	15	GAC	Asp	14	GGC	Gly	19	
GTA	Val	0	GCA	Ala	3	GAA	Glu	20	GGA	Gly	3	
GTG	Val	25	GCG	Ala	0	GAG	Glu	18	GGG	Gly	0	

Figure 5I

Code	on Usage:	RDver	4								
TTT	Phe	13	TCT	Ser	11	TAT	Tyr	7	TGT	Cys	7
TTÇ	Phe	12	TCC	Ser	2	TAC	Tyr	13	TGC	Cys	4
TTA	Leu	0	TCA	Ser	2	TAA	***	0	TGA	***	0
TTG	Leu	28	TCG	Ser	0	TAG	***	0	TGG	Trp	2
CTT	Leu	0	CCT	Pro	16	CAT	His	11	CGT	Arg	15
CTC	Leu	7	CCC	Pro	2	CAC	His	2	CGC	Arg	11
CTA	Leu	0	CCA	Pro	10	CAA	Gln	7	CGA	Arg	0
CTG	Leu	20	CCG	Pro	0	CAG	Gln	8	CGG	Arg	ŏ
ATT	Ile	21	ACT	Thr	11	AAT	Asn	10	AGT	Ser	1
ATC	Ile	18	ACC	Thr	11	AAC	Asn	11	AGC	Ser	14
ATA	Ile	0	ACA	Thr	0	AAA	Lys	13	AGA	Arg	0
ATG	Met	11	ACG	Thr	0	AAG	Lys	22	AGG	Arg	ŏ
GTT	Val	3	GCT	Ala	22	GAT	Asp	15	GGT	Gly	14
GTC	Val	27	GCC	Ala	11	GAC	Asp	11	GGC	Gly	21
GTA	Val	0	GCA	Ala	4	GAA	Glu	18	GGA	Gly	4
GTG	Val	19	GCG	Ala	0	GAG	Glu	20	GGG	Gly	0

Figure 5J

GTA Val

GTG Val

Code	on Usage	: GRve	r5								
TTT	Phe	10	TCT	Ser	11	TAT	Tyr	7	TGT	Сув	8
TTC	Phe	15	TCC	Ser	4	TAC	Tyr	12	TGC	Cys	3
TTA	Leu	0	TCA	Ser	1	TAA	***	0	TGA	***	0
TTG	Leu	23	TCG	Ser	0	TAG	***	Ó	TGG	Trp	2
CTT	Leu	1	CCT	Pro	17	CAT	His	6	CGT	Arg	13
CTC	Leu	12	CCC	Pro	2	CAC	His	7	CGC	Arg	11
CTA	Leu	0	CCA	Pro	9	CAA	Gln	11	CGA	Arg	2
CTG	Leu	19	CCG	Pro	ō	CAG	Gln	3	CGG	Arg	0
ATT	Ile	15	ACT	Thr	14	AAT	Asn	9	AGT	Ser	1
ATC	Ile	23	ACC	Thr	8	AAC	Asn	13	AGC	Ser	14
ATA	Ile	0	ACA	Thr	0	AAA	Lys	19	AGA	Arg	0
ATG	Met	11	ACG	Thr	o	AAG	r\a -1	16	AGG	Arg	0
GTT	Val	3	GCT	Ala	18	GAT	Asp	12	GGT	Gly	16
GTC	Val	21	GCC	Ala	14	GAC	Asp	14	GGC	Gly	21
GTA	Val	7	CCA	A 1 -	-	~	<u>_</u>			1	

5

0

GAA Glu

GAG Glu

19

19

GGA Gly

GGG Gly

GCA Ala

GCG Ala

1

25

1

1

Figure 5K

Code	n Usage:	RDve	r5								
TTT	Phe	13	TCT	ser	12	TAT	Tyr	7	TGT	Cys	7
TTC	Phe	12	TCC	Ser	2	TAC	Tyr	13	TGC	Cys	4
TTA	Leu	0	TCA	Ser	2	TAA	***	0	TGA	***	ō
TTG	Leu	25	TCG	Ser	0	TAG	***	0	TGG	Trp	2
CTT	Leu	1	CCT	Pro	15	CAT	His	9	CGT	Arg	14
CTC	Leu	11	CCC	Pro	1	CAC	His	4	CGC	Arg	12
CTA	Leu	0	CCA	Pro	12	CAA	Gln	7	CGA	Arg	0
CTG	Leu	18	CCG	Pro	0	CAG	Gln	8	CGG	Arg	ō
ATT	Ile	19	ACT	Thr	10	AAT	Asn	9	AGT	Ser	2
ATC	Ile	20	ACC	Thr	11	AAC	Asn	12	AGC	Ser	12
ATA	Ile	0	ACA	Thr	1	AAA	Lys	13	AGA	Arg	0
ATG	Met	11	ACG	Thr	0	AAG	Lys	22	AGG	Arg	0
GTT	Val	5	GCT	Ala	21	GAT	Asp	14	GGT	Gly	14
GTC	Val	26	GCC	Ala	12	GAC	Asp	12	GGC	Gly	21
GTA	Val	1	GCA	Ala	4	GAA	Glu	18	GGA	Gly	3
GTG	Val	17	GCG	Ala	0	GAG	Glu	20	GGG	Gly	1

Figure 6

Synthetic oligos for engineered GR/RD genes (All oligos listed 5'to 3') Coding strand: 5'_____(___)n____3'
Non-coding strand: 3'_____(____)n___5' Oligos with pRAM flanking sequence identical for GR/RD 1) coding strand upstream flanking RAM-C1: ACGCCAGCCCAAGCTTAGGCCTGAGTGGC (SEQ ID NO:35) RAM-C2: CTTAATTCTCCCCATCCCCCTGTTGACAATTAATCATCGGCTCG (SEQ ID NO:36) RAM-C3: TATAATGTGAGGAATTGCGAGCGGATAACAATTTCACACA (SEQ ID NO:37) 2) coding strand downstream flanking RAM-C4: ATGGGATGTTACCTAGACCAATATGAAATATTTGGTAAAT (SEQ ID NO:38) RAM-C5: AAATGCTTAATGAATTTCAAAAAAAAAAAAAAAAAGGAATTC (SEQ ID NO:39) RAM-C6: GATATCAAGCTTATCGATACCGTCGACCTCGAGGATTATA (SEQ ID NO:40) RAM-C7: TAGAAAAAGGCCTCGGCGGCCGCTAGTTCAGTCAGTT (SEQ ID NO:41) 3) non-coding strand downstream flanking RAM-N1: AACTGACTGAACTAGCG (SEQ ID NO:42) RAM-N2: GCCGCCGAGGCCTTTTTCTATATAATCCTCGAGGTCGACG (SEQ ID NO:43) RAM-N4: TTGAAATTCATTAAGCATTTATTTACCAAATATTTCATAT (SEQ ID NO:46) RAM-N5: TGGTCTAGGTAACATCCCATCACTAGCTTTTTTTTCTATA (SEQ ID NO:47) 4) non-coding strand upstream flanking RAM-N6: TCGCAATTCCTCACATTATACGAGCCGATGATTAATTGTC (SEQ ID NO:48) RAM-N7: AACAGGGGGATGGGGAGAATTAAGGCCACTCAGGCCTAAGCTTGGGCTGGCGT (SEQ ID NO:49) GRver5 with flanking seq. of pRAM to end of Sfi I primers 1) Coding strand (Start and stop codons are underlined) GR-C1: GGAAACAGGATCCCATGATGAAACGCGAAAAGAACGTGAT (SEQ ID NO:50) GR-C2: CTACGGCCCAGAACCACTGCATCCACTGGAAGACCTCACC (SEQ ID NO:51) GR-C3: GCTGGTGAGATGCTCTTCCGAGCACTGCGTAAACATAGTC (SEQ ID NO:52) GR-C4: ACCTCCTCAAGCACTCGTGGACGTCGTGGGAGACGAGAG (SEQ ID NO:53) GR-C5: CCTCTCCTACAAAGAATTTTTCGAAGCTACTGTGCTGTTG (SEQ ID NO:54) GR-C6: GCCCAAAGCCTCCATAATTGTGGGTACAAAATGAACGATG (SEQ ID NO:55) GR-C7: TGGTGAGCATTTGTGCTGAGAATAACACTCGCTTCTTTAT (SEQ ID NO:56) GR-C8: TCCTGTAATCGCTGCTTGGTACATCGGCATGATTGTCGCC (SEQ ID NO:57) GR-C9: CCTGTGAATGAATCTTACATCCCAGATGAGCTGTGTAAGG (SEQ ID NO:58) GR-C10:TTATGGGTATTAGCAAACCTCAAATCGTCTTTACTACCAA (SEQ ID NO:59) GR-C11: AAACATCTTGAATAAGGTCTTGGAAGTCCAGTCTCGTACT (SEQ ID NO:60) GR-C12: AACTTCATCAAACGCATCATTATTCTGGATACCGTCGAAA (SEQ ID NO:61) GR-C13: ACATCCACGGCTGTGAGAGCCTCCCTAACTTCATCTCTCG (SEQ ID NO:62) GR-C14:TTACAGCGATGGTAATATCGCTAATTTCAAGCCCTTGCAT (SEQ ID NO:63) GR-Cl5:TTTGATCCAGTCGAGCAAGTGGCCGCTATTTTGTGCTCCT (SEQ ID NO:64) GR-C16:CCGGCACCACTGGTTTGCCTAAAGGTGTCATGCAGACTCA (SEQ ID NO:65) GR-C17: CCAGAATATCTGTGTGCGTTTGATCCACGCTCTCGACCCT (SEQ ID NO:66) GR-C18: CGTGTGGGTACTCAATTGATCCCTGGCGTGACTGTGCTGG (SEQ ID NO:67) GR-C19: TGTATCTGCCTTTCTTTCACGCCTTTGGTTTCTCTATTAC (SEQ ID NO:68) GR-C20:CCTGGGCTATTTCATGGTCGGCTTGCGTGTCATCATGTTT (SEQ ID NO:69)

Figure 6 (Cont.)

```
GR-C21: CGTCGCTTCGACCAAGAAGCCTTCTTGAAGGCTATTCAAG
                                                      (SEQ ID NO:70)
GR-C22: ACTACGAGGTGCGTTCCGTGATCAACGTCCCTTCAGTCAT
                                                      (SEQ ID NO:71)
GR-C23: TTTGTTCCTGAGCAAATCTCCTTTGGTTGACAAGTATGATCTG
                                                      (SEQ ID NO:72)
GR-C24: AGCAGCTTGCGTGAGCTGTGCTGCTGCTCCTT
                                                      (SEQ ID NO:73)
GR-C25: TGGCCAAAGAAGTGGCCGAGGTCGCTAAGCGTCTGAA
                                                      (SEQ ID NO:74)
GR-C26: CCTCCCTGGTATCCGCTGCGGTTTTGGTTTGACTGAGAGC
                                                      (SEQ ID NO:75)
GR-C27: ACTTCTGCTAACATCCATAGCTTGCGAGACGAGTTTAAGT
                                                      (SEQ ID NO:76)
GR-C28: CTGGTAGCCTGGGTCGCGTGACTCCTCTTATGGCTGCAAA
                                                      (SEQ ID NO:77)
GR-C29: GATCGCCGACCGTGAGACCGGCAAAGCACTGGGCCCAAAT
                                                      (SEQ ID NO:78)
GR-C30: CAAGTCGGTGAATTGTGTATTAAGGGCCCTATGGTCTCTA
                                                      (SEQ ID NO:79)
GR-C31: AAGGCTACGTGAACAATGTGGAGGCCACTAAAGAAGCCAT
                                                      (SEQ ID NO:80)
GR-C32: TGATGATGATGGCTGGCTCCATAGCGGCGACTTCGGTTAC
                                                      (SEQ ID NO:81)
GR-C33: TATGATGAGGACGAACACTTCTATGTGGTCGATCGCTACA
                                                      (SEQ ID NO:82)
GR-C34: AAGAATTGATTAAGTACAAAGGCTCTCAAGTCGCACCAGC
                                                      (SEQ ID NO:83)
GR-C35: CGAACTGGAAGAAATTTTGCTGAAGAACCCTTGTATCCGC
                                                      (SEQ ID NO:84)
GR-C36:GACGTGGCCGTCGTGGGTATCCCAGACTTGGAAGCTGGCG
                                                      (SEQ ID NO:85)
GR-C37: AGTTGCCTAGCGCCTTTGTGGTGAAACAACCCGGCAAGGA
                                                      (SEQ ID NO:86)
GR-C38: GATCACTGCTAAGGAGGTCTACGACTATTTGGCCGAGCGC
                                                      (SEQ ID NO:87)
GR-C39:GTGTCTCACACCAAATATCTGCGTGGCGGCGTCCGCTTCG
                                                      (SEQ ID NO:88)
GR-C40: TCGATTCTATTCCACGCAACGTTACCGGTAAGATCACTCG
                                                      (SEQ ID NO:89)
GR-C41: TAAAGAGTTGCTGAAGCAACTCCTCGAAAAAGCTGGCGGC
                                                      (SEQ ID NO:90)
GR-C42: TAGTAAAGTCTTCATGATTATATAGAAAAAAAAGCTAGTG
                                                      (SEQ ID NO:91)
2) non-coding strand
GR-N1: TAATCATGAAGACT<u>TTACTA</u>GCCGCCAGCTTTTTCGAGGA
                                                      (SEQ ID NO:92)
GR-N2: GTTGCTTCAGCAACTCTTTACGAGTGATCTTACCGGTAAC
                                                      (SEQ ID NO:93)
GR-N3: GTTGCGTGGAATAGAATCGACGAAGCGGACGCCGCCACG
                                                      (SEQ ID NO:94)
GR-N4: CAGATATTTGGTGTGAGACACGCGCTCGGCCAAATAGTCGT (SEQ ID NO:95)
GR-N5: AGACCTCCTTAGCAGTGATCTCCTTGCCGGGTTGTTTCAC
                                                      (SEQ ID NO:96)
GR-N6: CACAAAGGCGCTAGGCAACTCGCCAGCTTCCAAGTCTGGG
                                                      (SEQ ID NO:97)
GR-N7: ATACCCACGACGCCACGTCGCGGATACAAGGGTTCTTCA
                                                      (SEQ ID NO:98)
GR-N8: GCAAAATTTCTTCCAGTTCGGCTGGTGCGACTTGAGAGCC
                                                      (SEQ ID NO:99)
GR-N9: TTTGTACTTAATCAATTCTTTGTAGCGATCGACCACATAG
                                                      (SEQ ID NO:100)
GR-N10: AAGTGTTCGTCCTCATCATAGTAACCGAAGTCGCCGCTAT
                                                      (SEQ ID NO:101)
GR-N11:GGAGCCAGCCATCATCATCAATGGCTTCTTTAGTGGCCTC
                                                      (SEQ ID NO:102)
GR-N12: CACATTGTTCACGTAGCCTTTAGAGACCATAGGGCCCTTA
                                                      (SEQ ID NO:103)
GR-N13: ATACACAATTCACCGACTTGATTTGGGCCCAGTGCTTTGC
                                                      (SEQ ID NO:104)
GR-N14: CGGTCTCACGGTCGGCGATCTTTGCAGCCATAAGAGGGGT
                                                      (SEQ ID NO:105)
GR-N15: CACGCGACCCAGGCTACCAGACTTAAACTCGTCTCGCAAG
                                                      (SEQ ID NO:106)
GR-N16: CTATGGATGTTAGCAGAAGTGCTCTCAGTCAAACCAAAAC
                                                      (SEQ ID NO:107)
GR-N17: CGCAGCGGATACCAGGGAGGTTCAGACGCTTAGCAGCGAC
                                                      (SEQ ID NO:108)
GR-N18: CTCGGCCACTTCTTTGGCCAAAGGAGCAGCGCCACAGCAC
                                                      (SEQ ID NO:109)
GR-N19: AGCTCACGCAAGCTGCTCAGATCATACTTGTCAACCAAAG
                                                      (SEQ ID NO:110)
GR-N20: GAGATTTGCTCAGGAACAAAATGACTGAAGGGACGTTGAT
                                                      (SEQ ID NO:111)
GR-N21: CACGGAACGCACCTCGTAGTCTTGAATAGCCTTCAA
                                                      (SEQ ID NO:112)
GR-N22:GAAGGCTTCTTGGTCGAAGCGACGAAACATGATGACACGCAAGC (SEQ ID NO:113)
GR-N23: CGACCATGAAATAGCCCAGGGTAATAGAGAAACCAAAGGC
                                                     (SEQ ID NO:114)
GR-N24:GTGAAAGAAAGGCAGATACACCAGCACAGTCACGCCAGGG
                                                     (SEQ ID NO:115)
GR-N25: ATCAATTGAGTACCCACACGAGGGTCGAGAGCGTGGATCA
                                                     (SEQ ID NO:116)
GR-N26: AACGCACACAGATATTCTGGTGAGTCTGCATGACACCTTT
                                                     (SEQ ID NO:117)
GR-N27: AGGCAAACCAGTGGTGCCGGAGGAGCACAAAATAGCGGCC
                                                     (SEQ ID NO:118)
```

Figure 6 (Cont.)

```
GR-N28: ACTTGCTCGACTGGATCAAAATGCAAGGGCTTGAAATTAG
                                                    (SEQ ID NO:119)
GR-N29: CGATATTACCATCGCTGTAACGAGAGATGAAGTTAGGGAG
                                                    (SEQ ID NO:120)
GR-N30:GCTCTCACAGCCGTGGATGTTTTCGACGGTATCCAGAATA
                                                    (SEQ ID NO:121)
GR-N31: ATGATGCGTTTGATGAAGTTAGTACGAGACTGGACTTCCA
                                                    (SEQ ID NO:122)
GR-N32: AGACCTTATTCAAGATGTTTTTGGTAGTAAAGACGATTTG
                                                    (SEQ ID NO:123)
GR-N33: AGGTTTGCTAATACCCATAACCTTACACAGCTCATCTGGG
                                                    (SEO ID NO:124)
GR-N34: ATGTAAGATTCATTCACAGGGGCGACAATCATGCCGATGT
                                                    (SEQ ID NO:125)
GR-N35: ACCAAGCAGCGATTACAGGAATAAAGAAGCGAGTGTTATT
                                                    (SEQ ID NO:126)
GR-N36: CTCAGCACAAATGCTCACCACATCGTTCATTTTGTACCCA
                                                    (SEQ ID NO:127)
GR-N37: CAATTATGGAGGCTTTGGGCCAACAGCACAGTAGCTTCGA
                                                    (SEQ ID NO:128)
GR-N38: AAAATTCTTTGTAGGAGAGGCTCTCGTCTCCCACGACGTC
                                                    (SEQ ID NO:129)
GR-N39: CACGAGTGCTTGAGGGAGGTGACTATGTTTACGCAGTGCT
                                                    (SEQ ID NO:130)
GR-N40: CGGAAGAGCATCTCACCAGCGGTGAGGTCTTCCAGTGGAT
                                                    (SEQ ID NO:131)
GR-N41:GCAGTGGTTCTGGGCCGTAGATCACGTTCTTTTCGCGTTT
                                                    (SEQ ID NO:132)
GR-N42: CATCATGGGATCCTGTTTCCTGTGTGAAATTGTTATCCGC
                                                    (SEQ ID NO:133)
```

RDver5 with flanking sequence of pRAM to end of Sfi I primers 1) coding strand

```
RD-C1: GGAAACAGGATCCCATGATGAAGCGTGAGAAAAATGTCAT
                                                       (SEQ ID NO:134)
RD-C2: CTATGGCCCTGAGCCTCTCCATCCTTTGGAGGATTTGACT
                                                       (SEQ ID NO:135)
RD-C3: GCCGGCGAAATGCTGTTTCGTGCTCTCCGCAAGCACTCTC
                                                       (SEQ ID NO:136)
RD-C4: ATTTGCCTCAAGCCTTGGTCGATGTGGTCGGCGATGAATC
                                                       (SEQ ID NO:137)
RD-C5: TTTGAGCTACAAGGAGTTTTTTGAGGCAACCGTCTTGCTG
                                                       (SEQ ID NO:138)
RD-C6: GCTCAGTCCCTCCACAATTGTGGCTACAAGATGAACGACG
                                                       (SEQ ID NO:139)
RD-C7: TCGTTAGTATCTGTGCTGAAAACAATACCCGTTTCTTCAT
                                                       (SEQ ID NO:140)
RD-C8: TCCAGTCATCGCCGCATGGTATATCGGTATGATCGTGGCT
                                                       (SEQ ID NO:141)
RD-C9: CCAGTCAACGAGAGCTACATTCCCGACGAACTGTGTAAAG
                                                       (SEQ ID NO:142)
RD-C10: TCATGGGTATCTCTAAGCCACAGATTGTCTTCACCACTAA
                                                       (SEQ ID NO:143)
RD-C11: GAATATTCTGAACAAAGTCCTGGAAGTCCAAAGCCGCACC
                                                       (SEQ ID NO:144)
RD-C12: AACTTTATTAAGCGTATCATCTTGGACACTGTGGAGA
                                                       (SEQ ID NO:145)
RD-C13:ATATTCACGGTTGCGAATCTTTGCCTAATTTCATCTCTCG
                                                       (SEQ ID NO:146)
RD-C14:CTATTCAGACGGCAACATCGCAAACTTTAAACCACTCCAC
                                                       (SEQ ID NO:147)
RD-C15:TTCGACCCTGTGGAACAAGTTGCAGCCATTCTGTGTAGCA
                                                       (SEQ ID NO:148)
RD-C16: GCGGTACTACTGGACTCCCAAAGGGAGTCATGCAGACCCA
                                                       (SEQ ID NO:149)
RD-C17: TCAAAACATTTGCGTGCGTCTGATCCATGCTCTCGATCCA
                                                       (SEQ ID NO:150)
RD-C18: CGCTACGGCACTCAGCTGATTCCTGGTGTCACCGTCTTGG
                                                       (SEQ ID NO:151)
RD-C19: TCTACTTGCCTTTCTTCCATGCTTTCGGCTTTCATATTAC
                                                       (SEQ ID NO:152)
RD-C20:TTTGGGTTACTTTATGGTCGGTGTCTCCGCGTGATTATGTTC
                                                       (SEQ ID NO:153)
RD-C21:CGCCGTTTTGATCAGGAGGCTTTCTTGAAAGCCATCCAAG
                                                       (SEQ ID NO:154)
RD-C22: ATTATGAAGTCCGCAGTGTCATCAACGTGCCTAGCGTGAT
                                                       (SEQ ID NO:155)
RD-C23: CCTGTTTTTGTCTAAGAGCCCACTCGTGGACAAGTACGAC
                                                       (SEQ ID NO:156)
RD-C24: TTGTCTTCACTGCGTGAATTGTGTTGCGGTGCCGCTCCAC
                                                       (SEQ ID NO:157)
RD-C25: TGGCTAAGGAGGTCGCTGAAGTGGCCGCCAAACGCTTGAA
                                                       (SEQ ID NO:158)
RD-C26: TCTTCCAGGGATTCGTTGTGGCCTCACCGAATCT
                                                       (SEQ ID NO:159)
RD-C27: ACCAGCGCTATTATTCAGTCTCTCCGCGATGAGTTTAAGA
                                                       (SEQ ID NO:160)
RD-C28: GCGGCTCTTTGGGCCGTGTCACTCCACTCATGGCTGCTAA
                                                       (SEQ ID NO:161)
RD-C29: GATCGCTGATCGCGAAACTGGTAAGGCTTTGGGCCCTAAC
                                                       (SEQ ID NO:162)
RD-C30: CAAGTGGGCGAGCTGTGTATCAAAGGCCCTATGGTGAGCA
                                                       (SEQ ID NO:163)
RD-C31:AGGGTTATGTCAATAACGTCGAAGCTACCAAGGAGGCCAT
                                                       (SEQ ID NO:164)
RD-C32: CGACGACGACGGCTGGTTGCATTCTGGTGATTTTGGATAT
                                                       (SEQ ID NO:165)
RD-C33: TACGACGAAGATGAGCATTTTTACGTCGTGGATCGTTACA
                                                       (SEQ ID NO:166)
RD-C34: AGGAGCTGATCAAATACAAGGGTAGCCAGGTTGCTCCAGC
                                                       (SEQ ID NO:167)
RD-C35: TGAGTTGGAGGAGATTCTGTTGAAAAATCCATGCATTCGC
                                                       (SEQ ID NO:168)
```

Figure 6 (Cont.)

```
RD-C36:GATGTCGCTGTGGTCGGCATTCCTGATCTGGAGGCCGGCG
                                                        (SEQ ID NO:169)
 RD-C37: AACTGCCTTCTGCTTTCGTTGTCAAGCAGCCTGGTAAAGA
                                                        (SEQ ID NO:170)
 RD-C38:AATTACCGCCAAAGAAGTGTATGATTACCTGGCTGAACGT
                                                        (SEQ ID NO:171)
 RD-C39:GTGAGCCATACTAAGTACTTGCGTGGCGGCGTGCGTTTTG
                                                        (SEQ ID NO:172)
 RD-C40:TTGACTCCATCCCTCGTAACGTAACAGGCAAAATTACCCG
                                                        (SEQ ID NO:173)
 RD-C41: CAAGGAGCTGTTGAAACAATTGTTGGAGAAGGCCGGCGGT
                                                        (SEQ ID NO:174)
 RD-C42: TAGTAAAGTCTTCATGATTATATAGAAAAAAAGCTAGTG
                                                        (SEQ ID NO:175)
 2) non-coding strand
 RD-N1: TAATCATGAAGACTTTACTAACCGCCGGCCTTCTCCAACA (SEQ ID NO:176)
 RD-N2: ATTGTTTCAACAGCTCCTTGCGGGTAATTTTGCCTGTTAC (SEQ ID NO:177)
 RD-N3: GTTACGAGGGATGGAGTCAACAAAACGCACGCCCACGC (SEQ ID NO:178)
 RD-N4: AAGTACTTAGTATGGCTCACACGTTCAGCCAGGTAATCAT (SEQ ID NO:179)
 RD-N5: ACACTTCTTTGGCGGTAATTTCTTTACCAGGCTGCTTGAC (SEQ ID NO:180)
 RD-N6: AACGAAAGCAGAAGGCAGTTCGCCGGCCTCCAGATCAGGA (SEQ ID NO:181)
RD-N7: ATGCCGACCACAGCGACATCGCGAATGCATGGATTTTTCA (SEQ ID NO:182)
RD-N8: ACAGAATCTCCTCCAACTCAGCTGGAGCAACCTGGCTACC (SEQ ID NO:183)
RD-N9: CTTGTATTTGATCAGCTCCTTGTAACGATCCACGACGTAA (SEQ ID NO:184)
RD-N10:AAATGCTCATCTTCGTCGTAATATCCAAAATCACCAGAAT (SEQ ID NO:185)
RD-N11:GCAACCAGCCGTCGTCGTCGATGGCCTCCTTGGTAGCTTC (SEQ ID NO:186)
RD-N12:GACGTTATTGACATAACCCTTGCTCACCATAGGGCCTTTG (SEQ ID NO:187)
RD-N13:ATACACAGCTCGCCCACTTGGTTAGGGCCCAAAGCCTTAC (SEQ ID NO:188)
RD-N14: CAGTTTCGCGATCAGCGATCTTAGCAGCCATGAGTGGAGT (SEQ ID NO:189)
RD-N15:GACACGGCCCAAAGAGCCGCTCTTAAACTCATCGCGGAGA (SEQ ID NO:190)
RD-N16:GACTGAATAATAGCGCTGGTAGATTCGGTGAGGCCGA
                                                 (SEQ ID NO:191)
RD-N17:AGCCACAACGAATCCCTGGAAGATTCAAGCGTTTGGCGGCCAC (SEQ ID NO:192)
RD-N18:TTCAGCGACCTCCTTAGCCAGTGGAGCGGCACCGCAACAC (SEQ ID NO:193)
RD-N19: AATTCACGCAGTGAAGACAAGTCGTACTTGTCCACGAGTG (SEQ ID NO:194)
RD-N20:GGCTCTTAGACAAAAACAGGATCACGCTAGGCACGTTGAT (SEQ ID NO:195)
RD-N21:GACACTGCGGACTTCATAATCTTGGATGGCTTTCAAGAAA (SEQ ID NO:196)
RD-N22:GCCTCCTGATCAAAACGGCGGAACATAATCACGCGGAGAC (SEQ ID NO:197)
RD-N23:CGACCATAAAGTAACCCAAAGTAATATGAAAGCCGAAAGC (SEQ ID NO:198)
RD-N24:ATGGAAGAAAGGCAAGTAGACCAAGACGGTGACACCAGGA (SEQ ID NO:199)
RD-N25:ATCAGCTGAGTGCCGTAGCGTGGATCGAGAGCATGGATCA (SEQ ID NO:200)
RD-N26:GACGCACACTTTTGATGGGTCTGCATGACTCCCTT (SEQ ID NO:201)
RD-N27:TGGGAGTCCAGTAGTACCGCTGCTACACAGAATGGCTGCA (SEQ ID NO:202)
RD-N28:ACTTGTTCCACAGGGTCGAAGTGGAGTGGTTTAAAGTTTG (SEQ ID NO:203)
RD-N29:CGATGTTGCCGTCTGAATAGCGAGAGATGAAATTAGGCAA (SEQ ID NO:204)
RD-N30:AGATTCGCAACCGTGAATATTCTCCACAGTGTCCAAGATG (SEQ ID NO:205)
RD-N31:ATGATACGCTTAATAAAGTTGGTGCGGCTTTGGACTTCCA (SEQ ID NO:206)
RD-N32:GGACTTTGTTCAGAATATTCTTAGTGGTGAAGACAATCTG (SEQ ID NO:207)
RD-N33:TGGCTTAGAGATACCCATGACTTTACACAGTTCGTCGGGA (SEQ ID NO:208)
RD-N34:ATGTAGCTCTCGTTGACTGGAGCCACGATCATACCGATAT (SEQ ID NO:209)
RD-N35:ACCATGCGGCGATGACTGGAATGAAGAAACGGGTATTGTT (SEQ ID NO:210)
RD-N36:TTCAGCACAGATACTAACGACGTCGTTCATCTTGTAGCCA (SEQ ID NO:211)
RD-N37:CAATTGTGGAGGGACTGAGCCAGCAAGACGGTTGCCTCAA (SEQ ID NO:212)
RD-N38:AAAACTCCTTGTAGCTCAAAGATTCATCGCCGACCACATC (SEQ ID NO:213)
RD-N39:GACCAAGGCTTGAGGCAAATGAGAGTGCTTGCGGAGAGCA (SEQ ID NO:214)
RD-N40:CGAAACAGCATTTCGCCGGCAGTCAAATCCTCCAAAGGAT (SEQ ID NO:215)
RD-N41:GGAGAGGCTCAGGGCCATAGATGACATTTTTCTCACGCTT (SEQ ID NO:216)
RD-N42: CATCATGGGATCCTGTTTCCTGTGTGAAATTGTTATCCGC (SEQ ID NO:217)
```

Figure 7

RELLUC.SEQ A T RLUCVER1.SEQ A T RLUCVER2.SEQ A T RLUCFINL.SEQ A T	GACTTCC GGCTTCC GGCTTCC	A A G G T T T A A A A G G T G T A G A A G G T G T	T G A T C C A G A C G A C C C C G A C G A C C C C G A C G A C C C C G A	A C A A A G G A A A C G G A G C G C A G C G C	40 40 40 40
RLUCVER1.SEQT G RLUCVER2.SEQT G	A T C A C C G A T C A C T G	GCCCTCAG GGCCTCAG	т с с т с с с с с т с с т с с с с Гт	A G A T G T A A A C A A A T C G C T G C A A G C A G A T C G C T G C A A G C A A A T C G C T G C A A G C A A A T	80 80 80
RLUCVER1.SEQG A	A C G T G C T	GIG AICIT CICIT 1	TCATCAACT	ATTATGATTCAGAA ACTACGACAGCGAG ACTATGATTCCGAG ACTATGATTCCGAG	120
RLUCVER1.SEQ A A RLUCVER2.SEQ A A	G C A C G C C	G A G A A C G C C G A G A A C G C C	CGTGATCTT	TTTACATGGTAACG CCTGCACGGCAACG TCTGCATGGTAACG TCTGCATGGTAACG	160 160 160 160
RLUCVER1.SEQCC RLUCVER2.SEQCT	G C C T C C A	G C T A C C T G T G C T A C C T G T	T G G A G G C A C	GTTGTGCCACATAT GTGGTGCCTCACAT GTCGTGCCTCACAT GTCGTGCCTCACAT	200 200 200 200
RLUCVER1.SEQC G	A G C C C G T	G G С С С G С Т G G G С Т С G С Т G	GCATCATCC	C A G A T C T T A T T G G T C T G A C C T G A T C G G C C T G A T C T G A T C G G A C T G A T C T G A T C G G A	240 240 240 240
RLUCVER1. SEQ A T (G G G C A A G G G T A A G	T C C G G C A A G T C C G G C A A G	GAGCIGGICAAI GAGCIGGIGAA	CGGCTCCTACCGCC	280 280 280 280
RLUCVER1. SEO T [G]	CITIGIG AICIC.	AICITACAAGIT	PACCTEACC	G C C T G G T T C G A G C T G C T T G G T T C G A G C T	320 320 320 320
RLUCVER1.SEQG C	T G A A C C T	GC CCAAGAA FC CAAAGAA	AGATCATCT AAATCATCT	TCGTGGGCCACGAC TTGTGGGCCACGAC	360 360 360 360
RLUCVER1.SEQT G (G G G G G C T	г <u>ссста</u> ссс г стста ссс	TTCCACTA TTTCACTA	CTCCTACGAGCACC CTCCTACGAGCACC	400 400 400 400
RLUCVER1. SEQ A G (GACAAGA: GACAAGA:	CAAGGCCA	TCG TGC A C	G CCG AGA GCG TGG T G C T G AGA G T G T C G T	440 440 440 440

Figure 7 (Cont.)

														_																		
RELLUC. SI RLUCVER1 RLUCVER2 RLUCFINL	. SEQ	GG	A	cla	T T	G.	АТ АТ	C	G .	A IG	T			G	G	G A	C	G	A G	T	G	3 C	C	T (G A	C	A	Ţſ	cG	A	3 4	80 80
RELLUC. SE RLUCVER1. RLUCVER2. RLUCFINL.	SEQ	3 A 3 A	G	G A G A	T	A,	rc	G	clo	: 0 : 0	T	G I	ат ат	C	A A	AG	A	G	CG	A[G (; A	G	G	3 C	G	A	G	A A		. 5	20 20 20 20
RELLUC.SE RLUCVER1. RLUCVER2. RLUCFINL.	SEO 1	G	G	r G	cl	T C	3 G	A.	G F	A	디	A /	, c	T	T (T	T	C	3 T	ر اما	G A		A	C	A	T	G	c) 2	G		5	60 60 60
RELLUC.SE RLUCVER1. RLUCVER2. RLUCFINL.	SEO A	Ā	G	A	A	G	T	C.	АТ Ат	G		GIO	A	V L	G C	T	G	G F	\ G	C	T	G	A	3 G	A	G	T :	rC	G	cC	6	00 00 00
RELLUC.SE RLUCVER1. RLUCVER2. RLUCFINL.	SEQ G	c		A		CI	G	G .	A G	C		T	T	C i	A A	G	G.	A G	A	AC	G	G	C	A	G	G	T (<u> </u>	G	c c	6	
RELLUC.SE RLUCVER1 RLUCVER2 RLUCFINL	SEO G	G	C (1	AI			T	T	C	C	r G	G	C	: lc]C	GK	C G	A	G A	T	C	cc	T	익	T !	G	T	G	A A A A A A A A		10
RELLUC.SEG RLUCVER1.S RLUCVER2.S RLUCFINL.S	SEQG	G	GA	G	GIO	A	A	GIO	: C	C	G F		G	TC	G	T	G	C A	G.	TA	叵	G :	r G	C	G	C	A A	C	T A	A C	72	0
RELLUC. SEG RLUCVER1.S RLUCVER2.S RLUCFINL.S	ひしいん	M I	_ G		CIT	. А	16:14	C 7	16		CIC		~ 1	~ 1 2			~ .		~		\sim	m 1 /	-1-	_	-			1 -		G T G T G T		_
RELLUC. SEG RLUCVER1.S RLUCVER2.S RLUCFINL.S	EO T	CZ	A T	C	GA	G	T		G	A		C	7	3 G		T '	rc	T	T	T	C	CA	A	c	G	ع اد] A	T		; T ; T ; T	80	0
RELLUC.SEC RLUCVER1.S RLUCVER2.S RLUCFINL.S	EQC	G I G I	A A	9 9	G C	G	C (A S A S A S	AAA	G I	A A A A	G G	T T	1 1	c	c (A T	A	T F	, c	Ī	G A G A	G	T	T C	r 6	T	C G	A A A A	G	844	0
RELLUC.SEQ RLUCVER1.S RLUCVER2.S RLUCFINL.S	G EQG EQG	T A	A	A A	A G	G	T (. T . T	T G C	C I	T C	T T	T I	T	င်	G (ACA	A G	G A G A	G	G . G .	A T	G	c c	A C		T	G	A T	G]G	880 880	0

Figure 7 (Cont.)

LLUC. SEQ AAATGGGAAAATATATCAAATCGTTCGTTGAGCGAGTTC	T 920
UCVER1. SEQAGAT G G G C A A G T A C A T C A A G A G C T T C G T G G A G C G C G T G C	T 920
UCVER2. SEQAAATGGGTAAGTACATCAAGAGCTTCGTGGAGCGCGTGC	
ucfinl.seqaaat g g g taagt aclat c a a <u>g a g c</u> t t c g t g g a g c g c g t g c	T 920
LLUC. SEQ CAAAATGAACAA	933
JCVER1. SEQGA AGA ACG AGC AG	933
JCVER2.SEQIGIA AIGIA AIGIG AIGIC AIGI	933
UCFINL. SEQGA AGA ACG AGC AG	933

Figure 8

			_	_	_		_		_	_	_	_	_			_						_			_																	
RELLUC. SEQ	M	T	S	K	V	Y	D	P	Ε	Q	R	K	R	М	I	T	G	P	Q	W	W	Α	R	С	ĸ	Q	М	N	v	L	D	s	F	I	N	Y	Y		, s	E	11	8
RLUCVER1.SEQ	M	A	s	K	ν	Y	D	P	E	Q	R	K	R	М	I	T	G	P	Q	W	W	Α	R	¢	K	Q	М	N	V	L	D	s	F	I	N	Y	Y	. 0) S	E	11/	8
RLUCVER2.SEQ	M	A	s	K	ν	Y	D	P	Ε	Q	R	K	R	М	Ι	T	G	₽	Q	W	W	A	R	С	K	Q	М	N	V	L	D	s	F	I	N	Y	Y	0	S	E	118	8
RLUCFINL. SEQ	M	A	s	K	V	Y	D	P	E	Q	R	K	R	M	I	T	G	₽	Q	W	W	A	R	С	K	Q	M	N	V	L	D	S	F	I	N	Y	Y	0	S	E	118	8
RELLUC. SEQ	ĸ	н	A	E	N	A	ν	I	F	Ł	н	G	N	А	A	s	s	Y	L	W	R	н	v	v	P	н	I	E	P	v	A	R	С	1	ı	P	0	L	. I	G	230	В
RLUCVER1.SEQ	K	H	A	Ε	N	A	٧	I	F	Ļ	Н	G	N	Α	Α	s	s	Y	L	W	R	н	٧	٧	P	н	I	E	₽	v	А	R	c	I	I	P	D	L	. I	G	230	A
RLUCVER2.SEQ	ĸ	Н	A	E	N	A	v	I	F	L	н	Ģ	N	A	A	s	s	Y	L	W	R	н	v	٧	P	н	I	E	₽	v	А	R	c	I	1	P	D	L	. I	G	238	Á
RLUCFINL. SEQ	K	Н	A	E	N	A	V	I	F	L	Н	G	N	A	A	s	s	Y	L	W	R	H	٧	V	₽	H	I	E	P	V	A	R	c	I	I	P	D	L	· I	G	238	8
RELLUC.SEQ	м	G	ĸ	\$	G	ĸ	s	G	N	G	s	Y	R	L	L	D	н	Y	к	Y	L	т	А	w	F	E	ı.	ı.	N	T.	P	ĸ	ĸ	T	T	F	v	· G	н	D	355	R
RLUCVER1.SEQ	м	G	ĸ	s	G	ĸ	s	G	N	G	s	Y	R	L	L	D	н	Υ	к	Y	L	T	A	W	F	E	L	L	N	L	P	ĸ	ĸ	T	ī	F	v	G	. н	D	350	2
RLUCVER2.SEQ	М	G	ĸ	s	G	ĸ	s	G	N	G	s	Y	R	L	L	D	н	Y	к	Y	L	т	A	w	F	E	т.	Τ.	N	7.	p	ĸ	ĸ	- T	T	F	v	۰	 H	n	356	,
RLUCFINL. SEQ	M	G	ĸ	s	G	ĸ	s	G	N	G	s	Y	R	L	L	D	Н	Y	ĸ	Y	L	T	A	W	F	Ε	L	L	N	L	P	ĸ	ĸ	ī	I	F	v	G	Н	D	358	3
מפונוור פפת	£a?	_		_			E,		v		·	_		_	_	.,			_				_	_	_			_		_	_	_		_	_							
RELLUC. SEQ	Tall	6	``	~		~			1	2	1	E	н.	Q	U	K	ī	K	A	1	٧	н	A	E	5	ν	٧	D -	v	I	E	S	W	D	E	W	P	D	I	Ē	478	3
RLUCVER1.SEQ	띪	9	,	~	T.				1	2	ı	-	н	Q	ט	K	1	к	A	1	٧	н	A	E	S	v 	٧	D -	V	1	E	S	W	D	E	W	P	D	Ι	Е	478	3
RLUCVER2.SEQ	e.	0	~	~			F	п	ı	٥	1	Ŀ	п	Q	ט	K	1	K	A	Ţ	٧	н	A	E	S	٧	٧	D	٧	1	E	5	W	D	Ε	W	P	D	I	E	478	}
RLUCFINL. SEQ		G	~	C	υ	*	F	п	1	3	1	E	н	Q	ט	ĸ	1	ĸ	A	I	٧	н	A	E	S	V	V	D	V	Ι	Е	5.	W	D	3	W	₽	D	I	Ε	478	ţ
RELLUC.SEQ	Ε	D	I	A	L	I	K	S	E	E	G	E	K	M	V	L	E	N	N	F	F	v	E	т	М	L	P	Ş	K	I	М	R	ĸ	L	E	P	Ε	E	F	A	598	ł
RLUCVER1.SEQ	Ε	D	Ι	A	L	1	K	S	E	E	G	E	K	М	٧	L	E	N	N	F	F	٧	E	T	M	L	P	S	K	I	M	R	K	L	E	P	Ε	E	F	Α	598	ţ
RLUCVER2.SEQ	Ε	D	Ι	A	L	I	K	5	E	Е	G	E	K	М	٧	L	E	N	N	F	F	٧	E	T	M	L	P	s	K	I	M	R	K	L	E	P	E	E	F	A	598	J
RLUCFINL.SEQ	Е	D	I	A	L	I	K	s	E	Ε	G	Ε	K	M	٧	L	E	N	N	F	F	٧	E	T	M	L	P	S	K	I	M :	R	K	L	E	P	E	E	F	A	598	l
RELLUC.SEQ	A	Y	L	E	P	F	ĸ	£	ĸ	G	E	v	R	R	P	т	L	s	W	P	R	Ε	I	P	L	v	K (G	G I	ĸ	P	D	v	v	0	I	v	R	N	Y	718	
RLUCVER1.SEQ	A	Y	L	E	P	F	K	E	ĸ	G	E	٧	R	R	₽	T	L	S	W	P	R	Е	τ	₽	L	V	K	G	G I	К	Р :	D	v	v	0	I	v	R	N	Y	718	1
LUCVER2.SEQ	A	Y	L	E	₽	F	K	E	K	G	E	٧	R	R	P	T	L	S	W	Р	R	E	1	P	L	V	K (G	G I	K	P	D	v	v	0	Ι	ν	R	N	Y	718	
RLUCFINL.SEQ	A	Y	L	E	P	F	ĸ	Ē	K	G	E	V	R	R	P	T	L	S	W	P	R	E	Į	P	L '	V	K (G (G	к	P	D	V	V	Q	I	v	R	N	Y	718	
RELLUC.SEQ	N	A	Y	L	R	A	s	D	D	L	P	ĸ	м	F	I	E	s	D	P	G	F	F	s	N	Δ.	т .	v i	F. 6	c i	Δ.	K 1	K	F	Þ	N	т	F		v	ĸ	838	
RLUCVER1.SEQ	N	A	Y	L	R	A	s	D	D	L	P	ĸ	М	F	ī	E	s	D	P	G	F	F	s	N.	A	1	v i	E (G i	Α.	K	K	F	P	N	Ť	E	F	v	ĸ	939	
RLUCVER2.SEQ	N	A	Y	L	R	A	s	D	D	L	₽	ĸ	м	F	I	E	s	D	P	G	F	F	s	N .	A	- T '	V 1	E (G /	Α.	KI	K	F	P	N	Т	E	F	v	ĸ	ยรย	
RLUCFINL.SEQ	N	A	Y	L	R	A	s	D	D	L	₽	ĸ	M	F	I	E	s	D	P	G	F	F	S	N .	A :	Ι,	v	E	Gi	١.	K	K	F	P	N	T	E	F	v	ĸ	838	
ELLUC. SEQ	v	ĸ	c	T.	ч	F	•	^	F	n	ь	ъ	_	F	u	c	v	v		v			.,	P 1							^											
LUCVER1.SEQ																																									931	
LUCVER2.SEQ	v	ĸ	G	ī.	H	F	S	ď	E	D	^ A	Þ	מ	Ē	M	G	r	i.	1 T	T.	2	E '	٧,,,	E .					N 2	، د د	ď										931	
LUCFINL. SEQ	v	ĸ	G	1.	H	r	S	0	F	Б	2	D	מ	£	M	9	v	· ·	1	r v		r r	٧.		r '	7 1			N B	، د	ď										931	
	•	••	-		**	£	J	~	٥		^	-	_	-	177	J	^	٠.	1.	•	3	£	٧	c 1	κ,	,	ь !		N I	٠ د	¥										931	
					_																																					

Figure 9A

Codon usage in RELLUC

(Renilla reniformis; Genbank ACCESSION:M63501; Medline:91239583)

TTT	Phe	11	TCT	Ser	5	TAT	Tyr	12	TGT	Cys	3
TTC	Phe	5	TCC	Ser	1	TAC	Tyr	1	TGC	Cys	0
TTA	Leu	8	TCA	Ser	6	TAA	***	0	TGA	***	0
TTG	Leu	4	TCG	Ser	4	TAG	***	0	TGG	Trp	8
CTT	Leu	8	CCT	Pro	5	CAT	His	9	CGT	Arq	4
CTC	Leu	1	CCC	Pro	0	CAC	His	1	CGC	Arg	ō
CTA	Leu	1	CCA	Pro	11	CAA	Gln	6	CGA	Arg	2
CTG	Leu	0	CCG	Pro	2	CAG	Gln	1	CGG	Arg	2
ATT	Ile	12	ACT	Thr	4	AAT	Asn	11	AGT	Ser	2
ATC	Ile	6	ACC	Thr	1	AAC	Asn	2	AGC	Ser	1
ATA	Ile	3	ACA	Thr	1	AAA	Lys	21	AGA	Arg	2
ATG	Met	9	ACG	Thr	0	AAG	Lys	6	AGG	Arg	3
GTT	Val	12	GCT	Ala	5	GAT	Asp	16	GGT	Gly	10
GTC	Val	2	GCC	Ala	3	GAC	Asp	1	GGC	Gly	4
GTA	Val	6	GCA	Ala	8	GAA	Glu	25	GGA	Gly	3
GTG	Val	3	GCG	Ala	3	GAG	Glu	5	GGG	Glv	o

Figure 9B
Codon Usage in Rluc-final

TTT	Phe	4	TCT	Ser	0	TAT	Tyr	2	TGT	Cys	1
TTC	Phe	12	TCC	Ser	10	TAC	Tyr	11	TGC	Cys	2
TTA	Leu	0	TCA	Ser	1	TAA	***	0	TGA	***	0
TTG	Leu	0	TCG	Ser	0	TAG	***	0	TGG	Trp	8
CTT	Leu	3	CCT	Pro	11	CAT	His	2	CGT	Arg	0
CTC	Leu	6	CCC	Pro	3	CAC	His	8	CGC	Arq	7
CTA	Leu	0	CCA	Pro	4	CAA	Gln	3	CGA	Arg	0
CTG	Leu	13	CCG	Pro	0	CAG	Gln	4	CGG	Arg	3
ATT	Ile	3	ACT	Thr	1	AAT	Asn	2	AGT	Ser	1
ATC	Ile	18	ACC	Thr	4	AAC	Asn	11 .	AGC	Ser	7
ATA	Ile	0	ACA	Thr	0	AAA	Lys	4	AGA	Arg	2
ATG	Met	9	ACG	Thr	0	AAG	Lys	23	AGG	Arg	1
GTT	Val	2	GCT	Ala	11	GAT	Asp	6	GGT	Gly	3
GTC	Val	8	GCC	Ala	9	GAC	Asp	11	GGC	Gly	7
GTA	Val	0	GCA	Ala	0	GAA	Glu	2	GGA	Gly	3
GTG	Val	13	GCG	Ala	0	GAG	Glu	28	GGG	Glv	4

Figure 10

Oligonucleotides for the assembly of synthetic Renilla luciferase gene

Sense Strand		
Oligo name	Oligo sequence from 5' to 3'	
RLS1 (1-40)		(050 m) (0 * (0
	AACCATGGCTTCCAAGGTGTACGACCCCGAGCAACGCAAA	(SEQ ID NO:246)
RLS2 (41-80)	CGCATGATCACTGGGCCTCAGTGGTGGGCTCGCTGCAAGC	(SEQ ID NO:247)
RLS3 (81-120)	AAATGAACGTGCTGGACTCCTTCATCAACTACTATGATTC	(SEQ ID NO:248)
RL\$4 (121-170)	CGAGAAGCACGCCGAGAACGCCGTGATTTTTCTGCATGGTAACGCT	
DI 66 (171 210)	COLOCTA COTOTO A COLOCTO DO COTO DO CARROL CARROL CAR	(SEQ ID NO:249)
RLS5 (171-210)	CCAGCTACCTGTGGAGGCACGTCGTGCCTCACATCGAGCC	(SEQ ID NO:250)
RLS6 (211-250)	CGTGGCTAGATGCATCATCCCTGATCTGATCGGAATGGGT	(SEQ ID NO:251)
RLS7 (251-290)	AAGTCCGGCAAGAGCGGGAATGGCTCATATCGCCTCCTGG	(SEQ ID NO:252)
RLS8 (291-330)	ATCACTACAAGTACCTCACCGCTTGGTTCGAGCTGCTGAA	(SEQ ID NO:253)
RLS9 (331-370)	CCTTCCAAAGAAATCATCTTTGTGGGCCACGACTGGGGG	(SEQ ID NO:254)
RLS10 (371-410)	GCTTGTCTGGCCTTTCACTACTCCTACGAGCACCAAGACA	(SEQ ID NO:255)
RLS11 (411-450)	AGATCAAGGCCATCGTCCATGCTGAGAGTGTCGTGGACGT	(SEQ ID NO:256)
RLS12 (451-495)	GATCGAGTCCTGGGACGAGTGGCCTGACATCGAGGAGGATATCGC	(SEQ ID NO:257)
RLS13 (496-535)	CCTGATCAAGAGCGAAGAGGGCGAGAAAATGGTGCTTGAG	(SEQ ID NO:258)
RLS14 (536-575)	AATAACTTCTTCGTCGAGACCATGCTCCCAAGCAAGATCA	(SEQ ID NO:259)
RLS15 (576-620)	TGCGGAAACTGGAGCCTGAGGAGTTCGCTGCCTACCTGGAGCCAT	(SEQ ID NO:260)
RLS16 (621-660)	TCAAGGAGAAGGGCGAGGTTAGACGGCCTACCCTCTCCTG	(SEQ ID NO:261)
RLS17 (661-700)	GCCTCGCGAGATCCCTCTCGTTAAGGGAGGCAAGCCCGAC	(SEQ ID NO:262)
RLS18 (701-740)	GTCGTCCAGATTGTCCGCAACTACAACGCCTACCTTCGGG	(SEQ ID NO:263)
RLS19 (741-780) RLS20 (781-820)	CCAGCGACGATCTGCCTAAGATGTTCATCGAGTCCGACCC	(SEQ ID NO:264)
RLS21 (821-860)	TGGGTTCTTTTCCAACGCTATTGTCGAGGGAGCTAAGAAG TTCCCTAACACCGAGTTCGTGAAGGTGAAGGGCCTCCACT	(SEQ ID NO:265)
RLS22 (861-900)		(SEQ ID NO:266)
RLS23 (901-949)	TCAGCCAGGAGGACGCTCCAGATGAAATGGGTAAGTACAT CAAGAGCTTCGTGGAGCGCGTGCTGAAGAACGAGCAGTAATTCTAG	(SEQ ID NO:267)
KL323 (901-949)	CANDAGCTICGTOGAGCGCGTGCTGAAGAACGAGCAGTAATTCTAG	(SEQ ID NO:268)
Anti-sansa Strand		(3EQ ID 110.200)
Anti-sense Strand	Olice Common from 57 to 27	(SEQ ID NO.200)
Oligo name	Oligo Sequence from 5' to 3'	, - ,
Oligo name RLASI (1-29)	GCTCTAGAATTACTGCTCGTTCTTCAGCA	(SEQ ID NO:269)
Oligo name RLAS1 (1-29) RLAS2 (30-69)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC	(SEQ ID NO:269) (SEQ ID NO:270)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTC	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGGTGTTAGGGAACTTCTTAGCTCCCTCGACAA	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:272)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGGTGTTAGGGAACTTCTTAGCTCCCTCGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:272) (SEQ ID NO:273)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGGTGTTAGGGAACTTCTTAGCTCCCTCGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGGCCCGAAGGTAGGCGTTGTAG	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:272) (SEQ ID NO:273) (SEQ ID NO:274)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS7 (230-269)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGGTGTTAGGGAACTTCTTAGCTCCTCGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGGCCCGAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGCCTTGAC	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:272) (SEQ ID NO:273) (SEQ ID NO:274) (SEQ ID NO:275)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS7 (230-269) RLAS8 (270-309)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGGTGTTAGGGAACTTCTTAGCTCCTCGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGGCCCGAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGGCTTGCCTCCCTTAA CGAGAGGGATCTCGCGAGGCCAGGAGAGGGTAGGCCGTCT	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:272) (SEQ ID NO:273) (SEQ ID NO:274) (SEQ ID NO:275) (SEQ ID NO:276)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS7 (230-269) RLAS8 (270-309) RLAS9 (310-349)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGGTGTTAGGGAACTTCTTAGCTCCCTCGACAA TAGCGTTGGAAAAGAACCACGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGGCCCGAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGGCTTGCCTCCCTTAA CGAGAGGGATCTCCCGAGGCCAGGAGAGGGTAGGCCGTCT AACCTCGCCCTTTCCCTTGAATGGCTCCAGGTAGGCAGCG	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:272) (SEQ ID NO:273) (SEQ ID NO:274) (SEQ ID NO:274) (SEQ ID NO:275) (SEQ ID NO:276) (SEQ ID NO:277)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS6 (190-229) RLAS7 (230-269) RLAS8 (270-309) RLAS9 (310-349) RLAS10 (350-394)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGGTGTTAGGGAACTTCTTAGCTCCCTCGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGCCCGAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGGCTTGCCTCCCTTAA CGAGAGGGATCTCGCGGAGGCCAGGAGAGGGTAGGCCGTCT AACCTCGCCCTTCTCCTTGAATGGCTCCAGGTAGGCAGCA	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:272) (SEQ ID NO:273) (SEQ ID NO:274) (SEQ ID NO:275) (SEQ ID NO:275) (SEQ ID NO:277) (SEQ ID NO:277) (SEQ ID NO:277)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS7 (230-269) RLAS8 (270-309) RLAS9 (310-349) RLAS10 (350-394) RLAS11 (395-434)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGGTGTTAGGGAACTTCTTAGCTCCCTCGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCCGCAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGCGTTGCCTCCCTTAA CGAGAGGGATCTCGCGGAGGCCAGGAGAGGGTAGGCCGTCT AACCTCGCCCTTCTCCTTGAATGGCTCCAGGTAGGCAGCG AACTCCTCAGGCTCCAGTTTCCCGCATGATCTTGCTTCGGACGACGATG	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:272) (SEQ ID NO:273) (SEQ ID NO:273) (SEQ ID NO:275) (SEQ ID NO:276) (SEQ ID NO:277) (SEQ ID NO:278) (SEQ ID NO:278) (SEQ ID NO:279)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS7 (230-269) RLAS8 (270-309) RLAS8 (270-309) RLAS9 (310-349) RLAS10 (350-394) RLAS11 (395-434) RLAS12 (435-474)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGGTGTTAGGGAACTTCTTAGCTCCCTCGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGGCCGAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGGCTTGCCTCCCTTAA CGAGAGGGATCTCGCGAGGCCAGGAGGGGTAGGCCGTCT AACCTCGCCCTTCTCCTTGAATGGCTCCAGGTAGGCAGCG AACTCCTCAGGCTCCAGTTTCCGCATGATCTTGCTTCGGAGCATG GTCTCGACGAAGAAGTTATTCTCAAGCACCATTTTCTCGC CCTCTTCGCTCTTGATCAGGGCGATATCCTCCTCGATGTC	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:272) (SEQ ID NO:273) (SEQ ID NO:274) (SEQ ID NO:275) (SEQ ID NO:276) (SEQ ID NO:277) (SEQ ID NO:278) (SEQ ID NO:279) (SEQ ID NO:279) (SEQ ID NO:280)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS7 (230-269) RLAS8 (270-309) RLAS8 (270-309) RLAS10 (350-394) RLAS11 (395-434) RLAS12 (435-474) RLAS13 (475-517)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGGTGTAAGGGAACTTCTTAGCTCCCTCGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGGCCCGAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGGCTTGCCTCCCTTAA CGAGAGGGATCTCGCGAGGCCAGGAGAGGTAGGCCGTCT AACCTCGCCCTTTCCTTGAATGGCTCCAGGTAGGCAGCG AACTCCTCAGGCTCCAGTTTCCGCATGATCTTGCTTGGAGGAGCATG GTCTCGACGAAGAAGTTATTCTCAAGCACCATTTTCTCGC CCTCTTCGCTCTTGATCAGGGCGATATCCTCCTCAATGTC AGGCCACTCGTCCCAGGACTCGATCTCCACGACACTCTTCA	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:272) (SEQ ID NO:273) (SEQ ID NO:274) (SEQ ID NO:275) (SEQ ID NO:276) (SEQ ID NO:277) (SEQ ID NO:278) (SEQ ID NO:279) (SEQ ID NO:280) (SEQ ID NO:281)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS7 (230-269) RLAS8 (270-309) RLAS9 (310-349) RLAS10 (350-394) RLAS11 (395-434) RLAS12 (435-474) RLAS13 (475-517) RLAS14 (518-559)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGTGTTAGGGAACTTCTTAGCTCCCTCGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGGCCCGAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGCCTTGCCTCCCTTAA CGAGAGGGATCTCGCGAGGCCAGGAGAGGGTAGGCCGTCT AACCTCGCCCTTCTCCTTGAATGGCTCCAGGTAGGCAGCG AACTCCTCAGGCTCCAGTTTCCGCATGATCTTGCTTGGGAGCATG GTCTCGACGAAGAAGTTATTCTCAAGCACCATTTTCTCGC CCTCTTCGCTCTTGATCAGGCGGATATCCTCCTCGATGTC AGGCCACTCGTCCCAGGACTCGATCACGACACTCTCA GCATGGACGACGATGATCTTTGTCTTOGTGCTCGAGGAG	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:271) (SEQ ID NO:273) (SEQ ID NO:273) (SEQ ID NO:274) (SEQ ID NO:275) (SEQ ID NO:276) (SEQ ID NO:276) (SEQ ID NO:277) (SEQ ID NO:279) (SEQ ID NO:280) (SEQ ID NO:281) (SEQ ID NO:282)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS7 (230-269) RLAS8 (270-309) RLAS8 (270-309) RLAS10 (350-394) RLAS11 (395-434) RLAS12 (435-474) RLAS13 (475-517) RLAS13 (475-517) RLAS14 (518-559) RLAS15 (560-599)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGTGTTAGGGAACTTCTTAGCTCCCTCGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGGCCCGAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGGCTTGCCTCCCTTAA CGAGAGGGATCTCGCGAGGCCAGGAGAGGGTAGGCCGTCT AACCTCGCCCTTCTCCTTGAATGGCTCCAGGTAGGCAGCA AACTCCTCAGGCTCCAGTTTCCGCATGATCTTGCTTCGGCAGCATG GTCTCGACGAAGAAGTTATTCTCAAGCACCATTTTCTCGC CCTCTTCGCTCTTGATCAGGGCGATATCCTCCTCGATGTC AGGCCACTCCTCCCAGGACTCGATCACGTCCACGACACTCTCA GCATGGACGATGGCCTTGATCTTGTCTTG	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:271) (SEQ ID NO:272) (SEQ ID NO:273) (SEQ ID NO:274) (SEQ ID NO:275) (SEQ ID NO:276) (SEQ ID NO:276) (SEQ ID NO:278) (SEQ ID NO:278) (SEQ ID NO:278) (SEQ ID NO:280) (SEQ ID NO:281) (SEQ ID NO:282) (SEQ ID NO:283)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS7 (230-269) RLAS7 (230-269) RLAS8 (270-309) RLAS9 (310-349) RLAS10 (350-394) RLAS11 (395-434) RLAS12 (435-474) RLAS13 (475-517) RLAS14 (518-559) RLAS15 (560-599) RLAS16 (600-639)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGTGTTAGGGAACTTCTTAGCTCCTCGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGGCCCGAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGGCTTGCCTCCCTTAA CGAGAGGGATCTCGCGAGGCCAGGAGAGGGTAGGCCGTCT AACCTCGCCCTTCTCCTTGAATGGCTCCAGGTAGGCAGCA AACTCCTCAGGCTCCAGTTTCCGCATGATCTTGCTTCGGCACCATG GTCTCGACGAAGAAGTTATTCTCAAGCACCATTTTCTCGC CCTCTTCGCTCTTGATCAGGGCGATATCCTCCTCGATGTC AGGCCACTCGTCCAGGACTCGATCACGTCCACGACACTCTCA GCATGGACGATGGCCTTGATCTTGTTTTGTT	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:271) (SEQ ID NO:272) (SEQ ID NO:273) (SEQ ID NO:274) (SEQ ID NO:275) (SEQ ID NO:276) (SEQ ID NO:276) (SEQ ID NO:277) (SEQ ID NO:278) (SEQ ID NO:279) (SEQ ID NO:280) (SEQ ID NO:281) (SEQ ID NO:282) (SEQ ID NO:283) (SEQ ID NO:284)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS7 (230-269) RLAS8 (270-309) RLAS9 (310-349) RLAS10 (350-394) RLAS11 (395-434) RLAS12 (435-474) RLAS13 (475-517) RLAS14 (518-559) RLAS15 (560-599) RLAS15 (560-639) RLAS16 (600-639) RLAS17 (640-679)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGGTGTTAGGGAACTTCTTAGCTCCCTCGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGGCCCGAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGGCTTGCCTCCCTTAA CGAGAGGGATCTCGCGAGGCCAGGAGAGGGGTAGGCCGTCT AACCTCGCCCTTCTCCTTGAATGGCTCCAGGTAGGCAGCG AACTCCTCAGGCTCCAGTTTCCGCATGATCTTGCTTGGGAGCATG GTCTCGACGAAGAAGTTATTCTCAAGCACCATTTTCTCGC CCTCTTCGCTCTTGATCAGGGCGATATCCTCCTCGATGTC AGGCCACTCGTCCCAGGACTCGATCACGTCCACGACACTCTCA GCATGGACGATGGCCTTGATCTTGTTTTTGTTCTTGTGCTCGTAGGAG TAGTGAAAGGCCAGACAAGCCCCCCAGTCGTCGCCCACAA AGATGATTTTCTTTGGAAGGTTCAGCAGCTCGAACCAAC GCTGAGGTACTTTTTTTTTT	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:271) (SEQ ID NO:273) (SEQ ID NO:273) (SEQ ID NO:275) (SEQ ID NO:275) (SEQ ID NO:276) (SEQ ID NO:276) (SEQ ID NO:277) (SEQ ID NO:279) (SEQ ID NO:279) (SEQ ID NO:280) (SEQ ID NO:281) (SEQ ID NO:282) (SEQ ID NO:283) (SEQ ID NO:284) (SEQ ID NO:285
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS7 (230-269) RLAS8 (270-309) RLAS9 (310-349) RLAS10 (350-394) RLAS11 (395-434) RLAS12 (435-474) RLAS13 (475-517) RLAS14 (518-559) RLAS15 (560-599) RLAS15 (6600-639) RLAS17 (640-679) RLAS18 (680-719)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGGTGTTAGGGAACTTCTTAGCTCCCTCGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGCCCGAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGGCTTGCCTCCCTTAA CGAGAGGGATCTCGCTGAGGCCAGGAGAGGGTAGGCCGTCT AACCTCGCCCTTCTCCTTGAATGGCTCCAGGTAGGCAGCG AACTCCTCAGGCTCCAGTTTCCGCATGATCTTGCTTGGGAGGATG GTCTCGACGAAGAAGTTATTCTCAAGCACCATTTTCTCGC CCTCTTCGCTCTTGATCAGGGCGATATCCTCCTCGATGTC AGGCCACTCGTCCCAGGACTCGATCACTCCACGACACTCTCA GCATGGACGATGGCCTTGATCTTGTTTTGTT	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:271) (SEQ ID NO:273) (SEQ ID NO:273) (SEQ ID NO:275) (SEQ ID NO:275) (SEQ ID NO:275) (SEQ ID NO:276) (SEQ ID NO:277) (SEQ ID NO:279) (SEQ ID NO:279) (SEQ ID NO:280) (SEQ ID NO:281) (SEQ ID NO:283) (SEQ ID NO:284) (SEQ ID NO:285) (SEQ ID NO:286)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS7 (230-269) RLAS8 (270-309) RLAS8 (270-309) RLAS9 (310-349) RLAS10 (350-394) RLAS11 (395-434) RLAS12 (435-474) RLAS13 (475-517) RLAS14 (518-559) RLAS15 (560-599) RLAS16 (600-639) RLAS17 (640-679) RLAS18 (680-719) RLAS18 (680-719) RLAS19 (720-764)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGGTGTAAGGGAACTTCTTAGCTCCCTCGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGGCCCGAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGCCTGACCCCTTAA CGAGAGGGATCTCGCGAGGCCAGGAGAGGGTAGGCCGTCT AACCTCGCCCTTCTCCTTGAATGGCTCCAGGTAGGCAGCG AACTCCTCAGGCTCCAGTTTCCGCATGATCTTGCTTGGAGGACATG GTCTCGACGAAGAAGTTATTCTCAAGCACCATTTTCTCGC CCTCTTCGCTCTTGATCAGGGCGATATCCTCCTCGATGTC AGGCCACTCGTCCCAGGACTCGATCACGTCACG	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:271) (SEQ ID NO:273) (SEQ ID NO:273) (SEQ ID NO:274) (SEQ ID NO:275) (SEQ ID NO:276) (SEQ ID NO:276) (SEQ ID NO:277) (SEQ ID NO:279) (SEQ ID NO:281) (SEQ ID NO:281) (SEQ ID NO:282) (SEQ ID NO:283) (SEQ ID NO:284) (SEQ ID NO:284) (SEQ ID NO:285) (SEQ ID NO:285) (SEQ ID NO:285) (SEQ ID NO:286) (SEQ ID NO:287)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS7 (230-269) RLAS8 (270-309) RLAS9 (310-349) RLAS10 (350-394) RLAS11 (395-434) RLAS12 (435-474) RLAS13 (475-517) RLAS14 (518-559) RLAS15 (560-599) RLAS16 (600-639) RLAS17 (640-679) RLAS18 (680-719) RLAS19 (720-764) RLAS20 (765-804)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCGGTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGTGTTAGGGAACTTCTTAGCTCCCTCGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGGCCCGAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGCCTGACCTCCCTTAA CGAGAGGGATCTCGCGAGGCCAGGAGAGGGTAGGCCGTCT AACCTCGCCCTTCTCCTTGAATGGCTCCAGGTAGGCAGCG AACTCCTCAGGCTCCAGTTTCCGCATGATCTTGCTTGGGAGCATG GTCTCGACGAAGAAGTTATTCTCAAGCACCATTTTCTCGC CCTCTTCGCTCTTGATCAGGCGGATATCCTCCAGACACTCTCA GCATGGACGATGACCTTGATCTTGTCTTG	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:271) (SEQ ID NO:273) (SEQ ID NO:273) (SEQ ID NO:275) (SEQ ID NO:275) (SEQ ID NO:276) (SEQ ID NO:276) (SEQ ID NO:277) (SEQ ID NO:277) (SEQ ID NO:279) (SEQ ID NO:280) (SEQ ID NO:281) (SEQ ID NO:281) (SEQ ID NO:283) (SEQ ID NO:284) (SEQ ID NO:284) (SEQ ID NO:286) (SEQ ID NO:286) (SEQ ID NO:286) (SEQ ID NO:287) (SEQ ID NO:288)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS7 (230-269) RLAS8 (270-309) RLAS9 (310-349) RLAS10 (350-394) RLAS11 (395-434) RLAS12 (435-474) RLAS13 (475-517) RLAS14 (518-559) RLAS15 (560-599) RLAS16 (600-639) RLAS17 (640-679) RLAS19 (720-764) RLAS20 (765-804) RLAS21 (805-849)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGTGTTAGGGAACTTCTTAGCTCCCTGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGGCCCGAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGCCTTGACTCCCTTAA CGAGAGGGATCTCGCGAGGCCAGGAGAGGGTAGGCCGTCT AACCTCGCCCTTCTCCTTGAATGGCTCCAGGTAGGCAGCA AACTCCTCAGGCTCCAGTTTCCGCATGATCTTGCTTGGGAGCATG GTCTCGACGAAGAAAGTTATTCTCAAGCACCATTTTCTCGC CCTCTTCGCTCTTGATCAGGGCGATATCCTCCTCGATGTC AGGCCACTCCTCCCAGGACTCGATCACCTCCACGACACTCTCA GCATGGACGATGGCCTTGATCTTGTCTTG	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:271) (SEQ ID NO:272) (SEQ ID NO:273) (SEQ ID NO:273) (SEQ ID NO:274) (SEQ ID NO:275) (SEQ ID NO:276) (SEQ ID NO:276) (SEQ ID NO:278) (SEQ ID NO:278) (SEQ ID NO:280) (SEQ ID NO:280) (SEQ ID NO:281) (SEQ ID NO:281) (SEQ ID NO:283) (SEQ ID NO:284) (SEQ ID NO:284) (SEQ ID NO:285) (SEQ ID NO:286) (SEQ ID NO:286) (SEQ ID NO:287) (SEQ ID NO:288) (SEQ ID NO:288)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS7 (230-269) RLAS7 (230-269) RLAS9 (310-349) RLAS10 (350-394) RLAS11 (395-434) RLAS12 (435-474) RLAS12 (435-474) RLAS13 (475-517) RLAS14 (518-559) RLAS15 (560-599) RLAS16 (600-639) RLAS17 (640-679) RLAS19 (720-764) RLAS19 (720-764) RLAS20 (765-804) RLAS21 (805-849) RLAS21 (805-849) RLAS22 (850-889)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGTGTTAGGGAACTTCTTAGCTCCTCGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGGCCCGAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGGCTTGCCTCCCTTAA CGAGAGGGATCTCGCGAGGCCAGGAGAGGGTAGGCCGTCT AACCTCGCCCTTCTCCTTGAATGGCTCCAGGTAGGCAGCATG GTCTCGACGAAGAAGTTATTCTCAAGCACCATTTTCTCGC CCTCTTCGCTCTTGATCAGGGCGATATCCTCCACGACACTCTCA GCATGGACGATGGCCTTGATCTTGTCTTCGTTCTCGAGGAG TAGTGAAAGGCCAGACAAGCCCCCCAGTCGTAGCCACAA AGATGATTTTCTTTTGGAAGGTTCAGCAGCTCGAACCAAC GCTGAGGGTACTTGTATCTAGCAGAGCTCGAACCAAC GGTGAGGTACTTTTTGTGTATCAGCAGCTCGAACCAAC GGTGAGGTACTTTTTTTTTT	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:271) (SEQ ID NO:272) (SEQ ID NO:273) (SEQ ID NO:274) (SEQ ID NO:275) (SEQ ID NO:276) (SEQ ID NO:276) (SEQ ID NO:277) (SEQ ID NO:278) (SEQ ID NO:278) (SEQ ID NO:281) (SEQ ID NO:281) (SEQ ID NO:282) (SEQ ID NO:282) (SEQ ID NO:283) (SEQ ID NO:284) (SEQ ID NO:284) (SEQ ID NO:285) (SEQ ID NO:286) (SEQ ID NO:287) (SEQ ID NO:288) (SEQ ID NO:289) (SEQ ID NO:289) (SEQ ID NO:290)
Oligo name RLAS1 (1-29) RLAS2 (30-69) RLAS3 (70-109) RLAS4 (110-149) RLAS5 (150-189) RLAS6 (190-229) RLAS7 (230-269) RLAS8 (270-309) RLAS9 (310-349) RLAS10 (350-394) RLAS11 (395-434) RLAS12 (435-474) RLAS13 (475-517) RLAS14 (518-559) RLAS15 (560-599) RLAS16 (600-639) RLAS17 (640-679) RLAS19 (720-764) RLAS20 (765-804) RLAS21 (805-849)	GCTCTAGAATTACTGCTCGTTCTTCAGCA CGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATC TGGAGCGTCCTCGGCTGAAGTGGAGGCCCTTCACCTTC ACGAACTCGTGTTAGGGAACTTCTTAGCTCCCTGACAA TAGCGTTGGAAAAGAACCCAGGGTCGGACTCGATGAACAT CTTAGGCAGATCGTCGCTGGCCCGAAGGTAGGCGTTGTAG TTGCGGACAATCTGGACGACGTCGGCCTTGACTCCCTTAA CGAGAGGGATCTCGCGAGGCCAGGAGAGGGTAGGCCGTCT AACCTCGCCCTTCTCCTTGAATGGCTCCAGGTAGGCAGCA AACTCCTCAGGCTCCAGTTTCCGCATGATCTTGCTTGGGAGCATG GTCTCGACGAAGAAAGTTATTCTCAAGCACCATTTTCTCGC CCTCTTCGCTCTTGATCAGGGCGATATCCTCCTCGATGTC AGGCCACTCCTCCCAGGACTCGATCACCTCCACGACACTCTCA GCATGGACGATGGCCTTGATCTTGTCTTG	(SEQ ID NO:269) (SEQ ID NO:270) (SEQ ID NO:271) (SEQ ID NO:271) (SEQ ID NO:272) (SEQ ID NO:273) (SEQ ID NO:273) (SEQ ID NO:274) (SEQ ID NO:275) (SEQ ID NO:276) (SEQ ID NO:276) (SEQ ID NO:278) (SEQ ID NO:278) (SEQ ID NO:280) (SEQ ID NO:280) (SEQ ID NO:281) (SEQ ID NO:281) (SEQ ID NO:283) (SEQ ID NO:284) (SEQ ID NO:284) (SEQ ID NO:285) (SEQ ID NO:286) (SEQ ID NO:286) (SEQ ID NO:287) (SEQ ID NO:288) (SEQ ID NO:288)

Figure 11

GRVER51.SEQ ATGATGAAACGCGAAAAGAACGTGATCTACGGCCCAGAAC 40 LUCPPLYG.SEQATGATGAAGAGAGAAAAATGTTATATATGGACCCGAAC 40 RD1561H9.SEQATGATAAAGCGTGAGAAAAATGTCATCTATGGCCCTGAGC 40	
GRVER51.SEQ CACTGCATCCACTGGAAGACCTCACCGCTGGTGAGATGCT 80 LUCPPLYG.SEQCCCTACACCCCTTGGAAGACTTAACAGCAGGAGAAATGCT 80 RD1561H9.SEQCTCTCCATCCTTTGGAGGATTTGACTGCCGGCGAAATGCT 80	
GRVER51.SEQ CTTCCGAGCACTGCGTAAACATAGTCACCTCCCTCAAGCA 120 LUCPPLYG.SEQCTTCAGGGCCCTTCGAAAACATTCTCATTTACCGCAGGCT 120 RD1561H9.SEQGTTTCGTGCTCTCCGCAAGCACTCTCATTTGCCTCAAGCC 120	
GRVER51.SEQ CTCGTGGACGTCGTGGGAGACGAGACCTCTCCTACAAAG 160 LUCPPLYG.SEQTTAGTAGATGTGTTTGGTGACGAATCGCTTTCCTATAAAG 160 RD1561H9.SEQTTGGTCGATGTGGTCGGCGATGAATCTTTGAGCTACAAGC 160	
GRVER51.SEQ AATTTTTCGAAGCTACTGTGCTGTTGGCCCAAAGCCTCCA 200 LUCPPLYG.SEQAGTTTTTTGAAGCTACATGCCTCCTAGCGCAAAGTCTCCA 200 R01561H9.SEQAGTTTTTTGAGGCAACCGTCTTGCTGGCTCAGTCCCA 200	
GRVER51. SEQ TAATTGTGGGTACAAAATGAACGATGTGGTGAGCATTTGT 240 LUCPPLYG. SEQCAATTGTGGATACAAGATGAATGTAGTGTCGATCTGC 240 RD1561H9. SEQCAATTGTGGCTACAAGATGAACGACGTCGTTAGTATCTGT 240	•
GRVER51.SEQ G CTG A G A A T A A C A C T C G C T T C T T T A T T C C T G T A A T C G C T G 280 LUCPPLYG.SEQG C C G A G A A T A A T A A A G A T T T T T T	
GRVER51.SEQ CTTGGTACATCGGCCATGATTGTCGCCCCCTGTGAATGAA	
GRVER51.SEQ TTACATCCCAGATGAGCTGTGTGAAGGTTATGGGTAT <u>TAGC</u> 360 LUCFPLYG.SEQTTACATCCCAGATGAACTCTGTAAGGTCATGGGTATATCG 360 RD1561H9.SEQCTACATTCCCGACGAACTGTGTAAAGTCATGGGTATCTCT 360	
GRVER51.SEQ AAACCTCAAATCGTCTTTACTACCAAAAACATCTTGAATA 400 LUCPPLYG.SEQAAACCACAAATAGTTTTTTGTACAAAGAACATTTTAAATA 400 RD1561H9.SEQAAGCCACAGATTGTCTTCACCACTAAGAATATTCTGAACA 400	
GRVERS1.SEQ AGGTCTTGGAAGTCCAG <mark>TCTC</mark> GTACTAACTTCATCAAACG 440 LUCPPLYG.SEQAGGTATTGGAGGTACAGAGCAGAACTAATTTCATAAAAAG 440 RD1561H9.SEQAAGTCCTGGAAGTCCAAAGCCGCACCAACTTTATTAAGCG 440	
GRVERS1. SEQ CATCATTATTCTGGATACCGTCGAAAACATCCACGGCTGT 480 LUCPPLYG. SEQGATCATCATACTTGATACTGTAGAAAACATACACGGTTGT 480 RD1561H9. SEQTATCATCATCTTGGACACTGTGGAGAATATTCACGGTTGC 480	
GRVER51.SEQ GAGAGCCTCCCTAACTTCATCTCTCGTTACAGCGATGGTA 520 LUCPPLYG.SEQGAAAGTCTTCCCAATTTTATTTCTCGTTATTCGGATGGAA 520 RD1561H9.SEQGAATCTTTGCCTAATTTCATCTCTCGCTATTCAGACGGCA 520	1
GRVER51.SEQ ATATEGETA ATTTCA AGC CETTGCATTTTGATCCAGTEGA 560 LUCPPLYG. SEQATATTGCCA ACTTCA AACCTTTA CATTACGATCCTGTTGA 560 RD1561H9. SEQAEATCGCAAACTTTA AACCAETCCAETTCGAECCTGTGGA 560	

Figure 11 (Cont.)

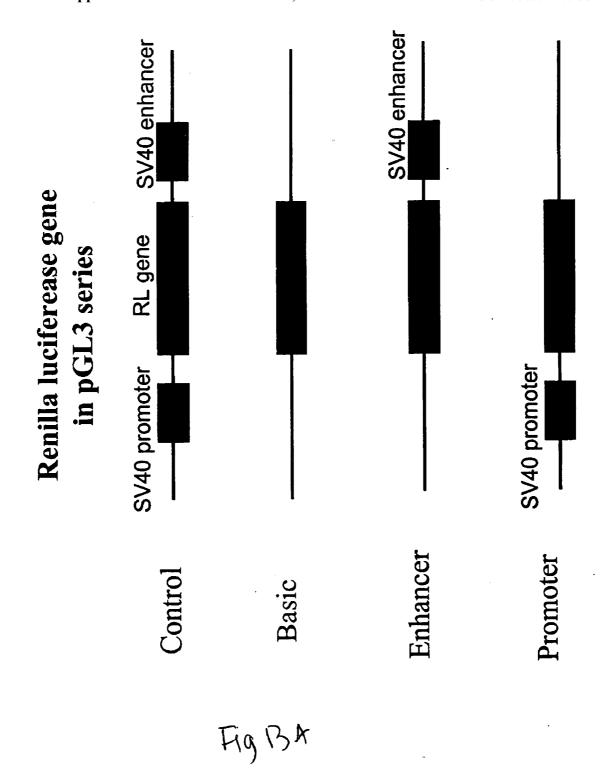
·
GRVER51.SEQ G C A A G T G G C C G C T A T T T T G T G C T C C T C C G G C A C C A C T G G T 600 LUCPPLYG.SEQG C A A G T G G C A G C T A T C T T A T G T T C G T C A G G C A C T A C T G G A 600 RD1561H9.SEQA C A A G T T G C A G C C A T T C T G T G T A G C A G C G T A C T G G A 600
GRVER51. SEQ T TGC CTAAAGGTGTCATGCAGACTCACCAGAATATCTGTG 640 LUCPPLYG. SEQTTACCGAAAGGTGTAATGCAAACTCACCAAAATATTTGTG 640 RD1561H9. SEQCTCCCAAAGGGAATCATGCAGACCCATCAAAACATTTGCG 640
GRVER51.SEQ TGC GTTTGATCC ACGCTCTCGACCCTCGTGTGGGTACTCA 680 LUCPPLYG.SEQTCCGACTTATACATGCTTTAGACCCCAGGGCAGGAACGCA 680 RD1561H9.SEQTGCGTCTGATCCATGCTCTCGATCCACGCTACGGCACTCA 680
GRVER51.SEQ ATTGATCCCTGGCGTGACTGTGCTGGTGTATCTGCCTTTC 720 LUCPPLYG.SEQACTTATTCCTGGTGTGACAGTCTTAGTATATCTGCCTTTT 720 RD1561H9.SEQGCTGATTCCTGGTGTCACCGTCTTGGTCTACTTGCCTTTC 720
GRVER51.SEQ T TTC ACG CCT T T G GTT T C T C T A TTACCCT G G GCT ATT T C A 760 LUCPPLYG.SEQT T C C A T G C T T T T G G G T T C T C T A T A A A C T T G G G A T A C T T C A 760 RD1561H9.SEQT T C C A T G C T T T C G G C T T T C A T A C T T T T
GRVER51.SEQ T G G T C G G C T T G C G T G T C A T G T T T C G T C G C T T C G A C C A 800 LUCPPLYG.SEQT G G T G G T C T T C G T G T T A T C A T G T T A A G A C G A T T T G A T C A 800 RD1561H9.SEQT G G T C G G G T C T C C G C G T G A T T A T G T T C C G C C G T T T T G A T C A 800
GRVER51.SEQ A G A A G C C T T C T T G A A G G C T A T T C A A G A C T A C G A G G T G C G T 840 LUCPPLYG.SEQA G A A G C A T T T C T A A A A G C T A T T C A G G A T T A T G A A G T T C G A 840 RD1561H9.SEQGG A G G C T T T C T T G A A G C C A T C C A A G A T T A T G A A G T C C G C 840
GRVERS1. SEQ T C C G T G A T C A A C G T C C C T T C A G T C A T T T T G T T C C T G A G C A 880 LUCPPLYG. SEQA G T G T A A T T A A C G T T C C A G C A A T A A T A T T G T T C T T A T C G A 880 RD1561H9. SEQA G T G T C A T C A A C G T G C C T A G C G T G A T C C T G T T T T T G T C T A 880
GRVER51.SEQ A ATCTCCTTTGGTTGACAAGTATGATCTGAGCAGCTTGCG 920 LUCPPLYG.SEQAAAGTCCTTTGGTTGACAAATACGATTTATCAAGTTTAAG 920 RD1561H9.SEQAGAGCCCACTCGTGGGACAAGTACGACTTGTCTTCACTGCG 920
GRVER51.SEQ TG AGCTGTGCTGTGGCGCTCCTTTGGCCAAAGAAGTG 960 LUCPPLYG.SEQGGAATTGTGTTGCGGTGCGGCACCATTAGCAAAAGAAGTT 960 RD1561H9.SEQTGAATTGTGTTGCGGTGCCGCTCCACTGGCTAAGGAGGTC 960
GRVER51.SEQ G CCG A G G T CG CTG CTA A G C G T CTG A A C CTC C CTG G T A T C C 1000 LUCPPLYG.SEQG C T G A G G T T A A A C G A T T A A A C T T G C C A G G A A T T C 1000 RD1561H9.SEQG C T G A A G T G G C C G C C A A A C G C T T G A A T C T T C C A G G G A T T C 1000
GRVER51.SEQ GCTGCGGTTTTGGTTTGACTGAGAGCACTTCTGCTAACAT 1040 LUCPPLYG.SEQGCTGTGGGATTTGGTTTGACAGATCTACTTCAGCTAATAT 1040 RD1561H9.SEQGTTGTGGCTTCGGCCTCACCGAATCTACCAGTGCGATTAT 1040
GRVER51.SEQ CCATAGCTTGCGAGACGAGTTTAAGTCTGGTAGCCTGGGT 1080 LUCPPLYG.SEQACACAGTCTTGGGGTGAATTTAAATCAGGATCACTTGGA 1080 RD1561H9.SEQCCAGACTCTCGGGGGATGAGTTTAAGAGCGGCTCTTTGGGC 1080
GRVERS1.SEQ CGCGTGACTCCTCTTATGGCTGCAAAGATCGCCGACCGTG 1120 LUCPPLYG.SEQAGAGTTACTCCTTTAATGGCAGCTAAAATAGCAGATAGGG 1120 RD1561H9.SEQCGTGTCACTCCACTCCATGGCTGCTAAGATCGCTGATCGCCG 1120

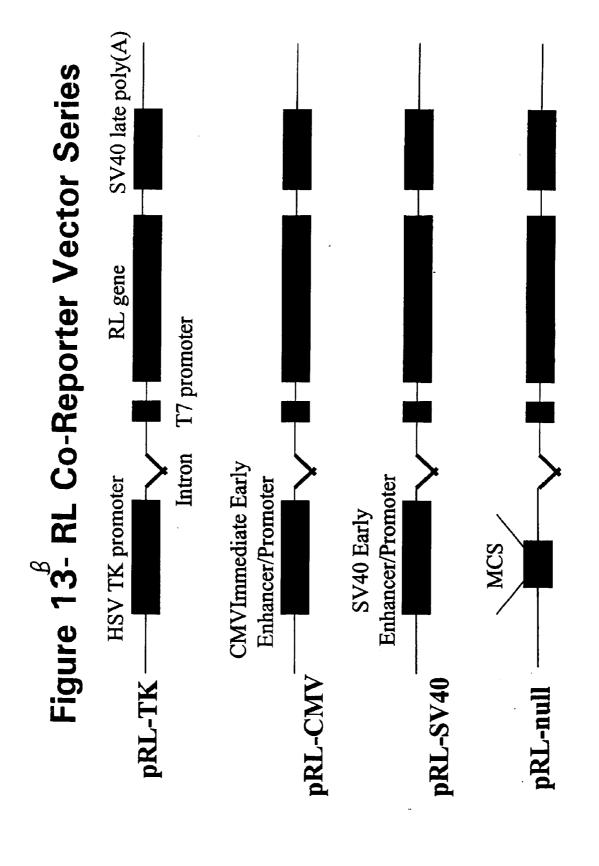
Figure 11 (Cont.)

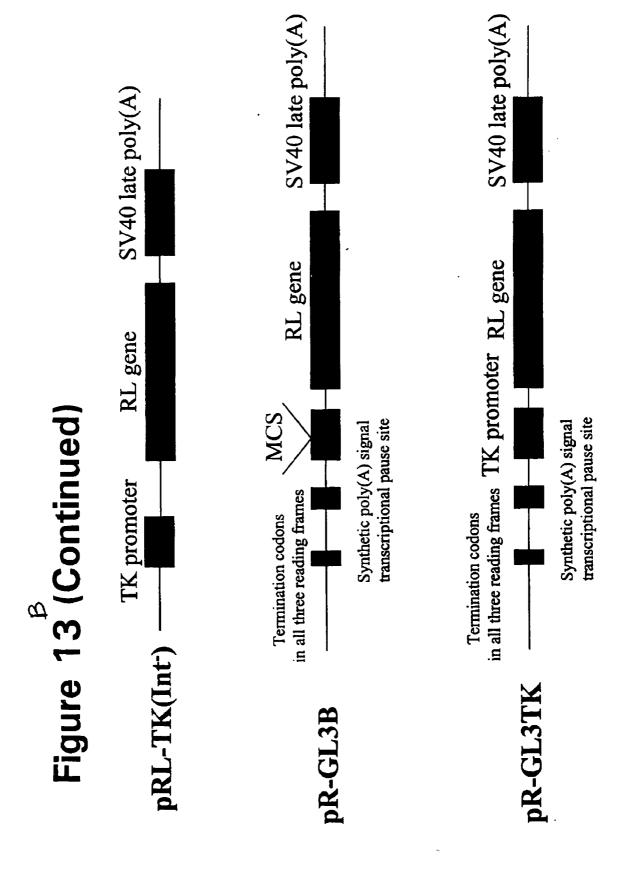
```
GRVER51. SEQ AGACCGGCAAAGCACTGGGCCCAAATCAAGTCGGTGAATT 1160
LUCPPLYG. SEQA A A C T G G T A A A G C A T T G G G A C C A A A T C A A G T T G G T G A A T T 1160
RD1561H9. SEQA A A C T G G T A A G G C T T T G G G C C C G A A C C A A G T G G G C G A G C T 1160
GRVER51. SEQ GT GTATTAAGG GCC CTATGGTCT CTAAAGGCTACGTGAAC 1200
LUCPPLYG. SEQATGCGTTAAAGGTCCCATGGTATCGAAAGGTTACGTGAAC 1200
RD1561H9. SEQGT GTATCAAAGGCCCTATGGTGAGCAAGGGTTATGTCAAT 1200
GRVER51. SEQ A A T G T G G A G G C C A C T A A A G A A G C C A T T G A T G A T G G C T 1240
LUCPPLYG. SEQA A T G T A G A A G C T A C C A A A G A A G C T A T T G A T G A T G A T G G T T 1240
RD1561H9. SEQN ACGTTGA AGCTACCA AGGAGGCCATCGACGACGACGGCT 1240
GRVER51.SEQ GGCTCCATAGCGGCGACTTCGGTTACTATGATGAGGACGA 1280
LUCPPLYG. SEQG G C T T C A C T C T G G A G A C T T T G G A T A C T A T G A T G A G G A T G A 1280
RD1561H9. SEQG G T T G C A T T C T G G T G A T T T T T T T T A C G A C G A A G A T G A 1280
GRVER51.SEQ ACACTTCTATGTGGTCGATCGCTACAAAGAATTGATTAAG 1320
LUCPPLYG. SEQG CATTTCTATGTGGTGGACCGTTACAAGGAATTGATTAAA 1320
RD1561H9.SEQG CATTTTTACGTCGTGGATCGTTACAAGGAGCTGATCAAA 1320
GRVER51. SEQ TACAAAGGCTCTCAAGTCGCACCAGCCGAACTGGAAGAAA 1360
LUCPPLYG. SEQT A T A A G G G C T C T C A G G T A G C A C C T G C A G A A C T A G A A G A G A G A 1360
RD1561H9.SEQTACAAGGGTAGCCAGGTTGCTCCAGCTGAGTTGGAGGAGA 1360
GRVER51.SEQ TTTTGCTGAAGAACCCTTGTATCCGCGACGTGGCCGTCGT 1400
LUCPPLYG. SEQT T T T A T T G A A A A T C C A T G T A T C A G A G A T G T T G C T G T G G T 1400 RD1561H9. SEQT T C T G T G A A A A A T C C A T G C A T T C G C G A T G T C G C T G T G G T 1400
GRVER51. SEQ GGGTATCC CAGACTTGGAAGCTGGCGAGTTGCCTAGCGCC 1440
LUCPPLYG. SEQT G G T A T T C C T G A T C T A G A A G C T G G A G A A C T G C C A T C T G C G 1440
RD1561H9. SEQC G G C A T T C C T G A T C T G G A G G C C G G C G A A C T G C C T T C T G C T 1440
GRVER51.SEQ TTTGTGGTGAAACAACCCGGCAAGGAGATCACTGCTAAGG 1480
LUCPPLYG. SEQT TT GT GGT TA AA CAGCCCGGAAAGGAGATTA CAGCTAAAG 1480
RD1561H9. SEQT TCGTTGTCAAGCAGCCTTGGTACAGAAATTA CCGGCCAAAG 1480
GRVER51.SEQ AGG TCT A C G ACT A TTTGG C C G A GCGCG TGT CTC ACA CCA A 1520
LUCPPLYG. SEQNAGTGTACGATTATCTTGCCGAGAGGGTCTCCCATACAAA 1520
RD1561H9. SEQNAGTGTATGATTACCTGGCTGAACGTGTGAGCCATACTAA 1520
GRVER51.SEQ AT A TOT G C G T G G C G G C G T C C G C T C G A T T C T A T T C C A 1560
LUCPPLYG. SEQG T A T T T G C G T G G A G G G G T T C G A T T C G T T G A T A G C A T A C C A 1560
RD1561H9. SEQG T ACT T G C G T G G C G G C G T G C G T T T G T T G A C T C C A T C C C T 1560
GRVER51.SEQ CGCAACGTTACCGGTAAGATCACTCGTAAAGAGTTGCTGA 1600
LUCPPLYG. SEQAGGAATGTTACAGGTAAAATTACAAGAAAGGAACTTCTGA 1600
RD1561H9. SEQCGTAACGTAACAGGCAAAATTACCCGGCAAGGAACTTCTGA 1600
                                                                                        1626
GRVER51.SEQ A G C A A C T C C T C G A A A A G C T G G C G G C
LUCPPLYG. SEQAGCAGTTGCTGGAGAAGAGTTCTAAACTT
RD1561H9. SEQAACAATTGTTGGTTGAAGGCCGGCGGT
                                                                                        1629
                                                                                        1626
```

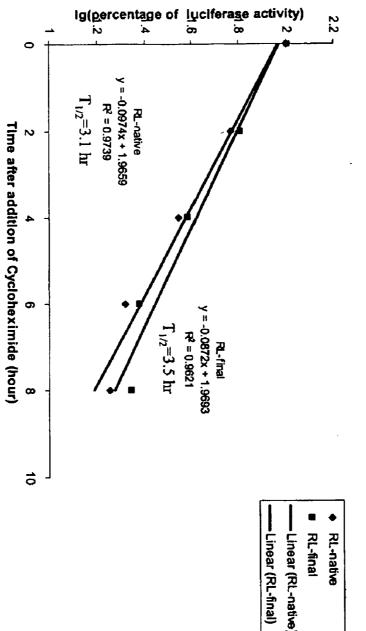
Figure 12

GRVER51.SEQ MMKREKNVIYGPEPLHPLEDLTAGEMLFRALRKHSHLPQA	118
LUCPPLYG. SEQM M K R E K N V I Y G P E P L H P L E D L T A G E M L F R A L R K H S H L P Q A	118
RD1561H9.SEQMIKREKNVIYGPEPLHPLEDLTAGEMLFRALRKHSHLPQA	110
KD1261H3.25GWILK K S K M A I I G E E P P L F P P F I M G P P F W W W 2 W P F A W	110
GRVER51.SEQ L V D V V G D E S L S Y K E F F E A T V L L A Q S L H N C G Y K M N D V V S I C	238
LUCPPLYG. SEQL V D V F G D E S L S Y K E F F E A T C L L A Q S L H N C G Y K M N D V V S I C	238
RD1561H9.SEQLVDVVGDESLSYKEFFEATVLLAQSLHNCGYKMNDVVSIC	238
THE STATE OF THE S	358
GRVER51.SEQ AENNTRFFIPVIAAWYIGMIVAPVNESYIPDELCKVMGIS	
LUCPPLYG. SEQAENNKR FFIPIIAAW YIGMIVAP VNE SYIP DELCK VM GIS	358
RD1561H9.SEQAENNTRFFIPVIAAWYIGMIVAPVNESYIPDELCKVMGIS	358
	
GRVER51.SEQ K P Q I V F T T K N I L N K V L E V Q S R T N F I K R I I I L D T V E N I H G C	478
LUCPPLYG. SEQK PQIVFCTKNILNKVLEVQSRTNFIKRIIILDTVENIHGC	478
RD1561H9.SEQKPQIVFTTKNILNKVLEVQSRTNFIKRIIILDTVENIHGC	478
_	
GRVER51.SEQ E S L P N F I S R Y S D G N I A N F K P L H F D P V E Q V A A I L C S S G T T G	598
LUCPPLYG. SEQE S L P N F I S R Y S D G N I A N F K P L H Y D P V E Q V A A I L C S S G T T G	
RD1561H9.SEQESLPNFISRYSDGNIANFKPLHFDPVEQVAAILCSSGTTG	
RUISOTAS.SEQUES LENGTES KIS DON'T MARKE LINE DE VE QUARTE COSSITIO	230
GRVER51.SEQ LPKGVMQTHQNICVRLIHALDPRVGTQLIPGVTVLVYLPF	718
LUCPPLYG. SEQL PKG VMQTHQNIC VRLIHALDPRAGTQLIPGVTVLVYLPF	718
RD1561H9.SEQLPKGVMQTHQNICVRLIHALDPRYGTQLIPGVTVLVYLPF	
GRVER51.SEQ FHAFGFS ITLGYFM V G L R V I M FRR F D Q E A F L K A I Q D Y E V R	020
LUCPPLYG.SEQFHAFGFSINLGYFMVGLRVINLRRFDQEAFLKAIQDYEVR	838
RD1561H9. SEQFH A F G F H I T L G Y F M V G L R V I M F R R F D Q E A F L K A I Q D Y E V R	838
GRVER51.SEQ S V I N V PS V I L F L S K S P L V D K Y D L S S L R E L C C G A A P L A K E V	958
LUCPPLYG. SEQS V I N V P A I I L F L S K S P L V D K Y D L S S L R E L C C G A A P L A K E V	
RD1561H9.SEQS V I N V PS V I L F L S K S P L V D K Y D L S S L R E L C C G A A P L A K E V	930
GRVER51.SEQ A E V A A K R L N L P G I R C G F G L T E S T S A N I H S L R D E F K S G S L G	1078
LUCPPLYG, SEQAE V A V K R L N L P G I R C G F G L T E S T S A N I H S L G D E F K S G S L G	1078
RD1561H9.SEQA E V AAKR L N L P G I R C G F G L T E S T S A I I Q T L G D E F K S G S L G	1078
COMPANDE AND DUMBER OF A CONTRACT OF A CONTR	1100
GRVER51.SEQ R V T P L M A A K I A D R E T G K A L G P N Q V G E L C I K G P M V S K G Y V N	
LUCPPLYG. SEQR V T P L M A A K I A D R E T G K A L G P N Q V G E L C V K G P M V S K G Y V N	
RD1561H9.SEQRVTPLMAAKIADRETGKALGPNQVGELCIKGPMVSKGYVN	1198
_	
GRVER51.SEQ NVEATKEAIDDDGWLHSGDFGYYDEDEHFYVVDRYKELIK	1318
LUCPPLYG. SEQN V E A T K E A I D D D G W L H S G D F G Y Y D E D E H F Y V V D R Y K E L I K	
RD1561H9.SEQN V E A T K E A I D D D G W L H S G D F G Y Y D E D E H F Y V V D R Y K E L I K	1310
GRVER51.SEQ YKGSQVAPAELEEILLKNPCIRDVAVVGIPDLEAGELPSA	
LUCPPLYG.SEQY K G S Q V A P A E L E E I L L K N P C I R D V A V V G I P D L E A G E L P S A	
RD1561H9.SEQYKGSQVAPAELEEILLKNPCIRDVAVVGIPDLEAGELPSA	
	1650
GRVER51.SEQ F V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P	
LUCPPLYG. SEQF V V K Q P G K E I T A K E V Y D Y L A E R V S H T K Y L R G G V R F V D S I P	
RD1561H9.SEQFVVKQPGTEITAKEVYDYLAERVSHTKYLRGGVRFVDSIP	1558
GRVER51. SEO. B.N.V.T.G.K.I.T.B.K.E.L.I.K.O.T.I.E.K.A.G.G.	1624
GRVER51.SEQ R N V T G K I T R K E L L K Q L L E K A G G	1624
LUCPPLYG. SEQR N V T G K I T R K E L L K Q L L E K S S K L	1627









Halflife of RL-synthetic and RL-native in CHO Cells

Fig 14

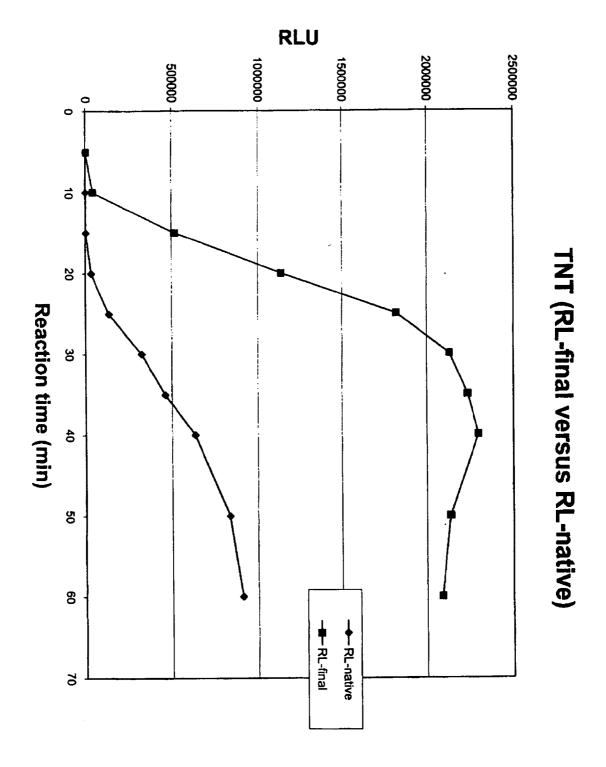


Fig 15 A

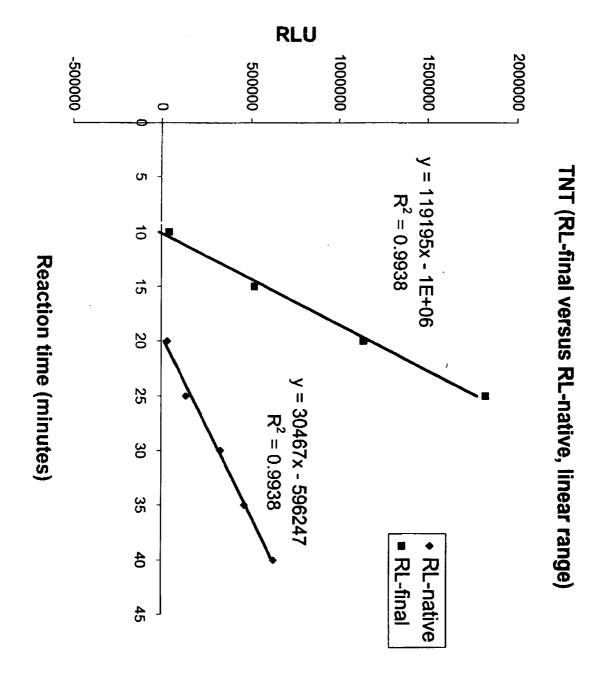


Fig15B

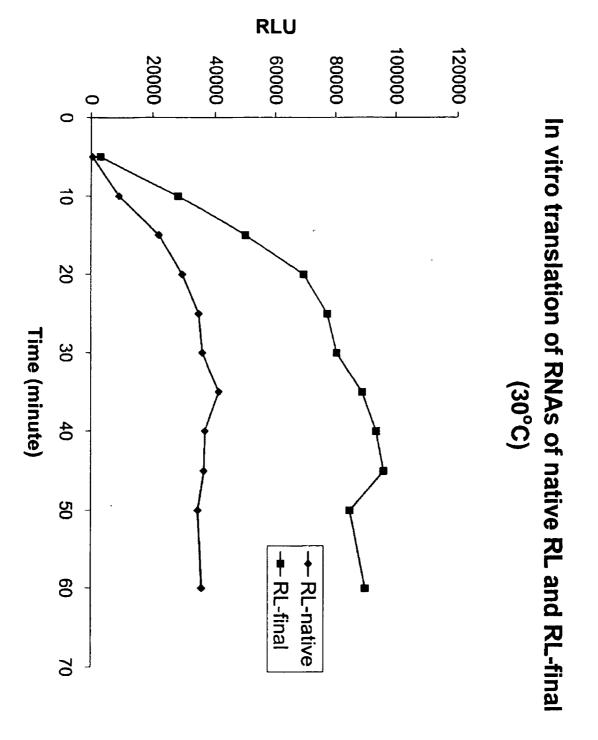
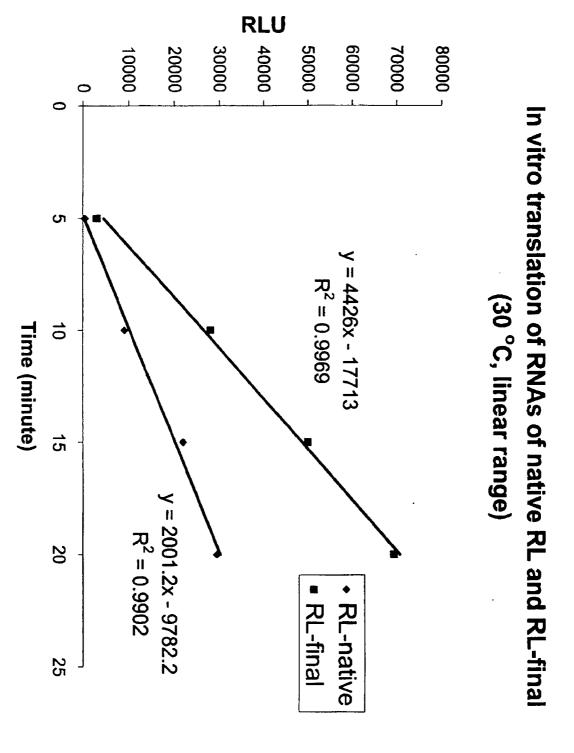


FIG15C



T19 150

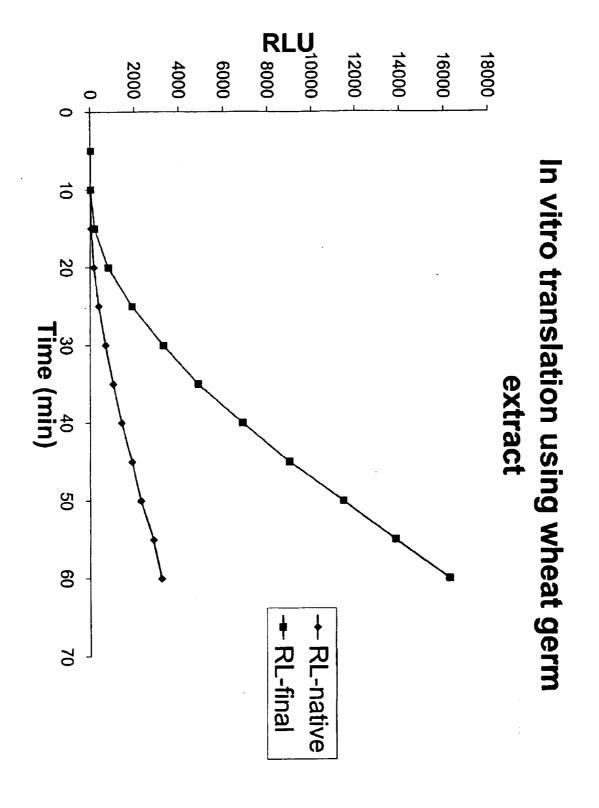
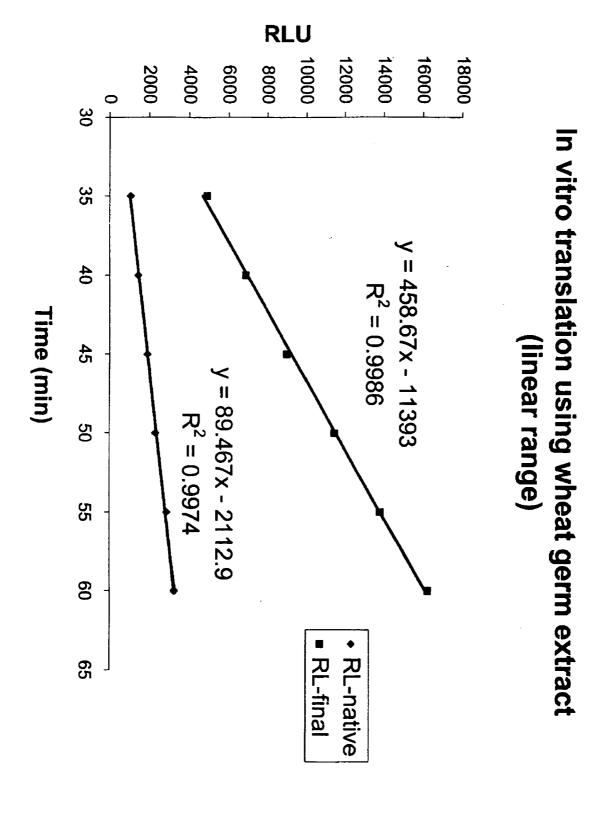
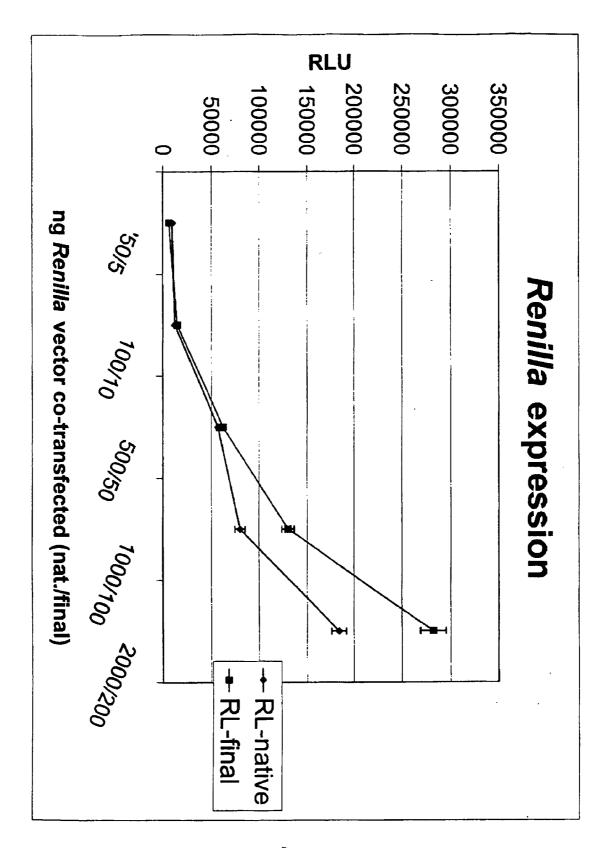


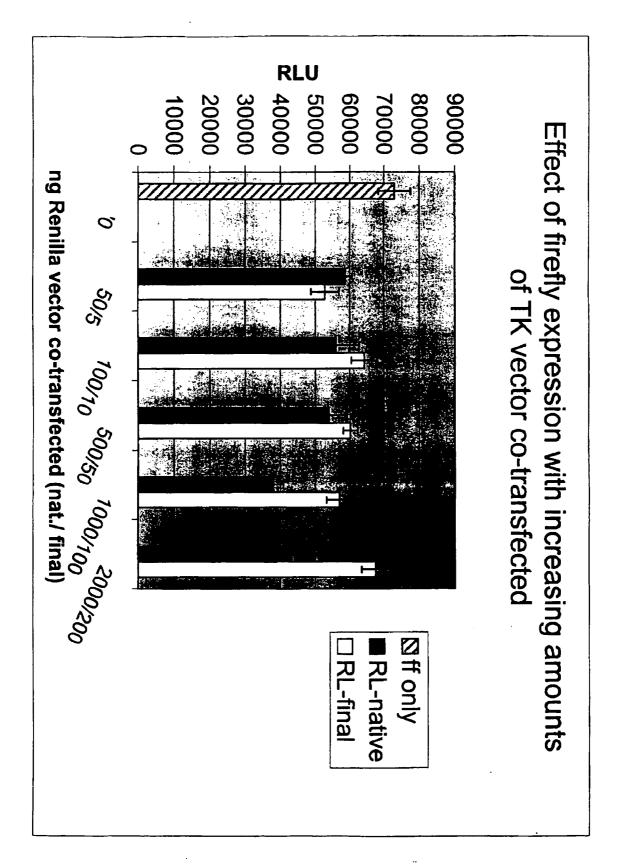
Fig 15E



F19151=



Figlist



Fialler

Figure 17 A

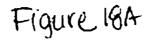
Beetle Luciferin

Figure 7 173

GRver5.1 DNA sequence of pGL3 vectors

ATGGTGAAACGCGAAAAGAACGTGATCTACGGCCCAGAACCACTGCATCC	50
${\tt ACTGGAAGACCTCACCGCTGGTGAGATGCTCTTCCGAGCACTGCGTAAAC}$	100
${\tt ATAGTCACCTCCAAGCACTCGTGGACGTCGTGGGAGACGAGAGCCTC}$	150
TCCTACAAAGAATTTTTCGAAGCTACTGTGCTGTTGGCCCAAAGCCTCCA	200
TAATTGTGGGTACAAAATGAACGATGTGGTGAGCATTTGTGCTGAGAATA	250
${\tt ACACTCGCTTCTTTATTCCTGTAATCGCTGCTTGGTACATCGGCATGATT}$	300
GTCGCCCCTGTGAATGAATCTTACATCCCAGATGAGCTGTGTAAGGTTAT	350
GGGTATTAGCAAACCTCAAATCGTCTTTACTACCAAAAACATCTTGAATA	400
${\tt AGGTCTTGGAAGTCCAGTCTCGTACTAACTTCATCAAACGCATCATTATT}$	450
CTGGATACCGTCGAAAACATCCACGGCTGTGAGAGCCTCCCTAACTTCAT	500
${\tt CTCTCGTTACAGCGATGGTAATATCGCTAATTTCAAGCCCTTGCATTTTG}$	550
${\tt ATCCAGTCGAGCAAGTGGCCGCTATTTTGTGCTCCTCCGGCACCACTGGT}$	600
${\tt TTGCCTAAAGGTGTCATGCAGACTCACCAGAATATCTGTGTGCGTTTGAT}$	650
${\tt CCACGCTCTCGACCCTCGTGTGGGTACTCAATTGATCcCTGGCGTGACTG}$	700
${\tt TGCTGGTGTATCTGCCTTTCTTTCACGCCTTTGGTTTCTCTATTACCCTG}$	750
${\tt GGCTATTTCATGGTCGGCTTGCGTGTCATCATGTTTCGTCGCTTCGACCA}$	800
${\tt AGAAGCCTTCTTGAAGGCTATTCAAGACTACGAGGTGCGTTCCGTGATCA}$	850
${\tt ACGTCCCTTCAGTCATTTTGTTCCTGAGCAAATCTCCTTTGGTTGACAAG}$	900
${\tt TATGATCTGAGCAGCTTGCGTGAGCTGTTGCTGTGCTGCTCCTTTGGC}$	950
CAAAGAAGTGGCCGAGGTCGCTGCTAAGCGTCTGAACCTCCCTGGTATCC	1000
GCTGCGGTTTTGGTTTGACTGAGAGCACTTCTGCTAACATCCATAGCTTG	1050
CGAGACGAGTTTAAGTCTGGTAGCCTGGGTCGCGTGACTCCTCTTATGGC	1100
TGCAAAGATCGCCGACCGTGAGACCGGCAAAGCACTGGGCCCAAATCAAG	1150
TCGGTGAATTGTGTATTAAGGGCCCTATGGTCTCTAAAGGCTACGTGAAC	1200
AATGTGGAGGCCACTAAAGAAGCCATTGATGATGATGGCTGGC	1250
CGGCGACTTCGGTTACTATGATGAGGACGAACACTTCTATGTGGTCGATC	1300
GCTACAAAGAATTGATTAAGTACAAAGGCTCTCAAGTCGCACCAGCCGAA	1350
CTGGAAGAATTTTGCTGAAGAACCCTTGTATCCGCGACGTGGCCGTCGT	1400
GGGTATCCCAGACTTGGAAGCTGGCGAGTTGCCTAGCGCCTTTGTGGTGA	1450
AACAACCCGGCAAGGAGATCACTGCTAAGGAGGTCTACGACTATTTGGCC	1500
GAGCGCGTGTCTCACACCAAATATCTGCGTGGCGGCGTCCGCTTCGTCGA	1550
TTCTATTCCACGCAACGTTACCGGTAAGATCACTCGTAAAGAGTTGCTGA	1600
AGCAACTCCTCGAAAAAGCTGGCGGC	1626

SEQIDNU 297



RDver5.1 DNA sequence of pGL3 vectors

ATGGTGAAGCGTGAGAAAAATGTCATCTATGGCCCTGAGCCTCTCCATCC	50
TTTGGAGGATTTGACTGCCGGCGAAATGCTGTTTCGTGCTCTCCGCAAGC	100
ACTCTCATTTGCCTCAAGCCTTGGTCGATGTGGTCGGCGATGAATCTTTG	150
AGCTACAAGGAGTTTTTTGAGGCAACCGTCTTGCTGGCTCAGTCCCTCCA	200
CAATTGTGGCTACAAGATGAACGACGTCGTTAGTATCTGTGCTGAAAACA	250
ATACCCGTTTCTTCATTCCAGTCATCGCCGCATGGTATATCGGTATGATC	300
GTGGCTCCAGTCAACGAGAGCTACATTCCCGACGAACTGTGTAAAGTCAT	350
GGGTATCTCTAAGCCACAGATTGTCTTCACCACTAAGAATATTCTGAACA	400
AAGTCCTGGAAGTCCAAAGCCGCACCAACTTTATTAAGCGTATCATC	450
TTGGACACTGTGGAGAATATTCACGGTTGCGAATCTTTGCCTAATTTCAT	500
CTCTCGCTATTCAGACGGCAACATCGCAAACTTTAAACCACTCCACTTCG	550
ACCCTGTGGAACAAGTTGCAGCCATTCTGTGTAGCAGCGGTACTACTGGA	600
CTCCCAAAGGGAGTCATGCAGACCCATCAAAACATTTGCGTGCG	650
CCATGCTCTCGATCCACGCTACGGCACTCAGCTGATTCCTGGTGTCACCG	700
TCTTGGTCTACTTGCCTTTCTTCCATGCTTTCGGCTTTCATATTACTTTG	75Ō
GGTTACTTTATGGTCGGTCTCCGCGTGATTATGTTCCGCCGTTTTGATCA	800
GGAGGCTTTCTTGAAAGCCATCCAAGATTATGAAGTCCGCAGTGTCATCA	850
ACGTGCCTAGCGTGATCCTGTTTTTGTCTAAGAGCCCACTCGTGGACAAG	900
TACGACTTGTCTTCACTGCGTGAATTGTGTTGCGGTGCCGCTCCACTGGC	950
TAAGGAGGTCGCTGAAGTGGCCGCCAAACGCTTGAATCTTCCAGGGATTC	1000
GTTGTGGCTTCGGCCTCACCGAATCTACCAGCGCTATTATTCAGTCTCTC	1050
CGCGATGAGTTTAAGAGCGGCTCTTTGGGCCGTGTCACTCCACTCATGGC	1100
TGCTAAGATCGCTGATCGCGAAACTGGTAAGGCTTTGGGCCCGAACCAAG	1150
TGGGCGAGCTGTGTATCAAAGGCCCTATGGTGAGCAAGGGTTATGTCAAT	1200
AACGTTGAAGCTACCAAGGAGGCCATCGACGACGACGGCTGGTTGCATTC	1250
TGGTGATTTTGGATATTACGACGAAGATGAGCATTTTTACGTCGTGGATC	1300
GTTACAAGGAGCTGATCAAATACAAGGGTAGCCAGGTTGCTCCAGCTGAG	1350
TTGGAGGAGATTCTGTTGAAAAATCCATGCATTCGCGATGTCGCTGTGGT	1400
CGGCATTCCTGATCTGGAGGCCGGCGAACTGCCTTCTGCTTTCGTTGTCA	1450
AGCAGCCTGGTAAAGAAATTACCGCCAAAGAAGTGTATGATTACCTGGCT	1500
GAACGTGTGAGCCATACTAAGTACTTGCGTGGCGGCGTGCGT	1550
CTCCATCCCTCGTAACGTAACAGGCAAAATTACCCGCAAGGAGCTGTTGA	1600
AACAATTGTTGGAGAAGGCCGGCGGT	1626

SEQ ID NU: 299

RD1561H9 DNA sequence of pGL3 vectors

ATGGTAAAGCGTGAGAAAAATGTCATCTATGGCCCTGAGCCTCTCCATCC 50 TTTGGAGGATTTGACTGCCGGCGAAATGCTGTTTCGTGCTCTCCGCAAGC 100 150 **ACTCTCATTTGCCTCAAGCCTTGGTCGATGTGGTCGGCGATGAATCTTTG** 200 AGCTACAAGGAGTTTTTTTGAGGCAACCGTCTTGCTGGCTCAGTCCCTCCA CAATTGTGGCTACAAGATGAACGACGTCGTTAGTATCTGTGCTGAAAACA 250 ATACCCGTTTCTTCATTCCAGTCATCGCCGCATGGTATATCGGTATGATC 300 350 GTGGCTCCAGTCAACGAGAGCTACATTCCCGACGAACTGTGTAAAGTCAT GGGTATCTCTAAGCCACAGATTGTCTTCACCACTAAGAATATTCTGAACA 400 **AAGTCCTGGAAGTCCAAAGCCGCACCAACTTTATTAAGCGTATCATCATC** 450 TTGGACACTGTGGAGAATATTCACGGTTGCGAATCTTTGCCTAATTTCAT 500 550 CTCTCGCTATTCAGACGGCAACATCGCAAACTTTAAACCACTCCACTTCG ACCCTGTGGAACAAGTTGCAGCCATTCTGTGTAGCAGCGGTACTACTGGA 600 650 CCATGCTCTCGATCCACGCTACGGCACTCAGCTGATTCCTGGTGTCACCG 700 TCTTGGTCTACTTGCCTTTCTTCCATGCTTTCGGCTTTCATATTACTTTG 750 800 GGTTACTTTATGGTCGGTCTCCGCGTGATTATGTTCCGCCGTTTTGATCA GGAGGCTTTCTTGAAAGCCATCCAAGATTATGAAGTCCGCAGTGTCATCA 850 ACGTGCCTAGCGTGATCCTGTTTTTTGTCTAAGAGCCCACTCGTGGACAAG 900 TACGACTTGTCTTCACTGCGTGAATTGTGTTGCGGTGCCGCTCCACTGGC TAAGGAGGTCGCTGAAGTGGCCGCCAAACGCTTGAATCTTCCAGGGATTC 1000 GTTGTGGCTTCGGCCTCACCGAATCTACCAGTGCGATTATCCAGACTCTC 1050 GGGGATGAGTTTAAGAGCGGCTCTTTGGGCCGTGTCACTCCACTCATGGC 1100 TGCTAAGATCGCTGATCGCGAAACTGGTAAGGCTTTGGGCCCGAACCAAG 1150 TGGGCGAGCTGTGTATCAAAGGCCCTATGGTGAGCAAGGGTTATGTCAAT 1200 AACGTTGAAGCTACCAAGGAGGCCATCGACGACGACGCTGGTTGCATTC 1250 TGGTGATTTTGGATATTACGACGAAGATGAGCATTTTTACGTCGTGGATC 1300 GTTACAAGGAGCTGATCAAATACAAGGGTAGCCAGGTTGCTCCAGCTGAG 1350 TTGGAGGAGATTCTGTTGAAAAATCCATGCATTCGCGATGTCGCTGTGGT 1400 CGGCATTCCTGATCTGGAGGCCGGCGAACTGCCTTCTGCTTTCGTTGTCA 1450 AGCAGCCTGGTACAGAAATTACCGCCAAAGAAGTGTATGATTACCTGGCT 1500 CTCCATCCCTCGTAACGTAACAGGCAAAATTACCCGCAAGGAGCTGTTGA 1600 AACAATTGTTGGTGAAGGCCGGCGGT 1626

SEQ ID NO. 30)

GRver5.1 protein sequence of pGL3 vectors

MVKREKNVIYGPEPLHPLEDLTAGEMLFRALRKHSHLPQALVDVVGDESL	50
SYKEFFEATVLLAQSLHNCGYKMNDVVSICAENNTRFFIPVIAAWYIGMI	100
VAPVNESYIPDELCKVMGISKPQIVFTTKNILNKVLEVQSRTNFIKRIII	150
LDTVENIHGCESLPNFISRYSDGNIANFKPLHFDPVEQVAAILCSSGTTG	200
LPKGVMQTHQNICVRL1HALDPRVGTQLIPGVTVLVYLPFFHAFGFSITL	250
GYFMVGLRVIMFRRFDQEAFLKAIQDYEVRSVINVPSVILFLSKSPLVDK	300
YDLSSLRELCCGAAPLAKEVAEVAAKRLNLPGIRCGFGLTESTSANIHSL	350
RDEFKSGSLGRVTPLMAAKIADRETGKALGPNQVGELCIKGPMVSKGYVN	400
NVEATKEAIDDDGWLHSGDFGYYDEDEHFYVVDRYKELIKYKGSQVAPAE	450
LEEILLKNPCIRDVAVVGIPDLEAGELPSAFVVKQPGKEITAKEVYDYLA	500
ERVSHTKYLRGGVRFVDSIPRNVTGKITRKELLKQLLEKAGG	542

SEG ID WU 298

RDver5.1 protein sequence of pGL3 vectors

MVKREKNVIYGPEPLHPLEDLTAGEMLFRALRKHSHLPQALVDVVGDESL	50
SYKEFFEATVLLAQSLHNCGYKMNDVVSICAENNTRFFIPVIAAWYIGMI	100
VAPVNESYIPDELCKVMGISKPQIVFTTKNILNKVLEVQSRTNFIKRIII	150
LDTVENIHGCESLPNFISRYSDGNIANFKPLHFDPVEQVAAILCSSGTTG	200
LPKGVMQTHQNICVRLIHALDPRYGTQLIPGVTVLVYLPFFHAFGFHITL	250
GYFMVGLRVIMFRRFDQEAFLKAIQDYEVRSVINVPSVILFLSKSPLVDK	300
YDLSSLRELCCGAAPLAKEVAEVAAKRLNLPGIRCGFGLTESTSAIIQSL	350
RDEFKSGSLGRVTPLMAAKIADRETGKALGPNQVGELCIKGPMVSKGYVN	400
NVEATKEAIDDDGWLHSGDFGYYDEDEHFYVVDRYKELIKYKGSQVAPAE	450
LEEILLKNPCIRDVAVVGIPDLEAGELPSAFVVKQPGKEITAKEVYDYLA	500
ERVSHTKYLRGGVRFVDSIPRNVTGKITRKELLKQLLEKAGG	542

SECTONO 300

RD1561H9 protein sequence of pGL3 vectors

MVKREKNVIYGPEPLHPLEDLTAGEMLFRALRKHSHLPQALVDVVGDESL	50
SYKEFFEATVLLAQSLHNCGYKMNDVVSICAENNTRFFIPVIAAWYIGMI	100
VAPVNESYIPDELCKVMGISKPQIVFTTKNILNKVLEVQSRTNFIKRIII	150
LDTVENIHGCESLPNFISRYSDGNIANFKPLHFDPVEQVAAILCSSGTTG	200
LPKGVMQTHQNICVRLIHALDPRYGTQLIPGVTVLVYLPFFHAFGFHITL	250
GYFMVGLRVIMFRRFDQEAFLKAIQDYEVRSVINVPSVILFLSKSPLVDK	300
YDLSSLRELCCGAAPLAKEVAEVAAKRLNLPGIRCGFGLTESTSAIIQTL	350
GDEFKSGSLGRVTPLMAAKIADRETGKALGPNQVGELCIKGPMVSKGYVN	400
NVEATKEAIDDDGWLHSGDFGYYDEDEHFYVVDRYKELIKYKGSQVAPAE	450
LEEILLKNPCIRDVAVVGIPDLEAGELPSAFVVKQPGTEITAKEVYDYLA	500
ERVSHTKYLRGGVRFVDSIPRNVTGKITRKELLKQLLVKAGG	542

SEA ID NU: 303

SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

STATEMENT OF GOVERNMENT RIGHTS

[0001] The invention was made at least in part with a grant from the Government of the United States of America (grant DMI-9402762 from the National Science Foundation). The Government has certain rights to the invention.

BACKGROUND OF THE INVENTION

[0002] Transcription, the synthesis of an RNA molecule from a sequence of DNA is the first step in gene expression. Sequences which regulate DNA transcription include promoter sequences, polyadenylation signals, transcription factor binding sites and enhancer elements. A promoter is a DNA sequence capable of specific initiation of transcription and consists of three general regions. The core promoter is the sequence where the RNA polymerase and its cofactors bind to the DNA. Immediately upstream of the core promoter is the proximal promoter which contains several transcription factor binding sites that are responsible for the assembly of an activation complex that in turn recruits the polymerase complex. The distal promoter, located further upstream of the proximal promoter also contains transcription factor binding sites. Transcription termination and polyadenylation, like transcription initiation, are site specific and encoded by defined sequences. Enhancers are regulatory regions, containing multiple transcription factor binding sites, that can significantly increase the level of transcription from a responsive promoter regardless of the enhancer's orientation and distance with respect to the promoter as long as the enhancer and promoter are located within the same DNA molecule. The amount of transcript produced from a gene may also be regulated by a post-transcriptional mechanism, the most important being RNA splicing that removes intervening sequences (introns) from a primary transcript between splice donor and splice acceptor sequences.

[0003] Natural selection is the hypothesis that genotypeenvironment interactions occurring at the phenotypic level lead to differential reproductive success of individuals and therefore to modification of the gene pool of a population. Some properties of nucleic acid molecules that are acted upon by natural selection include codon usage frequency, RNA secondary structure, the efficiency of intron splicing, and interactions with transcription factors or other nucleic acid binding proteins. Because of the degenerate nature of the genetic code, these properties can be optimized by natural selection without altering the corresponding amino acid sequence.

[0004] Under some conditions, it is useful to synthetically alter the natural nucleotide sequence encoding a polypeptide to better adapt the polypeptide for alternative applications. A common example is to alter the codon usage frequency of a gene when it is expressed in a foreign host cell. Although redundancy in the genetic code allows amino acids to be encoded by multiple codons, different organisms favor some codons over others. It has been found that the efficiency of protein translation in a non-native host cell can be substantially increased by adjusting the codon usage frequency but maintaining the same gene product (U.S. Pat. Nos. 5,096, 825, 5,670,356, and 5,874,304).

[0005] However, altering codon usage may, in turn, result in the unintentional introduction into a synthetic nucleic acid molecule of inappropriate transcription regulatory sequences. This may adversely effect transcription, resulting in anomalous expression of the synthetic DNA. Anomalous expression is defined as departure from normal or expected levels of expression. For example, transcription factor binding sites located downstream from a promoter have been demonstrated to effect promoter activity (Michael et al., 1990; Lamb et al., 1998; Johnson et al., 1998; Jones et al., 1997). Additionally, it is not uncommon for an enhancer element to exert activity and result in elevated levels of DNA transcription in the absence of a promoter sequence or for the presence of transcription regulatory sequences to increase the basal levels of gene expression in the absence of a promoter sequence.

[0006] Thus, what is needed is a method for making synthetic nucleic acid molecules with altered codon usage without also introducing inappropriate or unintended transcription regulatory sequences for expression in a particular host cell.

SUMMARY OF THE INVENTION

[0007] The invention provides a synthetic nucleic acid molecule comprising at least 300 nucleotides of a coding region for a polypeptide, having a codon composition differing at more than 25% of the codons from a wild type nucleic acid sequence encoding a polypeptide, and having at least 3-fold fewer, preferably at least 5-fold fewer, transcription regulatory sequences than would result if the differing codons were randomly selected. Preferably, the synthetic nucleic acid molecule encodes a polypeptide that has an amino acid sequence that is at least 85%, preferably 90%, and most preferably 95% or 99% identical to the amino acid sequence of the naturally-occurring (native or wild type) polypeptide (protein) from which it is derived. Thus, it is recognized that some specific amino acid changes may also be desirable to alter a particular phenotypic characteristic of the polypeptide encoded by the synthetic nucleic acid molecule. Preferably, the amino acid sequence identity is over at least 100 contiguous amino acid residues. In one embodiment of the invention, the codons in the synthetic nucleic acid molecule that differ preferably encode the same amino acids as the corresponding codons in the wild type nucleic acid sequence.

[0008] The transcription regulatory sequences which are reduced in the synthetic nucleic acid molecule include, but are not limited to, any combination of transcription factor binding sequences, intron splice sites, poly(A) addition sites, enhancer sequences and promoter sequences. Transcription regulatory sequences are well known in the art.

[0009] It is preferred that the synthetic nucleic acid molecule of the invention has a codon composition that differs from that of the wild type nucleic acid sequence at more than 30%, 35%, 40% or more than 45%, e.g., 50%, 55%, 60% or more of the codons. Preferred codons for use in the invention are those which are employed more frequently than at least one other codon for the same amino acid in a particular organism and, more preferably, are also not low-usage codons in that organism and are not low-usage codons in the organism used to clone or screen for the expression of the synthetic nucleic acid molecule (for example, *E. coli*).

Moreover, preferred codons for certain amino acids (i.e., those amino acids that have three or more codons,), may include two or more codons that are employed more frequently than the other (non-preferred) codon(s). The presence of codons in the synthetic nucleic acid molecule that are employed more frequently in one organism than in another organism results in a synthetic nucleic acid molecule which, when introduced into the cells of the organism that employs those codons more frequently, is expressed in those cells at a level that is greater than the expression of the wild type or parent nucleic acid sequence in those cells. For example, the synthetic nucleic acid molecule of the invention is expressed at a level that is at least about 110%, e.g., 150%, 200%, 500% or more (1000%, 5000%, or 10000%) of that of the wild type nucleic acid sequence in a cell or cell extract under identical conditions (such as cell culture conditions, vector backbone, and the like).

[0010] In one embodiment of the invention, the codons that are different are those employed more frequently in a mammal, while in another embodiment the codons that are different are those employed more frequently in a plant. A particular type of mammal, e.g., human, may have a different set of preferred codons than another type of mammal. Likewise, a particular type of plant may have a different set of preferred codons than another type of plant. In one embodiment of the invention. the majority of the codons which differ are ones that are preferred codons in a desired host cell. Preferred codons for mammals (e.g., humans) and plants are known to the art (e.g., Wada et al., 1990). For example, preferred human codons include, but are not limited to, CGC (Arg), CTG (Leu), TCT (Ser), AGC (Ser), ACC (Thr), CCA (Pro), CCT (Pro), GCC (Ala), GGC (Gly), GTG (Val), ATC (Ile), ATT (Ile), AAG (Lys), AAC (Asn), CAG (Gln), CAC (His), GAG (Glu), GAC (Asp), TAC (Tyr), TGC (Cys) and TTC (Phe) (Wada et al., 1990). Thus, preferred "humanized" synthetic nucleic acid molecules of the invention have a codon composition which differs from a wild type nucleic acid sequence by having an increased number of the preferred human codons, e.g. CGC, CTG, TCT, AGC, ACC, CCA, CCT, GCC, GGC, GTG, ATC, ATT, AAG, AAC, CAG, CAC, GAG, GAC, TAC, TGC, TTC, or any combination thereof. For example, the synthetic nucleic acid molecule of the invention may have an increased number of CTG or TTG leucine-encoding codons, GTG or GTC valine-encoding codons, GGC or GGT glycine-encoding codons, ATC or ATT isoleucine-encoding codons, CCA or CCT proline-encoding codons, CGC or CGT arginineencoding codons, AGC or TCT serine-encoding codons, ACC or ACT threonine-encoding codon, GCC or GCT alanine-encoding codons, or any combination thereof, relative to the wild type nucleic acid sequence. Similarly, synthetic nucleic acid molecules having an increased number of codons that are employed more frequently in plants, have a codon composition which differs from a wild type or parent nucleic acid sequence by having an increased number of the plant codons including, but not limited to, CGC (Arg), CTT (Leu), TCT (Ser), TCC (Ser), ACC (Thr), CCA (Pro), CCT (Pro), GCT (Ser), GGA (Gly), GTG (Val), ATC (Ile), ATT (Ile), AAG (Lys), AAC (Asn), CAA (Gln), CAC (His), GAG (Glu), GAC (Asp), TAC (Tyr), TGC (Cys), TTC (Phe), or any combination thereof (Murray et al., 1989). Preferred codons may differ for different types of plants (Wada et al., 1990).

[0011] The choice of codon may be influenced by many factors such as, for example, the desire to have an increased number of nucleotide substitutions or decreased number of transcription regulatory sequences. Under some circumstances (e.g. to permit removal of a transcription factor binding site) it may be desirable to replace a non-preferred codon with a codon other than a preferred codon or a codon other than the most preferred codon. Under other circumstances, for example, to prepare codon distinct versions of a synthetic nucleic acid molecule, preferred codon pairs are selected based upon the largest number of mismatched bases, as well as the criteria described above.

[0012] The presence of codons in the synthetic nucleic acid molecule that are employed more frequently in one organism than in another organism, results in a synthetic nucleic acid molecule which, when introduced into a cell of the organism that employs those codons, is expressed in that cell at a level which is greater than the level of expression of the wild type or parent nucleic acid sequence.

[0013] A synthetic nucleic acid molecule of the invention may encode a selectable marker protein or a reporter molecule. However, the invention applies to any gene and is not limited to synthetic reporter genes or synthetic selectable marker genes. In one embodiment of a synthetic nucleic acid molecule of the invention that is a reporter molecule, the synthetic nucleic acid molecule encodes a luciferase having a codon composition different than that of a wild type or parent Renilla luciferase or a beetle luciferase nucleic acid sequence. A synthetic click beetle luciferase nucleic acid molecule of the invention may optionally encode the amino acid valine at position 224 (i.e., it emits green light), or may optionally encode the amino acid histidine at position 224, histidine at position 247, isoleucine at position 346, glutamine at position 348 or combination thereof (i.e., it emits red light). Preferred synthetic luciferase nucleic acid molecules that are related to a wild type Renilla luciferase nucleic acid sequence include, but are not limited to, SEQ ID NO:21 (Rlucver2) or SEQ ID NO:22 (Rluc-final). Preferred synthetic luciferase nucleic acid molecules that are related to click beetle luciferase nucleic acid sequences include, but are not limited to, SEQ ID NO:7 (GRver5), SEQ ID NO:8 (GR6), SEQ ID NO:9 (GRver5.1), SEQ ID NO:14 (RDver5), SEQ ID NO:15 (RD7), SEQ ID NO:16 (RDver5.1), SEQ ID NO:17 (RDver5.2) or SEQ ID NO:18 (RD156-1H9).

[0014] The invention also provides an expression cassette. The expression cassette of the invention comprises a synthetic nucleic acid molecule of the invention operatively linked to a promoter that is functional in a cell. Preferred promoters are those functional in mammalian cells and those functional in plant cells. Optionally, the expression cassette may include other sequences, e.g., restriction enzyme recognition sequences and a Kozak sequence, and be a part of a larger polynucleotide molecule such as a plasmid, cosmid, artificial chromosome or vector, e.g., a viral vector.

[0015] Also provided is a host cell comprising the synthetic nucleic acid molecule of the invention, an isolated polypeptide (e.g., a fusion polypeptide encoded by the synthetic nucleic acid molecule of the invention), and compositions and kits comprising the synthetic nucleic acid molecule of the invention or the polypeptide encoded thereby in suitable container means and, optionally, instruc-

tion means. Preferred isolated polypeptides include, but are not limited to, those comprising SEQ ID NO:31 (GRver5.1), SEQ ID NO:226 (Rluc-final), or SEQ ID NO:223 (RD156-1H9).

[0016] The invention also provides a method to prepare a synthetic nucleic acid molecule of the invention by genetically altering a parent (either a wild type or another synthetic) nucleic acid sequence. The method may be used to prepare a synthetic nucleic acid molecule encoding a polypeptide comprising at least 100 amino acids. One embodiment of the invention is directed to the preparation of synthetic genes encoding reporter or selectable marker proteins. The method of the invention may be employed to alter the codon usage frequency and decrease the number of transcription regulatory sequences in any open reading frame or to decrease the number of transcription regulatory sites in a vector backbone. Preferably, the codon usage frequency in the synthetic nucleic acid molecule is altered to reflect that of the host organism desired for expression of that nucleic acid molecule while also decreasing the number of potential transcription regulatory sequences relative to the parent nucleic acid molecule.

[0017] Thus, the invention provides a method to prepare a synthetic nucleic acid molecule comprising an open reading frame. The method comprises altering (e.g., decreasing or eliminating) a plurality of transcription regulatory sequences in a parent (wild type or a synthetic) nucleic acid sequence that encodes a polypeptide having at least 100 amino acids to yield a synthetic nucleic acid molecule which has a decreased number of transcription regulatory sequences and which preferably encodes the same amino acids as the parent nucleic acid molecule. The transcription regulatory sequences are selected from the group consisting of transcription factor binding sequences, intron splice sites, poly(A) addition sites, enhancer sequences and promoter sequences, and the resulting synthetic nucleic acid molecule has at least 3-fold fewer, preferably 5-fold fewer, transcription regulatory sequences relative to the parent nucleic acid sequence. The method also comprises altering greater than 25% of the codons in the synthetic nucleic acid sequence which has a decreased number of transcription regulatory sequences to yield a further synthetic nucleic acid molecule, wherein the codons that are altered encode the same amino acids as those in the corresponding position in the synthetic nucleic acid molecule which has a decreased number of transcription regulatory sequences and/or in the parent nucleic acid sequence. Preferably, the codons which are altered do not result in an increase in transcriptional regulatory sequences. Preferably, the further synthetic nucleic acid molecule encodes a polypeptide that has at least 85%, preferably 90%, and most preferably 95% or 99% contiguous amino acid sequence identity to the amino acid sequence of the polypeptide encoded by the parent nucleic acid sequence.

[0018] Alternatively, the method comprises altering greater than 25% of the codons in a parent nucleic acid sequence which encodes a polypeptide having at least 100 amino acids to yield a codon-altered synthetic nucleic acid molecule, wherein the codons that are altered encode the same amino acids as those present in the corresponding positions in the parent nucleic acid sequence. Then, a plurality of transcription regulatory sequences in the codon-altered synthetic nucleic acid molecule are altered to yield a

further synthetic nucleic acid molecule. Preferably, the codons which are altered do not result in an increase in transcriptional regulatory sequences. Also, preferably, the further synthetic nucleic acid molecule encodes a polypeptide that has at least 85%, preferably 90%, and most preferably 95% or 99% contiguous amino acid sequence identity to the amino acid sequence of the polypeptide encoded by the parent nucleic acid sequence. Also provided is a synthetic (including a further synthetic) nucleic acid molecule prepared by the methods of the invention.

[0019] As described hereinbelow, the methods of the invention were employed with click beetle luciferase and Renilla luciferase nucleic acid sequences. While both of these nucleic acid molecules encode luciferase proteins, they are from entirely different families and are widely separated evolutionarily. These proteins have unrelated amino acid sequences, protein structures, and they utilize dissimilar chemical substrates. The fact that they share the name "luciferase" should not be interpreted to mean that they are from the same family, or even largely similar families. The methods produced synthetic luciferase nucleic acid molecules which exhibited significantly enhanced levels of mammalian expression without negatively effecting other desirable physical or biochemical properties (including protein half-life) and which were also largely devoid of known transcription regulatory elements.

[0020] The invention also provides at least two synthetic nucleic acid molecules that encode highly related polypeptides, but which synthetic nucleic acid molecules have an increased number of nucleotide differences relative to each other. These differences decrease the recombination frequency between the two synthetic nucleic acid molecules when those molecules are both present in a cell (i.e., they are "codon distinct" versions of a synthetic nucleic acid molecule). Thus, the invention provides a method for preparing at least two synthetic nucleic acid molecules that are codon distinct versions of a parent nucleic acid sequence that encodes a polypeptide. The method comprises altering a parent nucleic acid sequence to yield a first synthetic nucleic acid molecule having an increased number of a first plurality of codons that are employed more frequently in a selected host cell relative to the number of those codons present in the parent nucleic acid sequence. Optionally, the first synthetic nucleic acid molecule also has a decreased number of transcription regulatory sequences relative to the parent nucleic acid sequence. The parent nucleic acid sequence is also altered to yield a second synthetic nucleic acid molecule having an increased number of a second plurality of codons that are employed more frequently in the host cell relative to the number of those codons in the parent nucleic acid sequence, wherein the first plurality of codons is different than the second plurality of codons, and wherein the first and the second synthetic nucleic acid molecules preferably encode the same polypeptide. Optionally, the second synthetic nucleic acid molecule has a decreased number of transcription regulatory sequences relative to the parent nucleic acid sequence. Either or both synthetic molecules can then be further modified.

[0021] Clearly, the present invention has applications with many genes and across many fields of science including, but not limited to, life science research, agrigenetics, genetic therapy, developmental science and pharmaceutical development.

BRIEF DESCRIPTION OF THE FIGURES

[0022] FIG. 1. Codons and their corresponding amino acids.

[0023] FIG. 2. A nucleotide sequence comparison of a yellow-green (YG) click beetle luciferase nucleic acid sequence (YG #81-6G01; SEQ ID NO:2) and various synthetic green (GR) click beetle luciferase nucleic acid sequences (GRver1, SEQ ID NO:3; GRver2, SEQ ID NO:4; GRver3, SEQ ID NO:5; GRver4, SEQ ID NO:6; GRver5, SEQ ID NO:7; GR6, SEQ ID NO:8; GRver5.1, SEQ ID NO:9) and various red (RD) click beetle luciferase nucleic acid sequences (RDver1, SEQ ID NO: 10; RDver2, SEQ ID NO:11; RDver3, SEQ ID NO:12; RDver4, SEQ ID NO:13; RDver5, SEQ ID NO:14; RD7, SEQ ID NO:15; RDver5.1, SEQ ID NO:16; RDver5.2, SEQ ID NO:17; RD156-1H9, SEQ ID NO:18). The nucleotides enclosed in boxes are nucleotides that differ from the nucleotide present at the homologous position in SEQ ID NO:2.

[0024] FIG. 3. An amino acid sequence comparison of a YG click beetle luciferase amino acid sequence (YG#81-6G01, SEQ ID NO:24) and various synthetic GR click beetle luciferase amino acid sequences (GRver1, SEQ ID NO:25; GRver2, SEQ ID NO:26; GRver3, SEQ ID NO:27; GRver4, SEQ ID NO:28; GRver5, SEQ ID NO:29; GR6, SEQ ID NO:30; GRver5.1, SEQ ID NO:31) and various red (RD) click beetle luciferase amino acid sequences (RDver1, SEQ ID NO:32; RDver2, SEQ ID NO:33; RDver3, SEQ ID NO:34; RDver4, SEQ ID NO:218; RDver5, SEQ ID NO:219; RD7, SEQ ID NO:220; RDver5.1, SEQ ID NO:221; RDver5.2, SEQ ID NO:222; RD156-1H9, SEQ ID NO:223). All amino acid sequences are inferred from the corresponding nucleotide sequence. The amino acids enclosed in boxes are amino acids that differ from the amino acid present at the homologous position in SEQ ID NO:24.

[0025] FIG. 4. Codon usage in YG#81-6G01, GRver1, RDver1, GRver5, and RDver5, and humans (HUM) and relative codon usage in YG#81-6G01, GRver5, RDver5, and humans.

[0026] FIG. 5. Codon usage summaries for YG#81-6G01 (FIG. 5A), and GR/RD synthetic nucleic acid sequences, GRver1 (FIG. 5B), RDver1 (FIG. 5C), GRver2 (FIG. 5D), RDver2 (FIG. 5E), GRver3 (FIG. 5F), RDver3 (FIG. 5G), GRver4 (FIG. 5H), RDver4 (FIG. 5I), GRver5 (FIG. 5J), RDver5 (5K).

[0027] FIG. 6. Oligonucleotides employed to prepare synthetic GR/RD luciferase genes (SEQ ID Nos. 35-245).

[0028] FIG. 7. A nucleotide sequence comparison of a wild type *Renilla reniformis* luciferase nucleic acid sequence Genbank Accession No. M63501 (RELLUC, SEQ ID NO:19) and various synthetic *Renilla* luciferase nucleic acid sequences (Rlucver1, SEQ ID NO:20; Rlucver2, SEQ ID NO:21; Rluc-final, SEQ ID NO:22). The nucleotides enclosed in boxes are nucleotides that differ from the nucleotide present at the homologous position in SEQ ID NO:19.

[0029] FIG. 8. An amino acid sequence comparison of a wild type *Renilla reniformis* luciferase amino acid sequence (RELLUC, SEQ ID NO:224) and various synthetic *Renilla reniformis* luciferase amino acid sequences (Rlucver1, SEQ ID NO:225; Rlucver2, SEQ ID NO:226; Rluc-final, SEQ ID NO:227). All amino acid sequences are inferred from the

corresponding nucleotide sequence. The amino acids enclosed in boxes are amino acids that differ from the amino acid present at the homologous position in SEQ ID NO:224.

[0030] FIG. 9. Codon usage in wild-type (A) versus synthetic (B) *Renilla* luciferase genes. For codon usage in selected organisms, see, e.g., Wada et al., 1990; Sharp et al., 1988; Aota et al., 1988; and Sharp et al., 1987, and for plant codons, Murray et al. 1989.

[0031] FIG. 10. Oligonucleotides employed to prepare synthetic *Renilla* luciferase gene (SEQ ID Nos. 246-292).

[0032] FIG. 11. A nucleotide sequence comparison of a wild type yellow-green (YG) click beetle luciferase nucleic acid sequence (LUCPPLYG, SEQ ID NO:1) and the synthetic green click beetle luciferase nucleic acid sequences (GRver5.1, SEQ ID NO:9) and the synthetic red click beetle luciferase nucleic acid sequences (RD156-1H9, SEQ ID NO:18). The nucleotides enclosed in boxes are nucleotides that differ from the nucleotide present at the homologous position in SEQ ID NO:1. Both synthetic sequences have a codon composition that differs from LUCPPLYG at more than 25% of the codons and have at least 3-fold fewer transcription regulatory sequences relative to a random selection of codons at the codons which differ.

[0033] FIG. 12. An amino acid sequence comparison of a wild type YG click beetle luciferase amino acid sequence (LUCPPLYG, SEQ ID NO:23) and the synthetic GR click beetle luciferase amino acid sequences (GRver5. 1, SEQ ID NO:31) and the red (RD) click beetle luciferase amino acid sequences (RD156-1H9, SEQ ID NO:223). All amino acid sequences are inferred from the corresponding nucleotide sequence. The amino acids enclosed in boxes are amino acids that differ from the amino acid present at the homologous position in SEQ ID NO:23.

[0034] FIG. 13. pRL vector series. All of the vectors contain the *Renilla* wild type or synthetic gene as further described herein. FIG. 13A illustrates the *Renilla* luciferase gene in the pGL3 vectors (Promega Corp.) FIG. 13B illustrates the *Renilla* luciferase co-reporter vector series. pRL-TK has the herpes simplex virus (HSV) tk promoter; pRL-SV40 has the SV40 virus early enhancer/promoter; pRL-CMV has the cytomegalovirus (CMV) enhancer and immediate early promoter; pRL-null has MCS (multiple cloning sites) but no promoter or enhancer; pRL-TK(Int⁻) has HSV/tk promoter without an intron that is present in the other plasmids; pR-GL3B has the pGL-3 Basic backbone (Promega Corp.); pR-GL3 TK has the pGL3-Basic backbone with an HSV tk promoter.

[0035] FIG. 14. Half-life of synthetic (Rluc-final) and native *Renilla* luciferases in CHO cells.

[0036] FIGS. 15A-B. In vitro transcription/translation of *Renilla* luciferase nucleic acid sequences. A) t=0-60 minutes; B) linear range.

[0037] FIGS. 15C-D. In vitro translation of native and synthetic (Rluc-final) *Renilla* luciferase RNAs in a rabbit reticulocyte lysate. RNA was quantitated and the same amount was employed as in the translation reaction shown in FIGS. 15A-B. C) t=0-60 minutes; D) linear range.

[0038] FIGS. 15E-F. Translation of native and synthetic (Rluc-final) *Renilla* RNAs in a wheat germ extract. E) t=0-60 minutes; F) linear range.

[0039] FIG. 16. High expression from a synthetic *Renilla* nucleic acid sequence reduces the risk of promoter interference in a co-transfection assay. CHO cells were co-transfected with a constant amount (50 ng) of firefly luciferase expression vector (pGL3 control vector, with SV40 promoter and enhancer; Luc+) and a pRL vector having a native (0 ng, 50 ng, 100 ng, 500 ng, 1 µg or 2 µg) or synthetic (0 ng, 5 ng, 10 ng, 50 ng, 100 ng or 200 ng) *Renilla* luciferase gene.

[0040] FIGS. 17A-B. Illustrates the reactions catalyzed by firefly and click beetle (17A), and *Renilla* (17B) luciferases.

[0041] FIG. 18. Nucleotide and inferred amino acid sequence of click beetle luciferases in pGL3 vectors (GRver5.1 in pGL3, SEQ ID NO:297 encoding SEQ ID NO:298; RDver5.1 in pGL3, SEQ ID NO:299 encoding SEQ ID NO:300; and RD156-1H9 in pGL3, SEQ ID NO:301 encoding SEQ ID NO:302). To clone GRver5.1, RDver5.1, and RD156-1H9 nucleic acid sequences into pGL3 vectors, an oligonucleotide having an Nco I site at the initiation codon was employed, which resulted in an amino acid substitution at position 2 to valine.

DETAILED DESCRIPTION OF THE INVENTION

Definitions

[0042] The term "gene" as used herein, refers to a DNA sequence that comprises coding sequences necessary for the production of a polypeptide or protein precursor.

[0043] The polypeptide can be encoded by a full length coding sequence or by any portion of the coding sequence, as long as the desired protein activity is retained.

[0044] A "nucleic acid", as used herein, is a covalently linked sequence of nucleotides in which the 3' position of the pentose of one nucleotide is joined by a phosphodiester group to the 5' position of the pentose of the next, and in which the nucleotide residues (bases) are linked in specific sequence, i.e., a linear order of nucleotides. A "polynucleotide", as used herein, is a nucleic acid containing a sequence that is greater than about 100 nucleotides in length. An "oligonucleotide", as used herein, is a short polynucleotide or a portion of a polynucleotide. An oligonucleotide typically contains a sequence of about two to about one hundred bases. The word "oligo" is sometimes used in place of the word "oligonucleotide".

[0045] Nucleic acid molecules are said to have a "5'-terminus" (5' end) and a "3'-terminus" (3' end) because nucleic acid phosphodiester linkages occur to the 5' carbon and 3' carbon of the pentose ring of the substituent mononucleotides. The end of a polynucleotide at which a new linkage would be to a 5' carbon is its 5' terminal nucleotide. The end of a polynucleotide at which a new linkage would be to a 3' carbon is its 3' terminal nucleotide. A terminal nucleotide, as used herein, is the nucleotide at the end position of the 3'- or 5'-terminus.

[0046] DNA molecules are said to have "5' ends" and "3' ends" because mononucleotides are reacted to make oligonucleotides in a manner such that the 5' phosphate of one mononucleotide pentose ring is attached to the 3' oxygen of its neighbor in one direction via a phosphodiester linkage. Therefore, an end of an oligonucleotides referred to as the

"5' end" if its 5' phosphate is not linked to the 3' oxygen of a mononucleotide pentose ring and as the "3' end" if its 3' oxygen is not linked to a 5' phosphate of a subsequent mononucleotide pentose ring.

[0047] As used herein, a nucleic acid sequence, even if internal to a larger oligonucleotide or polynucleotide, also may be said to have 5' and 3' ends. In either a linear or circular DNA molecule, discrete elements are referred to as being "upstream" or 5' of the "downstream" or 3' elements. This terminology reflects the fact that transcription proceeds in a 5' to 3' fashion along the DNA strand. Typically, promoter and enhancer elements that direct transcription of a linked gene are generally located 5' or upstream of the coding region. However, enhancer elements can exert their effect even when located 3' of the promoter element and the coding region. Transcription termination and polyadenylation signals are located 3' or downstream of the coding region.

[0048] The term "codon" as used herein, is a basic genetic coding unit, consisting of a sequence of three nucleotides that specify a particular amino acid to be incorporation into a polypeptide chain, or a start or stop signal. FIG. 1 contains a codon table. The term "coding region" when used in reference to structural gene refers to the nucleotide sequences that encode the amino acids found in the nascent polypeptide as a result of translation of a mRNA molecule. Typically, the coding region is bounded on the 5' side by the nucleotide triplet "ATG" which encodes the initiator methionine and on the 3' side by a stop codon (e.g., TAA, TAG, TGA). In some cases the coding region is also known to initiate by a nucleotide triplet "TTG".

[0049] By "protein" and "polypeptide" is meant any chain of amino acids, regardless of length or post-translational modification (e.g., glycosylation or phosphorylation). The synthetic genes of the invention may also encode a variant of a naturally-occurring protein or polypeptide fragment thereof. Preferably, such a protein polypeptide has an amino acid sequence that is at least 85%, preferably 90%, and most preferably 95% or 99% identical to the amino acid sequence of the naturally-occurring (native) protein from which it is derived.

[0050] Polypeptide molecules are said to have an "amino terminus" (N-terminus) and a "carboxy terminus" (C-terminus) because peptide linkages occur between the backbone amino group of a first amino acid residue and the backbone carboxyl group of a second amino acid residue. The terms "N-terminal" and "C-terminal" in reference to polypeptide sequences refer to regions of polypeptides including portions of the N-terminal and C-terminal regions of the polypeptide, respectively. A sequence that includes a portion of the N-terminal region of polypeptide includes amino acids predominantly from the N-terminal half of the polypeptide chain, but is not limited to such sequences. For example, an N-terminal sequence may include an interior portion of the polypeptide sequence including bases from both the N-terminal and C-terminal halves of the polypeptide. The same applies to C-terminal regions. N-terminal and C-terminal regions may, but need not, include the amino acid defining the ultimate N-terminus and C-terminus of the polypeptide, respectively.

[0051] The term "wild type" as used herein, refers to a gene or gene product that has the characteristics of that gene

or gene product isolated from a naturally occurring source. A wild type gene is that which is most frequently observed in a population and is thus arbitrarily designated the "wild type" form of the gene. In contrast, the term "mutant" refers to a gene or gene product that displays modifications in sequence and/or functional properties (i.e., altered characteristics) when compared to the wild type gene or gene product. It is noted that naturally-occurring mutants can be isolated; these are identified by the fact that they have altered characteristics when compared to the wild type gene or gene product.

[0052] The terms "complementary" or "complementarity" are used in reference to a sequence of nucleotides related by the base-pairing rules. For example, for the sequence 5"A-G-T" 3', is complementary to the sequence 3"T-C-A" 5'. Complementarity may be "partial," in which only some of the nucleic acids' bases are matched according to the base pairing rules. Or, there may be "complete" or "total" complementarity between the nucleic acids. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, as well as detection methods which depend upon hybridization of nucleic acids.

[0053] The term "recombinant protein" or "recombinant polypeptide" as used herein refers to a protein molecule expressed from a recombinant DNA molecule. In contrast, the term "native protein" is used herein to indicate a protein isolated from a naturally occurring (i.e., a nonrecombinant) source. Molecular biological techniques may be used to produce a recombinant form of a protein with identical properties as compared to the native form of the protein.

[0054] The terms "fusion protein" and "fusion partner" refer to a chimeric protein containing the protein of interest (e.g., luciferase) joined to an exogenous protein fragment (e.g., a fusion partner which consists of a non-luciferase protein). The fusion partner may enhance the solubility of protein as expressed in a host cell, may, for example, provide an affinity tag to allow purification of the recombinant fusion protein from the host cell or culture supernatant, or both. If desired, the fusion partner may be removed from the protein of interest by a variety of enzymatic or chemical means known to the art.

[0055] The terms "cell," "cell line," "host cell," as used herein, are used interchangeably, and all such designations include progeny or potential progeny of these designations. By "transformed cell" is meant a cell into which (or into an ancestor of which) has been introduced a DNA molecule comprising a synthetic gene. Optionally, a synthetic gene of the invention may be introduced into a suitable cell line so as to create a stably-transfected cell line capable of producing the protein or polypeptide encoded by the synthetic gene. Vectors, cells, and methods for constructing such cell lines are well known in the art, e.g. in Ausubel, et al. (infra). The words "transformants" or "transformed cells" include the primary transformed cells derived from the originally transformed cell without regard to the number of transfers. All progeny may not be precisely identical in DNA content, due to deliberate or inadvertent mutations. Nonetheless, mutant progeny that have the same functionality as screened for in the originally transformed cell are included in the definition of transformants.

[0056] Nucleic acids are known to contain different types of mutations. A "point" mutation refers to an alteration in the sequence of a nucleotide at a single base position from the wild type sequence. Mutations may also refer to insertion or deletion of one or more bases, so that the nucleic acid sequence differs from the wild-type sequence.

[0057] The term "homology" refers to a degree of complementarity. There may be partial homology or complete homology (i.e., identity). Homology is often measured using sequence analysis software (e.g., Sequence Analysis Software Package of the Genetics Computer Group. University of Wisconsin Biotechnology Center. 1710 University Avenue. Madison, Wis. 53705). Such software matches similar sequences by assigning degrees of homology to various substitutions, deletions, insertions, and other modifications. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid, asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine.

[0058] A "partially complementary" sequence is one that at least partially inhibits a completely complementary sequence from hybridizing to a target nucleic acid is referred to using the functional term "substantially homologous." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or Northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or probe will compete for and inhibit the binding (i.e., the hybridization) of a completely homologous to a target under conditions of low stringency. This is not to say that conditions of low stringency are such that non-specific binding is permitted; low stringency conditions require that the binding of two sequences to one another be a specific (i.e., selective) interaction. The absence of non-specific binding may be tested by the use of a second target which lacks even a partial degree of complementarity (e.g., less than about 30% identity). In this case, in the absence of non-specific binding, the probe will not hybridize to the second non-complementary

[0059] When used in reference to a double-stranded nucleic acid sequence such as a cDNA or a genomic clone, the term "substantially homologous" refers to any probe which can hybridize to either or both strands of the double-stranded nucleic acid sequence under conditions of low stringency as described herein.

[0060] "Probe" refers to an oligonucleotide designed to be sufficiently complementary to a sequence in a denatured nucleic acid to be probed (in relation to its length) to be bound under selected stringency conditions.

[0061] "Hybridization" and "binding" in the context of probes and denature melted nucleic acid are used interchangeably. Probes which are hybridized or bound to denatured nucleic acid are base paired to complementary sequences in the polynucleotide. Whether or not a particular probe remains base paired with the polynucleotide depends on the degree of complementarity, the length of the probe, and the stringency of the binding conditions. The higher the stringency, the higher must be the degree of complementarity and/or the longer the probe.

[0062] The term "hybridization" is used in reference to the pairing of complementary nucleic acid strands. Hybridiza-

tion and the strength of hybridization (i.e., the strength of the association between nucleic acid strands) is impacted by many factors well known in the art including the degree of complementarity between the nucleic acids, stringency of the conditions involved affected by such conditions as the concentration of salts, the Tm (melting temperature) of the formed hybrid, the presence of other components (e.g., the presence or absence of polyethylene glycol), the molarity of the hybridizing strands and the G:C content of the nucleic acid strands.

[0063] The term "stringency" is used in reference to the conditions of temperature, ionic strength, and the presence of other compounds, under which nucleic acid hybridizations are conducted. With "high stringency" conditions, nucleic acid base pairing will occur only between nucleic acid fragments that have a high frequency of complementary base sequences. Thus, conditions of "medium" or "low" stringency are often required when it is desired that nucleic acids which are not completely complementary to one another be hybridized or annealed together. The art knows well that numerous equivalent conditions can be employed to comprise medium or low stringency conditions. The choice of hybridization conditions is generally evident to one skilled in the art and is usually guided by the purpose of the hybridization, the type of hybridization (DNA-DNA or DNA-RNA), and the level of desired relatedness between the sequences (e.g., Sarnbrook et al., 1989; Nucleic Acid Hybridization, A Practical Approach, IRL Press, Washington D.C., 1985, for a general discussion of the methods).

[0064] The stability of nucleic acid duplexes is known to decrease with an increased number of mismatched bases, and further to be decreased to a greater or lesser degree depending on the relative positions of mismatches in the hybrid duplexes. Thus, the stringency of hybridization can be used to maximize or minimize stability of such duplexes. Hybridization stringency can be altered by: adjusting the temperature of hybridization; adjusting the percentage of helix destabilizing agents, such as formamide, in the hybridization mix; and adjusting the temperature and/or salt concentration of the wash solutions. For filter hybridizations, the final stringency of hybridizations often is determined by the salt concentration and/or temperature used for the post-hybridization washes.

[0065] "High stringency conditions" when used in reference to nucleic acid hybridization comprise conditions equivalent to binding or hybridization at 42° C. in a solution consisting of 5×SSPE (43.8 g/l NaCl, 6.9 g/l NaH₂PO₄ H₂O and 1.85 g/l EDTA, pH adjusted to 7.4 with NaOH), 0.5% SDS, 5× Denhardt's reagent and 100 μg/ml denatured salmon sperm DNA followed by washing in a solution comprising 0.1×SSPE, 1.0% SDS at 42° C. when a probe of about 500 nucleotides in length is employed.

[0066] "Medium stringency conditions" when used in reference to nucleic acid hybridization comprise conditions equivalent to binding or hybridization at 42° C. in a solution consisting of 5×SSPE (43.8 g/l NaCl, 6.9 g/l NaH₂PO₄ H₂O and 1.85 g/l EDTA, pH adjusted to 7.4 with NaOH), 0.5% SDS, 5× Denhardt's reagent and 100 μg/ml denatured salmon sperm DNA followed by washing in a solution comprising 0.1×SSPE, 1.0% SDS at 42° C. when a probe of about 500 nucleotides in length is employed.

[0067] "Low stringency conditions" comprise conditions equivalent to binding or hybridization at 42° C. in a solution

consisting of 5xSSPE (43.8 g/l NaCl, 6.9 g/l NaH₂PO₄ H₂O and 1.85 g/l EDTA, pH adjusted to 7.4 with NaOH), 0.1% SDS, 5x Denhardt's reagent [50x Denhardt's contains per 500 ml: 5 g Ficoll (Type 400, Pharmacia), 5 g BSA (Fraction V; Sigma)] and 100 g/ml denatured salmon sperm DNA followed by washing in a solution comprising 5xSSPE, 0.1% SDS at 42° C. when a probe of about 500 nucleotides in length is employed.

[0068] The term " T_m " is used in reference to the "melting temperature". The melting temperature is the temperature at which 50% of a population of double-stranded nucleic acid molecules becomes dissociated into single strands. The equation for calculating the T_m of nucleic acids is wellknown in the art. The Tm of a hybrid nucleic acid is often estimated using a formula adopted from hybridization assays in 1 M salt, and commonly used for calculating Tm for PCR primers: [(number of A+T)×2° C.+(number of G+C)×4° C.]. (C. R. Newton et al., PCR, 2nd Ed., Springer-Verlag (New York, 1997), p. 24). This formula was found to be inaccurate for primers longer than 20 nucleotides. (Id.) Another simple estimate of the $T_{\rm m}$ value may be calculated by the equation: T_m =81.5+0.41(% G+C), when a nucleic acid is in aqueous solution at 1 M NaCl. (e.g., Anderson and Young, Quantitative Filter Hybridization, in Nucleic Acid Hybridization, 1985). Other more sophisticated computations exist in the art which take structural as well as sequence characteristics into account for the calculation of T_m . A calculated T_m is merely an estimate; the optimum temperature is commonly determined empirically.

[0069] The term "isolated" when used in relation to a nucleic acid, as in "isolated oligonucleotide" or "isolated polynucleotide" refers to a nucleic acid sequence that is identified and separated from at least one contaminant with which it is ordinarily associated in its source. Thus, an isolated nucleic acid is present in a form or setting that is different from that in which it is found in nature. In contrast, non-isolated nucleic acids (e.g., DNA and RNA) are found in the state they exist in nature. For example, a given DNA sequence (e.g., a gene) is found on the host cell chromosome in proximity to neighboring genes; RNA sequences (e.g., a specific mRNA sequence encoding a specific protein), are found in the cell as a mixture with numerous other mRNAs that encode a multitude of proteins. However, isolated nucleic acid includes, by way of example, such nucleic acid in cells ordinarily expressing that nucleic acid where the nucleic acid is in a chromosomal location different from that of natural cells, or is otherwise flanked by a different nucleic acid sequence than that found in nature. The isolated nucleic acid or oligonucleotide may be present in single-stranded or double-stranded form. When an isolated nucleic acid or oligonucleotide is to be utilized to express a protein, the oligonucleotide contains at a minimum, the sense or coding strand (i.e., the oligonucleotide may single-stranded), but may contain both the sense and anti-sense strands (i.e., the oligonucleotide may be double-stranded).

[0070] The term "isolated" when used in relation to a polypeptide, as in "isolated protein" or "isolated polypeptide" refers to a polypeptide that is identified and separated from at least one contaminant with which it is ordinarily associated in its source. Thus, an isolated polypeptide is present in a form or setting that is different from that in

which it is found in nature. In contrast, non-isolated polypeptides (e.g., proteins and enzymes) are found in the state they exist in nature.

[0071] The term "purified" or "to purify" means the result of any process that removes some of a contaminant from the component of interest, such as a protein or nucleic acid. The percent of a purified component is thereby increased in the sample.

[0072] The term "operably linked" as used herein refer to the linkage of nucleic acid sequences in such a manner that a nucleic acid molecule capable of directing the transcription of a given gene and/or the synthesis of a desired protein molecule is produced. The term also refers to the linkage of sequences encoding amino acids in such a manner that a functional (e.g., enzymatically active, capable of binding to a binding partner, capable of inhibiting, etc.) protein or polypeptide is produced.

[0073] The term "recombinant DNA molecule" means a hybrid DNA sequence comprising at least two nucleotide sequences not normally found together in nature.

[0074] The term "vector" is used in reference to nucleic acid molecules into which fragments of DNA may be inserted or cloned and can be used to transfer DNA segment(s) into a cell and capable of replication in a cell. Vectors may be derived from plasmids, bacteriophages, viruses, cosmids, and the like.

[0075] The terms "recombinant vector" and "expression vector" as used herein refer to DNA or RNA sequences containing a desired coding sequence and appropriate DNA or RNA sequences necessary for the expression of the operably linked coding sequence in a particular host organism. Prokaryotic expression vectors include a promoter, a ribosome binding site, an origin of replication for autonomous replication in a host cell and possibly other sequences, e.g. an optional operator sequence, optional restriction enzyme sites. A promoter is defined as a DNA sequence that directs RNA polymerase to bind to DNA and to initiate RNA synthesis. Eukaryotic expression vectors include a promoter, optionally a polyadenlyation signal and optionally an enhancer sequence.

[0076] The term "a polynucleotide having a nucleotide sequence encoding a gene," means a nucleic acid sequence comprising the coding region of a gene, or in other words the nucleic acid sequence which encodes a gene product. The coding region may be present in either a cDNA, genomic DNA or RNA form. When present in a DNA form, the oligonucleotide may be single-stranded (i.e., the sense strand) or double-stranded. Suitable control elements such as enhancers/promoters, splice junctions, polyadenylation signals, etc. may be placed in close proximity to the coding region of the gene if needed to permit proper initiation of transcription and/or correct processing of the primary RNA transcript. Alternatively, the coding region utilized in the expression vectors of the present invention may contain endogenous enhancers/promoters, splice junctions, intervening sequences, polyadenylation signals, etc. In further embodiments, the coding region may contain a combination of both endogenous and exogenous control elements.

[0077] The term "transcription regulatory element" or "transcription regulatory sequence" refers to a genetic element or sequence that controls some aspect of the expression

of nucleic acid sequence(s). For example, a promoter is a regulatory element that facilitates the initiation of transcription of an operably linked coding region. Other regulatory elements include, but are not limited to, transcription factor binding sites, splicing signals, polyadenylation signals, termination signals and enhancer elements.

[0078] Transcriptional control signals in eukaryotes comprise "promoter" and "enhancer" elements. Promoters and enhancers consist of short arrays of DNA sequences that interact specifically with cellular proteins involved in transcription (Maniatis et al., 1987). Promoter and enhancer elements have been isolated from a variety of eukaryotic sources including genes in yeast, insect and mammalian cells. Promoter and enhancer elements have also been isolated from viruses and analogous control elements, such as promoters, are also found in prokaryotes. The selection of a particular promoter and enhancer depends on the cell type used to express the protein of interest. Some eukaryotic promoters and enhancers have a broad host range while others are functional in a limited subset of cell types (for review, see Voss et al., 1986; and Maniatis et al., 1987. For example, the SV40 early gene enhancer is very active in a wide variety of cell types from many mammalian species and has been widely used for the expression of proteins in mammalian cells (Dijkema et al., 1985). Two other examples of promoter/enhancer elements active in a broad range of mammalian cell types are those from the human elongation factor 1 gene (Uetsuki et al., 1989; Kim, et al., 1990; and Mizushima and Nagata, 1990) and the long terminal repeats of the Rous sarcoma virus (Gorman et al., 1982); and the human cytomegalovirus (Boshart et al., 1985).

[0079] The term "promoter/enhancer" denotes a segment of DNA containing sequences capable of providing both promoter and enhancer functions (i.e., the functions provided by a promoter element and an enhancer element as described above). For example, the long terminal repeats of retroviruses contain both promoter and enhancer functions. The enhancer/promoter may be "endogenous" or "exogenous" or "heterologous." An "endogenous" enhancer/promoter is one that is naturally linked with a given gene in the genome. An "exogenous" or "heterologous" enhancer/promoter is one that is placed in juxtaposition to a gene by means of genetic manipulation (i.e., molecular biological techniques) such that transcription of the gene is directed by the linked enhancer/promoter.

[0080] The presence of "splicing signals" on an expression vector often results in higher levels of expression of the recombinant transcript in eukaryotic host cells. Splicing signals mediate the removal of introns from the primary RNA transcript and consist of a splice donor and acceptor site (Sambrook, et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York, 1989, pp. 16.7-16.8). A commonly used splice donor and acceptor site is the splice junction from the 16S RNA of SV40.

[0081] Efficient expression of recombinant DNA sequences in eukaryotic cells requires expression of signals directing the efficient termination and polyadenylation of the resulting transcript. Transcription termination signals are generally found downstream of the polyadenylation signal and are a few hundred nucleotides in length. The term

"poly(A) site" or "poly(A) sequence" as used herein denotes a DNA sequence which directs both the termination and polyadenylation of the nascent RNA transcript. Efficient polyadenylation of the recombinant transcript is desirable, as transcripts lacking a poly(A) tail are unstable and are rapidly degraded. The poly(A) signal utilized in an expression vector may be "heterologous" or "endogenous." An endogenous poly(A) signal is one that is found naturally at the 3' end of the coding region of a given gene in the genome. A heterologous poly(A) signal is one which has been isolated from one gene and positioned 3' to another gene. A commonly used heterologous poly(A) signal is the SV40 poly(A) signal. The SV40 poly(A) signal is contained on a 237 bp BamH I/Bcl I restriction fragment and directs both termination and polyadenylation (Sambrook, supra, at 16.6-16.7).

[0082] Eukaryotic expression vectors may also contain "viral replicons" or "viral origins of replication." Viral replicons are viral DNA sequences which allow for the extrachromosomal replication of a vector in a host cell expressing the appropriate replication factors. Vectors containing either the SV40 or polyoma virus origin of replication replicate to high copy number (up to 10⁴ copies/cell) in cells that express the appropriate viral T antigen. In contrast, vectors containing the replicons from bovine papillomavirus or Epstein-Barr virus replicate extrachromosomally at low. copy number (about 100 copies/cell).

[0083] The term "in vitro" refers to an artificial environment and to processes or reactions that occur within an artificial environment. In vitro environments include, but are not limited to, test tubes and cell lysates. The term "in situ" refers to cell culture. The term "in vivo" refers to the natural environment (e.g., an animal or a cell) and to processes or reaction that occur within a natural environment.

[0084] The term "expression system" refers to any assay or system for determining (e.g., detecting) the expression of a gene of interest. Those skilled in the field of molecular biology will understand that any of a wide variety of expression systems may be used. A wide range of suitable mammalian cells are available from a wide range of source (e.g., the American Type Culture Collection, Rockland, Md.). The method of transformation or transfection and the choice of expression vehicle will depend on the host system selected. Transformation and transfection methods are described, e.g., in Ausubel, et al., Current Protocols in Molecular Biology. John Wiley & Sons, New York. 1992. Expression systems include in vitro gene expression assays where a gene of interest (e.g., a reporter gene) is linked to a regulatory sequence and the expression of the gene is monitored following treatment with an agent that inhibits or induces expression of the gene. Detection of gene expression can be through any suitable means including, but not limited to, detection of expressed mRNA or protein (e.g., a detectable product of a reporter gene) or through a detectable change in the phenotype of a cell expressing the gene of interest. Expression systems may also comprise assays where a cleavage event or other nucleic acid or cellular change is detected.

[0085] The term "enzyme" refers to molecules or molecule aggregates that are responsible for catalyzing chemical and biological reactions. Such molecules are typically proteins, but can also comprise short peptides, RNAs,

ribozymes, antibodies, and other molecules. A molecule that catalyzes chemical and biological reactions is referred to as "having enzyme activity" or "having catalytic activity."

[0086] All amino acid residues identified herein are in the natural L-configuration. In keeping with standard polypeptide nomenclature (see *J. Biol. Chem.*, 243, 3557 (1969)), abbreviations for amino acid residues are as shown in the following Table of Correspondence.

TAI	TABLE OF CORRESPONDENCE					
1-Letter	3-Letter	AMINO ACID				
Y	Tyr	L-tyrosine				
G	Gly	glycine				
F	Phe	L-phenylalanine				
M	Met	L-methionine				
A	Ala	L-alanine				
S	Ser	L-serine				
I	Ile	L-isoleucine				
L	Leu	L-leucine				
T	Thr	L-threonine				
V	Val	L-valine				
P	Pro	L-proline				
K	Lys	L-lysine				
H	His	L-histidine				
Q	Gln	L-glutamine				
E	Glu	L-glutamic acid				
W	Trp	L-tryptophan				
R	Arg	L-arginine				
D	Asp	L-aspartic acid				
N	Asn	L-asparagine				
С	Cys	L-cysteine				

[0087] The term "sequence homology" means the proportion of base matches between two nucleic acid sequences or the proportion of amino acid matches between two amino acid sequences. When sequence homology is expressed as a percentage, e.g., 50%, the percentage denotes the proportion of matches over the length of sequence from one sequence that is compared to some other sequence. Gaps (in either of the two sequences) are permitted to maximize matching; gap lengths of 15 bases or less are usually used, 6 bases or less are preferred with 2 bases or less more preferred. When using oligonucleotides as probes or treatments, the sequence homology between the target nucleic acid and the oligonucleotide sequence is generally not less than 17 target base matches out of 20 possible oligonucleotide base pair matches (85%); preferably not less than 9 matches out of 10 possible base pair matches (90%), and more preferably not less than 19 matches out of 20 possible base pair matches (95%).

[0088] Two amino acid sequences are homologous if there is a partial or complete identity between their sequences. For example, 85% homology means that 85% of the amino acids are identical when the two sequences are aligned for maximum matching. Gaps (in either of the two sequences being matched) are allowed in maximizing matching; gap lengths of 5 or less are preferred with 2 or less being more preferred. Alternatively and preferably, two protein sequences (or polypeptide sequences derived from them of at least 100 amino acids in length) are homologous, as this term is used herein, if they have an alignment score of at more than 5 (in standard deviation units) using the program ALIGN with the mutation data matrix and a gap penalty of 6 or greater. See Dayhoff, M. O., in Atlas of Protein Sequence and Structure,

1972, volume 5, National Biomedical Research Foundation, pp. 101-110, and Supplement 2 to this volume, pp. 1-10. The two sequences or parts thereof are more preferably homologous if their amino acids are greater than or equal to 85% identical when optimally aligned using the ALIGN program.

[0089] The following terms are used to describe the sequence relationships between two or more polynucleotides: "reference sequence", "comparison window", "sequence identity", "percentage of sequence identity", and "substantial identity". A "reference sequence" is a defined sequence used as a basis for a sequence comparison; a reference sequence may be a subset of a larger sequence, for example, as a segment of a full-length cDNA or gene sequence given in a sequence listing, or may comprise a complete cDNA or gene sequence. Generally, a reference sequence is at least 20 nucleotides in length, frequently at least 25 nucleotides in length, and often at least 50 nucleotides in length. Since two polynucleotides may each (1) comprise a sequence (i.e., a portion of the complete polynucleotide sequence) that is similar between the two polynucleotides, and (2) may further comprise a sequence that is divergent between the two polynucleotides, sequence comparisons between two (or more) polynucleotides are typically performed by comparing sequences of the two polynucleotides over a "comparison window" to identify and compare local regions of sequence similarity.

[0090] A "comparison window", as used herein, refers to a conceptual segment of at least 20 contiguous nucleotides and wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) of 20 percent or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences.

[0091] Methods of alignment of sequences for comparison are well known in the art. Thus, the determination of percent identity between any two sequences can be accomplished using a mathematical algorithm. Preferred, non-limiting examples of such mathematical algorithms are the algorithm of Myers and Miller (1988); the local homology algorithm of Smith and Waterman (1981); the homology alignment algorithm of Needleman and Wunsch (1970); the search-for-similarity-method of Pearson and Lipman (1988); the algorithm of Karlin and Altschul (1990), modified as in Karlin and Altschul (1993).

[0092] Computer implementations of these mathematical algorithms can be utilized for comparison of sequences to determine sequence identity. Such implementations include, but are not limited to: CLUSTAL in the PC/Gene program (available from Intelligenetics, Mountain View, Calif.); the ALIGN program (Version 2.0) and GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Version 8 (available from Genetics Computer Group (GCG), 575 Science Drive, Madison, Wis., USA). Alignments using these programs can be performed using the default parameters. The CLUSTAL program is well described by Higgins et al. (1988); Higgins et al. (1989); Corpet et al. (1988); Huang et al. (1992); and Pearson et al. (1994). The ALIGN program is based on the algorithm of Myers and Miller, supra. The BLAST programs of Altschul et al. (1990), are based on the algorithm of Karlin and Altschul supra. To obtain gapped alignments for comparison purposes, Gapped BLAST (in BLAST 2.0) can be utilized as described in Altschul et al. (1997). Alternatively, PSI-BLAST (in BLAST 2.0) can be used to perform an iterated search that detects distant relationships between molecules. See Altschul et al., supra. When utilizing BLAST, Gapped BLAST, PSI-BLAST, the default parameters of the respective programs (e.g. BLASTN for nucleotide sequences, BLASTX for proteins) can be used. See http://www.ncbi.n-lm.nih.gov. Alignment may also be performed manually by inspection

[0093] The term "sequence identity" means that two polynucleotide sequences are identical (i.e., on a nucleotide-bynucleotide basis) over the window of comparison. The term "percentage of sequence identity" means that two polynucleotide sequences are identical (i.e., on a nucleotide-by-nucleotide basis) for the stated proportion of nucleotides over the window of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (e.g., A, T, C, G, U, or I) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (i.e., the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The terms "substantial identity" as used herein denote a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 60%, preferably at least 65%, more preferably at least 70%, up to about 85%, and even more preferably at least 90 to 95%, more usually at least 99%, sequence identity as compared to a reference sequence over a comparison window of at least 20 nucleotide positions, frequently over a window of at least 20-50 nucleotides, and preferably at least 300 nucleotides, wherein the percentage of sequence identity is calculated by comparing the reference sequence to the polynucleotide sequence which may include deletions or additions which total 20 percent or less of the reference sequence over the window of comparison. The reference sequence may be a subset of a larger sequence.

[0094] As applied to polypeptides, the term "substantial identity" means that two peptide sequences, when optimally aligned, such as by the programs GAP or BESTFIT using default gap weights, share at least about 85% sequence identity, preferably at least about 90% sequence identity, more preferably at least about 95% sequence identity, and most preferably at least about 99% sequence identity.

The Synthetic Nucleic Acid Molecules and Methods of the Invention

[0095] The invention provides compositions comprising synthetic nucleic acid molecules, as well as methods for preparing those molecules which yield synthetic nucleic acid molecules that are efficiently expressed as a polypeptide or protein with desirable characteristics including reduced inappropriate or unintended transcription characteristics when expressed in a particular cell type.

[0096] Natural selection is the hypothesis that genotypeenvironment interactions occurring at the phenotypic level lead to differential reproductive success of individuals and hence to modification of the gene pool of a population. It is generally accepted that the amino acid sequence of a protein found in nature has undergone optimization by natural selection. However, amino acids exist within the sequence of a protein that do not contribute significantly to the activity of the protein and these amino acids can be changed to other amino acids with little or no consequence. Furthermore, a protein may be useful outside its natural environment or for purposes that differ from the conditions of its natural selection. In these circumstances, the amino acid sequence can be synthetically altered to better adapt the protein for its utility in various applications.

[0097] Likewise, the nucleic acid sequence that encodes a protein is also optimized by natural selection. The relationship between coding DNA and its transcribed RNA is such that any change to the DNA affects the resulting RNA. Thus, natural selection works on both molecules simultaneously. However, this relationship does not exist between nucleic acids and proteins. Because multiple codons encode the same amino acid, many different nucleotide sequences can encode an identical protein. A specific protein composed of 500 amino acids can theoretically be encoded by more than 10^{150} different nucleic acid sequences.

[0098] Natural selection acts on nucleic acids to achieve proper encoding of the corresponding protein. Presumably, other properties of nucleic acid molecules are also acted upon by natural selection. These properties include codon usage frequency, RNA secondary structure, the efficiency of intron splicing, and interactions with transcription factors or other nucleic acid binding proteins. These other properties may alter the efficiency of protein translation and the resulting phenotype. Because of the redundant nature of the genetic code, these other attributes can be optimized by natural selection without altering the corresponding amino acid sequence.

[0099] Under some conditions, it is useful to synthetically alter the natural nucleotide sequence encoding a protein to better adapt the protein for alternative applications. A common example is to alter the codon usage frequency of a gene when it is expressed in a foreign host. Although redundancy in the genetic code allows amino acids to be encoded by multiple codons, different organisms favor some codons over others. The codon usage frequencies tend to differ most for organisms with widely separated evolutionary histories. It has been found that when transferring genes between evolutionarily distant organisms, the efficiency of protein translation can be substantially increased by adjusting the codon usage frequency (see U.S. Pat. Nos. 5,096,825, 5,670, 356 and 5,874,304).

[0100] Because of the need for evolutionary distance, the codon usage of reporter genes often does not correspond to the optimal codon usage of the experimental cells. Examples include β -galactosidase (β -gal) and chloramphenicol acetyltransferase (cat) reporter genes that are derived from E. coli and are commonly used in mammalian cells; the β-glucuronidase (gus) reporter gene that is derived from E. coli and commonly used in plant cells; the firefly luciferase (luc) reporter gene that is derived from an insect and commonly used in plant and mammalian cells; and the Renilla luciferase, and green fluorescent protein (gfp) reporter genes which are derived from coelenterates and are commonly used in plant and mammalian cells. To achieve sensitive quantitation of reporter gene expression, the activity of the gene product must not be endogenous to the experimental host cells. Thus, reporter genes are usually selected from organisms having unique and distinctive phenotypes. Consequently, these organisms often have widely separated evolutionary histories from the experimental host cells.

[0101] Previously, to create genes having a more optimal codon usage frequency but still encoding the same gene product, a synthetic nucleic acid sequence was made by replacing existing codons with codons that were generally more favorable to the experimental host cell (see U.S. Pat. Nos. 5,096,825, 5,670,356 and 5,874,304.) The result was a net improvement in codon usage frequency of the synthetic gene. However, the optimization of other attributes was not considered and so these synthetic genes likely did not reflect genes optimized by natural selection.

[0102] In particular, improvements in codon usage frequency are intended only for optimization of a RNA sequence based on its role in translation into a protein. Thus, previously described methods did not address how the sequence of a synthetic gene affects the role of DNA in transcription into RNA. Most notably, consideration had not been given as to how transcription factors may interact with the synthetic DNA and consequently modulate or otherwise influence gene transcription. For genes found in nature, the DNA would be optimally transcribed by the native host cell and would yield an RNA that encodes a properly folded gene product. In contrast, synthetic genes have previously not been optimized for transcriptional characteristics. Rather, this property has been ignored or left to chance.

[0103] This concern is important for all genes, but particularly important for reporter genes, which are most commonly used to quantitate transcriptional behavior in the experimental host cells. Hundreds of transcription factors have been identified in different cell types under different physiological conditions, and likely more exist but have not yet been identified. All of these transcription factors can influence the transcription of an introduced gene. A useful synthetic reporter gene of the invention has a minimal risk of influencing or perturbing intrinsic transcriptional characteristics of the host cell because the structure of that gene has been altered. A particularly useful synthetic reporter gene will have desirable characteristics under a new set and/or a wide variety of experimental conditions. To best achieve these characteristics, the structure of the synthetic gene should have minimal potential for interacting with transcription factors within a broad range of host cells and physiological conditions. Minimizing potential interactions between a reporter gene and a host cell's endogenous transcription factors increases the value of a reporter gene by reducing the risk of inappropriate transcriptional characteristics of the gene within a particular experiment, increasing applicability of the gene in various environments, and increasing the acceptance of the resulting experimental data.

[0104] In contrast, a reporter gene comprising a native nucleotide sequence, based on a genomic or cDNA clone from the original host organism, may interact with transcription factors when expressed in an exogenous host. This risk stems from two circumstances. First, the native nucleotide sequence contains sequences that were optimized through natural selection to influence gene transcription within the native host organism. However, these sequences might also influence transcription when the gene is expressed in exogenous hosts, i.e., out of context, thus interfering with its performance as a reporter gene. Second, the nucleotide sequence may inadvertently interact with transcription fac-

tors that were not present in the native host organism, and thus did not participate in its natural selection. The probability of such inadvertent interactions increases with greater evolutionary separation between the experimental cells and the native organism of the reporter gene.

[0105] These potential interactions with transcription factors would likely be disrupted when using a synthetic reporter gene having alterations in codon usage frequency. However, a synthetic reporter gene sequence, designed by choosing codons based only on codon usage frequency, is likely to contain other unintended transcription factor binding sites since the synthetic gene has not been subjected to the benefit of natural selection to correct inappropriate transcriptional activities. Inadvertent interactions with transcription factors could also occur whenever the encoded amino acid sequence is artificially altered, e.g., to introduce amino acid substitutions. Similarly, these changes have not been subjected to natural selection, and thus may exhibit undesired characteristics.

[0106] Thus, the invention provides a method for preparing synthetic nucleic acid sequences that reduce the risk of undesirable interactions of the nucleic acid with transcription factors when expressed in a particular host cell, thereby reducing inappropriate or unintended transcriptional characteristics. Preferably, the method yields synthetic genes containing improved codon usage frequencies for a particular host cell and with a reduced occurrence of transcription factor binding sites. The invention also provides a method of preparing synthetic genes containing improved codon usage frequencies with a reduced occurrence of transcription factor binding sites and additional beneficial structural attributes. Such additional attributes include the absence of inappropriate RNA splicing junctions, poly(A) addition signals, undesirable restriction sites, ribosomal binding sites, and secondary structural motifs such as hairpin loops.

[0107] Also provided is a method for preparing two synthetic genes encoding the same or highly similar proteins ("codon distinct" versions). Preferably, the two synthetic genes have a reduced ability to hybridize to a common polynucleotide probe sequence, or have a reduced risk of recombining when present together in living cells. To detect recombination, PCR amplification of the reporter sequences using primers complementary to flanking sequences and sequencing of the amplified sequences may be employed.

[0108] To select codons for the synthetic nucleic acid molecules of the invention, preferred codons have a relatively high codon usage frequency in a selected host cell, and their introduction results in the introduction of relatively few transcription factor binding sites, relatively few other undesirable structural attributes, and optionally a characteristic that distinguishes the synthetic gene from another gene encoding a highly similar protein. Thus, the synthetic nucleic acid product obtained by the method of the invention is a synthetic gene with improved level of expression due to improved codon usage frequency, a reduced risk of inappropriate transcriptional behavior due to a reduced number of undesirable transcription regulatory sequences, and optionally any additional characteristic due to other criteria that may be employed to select the synthetic sequence.

[0109] The invention may be employed with any nucleic acid sequence, e.g., a native sequence such as a cDNA or one which has been manipulated in vitro, e.g., to introduce

specific alterations such as the introduction or removal of a restriction enzyme recognition site, the alteration of a codon to encode a different amino acid or to encode a fusion protein, or to alter GC or AT content (% of composition) of nucleic acid molecules. Moreover, the method of the invention is useful with any gene, but particularly useful for reporter genes as well as other genes associated with the expression of reporter genes, such as selectable markers. Preferred genes include, but are not limited to, those encoding lactamase (β-gal), neomycin resistance (Neo), CAT, GUS, galactopyranoside, GFP, xylosidase, thymidine kinase, arabinosidase and the like. As used herein, a "marker gene" or "reporter gene" is a gene that imparts a distinct phenotype to cells expressing the gene and thus permits cells having the gene to be distinguished from cells that do not have the gene. Such genes may encode either a selectable or screenable marker, depending on whether the marker confers a trait which one can 'select' for by chemical means, i.e., through the use of a selective agent (e.g., a herbicide, antibiotic, or the like), or whether it is simply a "reporter" trait that one can identify through observation or testing, i.e., by 'screening'. Elements of the present disclosure are exemplified in detail through the use of particular marker genes. Of course, many examples of suitable marker genes or reporter genes are known to the art and can be employed in the practice of the invention. Therefore, it will be understood that the following discussion is exemplary rather than exhaustive. In light of the techniques disclosed herein and the general recombinant techniques which are known in the art, the present invention renders possible the alteration of any gene.

[0110] Exemplary marker genes include, but are not limited to, a neo gene, a β-gal gene, a gus gene, a cat gene, a gpt gene, a hyg gene, a hisD gene, a ble gene, a mprt gene, a bar gene, a nitrilase gene, a mutant acetolactate synthase gene (ALS) or acetoacid synthase gene (AAS), a methotrexate-resistant dhfr gene, a dalapon dehalogenase gene, a mutated anthranilate synthase gene that confers resistance to 5-methyl tryptophan (WO 97/26366), an R-locus gene, a β-lactamase gene, a xylE gene, an α-amylase gene, a tyrosinase gene, a luciferase (luc) gene, (e.g., a Renilla reniformis luciferase gene, a firefly luciferase gene, or a click beetle luciferase (Pyrophorus plagiophthalamus) gene), an aequorin gene, or a green fluorescent protein gene. Included within the terms selectable or screenable marker genes are also genes which encode a "secretable marker" whose secretion can be detected as a means of identifying or selecting for transformed cells. Examples include markers which encode a secretable antigen that can be identified by antibody interaction, or even secretable enzymes which can be detected by their catalytic activity. Secretable proteins fall into a number of classes, including small, diffusible proteins detectable, e.g., by ELISA, and proteins that are inserted or trapped in the cell membrane.

[0111] The method of the invention can be performed by, although it is not limited to, a recursive process. The process includes assigning preferred codons to each amino acid in a target molecule, e.g., a native nucleotide sequence, based on codon usage in a particular species, identifying potential transcription regulatory sequences such as transcription factor binding sites in the nucleic acid sequence having preferred codons, e.g., using a database of such binding sites, optionally identifying other undesirable sequences, and substituting an alternative codon (i.e., encoding the same amino

acid) at positions where undesirable transcription factor binding sites or other sequences occur. For codon distinct versions, alternative preferred codons are substituted in each version. If necessary, the identification and elimination of potential transcription factor or other undesirable sequences can be repeated until a nucleotide sequence is achieved containing a maximum number of preferred codons and a minimum number of undesired sequences including transcription regulatory sequences or other undesirable sequences. Also, optionally, desired sequences, e.g., restriction enzyme recognition sites, can be introduced. After a synthetic nucleic acid molecule is designed and constructed, its properties relative to the parent nucleic acid sequence can be determined by methods well known to the art. For example, the expression of the synthetic and target nucleic acid molecules in a series of vectors in a particular cell can be compared.

[0112] Thus, generally, the method of the invention comprises identifying a target nucleic acid sequence, such as a vector backbone, a reporter gene or a selectable marker gene, and a host cell of interest, for example, a plant (dicot or monocot), fungus, yeast or mammalian cell. Preferred host cells are mammalian host cells such as CHO, COS, 293, Hela, CV-1 and NIH3T3 cells. Based on preferred codon usage in the host cell(s) and, optionally, low codon usage in the host cell(s), e.g., high usage mammalian codons and low usage E. coli and mammalian codons, codons to be replaced are determined. For codon distinct versions of two synthetic nucleic acid molecules, alternative preferred codons are introduced to each version. Thus, for amino acids having more than two codons, one preferred codon is introduced to one version and another preferred codon is introduced to the other version. For amino acids having six codons, the two codons with the largest number of mismatched bases are identified and one is introduced to one version and the other codon is introduced to the other version. Concurrent, subsequent or prior to selecting codons to be replaced, desired and undesired sequences, such as undesired transcriptional regulatory sequences, in the target sequence are identified. These sequences can be identified using databases and software such as EPD, NNPD, REBASE, TRANSFAC, TESS, GenePro, MAR (www.ncgr.org/MAR-search) and BCM Gene Finder, further described herein. After the sequences are identified, the modification(s) are introduced. Once a desired synthetic nucleic acid sequence is obtained, it can be prepared by methods well known to the art (such as PCR with overlapping primers), and its structural and functional properties compared to the target nucleic acid sequence, including, but not limited to, percent homology, presence or absence of certain sequences, for example, restriction sites, percent of codons changed (such as an increased or decreased usage of certain codons) and expression rates.

[0113] As described below, the method was used to create synthetic reporter genes encoding *Renilla reniformis* luciferase, and two click beetle luciferases (one emitting green light and the other emitting red light). For both systems, the synthetic genes support much greater levels of expression than the corresponding native or parent genes for the protein. In addition, the native and parent genes demonstrated anomalous transcription characteristics when expressed in mammalian cells, which were not evident in the synthetic genes. In particular, basal expression of the native or parent genes is relatively high. Furthermore, the expres-

sion is induced to very high levels by an enhancer sequence in the absence of known promoters. The synthetic genes show lower basal expression and do not show the anomalous enhancer behavior. Presumably, the enhancer is activating transcriptional elements found in the native genes that are absent in the synthetic genes. The results clearly show that the synthetic nucleic acid sequences exhibit superior performance as reporter genes.

Exemplary Uses of the Molecules of the Invention

[0114] The synthetic genes of the invention preferably encode the same proteins as their native counterpart (or nearly so), but have improved codon usage while being largely devoid of known transcription regulatory elements in the coding region. (It is recognized that a small number of amino acid changes may be desired to enhance a property of the native counterpart protein, e.g. to enhance luminescence of a luciferase.) This increases the level of expression of the protein the synthetic gene encodes and reduces the risk of anomalous expression of the protein. For example, studies of many important events of gene regulation, which may be mediated by weak promoters, are limited by insufficient reporter signals from inadequate expression of the reporter proteins. The synthetic luciferase genes described herein permit detection of weak promoter activity because of the large increase in level of expression, which enables increased detection sensitivity. Also, the use of some selectable markers may be limited by the expression of that marker in an exogenous cell. Thus, synthetic selectable marker genes which have improved codon usage for that cell, and have a decrease in other undesirable sequences, (e.g., transcription factor binding sites), can permit the use of those markers in cells that otherwise were undesirable as hosts for those markers.

[0115] Promoter crosstalk is another concern when a coreporter gene is used to normalize transfection efficiencies. With the enhanced expression of synthetic genes, the amount of DNA containing strong promoters can be reduced, or DNA containing weaker promoters can be employed, to drive the expression of the co-reporter. In addition, there may be a reduction in the background expression from the synthetic reporter genes of the invention. This characteristic makes synthetic reporter genes more desirable by minimizing the sporadic expression from the genes and reducing the interference resulting from other regulatory pathways.

[0116] The use of reporter genes in imaging systems, which can be used for in vivo biological studies or drug screening, is another use for the synthetic genes of the invention. Due to their increased level of expression, the protein encoded by a synthetic gene is more readily detectable by an imaging system. In fact, using a synthetic *Renilla* luciferase gene, luminescence in transfected CHO cells was detected visually without the aid of instrumentation.

[0117] In addition, the synthetic genes may be used to express fusion proteins, for example fusions with secretion leader sequences or cellular localization sequences, to study transcription in difficult-to-transfect cells such as primary cells, and/or to improve the analysis of regulatory pathways and genetic elements. Other uses include, but are not limited to, the detection of rare events that require extreme sensitivity (e.g., studying RNA recoding), use with IRES, to improve the efficiency of in vitro translation or in vitro

transcription-translation coupled systems such as TNT (Promega Corp., Madison, Wis.), study of reporters optimized to different host organisms (e.g., plants, fungus, and the like), use of multiple genes as co-reporters to monitor drug toxicity, as reporter molecules in multiwell assays, and as reporter molecules in drug screening with the advantage of minimizing possible interference of reporter signal by different signal transduction pathways and other regulatory mechanisms.

[0118] Additionally, uses for the nucleic acid molecules of the invention include fluorescence activated cell sorting (FACS), fluorescent microscopy, to detect and/or measure the level of gene expression in vitro and in vivo, (e.g., to determine promoter strength), subcellular localization or targeting (fusion protein), as a marker, in calibration, in a kit, (e.g., for dual assays), for in vivo imaging, to analyze regulatory pathways and genetic elements, and in multi-well formats.

[0119] With respect to synthetic DNA encoding luciferases, the use of synthetic click beetle luciferases provides advantages such as the measurement of dual reporters. As *Renilla* luciferase is better suited for in vivo imaging (because it does not depend on ATP or Mg²⁺ for reaction, unlike firefly luciferase, and because coelenterazine is more permeable to the cell membrane than luciferin), the synthetic *Renilla* luciferase gene can be employed in vivo. Further, the synthetic *Renilla* luciferase has improved fidelity and sensitivity in dual luciferase assays, e.g., for biological analysis or in drug screening platform.

Demonstration of the Invention Using Luciferase Genes

[0120] The reporter genes for click beetle luciferase and *Renilla* luciferase were used to demonstrate the invention because the reaction catalyzed by the protein they encode are significantly easier to quantify than the product of most genes. However, for the purposes of demonstrating the present invention they represent genes in general.

[0121] Although the click beetle luciferase and Renilla luciferase genes share the name "luciferase", this should not be interpreted to mean that they originate from the same family of genes. The two luciferase proteins are evolutionarily distinct; they have fundamentally different traits and physical structures, they use vastly different substrates (FIG. 17), and they evolved from completely different families of genes. The click beetle luciferase is 61 kD in size, uses luciferin as a substrate and evolved from the CoA synthetases. The Renilla luciferase originates from the sea pansy Renilla Reniformis, is 35 kD in size, uses coelenterazine as a substrate and evolved from the $\alpha\beta$ hydrolases. The only shared trait of these two enzymes is that the reaction they catalyze results in light output. They are no more similar for resulting in light output than any other two enzymes would be, for example, simply because the reaction they catalyze results in heat.

[0122] Bioluminescence is the light produced in certain organisms as a result of luciferase-mediated oxidation reactions. The luciferase genes, e.g., the genes from luminous beetles, sea pansy, and, in particular, the luciferase from *Photinus pyralis* (the common firefly of North America), are currently the most popular luminescent reporter genes. Reference is made to Bronstein et al. (1994) for a review of luminescent reporter gene assays and to Wood (1995) for a

review of the evolution of beetle bioluminescence. See **FIG.** 17 for an illustration of the reactions catalyzed by each of firefly and click beetle luciferases (17A) and *Renilla* luciferase (17B).

[0123] Firefly luciferase and *Renilla* luciferase are highly valuable as genetic reporters due to the convenience, sensitivity and linear range of the luminescence assay. Today, luciferase is used in virtually every type of experimental biological system, including, but not limited to, prokaryotic and eukaryotic cell culture, transgenic plants and animals, and cell-free expression systems. The firefly luciferase enzyme is derived from a specific North American beetle, Photinus pyralis. The firefly luciferase enzyme and the click beetle luciferase enzyme are monomeric proteins (61 kDa) which generate light through monooxygenation of beetle luciferin utilizing ATP and O2 (FIG. 17A). The Renilla luciferase is derived from the sea pansy Renilla reniformis. The Renilla luciferase enzyme is a 36 kDa monomeric protein that utilizes O2 and coelenterazine to generate light (FIG. 17B).

[0124] The gene encoding firefly luciferase was cloned from *Photinus pyralis*, and demonstrated to produce active enzyme in *E. coli* (de Wet et al., 1987). The cDNA encoding firefly luciferase (luc) continues to gain favor as the gene of choice for reporting genetic activity in animal, plant and microbial cells. The firefly luciferase reaction, modified by the addition of CoA to produce persistent light emission, provides an extremely sensitive and rapid in vitro assay for quantifying firefly luciferase expression in small samples of transfected cells or tissues.

[0125] To use firefly luciferase or click beetle luciferase as a genetic reporter, extracts of cells expressing the luciferase are mixed with substrates (beetle luciferin, Mg2+ ATP, and O_2), and luminescence is measured immediately. The assay is very rapid and sensitive, providing gene expression data with little effort. The conventional firefly luciferase assay has been further improved by including coenzyme A in the assay reagent to yield greater enzyme turnover and thus greater luminescence intensity (Promega Luciferase Assay Reagent, Cat.# E1500, Promega Corporation, Madison, Wis.). Using this reagent, luciferase activity can be readily measured in luminometers or scintillation counters. Firefly and click beetle luciferase activity can also be detected in living cells in culture by adding luciferin to the growth medium. This in situ luminescence relies on the ability of beetle luciferin to diffuse through cellular and peroxisomal membranes and on the intracellular availability of ATP and O_2 in the cytosol and peroxisome.

[0126] Further, although reporter genes are widely used to measure transcription events, their utility can be limited by the fidelity and efficiency of reporter expression. For example, in U.S. Pat. No. 5,670,356, a firefly luciferase gene (referred to as luc+) was modified to improve the level of luciferase expression. While a higher level of expression was observed, it was not determined that higher expression had improved regulatory control.

[0127] The invention will be further described by the following nonlimiting examples.

EXAMPLE 1

Synthetic Click Beetle (RD and GR) Luciferase Nucleic Acid Molecules

[0128] LucPplYG is a wild-type click beetle luciferase that emits yellow-green luminescence (Wood, 1989). A mutant of LucPplYG named YG#81-6G01 was envisioned. YG#81-6G01 lacks a peroxisome targeting signal, has a lower K_M for luciferin and ATP, has increased signal stability and increased temperature stability when compared to the wild type (PCT/WO9914336). YG #81-6G01 was mutated to emit green luminescence by changing Ala at position 224 to Val (A224V is a green-shifting mutation), or to emit red luminescence by simultaneously introducing the amino acid substitutions A224H, S247H, N346I, and H348Q (red-shifting mutation set) (PCT/WO9518853)

[0129] Using YG #81-6G01 as a parent gene, two synthetic gene sequences were designed. One codes for a luciferase emitting green luminescence (GR) and one for a luciferase emitting red luminescence (RD). Both genes were designed to 1) have optimized codon usage for expression in mammalian cells, 2) have a reduced number of transcriptional regulatory sites including mammalian transcription factor binding sites, splice sites, poly(A) addition sites and promoters, as well as prokaryotic (E. coli) regulatory sites, 3) be devoid of unwanted restriction sites, e.g., those which are likely to interfere with standard cloning procedures, and 4) have a low DNA sequence identity compared to each other in order to minimize genetic rearrangements when both are present inside the same cell. In addition, desired sequences, e.g., a Kozak sequence or restriction enzyme recognition sites, may be identified and introduced.

[0130] Not all design criteria could be met equally well at the same time. The following priority was established for reduction of transcriptional regulatory sites: elimination of transcription factor (TF) binding sites received the highest priority, followed by elimination of splice sites and poly(A) addition sites, and finally prokaryotic regulatory sites. When removing regulatory sites, the strategy was to work from the lesser important to the most important to ensure that the most important changes were made last. Then the sequence was rechecked for the appearance of new lower priority sites and additional changes made as needed. Thus, the process for designing the synthetic GR and RD gene sequences, using computer programs described herein, involved 5 optionally iterative steps that are detailed below

- [0131] 1. Optimized codon usage and changed A224V to create GRver1, separately changed A224H, S247H, H348Q and N346I to create RDver1. These particular amino acid changes were maintained throughout all subsequent manipulations to the sequence.
- [0132] 2. Removed undesired restriction sites, prokaryotic regulatory sites, splice sites, poly(A) sites thereby creating GRver2 and RDver2.
- [0133] 3. Removed transcription factor binding sites (first pass) and removed any newly created undesired sites as listed in step 2 above thereby creating GRver3 and RDver3.
- [0134] 4. Removed transcription factor binding sites created by step 3 above (second pass) and removed any

- newly created undesired sites as listed in step 2 above thereby creating GRver4 and RDver4.
- [0135] 5. Removed transcription factor binding sites created by step 4 above (third Pass) and confirmed absence of sites listed in step 2 above thereby creating GRver5 and RDver5.
- [0136] 6. Constructed the actual genes by PCR using synthetic oligonucleotides corresponding to fragments of GRver5 and RDver5 designed sequences (FIGS. 6 and 10) thereby creating GR6 and RD7. GR6, upon sequencing was found to have the serine residue at amino acid position 49 mutated to an asparagine and the proline at amino acid position 230 mutated to a serine (S49N, P230S). RD7, upon sequencing was found to have the histidine at amino acid position 36 mutated to a tyrosine (H36Y). These changes occurred during the PCR process.
- [0137] 7. The mutations described in step 6 above (S49N, P230S for GR6 and H36Y for RD7) were reversed to create GRver5.1 and RDver5.1.
- [0138] 8. RDver5.1 was further modified by changing the arginine codon at position 351 to a glycine codon (R351G) thereby creating RDver5.2 with improved spectral properties compared to RDver5.1.
- [0139] 9. RDver5.2 was further mutated to increase luminescence intensity thereby creating RD156-1H9 which encodes four additional amino acid changes (M2I, S349T, K488T, E538V) and three silent single base changes (SEQ ID NO:18).
- Optimize Codon Usage and Introduce Mutations Determining Luminescence Color

[0140] The starting gene sequence for this design step was YG #81-6G01 (SEQ ID NO:2).

a) Optimize Codon Usage:

[0141] The strategy was to adapt the codon usage for optimal expression in human cells and at the same time to avoid *E. coli* low-usage codons. Based on these requirements, the best two codons for expression in human cells for all amino acids with more than two codons were selected (see Wada et al., 1990). In the selection of codon pairs for amino acids with six codons, the selection was biased towards pairs that have the largest number of mismatched bases to allow design of GR and RD genes with minimum sequence identity (codon distinction):

Arg: CGC/CGT Leu: CTG/TTG Ser: TCT/AGC
Thr: ACC/ACT Pro: CCA/CCT Ala: GCC/GCT
Gly: GGC/GGT Val: GTC/GTG Ile: ATC/ATT

Based on this selection of codons, two gene sequences encoding the YG#81-6G01 luciferase protein sequence were computer generated. The two genes were designed to have minimum DNA sequence identity and at the same time closely similar codon usage. To achieve this, each codon in the two genes was replaced by a codon from the limited list described above in an alternating fashion (e.g., $Arg_{(n)}$ is CGC in gene 1 and CGT in gene 2, $Arg_{(n+1)}$ is CGT in gene 1 and CGC in gene 2).

[0142] For subsequent steps in the design process it was anticipated that changes had to be made to this limited optimal codon selection in order to meet other design criteria, however, the following low-usage codons in mammalian cells were not used unless needed to meet criteria of higher priority:

Arg:	CGA	Leu:	CTA	Ser:	TCG
Pro:	CCG	Val:	GTA	Ile:	ATA

[0143] Also, the following low-usage codons in *E. coli* were avoided when reasonable (note that 3 of these match the low-usage list for mammalian cells):

Arg: CGA/CGG/AGA/AGG
Leu: CTA Pro: CCC Ile: ATA

- b) Introduce Mutations Determining Luminescence Color:
- [0144] Into one of the two codon-optimized gene sequences was introduced the single green-shifting mutation and into the other were introduced the 4 red-shifting mutations as described above.

[0145] The two output sequences from this first design step were named GRver1 (version 1 GR) and RDver1 (version 1 RD). Their DNA sequences are 63% identical (594 mismatches), while the proteins they encode differ only by the 4 amino acids that determine luminescence color (see FIGS. 2 and 3 for an alignment of the DNA and protein sequences).

[0146] Tables 1 and 2 show, as an example, the codon usage for valine and leucine in human genes, the parent gene YG#81-6G01, the codon-optimized synthetic genes GRver1 and RDver1, as well as the final versions of the synthetic genes after completion of step 5 in the design process (GRver5 and RDver5). For a complete summary of the codon changes, see FIGS. 4 and 5.

TABLE 1

			Valine			
Codon	Human	Parent	GR ver1	RD ver1	GR ver5	RD ver5
GTA	4	13	0	0	1	1
GTC	13	4	25	24	21	26
GTG	24	12	25	25	25	17
GTT	9	20	0	0	3	5

[0147]

TABLE 2

				Leucine	-		
(Codon	Human	Parent	GR ver1	RD ver1	GR ver5	RD ver5
	CTA	3	5	0	0	0	0
	CTC	12	4	0	1	12	11
	CTG	24	4	28	27	19	18
	CTT	6	12	0	0	1	1

TABLE 2-continued

			Leucii	<u>ie</u>		
Coc	lon Hı	uman Pare	ent GR ve	r1 RD ver	r1 GR ver	RD 5 ver5
TT		3 17 6 13	_	0 27	0 23	0 25

- 2. Remove Undesired Restriction Sites, Prokaryotic Regulatory Sites, Splice Sites and Poly(A) Addition Sites
- [0148] The starting gene sequences for this design step were GRver1 and RDver1.
- a) Remove Undesired Restriction Sites:
- [0149] To check for the presence and location of undesired restriction sites, the sequences of both synthetic genes were compared against a database of restriction enzyme recognition sequences (REBASE ver.712, http://www.neb.com/rebase) using standard sequence analysis software (GenePro ver 6.10, Riverside Scientific Ent.).

Specifically, the following restriction enzymes were classified as undesired:

- [0150] BamH I, Xho I, Sfi I, Kpn I, Sac I, Mlu I, Nhe I, Sma I, Xho I, Bgl II, Hind III, Nco I, Nar I, Xba I, Hpa I, Sal I,
- [0151] other cloning sites commonly used: EcoR I, EcoR V, Cla I,
- [0152] eight-base cutters (commonly used for complex constructs),
- [0153] BstE II (to allow N-terminal fusions),
- [0154] Xcm I (can generate A/T overhang used for T-vector cloning).
- To eliminate undesired restriction sites when found in a synthetic gene, one or more codons of the synthetic gene sequence were altered in accordance with the codon optimization guidelines described in 1a above.
- b) Remove Prokaryotic (E. coli) Regulatory Sequences:
- [0155] To check for the presence and location of prokaryotic regulatory sequences, the sequences of both synthetic genes were searched for the presence of the following consensus sequences using standard sequence analysis software (GenePro):
 - [0156] TATAAT (-10 Pribnow box of promoter)
 - [0157] AGGA or GGAG (ribosome binding site; only considered if paired with a methionine codon 12 or fewer bases downstream).
 - To eliminate such regulatory sequences when found in a synthetic gene, one or more codons of the synthetic gene at sequence were altered in accordance with the codon optimization guidelines described in 1a above.
 - c) Remove Splice Sites:

[0158] To check for the presence and location of splice sites, the DNA strand corresponding to the primary RNA transcript of each synthetic gene was searched for the

presence of the following consensus sequences (see Watson et al., 1983) using standard sequence analysis software (GenePro):

[0159] splice donor site: AG|GTRAGT (exon|intron), the search was performed for AGGTRAG and the lower stringency GGTRAGT;

[0160] splice acceptor site: $(Y)_nNCAG|G$ (intron|exon), the search was performed with n=1.

To eliminate splice sites found in a synthetic gene, one or more codons of the synthetic gene sequence were altered in accordance with the codon optimization guidelines described in 1a above. Splice acceptor sites were generally difficult to eliminate in one gene without introducing them into the other gene because they tended to contain one of the two only Gln codons (CAG); they were removed by placing the Gln codon CAA in both genes at the expense of a slightly increased sequence identity between the two genes.

d) Remove Poly(A) Addition Sites:

[0161] To check for the presence and location of poly(A) addition sites, the sequences of both synthetic genes were searched for the presence of the following consensus sequence using standard sequence analysis software (GenePro):

AATAAA.

To eliminate each poly(A) addition site found in a synthetic gene, one or more codons of the synthetic gene sequence were altered in accordance with the codon optimization guidelines described in 1a above. The two output sequences from this second design step were named GRver2 and RDver2. Their DNA sequences are 63% identical (590 mismatches) (FIGS. 2 and 3.

3. Remove Transcription Factor (TF) Binding Sites, then Repeat Steps 2 a-d

[0162] The starting gene sequences for this design step were GRver2 and RDver2. To check for the presence, location and identity of potential TF binding sites, the sequences of both synthetic genes were used as query sequences to search a database of transcription factor binding sites (TRANSFAC v3.2). The TRANSFAC database (http://transfac.gbf.de/TRANSFAC/index:html) holds information on gene regulatory DNA sequences (TF binding sites) and proteins (TFs) that bind to and act through them. The SITE table of TRANSFAC Release 3.2 contains 4,401 entries of individual (putative) TF binding sites (including TF binding sites in eukaryotic genes, in artificial sequences resulting from mutagenesis studies and in vitro selection procedures based on random oligonucleotide mixtures or specific theoretical considerations, and consensus binding sequences (from Faisst and Meyer, 1992)).

[0163] The software tool used to locate and display these TF binding sites in the synthetic gene sequences was TESS (Transcription Element Search Software, http://agave.hum-gen.upenn.edu/tess/index.html). The filtered string-based search option was used with the following user-defined search parameters:

[0164] Factor Selection Attribute: Organism Classifica-

[0165] Search Pattern: Mammalia

[0166] Max. Allowable Mismatch %: 0

[0167] Min. element length: 5

[0168] Min. log-likelihood: 10

This parameter selection specifies that only mammalian TF binding sites (approximately 1,400 of the 4,401 entries in the database) that are at least 5 bases long will be included in the search. It further specifies that only TF binding sites that have a perfect match in the query sequence and a minimum log likelihood (LLH) score of 10 will be reported. The LLH scoring method assigns 2 to an unambiguous match, 1 to a partially ambiguous match (e.g., A or T match W) and 0 to a match against 'N'. For example, a search with parameters specified above would result in a "hit" (positive result or match) for TATAA (SEQ ID NO:240) (LLH=10), STRATG (SEQ ID NO:241) (LLH=10), and MTTNCNNMA (SEQ ID NO:242) (LLH=10) but not for TRATG (SEQ ID NO: 243) (LLH=9) if these four TF binding sites were present in the query sequence. A lower stringency test was performed at the end of the design process to re-evaluate the search parameters.

[0169] When TESS was tested with a mock query sequence containing known TF binding sites it was found that the program was unable to report matches to sites ending with the 3' end of the query sequence. Thus, an extra nucleotide was added to the 3' end of all query sequences to eliminate this problem.

[0170] The first search for TF binding sites using the parameters described above found about 100 transcription factor binding sites (hits) for each of the two synthetic genes (GRver2 and RDver2). All sites were eliminated by changing one or more codons of the synthetic gene sequences in accordance with the codon optimization guidelines described in la above. However, it was expected that some these changes created new TF binding sites, other regulatory sites, and new restriction sites. Thus, steps 2 a-d were repeated as described, and 4 new restriction sites and 2 new splice sites were removed. The two output sequences from this third design step were named GRver3 and RDver3. Their DNA sequences are 66% identical (541 mismatches) (FIGS. 2 and 3).

4. Remove New Transcription Factor (TF) Binding Sites, then Repeat Steps 2 a-d

[0171] The starting gene sequences for this design step were GRver3 and RDver3. This fourth step is an iteration of the process described in step 3. The search for newly introduced TF binding sites yielded about 50 hits for each of the two synthetic genes. All sites were eliminated by changing one or more codons of the synthetic gene sequences in general accordance with the codon optimization guidelines described in 1a above. However, more high to medium usage codons were used to allow elimination of all TF binding sites. The lowest priority was placed on maintaining low sequence identity between the GR and RD genes. Then steps 2 a-d were repeated as described. The two output sequences from this fourth design step were named GRver4

and RDver4. Their DNA sequences are 68% identical (506 mismatches) (FIGS. 2 and 3).

5. Remove New Transcription Factor (TF) Binding Sites, then Repeat Steps 2 a-d

[0172] The starting gene sequences for this design step were GRver4 and RDver4. This fifth step is another iteration of the process described in step 3 above. The search for new TF binding sites introduced in step 4 yielded about 20 hits for each of the two synthetic genes. All sites were eliminated by changing one or more codons of the synthetic gene sequences in general accordance with the codon optimization guidelines described in 1a above. However, more high to medium usage codons were used (these are all considered "preferred") to allow elimination of all TF binding sites. The lowest priority was placed on maintaining low sequence identity between the GR and RD genes. Then steps 2 a-d were repeated as described. Only one acceptor splice site could not be eliminated. As a final step the absence of all TF binding sites in both genes as specified in step 3 was confirmed. The two output sequences from this fifth and last design step were named GRver5 and RDver5. Their DNA sequences are 69% identical (504 mismatches) (FIGS. 2 and 3).

Additional Evaluation of GRver5 and RDver5

a) Use Lower Stringency Parameters for TESS:

[0173] The search for TF binding sites was repeated as described in step 3 above, but with even less stringent user-defined parameters:

[0174] setting LLH to 9 instead of 10 did not result in new hits:

[0175] setting LLH to 0 through 8 (incl.) resulted in hits for two additional sites, MAMAG (22 hits) and CTKTK (24 hits);

[0176] setting LLH to 8 and the minimum element length to 4, the search yielded (in addition to the two sites above) different 4-base sites for AP-1, NF-1, and c-Myb that are shortened versions of their longer respective consensus sites which were eliminated in steps 3-5 above.

It was not realistic to attempt complete elimination of these sites without introduction of new sites, so no further changes were made.

b) Search Different Database:

[0177] The Eukaryotic Promoter Database (release 45) contains information about reliably mapped transcription start sites (1253 sequences) of eukaryotic genes. This database was searched using BLASTN 1.4.11 with default parameters (optimized to find nearly identical sequences rapidly; see Altschul et al, 1990) at the National Center for Biotechnology Information site (http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST). To test this approach, a portion of pGL3-Control vector sequence containing the SV40 promoter and enhancer was used as a query sequence, yielding the expected hits to SV40 sequences. No hits were found when using the two synthetic genes as query sequences.

Summary of GRver5 and RDver5 Synthetic Gene Properties

[0178] Both genes, which at this stage were still only "virtual" sequences in the computer, have a codon usage that

strongly favors mammalian high-usage codons and minimizes mammalian and *E. coli* low-usage codons. **FIG. 4** shows a summary of the codon usage of the parent gene and the various synthetic gene versions.

[0179] Both genes are also completely devoid of eukaryotic TF binding sites consisting of more than four unambiguous bases, donor and acceptor splice sites (one exception: GRver5 contains one splice acceptor site), poly(A) addition sites, specific prokaryotic (E. coli) regulatory sequences, and undesired restriction sites.

[0180] The gene sequence identity between GRver5 and RDver5 is only 69% (504 base mismatches) while their encoded proteins are 99% identical (4 amino acid mismatches), see FIGS. 2 and 3. Their identity with the parent sequence YG#81-6G1 is 74% (GRver5) and 73% (RDver5), see FIG. 2. Their base composition is 49.9% GC (GRver5) and 49.5% GC (RDver5), compared to 40.2% GC for the parent YG#81-6G01.

Construction of Synthetic Genes

[0181] The two synthetic genes were constructed by assembly from synthetic oligonucleotides in a thermocycler followed by PCR amplification of the full-length genes (similar to Stemmer et al. (1995) *Gene.* 164, pp. 49-53). Unintended mutations that interfered with the design goals of the synthetic genes were corrected.

a) Design of Synthetic Oligonucleotides:

[0182] The synthetic oligonucleotides were mostly 40mers that collectively code for both complete strands of each designed gene (1,626 bp) plus flanking regions needed for cloning (1,950 bp total for each gene; FIG. 6). The 5' and 3' boundaries of all oligonucleotides specifying one strand were generally placed in a manner to give an average offset/overlap of 20 bases relative to the boundaries of the oligonucleotides specifying the opposite strand.

[0183] The ends of the flanking regions of both genes matched the ends of the amplification primers (pRAMtailup: 5'-gtactgagacgacgagccaagcttaggcctgagtg SEQ ID NO:229, and pRAMtaildn: 5'-ggcatgagcgtgaactgactgaactagcggccgag SEQ ID NO:230) to allow cloning of the genes into our *E. coli* expression vector pRAM (WO99/14336).

[0184] A total of 183 oligonucleotides were designed (FIG. 6): fifteen oligonucleotides that collectively encode the upstream and downstream flanking sequences (identical for both genes; SEQ ID NOs: 35-49) and 168 oligonucleotides (4×42) that encode both strands of the two genes (SEQ ID NOs: 50-217).

[0185] All 183 oligonucleotides were run through the hairpin analysis of the OLIGO software (OLIGO 4.0 Primer Analysis Software © 1989-1991 by Wojciech Rychlik) to identify potentially detrimental intra-molecular loop formation. The guidelines for evaluating the analysis results were set according to recommendations of Dr. Sims (Sigma-Genosys Custom Gene Synthesis Department): oligos forming hairpins with $\Delta G < -10$ have to be avoided, those forming hairpins with $\Delta G \le -7$ involving the 3' end of the oligonucleotide should also be avoided, while those with an overall $\Delta G \le -5$ should not pose a problem for this application. The analysis identified 23 oligonucleotides able to form hairpins with a ΔG between -7.1 and -4.9. Of these, 5 had blocked

or nearly blocked 3' ends (0-3 free bases) and were redesigned by removing 1-4 bases at their 3' end and adding it to the adjacent oligonucleotide.

[0186] The 40mer oligonucleotide covering the sequence complementary to the poly(A) tail had a very low complexity 3' end (13 consecutive T bases). An additional 40mer was designed with a high complexity 3' end but a consequently reduced overlap with one of its complementary oligonucleotides (11 instead of 20 bases) on the opposite strand.

[0187] Even though the oligos were designed for use in a thermocycler-based assembly reaction, they could also be used in a ligation-based protocol for gene construction. In this approach, the oligonucleotides are annealed in a pairwise fashion and the resulting short double-stranded fragments are ligated using the sticky overhangs. However, this would require that all oligonucleotides be phosphorylated.

b) Gene Assembly and Amplification

[0188] In a first step, each of the two synthetic genes was assembled in a separate reaction from 98 oligonucleotides. The total volume for each reaction was 50 μ l:

[0189] $0.5 \mu M$ oligonucleotides (=0.25 pmoles of each oligo)

[0190] 1.0 U Taq DNA polymerase

[0191] 0.02 U Pfu DNA polymerase

[0192] 2 mM MgCl₂

[0193] 0.2 mM dNTPs (each)

[**0194**] 0.1% gelatin

[0195] Cycling conditions: (94° C. for 30 seconds, 52° C. for 30 seconds, and 72° C. for 30 seconds)×55 cycles.

[0196] In a second step, each assembled synthetic gene was amplified in a separate reaction. The total volume for each reaction was 50 µl:

[0197] 2.5 l assembly reaction

[0198] 5.0 U Taq DNA polymerase

[0199] 0.1 U Pfu DNA polymerase

[0200] 1 M each primer (pRAMtailup, pRAMtaildn)

[0201] 2 mM MgCl₂

[0202] 0.2 mM dNTPs (each)

[0203] Cycling conditions: (94° C. for 20 seconds, 65° C. for 60 seconds, 72° C. for 3 minutes)×30 cycles.

[0204] The assembled and amplified genes were subcloned into the pRAM vector and expressed in *E. coli*, yielding 1-2% luminescent GR or RD clones. Five GR and five RD clones were isolated and analyzed further. Of the five GR clones, three had the correct insert size, of which one was weakly luminescent and one had an altered restriction pattern. Of the five RD clones, two had the correct size insert with an altered restriction pattern and one of those was weakly luminescent. Overall, the analysis indicated the presence of a large number of mutations in the genes, most likely the result of errors introduced in the assembly and amplification reactions.

c) Corrective Assembly and Amplification

[0205] To remove the large number of mutations present in the full-length synthetic genes we performed an additional assembly and amplification reaction for each gene using the proof-reading DNA polymerase Tli. The assembly reaction contained, in addition to the 98 GR or RD oligonucleotides, a small amount of DNA from the corresponding full-length clones with mutations described above. This allows the oligos to correct mutations present in the templates.

[0206] The following assembly reaction was performed for each of the synthetic genes. The total volume for each reaction was 50 μ l:

[0207] $0.5 \mu M$ oligonucleotides (=0.25 pmoles of each oligo)

[0208] 0.016 pmol plasmid (mix of clones with correct insert size)

[0209] 2.5 U Tli DNA polymerase

[**0210**] 2 mM MgCl₂

[0211] 0.2 mM dNTPs (each)

[**0212**] 0.1% gelatin

[0213] Cycling conditions: 94° C. for 30 seconds, then (94° C. for 30 seconds, 52° C. for 30 seconds, 72° C. for 30 seconds) for 55 cycles, then 72° C. for 5 minutes.

[0214] The following amplification reaction was performed on each of the assembly reactions. The total volume for each amplification reaction was 50 μ l:

[0215] 1-5 µl of assembly reaction

[0216] 40 pmol each primer (pRAMtailup, pRAMtaildn)

[0217] 2.5 U Tli DNA polymerase

[**0218**] 2 mM MgCl₂

[**0219**] 0.2 mM dNTPs (each)

[0220] Cycling conditions: 94° C. for 30 seconds, then (94° C. for 20 seconds, 65° C. for 60 seconds and 72° C. for 3 minutes) for 30 cycles, then 72° C. for 5 minutes.

[0221] The genes obtained from the corrective assembly and amplification step were subcloned into the pRAM vector and expressed in *E. coli*, yielding 75% luminescent GR or RD clones. Forty-four GR and 44 RD clones were analyzed with our screening robot (WO99/14336). The six best GR and RD clones were manually analyzed and one best GR and RD clone was selected (GR6 and RD7). Sequence analysis of GR6 revealed two point mutations in the coding region, both of which resulted in an amino acid substitution (S49N and P230S). Sequence analysis of RD7 revealed three point mutations in the coding region, one of which resulted in an amino acid substitution (H36Y). It was confirmed that none of the silent point mutations introduced any regulatory or restriction sites conflicting with the overall design criteria for the synthetic genes.

d) Reversal of Unintended Amino Acid Substitutions

[0222] The unintended amino acid substitutions present in the GR6 and RD7 synthetic genes were reversed by site-directed mutagenesis to match the GRver5 and RDver5 designed sequences, thereby creating GRver5.1 and

RDver5.1. The DNA sequences of the mutated regions were confirmed by sequence analysis.

e) Improve Spectral Properties

[0223] The RDver5.1 gene was further modified to improve its spectral properties by introducing an amino change (R351G), thereby creating RDver5.2

pGL3 Vectors with RD and GR Genes

[0224] The parent click beetle luciferase YG#81-6G1 ("YG"), and the synthetic click beetle luciferase genes GRver5.1 ("GR"), RDver5.2 ("RD"), and RD156-1H9 were cloned into the four pGL3 reporter vectors (Promega Corp.):

[0225] pGL3-Basic=no promoter, no enhancer

[0226] pGL3-Control=SV40 promoter, SV40 enhancer

[0227] pGL3-Enhancer=SV40 enhancer (3' to luciferase coding sequences)

[0228] pGL3-Promoter=SV40 promoter.

[0229] The primers employed in the assembly of GR and RD synthetic genes facilitated the cloning of those genes into pRAM vectors. To introduce the genes into pGL3 vectors (Promega Corp., Madison, Wis.) for analysis in mammalian cells, each gene in a pRAM vector (pRAM RDver5.1, pRAM GRver5.1, and pRAM RD156-1H9) was amplified to introduce an Nco I site at the 5' end and an Xba I site at the 3' end of the gene. The primers for pRAM RDver5.1 and pRAM GRver5.1 were:

(SEQ ID NO:231) $\mbox{GR}{\rightarrow}5^{+}$ GGA TCC CAT GGT GAA GCG TGA GAA 3' or

 $\mbox{(SEQ ID NO:232)}$ $\mbox{RD}{\rightarrow}5^{\,\prime}$ GGA TCC CAT GGT GAA ACG CGA 3' and

(SEQ ID NO:233) 5' CTA GCT TTT TTT TCT AGA TAA TCA TGA AGA C 3'

[0230] The primers for pRAM RD156-1H9 were:

(SEQ ID NO:295)

 $5\,^{\circ}$ GCG TAG CCA TGG TAA AGC GTG AGA AAA ATG TC $3\,^{\circ}$ and

(SEQ ID NO:296)

5' CCG ACT CTA GAT TAC TAA CCG CCG GCC TTC ACC 3'

The PCR included:

[0231] 100 ng DNA plasmid

[0232] 1 μM primer upstream

[0233] 1 µM primer downstream

[0234] 0.2 mM dNTPs

[0235] 1× buffer (Promega Corp.)

[0236] 5 units Pfu DNA polymerase (Promega Corp.)

[0237] Sterile nanopure H₂O to 50 μl

[0238] The cycling parameters were: 94° C. for 5 minutes; (94° C. for 30 seconds; 55° C. for 1 minute; and 72° C. for 3 minutes)×15 cycles. The purified PCR product was

digested with Nco I and Xba I, ligated with pGL3-control that was also digested with Nco I and Xba I, and the ligated products introduced to *E. coli*. To insert the luciferase genes into the other pGL3 reporter vectors (basic, promoter and enhancer), the pGL3-control vectors containing each of the luciferase genes was digested with Nco I and Xba I, ligated with other pGL3 vectors that also were digested with Nco I and Xba I, and the ligated products introduced to *E. coli*. Note that the polypeptide encoded by GRver5.1 and RDver5.1 (and RD 156-1 H9, see below) nucleic acid sequences in pGL3 vectors has an amino acid substitution at position 2 to valine as a result of the Nco I site at the initiation codon in the oligonucleotide.

[0239] Because of internal Nco I and Xba I sites, the native gene in YG #81-6G01 was amplified from a Hind III site upstream to a Hpa I site downstream of the coding region and which included flanking sequences found in the GR and RD clones. The upstream primer (5'-CAA AAA GCT TGG CAT TCC GGT ACT GTT GGT AAA GCC ACC ATG GTG AAG CGA GAG-3'; SEQ ID NO:234) and a downstream primer (5'-CAATTGTTGTTGTTAACTTGT TTA TT-3'; SEQ ID NO:235) were mixed with YG#81-6G01 and amplified using the PCR conditions above. The purified PCR product was digested with Nco I and Xba I, ligated with pGL3-control that was also digested with Hind III and Hpa I, and the ligated products introduced into E. coli. To insert YG#81-6G01 into the other pGL3 reporter vectors (basic, promoter and enhancer), the pGL3-control vectors containing YG#81-6G01 were digested with Nco I and Xba I, ligated with the other pGL3 vectors that also were digested with Nco I and Xba I, and the ligated products introduced to E. coli. Note that the clone of YG#81-6G01 in the pGL3 vectors has a C instead of an A at base 786, which yields a change in the amino acid sequence at residue 262 from Phe to Leu (FIG. 2 shows the sequence of YG#81-6G01 prior to introduction into pGL3 vectors). To determine whether the altered amino acid at position 262 affected the enzyme biochemistry, the clone of YG#81-6G01 was mutated to resemble the original sequence. Both clones were then tested for expression in E. coli, physical stability, substrate binding, and luminescence output kinetics. No significant differences were found.

[0240] Partially purified enzymes expressed from the synthetic genes and the parent gene were employed to determine Km for luciferin and ATP (see Table 3).

TABLE 3

Enzyme	$K_{\mathbf{M}}\left(\mathrm{LH}_{2}\right)$	$K_{M}\left(ATP\right)$
YG parent	2 μM	17 μM
GR	1.3 μM	25 μM
RD	24.5 μM	46 μM

[0241] In vitro eukaryotic transcription/translation reactions were also conducted using Promega's TNT T7 Quick system according to manufacturer's instructions. Luminescence levels were 1 to 37-fold and 1 to 77-fold higher (depending on the reaction time) for the synthetic GR and RD genes, respectively, compared to the parent gene (corrected for luminometer spectral sensitivity).

[0242] To test whether the synthetic click beetle luciferase genes and the wild type click beetle gene have improved

expression in mammalian cells, each of the synthetic genes and the parent gene was cloned into a series of pGL3 vectors and introduced into CHO cells (Table 8). In all cases, the synthetic click beetle genes exhibited a higher expression than the native gene. Specifically, expression of the synthetic GR and RD genes was 1900-fold and 40-fold higher, respectively, than that of the parent (transfection efficiency normalized by comparison to native *Renilla* luciferase gene). Moreover, the data (basic versus control vector) show that the synthetic genes have reduced basal level transcription.

[0243] Further, in experiments with the enhancer vector where the percentage of activity in reference to the control is compared between the native and synthetic gene, the data showed that the synthetic genes have reduced risk of anomalous transcription characteristics. In particular, the parent gene appeared to contain one or more internal transcriptional regulatory sequences that are activated by the enhancer in the vector, and thus is hot suitable as a reporter gene while the synthetic GR and RD genes showed a clean reporter response (transfection efficiency normalized by comparison to native *Renilla* luciferase gene). See Table 9.

[0244] The clone names and their corresponding SEQ ID numbers for nucleotide sequence and amino acid sequence are listed below in Table 4.

TABLE 4

Clone name	Luciferase Type	SEQ ID NO.	SEQ ID NO.
LUCPPLYG	Wild type YG Click Beetle	1	23
YG#81-6G01	Mutant YG Click Beetle	2	24
GRver1	Synthetic Green Click Beetle	3	25
GRver2	Synthetic Green Click Beetle	4	26
GRver3	Synthetic Green Click Beetle	5	27
GRver4	Synthetic Green Click Beetle	6	28
GRver5	Synthetic Green Click Beetle	7	29
GR6	Synthetic Green Click Beetle	8	30
GRver5.1	Synthetic Green Click Beetle	9	31
RDver1	Synthetic Red Click Beetle	10	32
RDver2	Synthetic Red Click Beetle	11	33
RDver3	Synthetic Red Click Beetle	12	34
RDver4	Synthetic Red Click Beetle	13	218
RDver5	Synthetic Red Click Beetle	14	219
RD7	Synthetic Red Click Beetle	15	220
RDver5.1	Synthetic Red Click Beetle	16	221
RDver5.2	Synthetic Red Click Beetle	17	222
RD156-1H9	Synthetic Red Click Beetle	18	223
RELLUC	Wild type Renilla	19	224
Rlucver1	Synthetic Renilla	20	225
Rlucver2	Synthetic Renilla	21	226
Rluc-final	Synthetic Renilla	22	227

EXAMPLE 2

Evolution of the RD Luciferase Gene

[0245] RDver5.2 was mutated to increase its luminescence intensity, thereby creating RD156-1H9 which carries four additional amino acid changes (M2I, S349T, K488T, E538V) and three silent point mutations (SEQ ID NO:18).

a) Site-Directed Mutagenesis:

[0246] The initial strategy was to use site-directed mutagenesis. There are four amino acid differences between the GR and RD synthetic genes with H348Q providing the greatest contribution to red color. Thus, this substitution may also cause structural changes in the protein that could lead

to low light output. Optimization of positions near this area could increase light output. The following positions were selected for mutagenesis:

[0247] 1. S344 (at the edge of the binding pocket for luciferin)—randomize this codon.

[0248] 2. A245 (strictly conserved but closest to 348 and at the edge of the active site pocket)—randomize this codon.

[0249] 3. I347 (not conserved, next to 348 in sequence)—mutate to hydrophobic amino acids only.

[0250] 4. S349 (not conserved, next to 348 in sequence)—mutate to S, T, A, P only.

[0251] Oligonucleotides designed to mutate the above positions were used in a site-directed mutagenesis experiment (WO99/14336) and the resulting mutants were screened for luminescence intensity. There was little variation in light intensity and only about 25% were luminescent. For more detailed analysis, clones were picked and analyzed with the screening robot (PCT/WO9914336). None of the clones had a luminescence intensity (LI) higher than RDver5.2, but four of the clones had slightly lower composite Km for luciferin and ATP (Km).

b) Directed Evolution:

[0252] Protocols and procedures used for the directed evolution are detailed in see PCT/WO9914336. DNA from the four clones with lower Km was combined and three libraries of random mutants were produced. The libraries were screened with the robot and clones with the highest LI values were selected. These clones were shuffled together and another robotic screen was completed with an incubation temperature of 46° C. The three clones with the highest LI values were RD156-0B4, RD156-1A5, and RD156-1H9.

c) Analysis:

[0253] The three clones with the highest LI values were selected for manual analysis to confirm that their luminescence intensity was higher than that of RDver5.2 and to ensure that their spectral properties were not compromised. One of the clones was slightly green-shifted, all others maintained the spectral properties of RDver5.2 (Table 5).

TABLE 5

Clone	Peak (nm)	Width (nm)
RD156-0B4	616	68
RD156-1A5	614	70
RD156-1H9	618	69
RDver5.2 (prep #1)	617	70
RDver5.2 (prep #2)	618	69

[0254] The Km values for luciferin and the luminescence intensity relative to RDver5.2 were determined for all three clones in several independent experiments. All cells samples were processed with CCLR lysis buffer (E1483, Promega Corp., Madison, Wis.) and diluted 1:10 into buffer (25 mM HEPES pH 7.8, 5% glycerol, 1 mg/ml BSA, 150 mM NaCl). Table 7 summarizes the results (Lum: luminescence values were normalized to optical density; measurements for independent experiments are separated by forward slashes) from expression in bacterial cells. RD156-1H9, the clone with the

highest luminescence intensity (5 to 10-fold increase) also has an about 2-fold higher Km for luciferin.

TABLE 6

Clone	Km Luciferin [μM]	Lum (normalized to RDver5.2)
RD156-0B4	8/10	2.2/2.5
RD156-1A5	13/13	3.1/5.6
RD156-1H9	20/23/23	4/10.9/7.5
RDver5.2 (prep #1)	12/14/14	
RDver5.2 (prep #2)	40/50	
GRver5.1 (prep #1)	0.5	64
GRver5.1 (prep #2)	3	

[0255] Table 7 shows a comparison between the luminescence intensities of RD156-1H9, GRver5.1 and RDver5.2 normalized to GRver5.1 with and without correction for the spectral sensitivity of the luminometer photomultiplier tube. With correction, the luminescence intensity of clone RD156-1H9 was only about 2-fold lower than that of GRver5.1. The luciferin Km for clone RD156-1H9 is approximately 40-fold higher than GRver5.1. RD156-1H9 is thermostable at 50° C. for at least 2 hours.

TABLE 7

Name	No Correction	With Correction
RDver5.2	0.016	0.06
GRver5.1	1.000	1.00
RD156-1H9	0.116	0.45

[0256] Tables 8 and 9 show a comparison of luciferase expression levels in CHO cells. Table 8 shows the expression levels only from the control vectors in comparison to the firefly luciferase gene (RLU=relative light units). Table 9 shows a comparison of the expression levels in all four pGL3 vectors calculated as a percent of the expression level in pGL3-control.

TABLE 8

~	
Gene Expression	
rlu	
177	
343,417	
7,161	
20,802	
488,016	
	177 343,417 7,161 20,802

[0257]

TABLE 9

Synthetic Click Beet Vector	Percent of control vector
YG-control	100
RD-control	100
GR-control	100
RD156-1H9 control	100
YG-basic	3.3

TABLE 9-continued

Synthetic Click Beetle Gene Expression		
Vector	Percent of control vector	
RD-basic	1.0	
GR-basic	0.2	
RD156-1H9 basic	0.3	
YG-promoter	4.2	
RD-promoter	15.1	
GR-promoter	5.7	
RD156-1H9 promoter	15.5	
YG-enhancer	51.5	
RD-enhancer	2.8	
GR-enhancer	1.4	
RD156-1H9 enhancer	0.3	

EXAMPLE 3

Synthetic Renilla Luciferase Nucleic Acid Molecule

[0258] The synthetic *Renilla* luciferase genes prepared include 1) an introduced Kozak sequence, 2) codon usage optimized for mammalian (human) expression, 3) a reduction or elimination of unwanted restriction sites, 4) removal of prokaryotic regulatory sites (ribosome binding site and TATA box), 5) removal of splice sites and poly(A) addition sites, and 6) a reduction or elimination of mammalian transcriptional factor binding sequences.

[0259] The process of computer-assisted design of synthetic *Renilla* luciferase genes by iterative rounds of codon optimization and removal of transcription factor binding sites and other regulatory sites as well as restriction sites can be described in three steps:

[0260] 1. Using the wild type *Renilla* luciferase gene as the parent gene, codon usage was optimized, one amino acid was changed (T→A) to generate a Kozak consensus sequence, and undesired restriction sites were eliminated thereby creating synthetic gene Rlucver1.

[0261] 2. Remove prokaryotic regulatory sites, splice sites, poly(A) sites and transcription factor (TF) binding sites (first pass). Then remove newly created TF binding sites. Then remove newly created undesired restriction enzyme sites, prokaryotic regulatory sites, splice sites, and poly(A) sites without introducing new TF binding sites. This thereby created Rlucver2.

[0262] 3. Change 3 bases of Rlucver2 thereby creating Rluc-final.

[0263] 4. The actual gene was then constructed from synthetic oligonucleotides corresponding to the Rluc-final designed sequence. All mutations resulting from the assembly or PCR process were corrected. This gene is Rluc-final (SEQ ID NO:22) and encodes the amino acid sequence of SEQ ID NO:227.

Codon Selection

[0264] Starting with the *Renilla reniformis* luciferase sequence in Genbank (Accession No. M63501, SEQ ID NO: 19), codons were selected based on codon usage for optimal expression in human cells and to avoid *E. coli* low-usage codons. The best codon for expression in human cells (or the

best two codons if found at a similar frequency) was chosen for all amino acids with more than one codon (Wada et al., 1990):

Arg:	CGC	Lys:	AAG
Leu:	CTG	Asn:	AAC
Ser:	TCT/AGC	Gln:	CAG
Thr:	ACC	His:	CAC
Pro:	CCA/CCT	Glu:	GAG
Ala:	GCC	Asp:	GAC
Gly:	GGC	Tyr:	TAC
Val:	GTG	Cys:	TGC
Ile:	ATC/ATT	Phe:	TTC

[0265] In cases where two codons were selected for one amino acid, they were used in an alternating fashion. To meet other criteria for the synthetic gene, the initial optimal codon selection was modified to some extent later. For example, introduction of a Kozak sequence required the use of GCT for Ala at amino acid position 2 (see below).

[0266] The following low-usage codons in mammalian cells were not used unless needed: Arg: CGA, CGU; Leu: CTA, UUA; Ser: TCG; Pro: CCG; Val: GTA; and Ile: ATA. The following low-usage codons in *E. coli* were also avoided when reasonable (note that 3 of these match the low-usage list for mammalian cells): Arg: CGA/CGG/AGA/AGG, Leu: CTA; Pro: CCC; Ile: ATA.

Introduction of Kozak Sequences

[0267] The Kozak sequence: 5' aaccATGGCT 3' (SEQ ID NO: 293) (the Nco I site is underlined, the coding region is shown in capital letters) was introduced to the synthetic *Renilla* luciferase gene. The introduction of the Kozak sequence changes the second amino acid from Thr to Ala (GCT).

Removal of Undesired Restriction Sites

[0268] REBASE ver. 808 (updated Aug. 1, 1998; Restriction Enzyme Database; www.neb.com/rebase) was employed to identify undesirable restriction sites as described in Example 1. The following undesired restriction sites (in addition to those described in Example 1 were removed according to the process described in Example 1: EcoICR I, NdeI, NsiI, SphI, SpeI, XmaI, PstI.

[0269] The version of *Renilla* luciferase (Rluc) which incorporates all these changes is Rlucver1.

Removal of Prokarvotic ($E.\ coli$) Regulatory Sequences Splice Sites, and Poly(A) Sites

[0270] The priority and process for eliminating transcription regulation sites was as described in Example 1.

Removal of TF Binding Sites

[0271] The same process, tools, and criteria were used as described in Example 1, however, the newer version 3.3 of the TRANSFAC database was employed.

[0272] After removing prokaryotic regulatory sequences, splice sites and poly(A) sites from Rlucver1, the first search for TF binding sites identified about 60 hits. All sites were eliminated with the exception of three that could not be removed without altering the amino acid sequence of the synthetic *Renilla* gene:

[0273] 1. site at position 63 composed of two codons for W (TGGTGG), for CAC-binding protein T00076;

[0274] 2. site at position 522 composed of codons for KMV (AAN ATG GTN), for myc-DF1 T00517;

[0275] 3. site at position 885 composed of codons for EMG (GAR ATG GGN), for myc-DF1 T00517.

The subsequent second search for (newly introduced) TF binding sites yielded about 20 hits. All new sites were eliminated, leaving only the three sites described above. Finally, any newly introduced restriction sites, prokaryotic regulatory sequences, splice sites and poly(A) sites were removed without introducing new TF binding sites if possible.

[0276] Rlucver2 was obtained (SEQ ID Nos. 21 and 226).

[0277] As in Example 1, lower stringency search parameters were specified for the TESS filtered string search to further evaluate the synthetic *Renilla* gene.

[0278] With the LLH reduced from 10 to 9 and the minimum element length reduced from 5 to 4, the TESS filtered string search did not show any new hits. When, in addition to the parameter changes listed above, the organism classification was expanded from "mammalia" to "chordata", the search yielded only four more TF binding sites. When the Min LLH was further reduced to between 8 and 0, the search showed two additional 5-base sites (MAMAG and CTKTK) which combined had four matches in Rlucver2, as well as several 4-base sites. Also as in Example 1, Rlucver2 was checked for hits to entries in the EPD (Eukaryotic Promoter Database, Release 45). Three hits were determined (one to Mus musculus promoter H-2L^d (Cell, 44, 261 (1986), one to Herpes Simplex Virus type 1 promoter b'g'2.7 kb, and one to Homo sapiens DHFR promoter (J. Mol. Biol., 176, 169 (1984)). However, no further changes were made to Rlucver2.

Summary of Properties for Rlucver2

[0279] All 30 low usage codons were eliminated. The introduction of a Kozak sequence changed the second amino acid from Thr to Ala;

[0280] base composition: 55.7% GC (*Renilla* wild-type parent gene: 36.5%);

[0281] one undesired restriction site could not be eliminated: EcoR V at position 488;

[0282] the synthetic gene had no prokaryotic promoter sequence but one potentially functional ribosome binding site (RBS) at positions 867-73 (about 13 bases upstream of a Met codon) could not be eliminated;

[0283] all poly(A) addition sites were eliminated;

[0284] splice sites: 2 donor splice sites could not be eliminated (both share the amino acid sequence MGK);

[0285] TF sites: all sites with a consensus of>4 unambiguous bases were eliminated (about 280 TF binding

sites were removed) with 3 exceptions due to the preference to avoid changes to the amino acid sequence.

Synthetic *Renilla* luciferase sequences are shown in **FIGS. 7 and 8**. A codon usage comparison is shown in **FIG. 9**.

[0286] When introduced into pGL3, Rluc-final has a Kozak sequence (CACCATGGCT). The changes in Rluc-final relative to Rlucver2 were introduced during gene assembly. One change was at position 619, a C to an A, which eliminated a eukaryotic promoter sequence and reduced the stability of a hairpin structure in the corresponding oligonucleotide employed to assemble the gene. Other changes included a change from CGC to AGA at positions 218-220 (resulted in a better oligonucleotide for PCR).

Gene Assembly Strategy

[0287] The gene assembly protocol employed for the synthetic *Renilla* luciferase was similar to that described in Example 1. The oligonucleotides employed are shown in **FIG. 10**.

(SEQ ID NO:236)

Sense Strand primer:

5 ' AACCATGGCTTCCAAGGTGTACGACCCCGAGCAACGCAAA 3 '

(SEQ ID NO:237)

Anti-sense Strand primer:

5' GCTCTAGAATTACTGCTCGTTCTTCAGCACGCGCTCCACG 3'

[0288] The resulting synthetic gene fragment was cloned into a pRAM vector using Nco I and Xba I. Two clones having the correct size insert were sequenced. Four to six mutations were found in the synthetic gene from each clone. These mutations were fixed by site-directed mutagenesis (Gene Editor from Promega Corp., Madison, Wis.) and swapping the correct regions between these two genes. The corrected gene was confirmed by sequencing.

Other Vectors

[0289] To prepare an expression vector for the synthetic *Renilla* luciferase gene in a pGL-3 control vector backbone, 5 μ g of pGL3-control was digested with Nco I and Xba I in 50 μ l final volume with 2 μ l of each enzyme and 5 μ l 10× buffer B (nanopure water was used to fill the volume to 50 μ l). The digestion reaction was incubated at 37° C. for 2 hours, and the whole mixture was run on a 1% agarose gel in 1×TAE. The desired vector backbone fragment was purified using Qiagen's QIAquick gel extraction kit.

[0290] The native *Renilla* luciferase gene fragment was cloned into pGL3-control vector using two oligonucleotides, Nco I-RL-F and Xba I-RL-R, to PCR amplify native *Renilla* luciferase gene using pRL-CMV as the template. The sequence for Nco I-RL-F is 5'-CGCTAGCCATGGCTTC-GAAAGTTTATGATCC-3' (SEQ ID NO:238); the sequence for Xba I-RL-R is 5' GGCCAGTAACTCTAGAATTATTGTT-3' (SEQ ID NO:239). The PCR reaction was carried out as follows:

[0291] Reaction mixture (for 100 μl):

[0292] PCR Reaction: heat 94° C. for 2 minutes; (94° C. for 20 seconds; 65° C. for 1 minute; 72° C. for 2 minutes; then 72° C. for 5 minutes)×25 cycles, then incubate on ice. The PCR amplified fragment was cut from a gel, and the DNA purified and stored at -20° C.

[0293] To introduce native *Renilla* luciferase gene fragment into pGL3-control vector, 5 µg of the PCR product of the native *Renilla* luciferase gene (RAM-RL-synthetic) was digested with Nco I and Xba I. The desired *Renilla* luciferase gene fragment was purified and stored at -20° C.

[0294] Then 100 ng of insert and 100 ng of pGL3-control vector backbone were digested with restriction enzymes Nco I and Xba I and ligated together. Then 2 μ I of the ligation mixture was transformed into JM109 competent cells. Eight ampicillin resistance clones were picked and their DNA isolated. DNA from each positive clone of pGL3-controlnative and pGL3-control-synthetic was purified. The correct sequences for the native gene and the synthetic gene in the vectors were confirmed by DNA sequencing.

[0295] To determine whether the synthetic Renilla luciferase gene has improved expression in mammalian cells, the gene was cloned into the mammalian expression vector pGL3-control vector under the control of SV40 promoter and SV40 early enhancer (FIG. 13A). The native Renilla luciferase gene was also cloned into the pGL-3 control vector so that the expression from synthetic gene and the native gene could be compared. The expression vectors were then transfected into four common mammalian cell lines (CHO, NIH3T3, Hela and CV-1; Table 10), and the expression levels compared between the vectors with the synthetic gene versus the native gene. The amount of DNA used was at two different levels to ascertain that expression from the synthetic gene is consistently increased at different expression levels. The results show a 70-600 fold increase of expression for the synthetic Renilla luciferase gene in these cells (Table 10).

TABLE 10

Enhanced Synthetic Renilla Gene Expression			
Cell Type	Amount Vector	Fold Expression Increase	
СНО	0.2 μg	142	
	2.8 μg	145	
NIH3T3	0.2 μg	326	
	2.0 μg	593	
HeLa	0.2 μg	185	
	1.0 μg	103	
CV-1	0.2 μg	68	
	2.0 μg	72	

[0296] One important advantage of luciferase reporter is its short protein half-life. The enhanced expression could

also result from extended protein half-life and, if so, this gives an undesired disadvantage of the new gene. This possibility is ruled out by a cycloheximide chase ("CHX Chase") experiment (FIG. 14), which demonstrated that there was no increase of protein half-life resulted from the humanized *Renilla* luciferase gene.

[0297] To ensure that the increase in expression is not limited to one expression vector backbone, is promoter specific and/or cell specific, a synthetic *Renilla* gene (Rlucfinal) as well as native *Renilla* gene were cloned into different vector backbones and under different promoters (FIG. 13B). The synthetic gene always exhibited increased expression compared to its wild-type counterpart (Table 11).

TABLE 11

Renilla Gene Expression: native v. synthetic (Rluc-final)				
Vector	NIH-3T3	HeLa	СНО	
pRL-tk, native	3,834.6	922.4	7,671.9	
pRL-tk, synthetic	13,252.5	9,040.2	41,743.5	
pRL-CMV, native	168,062.2	842,482.5	153,539.5	
pRL-CMV, synthetic	2,168,129	8,440,306	2,532,576	
pRL-SV40, native	224,224.4	346,787.6	85,323.6	
pRL-SV40, synthetic	1,469,588	2,632,510	1,422,830	
pRL-null, native	2,853.8	431.7	2,434	
pRL-null, synthetic	9,151.17	2,439	28,317.1	
pRGL3b, native	12	21.8	17	
pRGL3b, synthetic	130.5	212.4	1,094.5	
pRGL3-tk, native	27.9	155.5	186.4	
pRGL3-tk, synthetic	6,778.2	8,782.5	9,685.9	
pRL-tk no intron, native	31.8	165	93.4	
pRL-tk no intron, synthetic	6,665.5	6,379	21,433.1	

[0298]

TABLE 12

Renilla Luciferase Expression in Mammalian Cells

	Percent of control vector				
Vector	CHO cells	NIH3T3 cells	HeLa cells		
pRL-control native	100	100	100		
pRL-control synthetic	100	100	100		
pRL-basic native	4.1	5.6	0.2		
pRL-basic synthetic	0.4	0.1	0.0		
pRL-promoter native	5.9	7.8	0.6		
pRL-promoter synthetic	15.0	9.9	1.1		
pRL-enhancer native	42.1	123.9	52.7		
pRL-enhancer synthetic	2.6	1.5	5.4		

(Vector Backbones Illustrated in FIG. 13A)

[0299] With reduced spurious expression the synthetic gene should exhibit less basal level transcription in a promoterless vector. The synthetic and native *Renilla* luciferase genes were cloned into the pGL3-basic vector to compare the basal level of transcription. Because the synthetic gene itself has increased expression efficiency, the activity from the promoterless vector cannot be compared directly to judge the difference in basal transcription, rather, this is taken into consideration by comparing the percentage of activity from the promoterless vector in reference to the control vector (expression from the basic vector divided by the expression in the fully functional expression vector with both promoter and enhancer elements). The data demon-

strate that the synthetic *Renilla* luciferase has a lower level of basal transcription than the native gene (Table 12)

[0300] It is well known to those skilled in the art that an enhancer can substantially stimulate promoter activity. To test whether the synthetic gene has reduced risk of inappropriate transcriptional characteristics, the native and synthetic gene were introduced into a vector with an enhancer element (pGL3-enhancer vector). Because the synthetic gene has higher expression efficiency, the activity of both cannot be compared directly to compare the level of transcription in the presence of the enhancer, however, this is taken into account by using the percentage of activity from enhancer vector in reference to the control vector (expression in the presence of enhancer divided by the expression in the fully functional expression vector with both promoter and enhancer elements). Such results show that when native gene is present, the enhancer alone is able to stimulate transcription from 42-124% of the control, however, when the native gene is replaced by the synthetic gene in the same vector, the activity only constitutes 1-5% of the value when the same enhancer and a strong SV40 promoter are employed. This clearly demonstrates that synthetic gene has reduced risk of spurious expression (Table 12).

[0301] The synthetic *Renilla* gene (Rluc-final) was used in in vitro systems to compare translation efficiency with the native gene. In a T7 quick coupled transcription/translation system (Promega Corp., Madison, Wis.), pRL-null native plasmid (having the native Renilla luciferase gene under the control of the T7 promoter) or the same amount of pRLnull-synthetic plasmid (having the synthetic Renilla luciferase gene under the control of the T7 promoter) was added to the TNT reaction mixture and luciferase activity measured every 5 minutes up to 60 minutes. Dual Luciferase assay kit (Promega Corp.) was used to measure Renilla luciferase activity. The data showed that improved expression was obtained from the synthetic gene (FIGS. 15A, B). To further evidence the increased translation efficiency of the synthetic gene, RNA was prepared by an in vitro transcription system, then purified. pRL-null (native or synthetic) vectors were linearized with BamH I. The DNA was purified by multiple phenol-chloroform extraction followed by ethanol precipitation. An in vitro T7 transcription system was employed by prepare RNAs. The DNA template was removed by using RNase-free DNase, and RNA was purified by phenol-chloroform extraction followed by multiple isopropanol precipitations. The same amount of purified RNA, either for the synthetic gene or the native gene, was then added to a rabbit reticulocyte lysate (FIGS. 15C, D) or wheat germ lysate (FIGS. 15E, F). Again, the synthetic Renilla luciferase gene RNA produced more luciferase than the native one. These data suggest that the translation efficiency is improved by the synthetic sequence. To determine why the synthetic gene was highly expressed in wheat germ, plant codon usage was determined. The lowest usage codons in higher plants coincided with those in mammals.

[0302] Reporter gene assays are widely used to study transcriptional regulation events. This is often carried out in co-transfection experiments, in which, along with the primary reporter construct containing the testing promoter, a second control reporter under a constitutive promoter is transfected into cells as an internal control to normalize experimental variations including transfection efficiencies between the samples. Control reporter signal, potential pro-

moter cross talk between the control reporter and primary reporter, as well as potential regulation of the control reporter by experimental conditions, are important aspects to consider for selecting a reliable co-reporter vector.

[0303] As described above, vector constructs were made by cloning synthetic *Renilla* luciferase gene into different vector backbones under different promoters. All the constructs showed higher expression in the three mammalian cell lines tested (Table 11). Thus, with better expression efficiency, the synthetic *Renilla* luciferase gives out higher signal when transfected into mammalian cells.

[0304] Because a higher signal is obtained, less promoter activity is required to achieve the same reporter signal, this reduced risk of promoter interference. CHO cells were transfected with 50 ng pGL3-control (firefly luc+) plus one of 5 different amounts of native pRL-TK plasmid (50, 100, 500, 1000, or 2000 ng) or synthetic pRL-TK (5, 10, 50, 100, or 200 ng). To each transfection, pUC19 carrier DNA was added to a total of 3 µg DNA. Shown in FIG. 16 is the experiment demonstrating that 10 fold less pRL-TK DNA gives similar or more signal as the native gene, with reduced risk of inhibiting expression from the primary reporter pGL3-control.

[0305] Experimental treatment sometimes may activate cryptic sites within the gene and cause induction or suppression of the co-reporter expression, which would compromise its function as co-reporter for normalization of transfection efficiencies. One example is that TPA induces expression of co-reporter vectors harboring the wild-type gene when transfecting MCF-7 cells. 500 ng pRL-TK (native), 5 µg native and synthetic pRG-B, 2.5 µg native and synthetic pRG-TK were transfected per well of MCF-7 cells. 100 ng/well pGL3-control (firefly luc+) was co-transfected with all RL plasmids. Carrier DNA, pUC19, was used to bring the total DNA transfected to 5.1 μg/well. 15.3 μl TransFast Transfection Reagent (Promega Corp., Madison, Wis.) was added per well. Sixteen hours later, cells were trypsinized, pooled and split into six wells of a 6-well dish and allowed to attach to the well for 8 hours. Three wells were then treated with the 0.2 nM of the tumor promoter, (phorbol-12-myristate-13-acetate, Calbiochem #524400-S), and three wells were mock treated with 20 μl DMSO. Cells were harvested with 0.4 ml Passive Lysis Buffer 24 hours post TPA addition. The results showed that by using the synthetic gene, undesirable change of coreporter expression by experimental stimuli can be avoided (Table 13). This demonstrates that using synthetic gene can reduce the risk of anomalous expression.

TABLE 13

TPA Induction			
Vector	Rlu	Fold Induction	
pRL-tk untreated (native)	184		
pRL-tk TPA treated (native)	812	4.4	
pRG-B untreated (native)	1		
pRG-B TPA treated (native)	8	8.0	
pRG-B untreated (final)	132		
pRG-B TPA treated (final)	195	1.47	
pRG-tk untreated (native)	44		
pRG-tk TPA treated (native)	192	4.36	
pRG-tk untreated (final)	12,816		
pRG-tk TPA treated (final)	11,347	0.88	

REFERENCES

- [0306] Altschul et al., Nucl. Acids Res., 25, 3389 (1997).
- [0307] Aota et al., Nucl. Acids Res., 16, 315 (1988).
- [0308] Boshart et al., Cell, 41, 521 (1985).
- [0309] Bronstein et al., Cal. Biochem., 219, 169 (1994).
- [0310] Corpet et al., Nucl. Acids Res., 16, 881 (1988).
- [0311] deWet et al., Mol. Cell. Biol., 7, 725 (1987).
- [0312] Dijkema et al., EMBO J., 4, 761 (1985).
- [0313] Faist and Meyer, Nucl. Acids Res., 20, 26 (1992).
- [0314] Gorman et al., *Proc. Natl. Acad. Sci. USA*, 79, 6777 (1982).
- [0315] Higgins et al., Gene, 73, 237 (1985).
- [0316] Higgins et al., CABIOS, 5, 151 (1989).
- [0317] Huang et al., CABIOS, 8, 155 (1992).
- [0318] Itolcik et al., PNAS, 94, 12410 (1997).
- [0319] Johnson et al., Mol. Reprod. Devel., 50, 377 (1998).
- [0320] Jones et al., Mol. Cell. Biol., 17, 6970 (1997).
- [0321] Karlin and Altschul, *Proc. Natl. Acad. Sci. USA*, 87, 2264 (1990).
- [0322] Karlin and Altschul, *Proc. Natl. Acad. Sci. USA*, 90, 5873 (1993).
- [0323] Keller et al., J. Cell Biol., 84, 3264 (1987).
- [0324] Kim et al., Gene, 91, 217 (1990).
- [0325] Lamb et al., Mol. Reprod. Devel., 51, 218 (1998).
- [0326] Mariatis et al., Science, 236, 1237 (1987).
- [0327] Michael et al., EMBO. J., 9, 481 (1990).
- [0328] Mizushima and Nagata, *Nucl. Acids Res.*, 18, 5322 (1990).
- [0329] Murray et al., Nucl. Acids Res., 17, 477 (1989).
- [0330] Myers and Miller, CABIOS, 4, 11 (1988).
- [0331] Needleman and Wunsen, *J. Mol. Biol.*, 48, 443 (1970).
- [0332] Pearson and Lipman, *Proc. Natl. Acad. Sci. USA*, 85, 2444 (1988).
- [0333] Pearson et al., Meth. Mol. Biol., 24, 307 (1994).
- [0334] Sharp et al., Nucl. Acids Res., 16, 8207 (1988).
- [0335] Sharp et al., Nucl. Acids Res., 15, 1281 (1987).
- [0336] Smith and Waterman, *Adv. Appl. Math.*, 2, 482 (1981).
- [0337] Stemmer et al., Gene, 164, 49 (1995).
- [0338] Uetsuki et al., J. Biol. Chem., 264, 5791 (1989).
- [0339] Voss et al., Trends Biochem. Sci., 11, 287 (1986).
- [0340] Wada et al., Nucl. Acids Res., 18, 2367 (1990).
- [0341] Watson et al, eds. *Recombinant DNA: A Short Course*, Scientific American Books, W. H. Freeman and Company, New York (1983).

[0342] Wood, K. Photochemistry and Photobiology, 62, 662 (1995).

[0343] Wood, K. Science 244, 700 (1989)

[0344] All publications, patents and patent applications are incorporated herein by reference. While in the foregoing specification, this invention has been described in relation to certain preferred embodiments thereof, and many details have been set forth for purposes of illustration, it will be apparent to those skilled in the art that the invention is susceptible to additional embodiments and that certain of the details herein may be varied considerably without departing from the basic principles of the invention.

SEQUENCE LISTING

```
<160> NUMBER OF SEQ ID NOS: 302
<210> SEQ ID NO 1
<211> LENGTH: 1629
<212> TYPE: DNA
<213> ORGANISM: Pyrophorus plagiophthalamus
<400> SEQUENCE: 1
atgatgaaga gagagaaaaa tgttatatat ggacccgaac ccctacaccc cttggaagac
ttaacagcag gagaaatgct cttcagggcc cttcgaaaac attctcattt accgcaggct
ttagtagatg tgtttggtga cgaatcgctt tcctataaag agttttttga agctacatgc
                                                                    180
ctcctagcgc aaagtctcca caattgtgga tacaagatga atgatgtagt gtcgatctgc
                                                                    240
gccgagaata ataaaagatt ttttattccc attattgcag cttggtatat tggtatgatt
                                                                    300
gtagcacctg ttaatgaaag ttacatccca gatgaactct gtaaggtcat gggtatatcg
                                                                    360
aaaccacaaa tagttttttg tacaaagaac attttaaata aggtattgga ggtacagagc
                                                                    420
agaactaatt tcataaaaag gatcatcata cttgatactg tagaaaacat acacggttgt
                                                                    480
gaaagtcttc ccaattttat ttctcgttat tcggatggaa atattgccaa cttcaaacct
                                                                   540
ttacattacg atcctgttga gcaagtggca gctatcttat gttcgtcagg cactactgga
                                                                    600
ttaccgaaag gtgtaatgca aactcaccaa aatatttgtg tccgacttat acatgcttta
                                                                    660
gaccccaggg caggaacgca acttattcct ggtgtgacag tcttagtata tctgcctttt
                                                                    720
                                                                    780
ttccatgctt ttgggttctc tataaacttg ggatacttca tggtgggtct tcgtgttatc
atgttaagac gatttgatca agaagcattt ctaaaaagcta ttcaggatta tgaagttcga
                                                                    840
agtgtaatta acgttccagc aataatattg ttcttatcga aaagtccttt ggttgacaaa
                                                                    900
960
gctgaggttg cagtaaaacg attaaacttg ccaggaattc gctgtggatt tggtttgaca
                                                                   1020
gaatctactt cagctaatat acacagtctt ggggatgaat ttaaatcagg atcacttgga
                                                                   1080
agagttactc ctttaatggc agctaaaata gcagataggg aaactggtaa agcattggga
                                                                   1140
ccaaatcaag ttggtgaatt atgcgttaaa ggtcccatgg tatcgaaagg ttacgtgaac
                                                                   1200
aatgtagaag ctaccaaaga agctattgat gatgatggtt ggcttcactc tggagacttt
                                                                   1260
ggatactatg atgaggatga gcatttctat gtggtggacc gttacaagga attgattaaa
                                                                   1320
tataagggct ctcaggtagc acctgcagaa ctagaagaga ttttattgaa aaatccatgt
                                                                  1380
atcagagatg ttgctgtggt tggtattcct gatctagaag ctggagaact gccatctgcg
                                                                   1440
tttgtggtta aacagcccgg aaaggagatt acagctaaag aagtgtacga ttatcttgcc
                                                                   1500
                                                                  1560
gagagggtct cccatacaaa gtatttgcgt ggaggggttc gattcgttga tagcatacca
                                                                   1620
```

aggaatgtta caggtaaaat tacaagaaag gaacttctga agcagttgct ggagaagagt

tctaaactt	1629
<210> SEQ ID NO 2 <211> LENGTH: 1626 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of clone YG#81-6G01	
<400> SEQUENCE: 2	
atgatgaagc gagagaaaaa tgttatatat ggacccgaac ccctacaccc cttggaagac	60
ttaacagctg gagaaatgct cttccgtgcc cttcgaaaac attctcattt accgcaggct	120
ttagtagatg tggttggcga cgaatcgctt tcctataaag agttttttga agcgacagtc	180
ctcctagcgc aaagtctcca caattgtgga tacaagatga atgatgtagt gtcgatctgc	240
gccgagaata atacaagatt ttttattccc gttattgcag cttggtatat tggtatgatt	300
gtagcacctg ttaatgaaag ttacatccca gatgaactct gtaaggtgat gggtatatcg	360
aaaccacaaa tagtttttac gacaaagaac attttaaata aggtattgga ggtacagagc	420
agaactaatt tcataaaaag gatcatcata cttgatactg tagaaaacat acacggttgt	480
gaaagtette ecaattttat ttetegttat teggatggaa atattgeeaa etteaaacet	540
ttacatttcg atcctgttga gcaagtggca gctatcttat gttcgtcagg cactactgga	600
ttaccgaaag gtgtaatgca aactcaccaa aatatttgtg tccgacttat acatgcttta	660
gaccccaggg caggaacgca acttattcct ggtgtgacag tcttagtata tctgcctttt	720
ttccatgctt ttgggttctc tataaccttg ggatacttca tggtgggtct tcgtgttatc	780
atgttcagac gatttgatca agaagcattt ctaaaagcta ttcaggatta tgaagttcga	840
agtgtaatta acgttccatc agtaatattg ttcttatcga aaagtccttt ggttgacaaa	900
tacgatttat caagtttaag ggaattgtgt tgcggtgcgg	960
gctgaggttg cagcaaaacg attaaacttg ccaggaattc gctgtggatt tggtttgaca	1020
gaatctactt cagctaatat acacagtctt agggatgaat ttaaatcagg atcacttgga	1080
agagttactc ctttaatggc agctaaaata gcagataggg aaactggtaa agcattggga	1140
ccaaatcaag ttggtgaatt atgcattaaa ggtcccatgg tatcgaaagg ttacgtgaac	1200
aatgtagaag ctaccaaaga agctattgat gatgatggtt ggcttcactc tggagacttt	1260
ggatactatg atgaggatga gcatttctat gtggtggacc gttacaagga attgattaaa	1320
tataagggct ctcaggtagc acctgcagaa ctagaagaga ttttattgaa aaatccatgt	1380
atcagagatg ttgctgtggt tggtattcct gatctagaag ctggagaact gccatctgcg	1440
tttgtggtta aacagcccgg aaaggagatt acagctaaag aagtgtacga ttatcttgcc	1500
gagagggtot cccatacaaa gtatttgcgt ggaggggttc gattcgttga tagcatacca	1560
aggaatgtta caggtaaaat tacaagaaag gaacttctga agcagttgct ggagaaggcg	1620
ggaggt	1626

<210> SEQ ID NO 3 <211> LENGTH: 1626 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence

<223> OTHER INFORMATION: Sequence of a synthetic luciferase

		-
<400> SEQUENCE: 3		
atgatgaaac gcgaaaagaa cgtcatctac ggcccagagc ctctgcaccc attggaagac	60	
ctgaccgccg gtgagatgtt gttccgtgct ctgcgtaaac attctcactt gcctcaagcc	120	
ctggtggatg tcgtgggcga cgaaagcttg tcttataagg agtttttcga agctactgtc	180	
ctgttggccc agtctctgca taattgcggt tacaaaatga acgatgtggt cagcatttgt	240	
gctgagaata acacccgctt tttcatccca gtgattgccg cttggtacat cggcatgatt	300	
gtcgcccctg tgaatgaatc ttatatccca gacgagttgt gcaaggtcat gggtattagc	360	
aaacctcaaa tcgtgtttac taccaagaac attctgaata aagtcttgga agtgcagtct	420	
cgtactaact tcatcaagcg cattatcatt ctggataccg tcgagaatat ccacggctgt	480	
gaaagcttgc caaactttat ttctcgttat agcgacggta atatcgctaa cttcaagcct	540	
ctgcattttg atccagtgga gcaagtcgcc gctattttgt gctctagcgg cactaccggt	600	
ctgcctaaag gcgtgatgca gactcaccaa aatatctgtg tccgcttgat tcatgccctg	660	
gacccacgtg tgggtaccca gttgatccct ggcgtgactg tcctggtgta cttgccattc	720	
tttcacgcct tcggtttttc tattaccctg ggctatttca tggtcggttt gcgcgtgatc	780	
atgtttcgtc gcttcgatca agaagctttt ctgaaggcca ttcaggacta cgaggtccgt	840	
agcgtgatca acgtcccttc tgtgattttg ttcctgagca aatctccatt ggtcgataag	900	
tatgacctga gctctttgcg cgaactgtgc tgtggcgctg cccctttggc taaagaggtg	960	
gccgaagtcg ctgccaagcg tctgaatttg ccaggtatcc gctgcggctt tggtctgact	1020	
gagagcacct ctgctaacat tcatagcttg cgtgatgaat tcaaatctgg cagcctgggt	1080	
cgcgtgactc ctttgatggc cgctaagatc gccgaccgtg agaccggcaa agctctgggt	1140	
ccaaatcaag tcggcgaatt gtgtattaag ggtcctatgg tgtctaaagg ctacgtcaac	1200	
aatgtggagg ccactaagga agctatcgat gacgatggtt ggctgcacag cggcgacttt	1260	
ggttattacg atgaggacga acatttctat gtcgtggatc gctacaaaga gttgattaag	1320	
tataaaggct ctcaggtcgc cccagctgag ctggaagaga tcttgctgaa gaacccttgc	1380	
attcgtgacg tggccgtcgt gggtatccca gatttggaag ctggcgagct gcctagcgcc	1440	
tttgtcgtga aacaaccagg taaggaaatt accgctaaag aggtctacga ctatttggcc	1500	
gaacgcgtgt ctcacactaa gtacctgcgt ggcggtgtcc gcttcgtgga tagcatccct	1560	
cgcaatgtca ccggcaaaat tactcgtaag gagttgctga aacagttgct ggaaaaggct	1620	
ggtggc	1626	
<210> SEQ ID NO 4 <211> LENGTH: 1626 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase		
<400> SEQUENCE: 4	<i>-</i> -	
atgatgaaac gcgaaaagaa cgtcatctac ggcccagagc ctctgcaccc attggaagac	60	
ctgaccgctg gtgagatgtt gttccgtgct ctgcgtaaac attctcactt gcctcaagcc	120	
ctggtcgatg tcgtgggcga cgagagcttg tcttataagg aatttttcga agctactgtc	180	

ctgttggccc aatctctgca taattgcggt tacaaaatga acgatgtggt cagcatttgt

-continued	
gctgagaata acacccgctt tttcatccca gtgattgccg cttggtacat cggcatgatt	300
gtcgcccctg tgaatgaatc ttatatccca gacgagttgt gcaaggtcat gggtattagc	360
aaacctcaaa tcgtgtttac taccaagaac attctgaata aggtcttgga agtgcagtct	420
cgtactaact tcatcaagcg cattatcatt ctggataccg tcgagaatat ccacggctgt	480
gagagettge caaactttat ttetegttat agegaeggta atategetaa etteaageet	540
ctgcattttg atccagtgga gcaagtcgcc gctattttgt gctctagcgg caccaccggt	600
ctgcctaaag gcgtgatgca gactcaccaa aatatctgtg tccgcttgat tcatgccctg	660
gacccacgtg tgggtactca gttgatccct ggcgtgactg tcctggtgta cttgccattc	720
tttcacgcct tcggtttttc tattaccctg ggctatttca tggtcggttt gcgcgtgatc	780
atgtttcgtc gcttcgatca agaagccttt ctgaaggcca ttcaagacta cgaggtccgt	840
agogtgatca acgtcccttc tgtgattttg ttcctgagca aatctccatt ggtcgataag	900
tatgacctga gcagcttgcg cgaactgtgc tgtggcgctg cccctttggc taaagaggtg	960
gccgaagtcg ctgccaagcg tctgaatttg ccaggtatcc gctgcggctt tggtctgact	1020
gagagcacct ctgctaacat tcatagcttg cgtgatgagt tcaaatctgg cagcctgggt	1080
cgcgtgactc ctttgatggc cgctaagatc gccgaccgtg agaccggcaa agctctgggt	1140
ccaaatcaag tcggcgaatt gtgtattaag ggtcctatgg tgtctaaagg ctacgtcaac	1200
aatgtggagg ccactaagga agctattgat gacgatggtt ggctgcacag cggcgacttt	1260
ggttattacg atgaggacga acatttctat gtcgtcgatc gctacaaaga gttgattaag	1320
tataaaggct ctcaagtcgc cccagctgag ctggaagaaa tcttgctgaa gaacccttgc	1380
attogtgacg tggccgtcgt gggtatccca gatttggaag ctggcgagct gcctagcgcc	1440
tttgtcgtga aacaaccagg caaggaaatt accgctaaag aggtctacga ctatttggcc	1500
gagogogtgt ctcacactaa gtacctgcgt ggcggtgtcc gcttcgtcga tagcatccct	1560
cgcaatgtca ccggcaaaat tactcgtaag gagttgctga aacagttgct ggaaaaggct	1620
ggtggc	1626
<210> SEQ ID NO 5 <211> LENGTH: 1626 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase	
<400> SEQUENCE: 5	
atgatgaaac gcgaaaagaa cgtgatctac ggcccagaac cactgcatcc actggaagac	60
ctcaccgctg gtgagatgct gttccgtgcc ctgcgtaaac atagccacct gcctcaagct	120
ctcgtggacg tcgtgggtga cgagagcctg tcttacaaag aatttttcga agctactgtg	180
ctgttggccc aaagcctgca taattgtggt tacaaaatga acgatgtggt gagcatctgt	240
gctgagaata acactcgctt ttttatccct gtgatcgctg cttggtacat cggcatgatt	300
gtcgcccctg tgaatgaatc ttacatccca gatgagttgt gtaaggtgat gggtattagc	360
aaacctcaaa tcgtctttac taccaaaaac atcctgaata aggtcttgga agtccagtct	420

cgtactaatt tcatcaaacg cattattatt ctggataccg tcgaaaacat ccacggctgt gagagcttgc ctaactttat ctctcgttac agcgatggta atatcgctaa tttcaagcca

-continued	
ctgcattttg atccagtcga gcaggtcgcc gccattttgt gctcttctgg caccactggt	600
ttgcctaaag gtgtcatgca gactcaccag aatatctgtg tgcgcttgat ccacgccctc	660
gaccetegtg tgggtactea attgatecet ggegtgactg tgetggtgta tttgeettte	720
tttcacgcct ttggtttttc tatcaccctg ggctatttca tggtcggctt gcgtgtgatc	780
atgtttcgtc gcttcgacca agaagccttc ctgaaggcta ttcaagacta cgaggtgcgt	840
tctgtgatca atgtcccatc tgtcattttg ttcctgagca aatctccttt ggttgacaag	900
tatgatctga gcagcttgcg tgaactgtgc tgtggcgctg ctcctttggc caaagaagtg	960
geogaggteg etgetaageg tetgaacete eetggtatee getgeggttt tggtttgaet	1020
gagageactt etgecaacat ecatagettg egtgaegagt ttaaatetgg tageetgggt	1080
cgcgtgaccc ctttgatggc tgcaaagatc gccgaccgtg agaccggcaa agccctgggc	1140
ccaaatcagg tcggtgaatt gtgcattaag ggccctatgg tctctaaagg ctacgtgaac	1200
aatgtggagg ccactaaaga agctattgat gatgatggtt ggttgcatag cggcgacttc	1260
ggttattatg atgaggacga acacttctat gtggtcgatc gctataaaga attgattaag	1320
tacaaagget etcaagtege eecagetgaa etggaagaaa ttttgetgaa gaaceettgt	1380
attcgcgacg tggccgtcgt gggtatccca gacttggaag ctggcgagtt gcctagcgcc	1440
tttgtggtga aacaacctgg caaggagatt actgctaagg aggtctacga ctatttggcc	1500
gagegegtgt eteacactaa atatetgegt ggeggegtee gettegtega ttetateeet	1560
cgcaacgtca ccggcaagat cactcgtaaa gagttgctga aacaattgct cgaaaaagct	1620
ggcggc	1626
<210> SEQ ID NO 6 <211> LENGTH: 1626 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase	
<400> SEQUENCE: 6	
atgatgaaac gcgaaaagaa cgtgatctac ggcccagaac cactgcatcc actggaagac	60
ctcaccgctg gtgagatgct cttccgtgca ctgcgtaaac atagtcacct ccctcaagct	120
ctcgtggacg tcgtgggaga cgagagcctc tcttacaaag aatttttcga agctactgtg	180
ctgttggccc aaagcctcca taattgtgga tacaaaatga acgatgtggt gagcatttgt	240
gctgagaata acactcgctt ctttatccct gttatcgctg cttggtacat cggcatgatt	300
gtcgcccctg tgaatgaatc ttacatccca gatgagctgt gtaaggttat gggtattagc	360
aaacctcaaa tcgtctttac taccaaaaat atcctgaata aggtcttgga agtccagtct	420
cgtactaact tcatcaaacg catcattatt ctggataccg tcgaaaacat ccatggctgt	480
gagagcctgc ctaacttcat ctctcgttac agcgatggta atatcgctaa tttcaaacca	540
ctgcattttg atccagtcga gcaagtggcc gctattttgt gctcttccgg caccactggt	600
ttgcctaaag gtgtcatgca gactcaccag aatatctgtg tgcgtttgat ccacgctctc	660
gaccotcgtg tgggtactca attgatccct ggcgtgactg tgctggtgta tctgcctttc	720

tttcacgcct ttggtttttc tattaccctg ggctatttca tggtcggctt gcgtgtcatc atgttcgtc gcttcgacca agaagccttc ttgaaggcta ttcaagacta cgaggtgcgt

-continued	
totgtcatca atgtcccttc agtcattttg ttcctgagca aatctccttt ggttgacaag	900
tatgatctga gcagcttgcg tgagctgtgc tgtggcgctg ctcctttggc caaagaagtg	960
gccgaggtcg ctgctaagcg tctgaacctc cctggtatcc gctgcggttt tggtttgact	1020
gagageactt ctgctaacat ccatagettg cgagacgagt ttaagtetgg tageetgggt	1080
cgcgtgactc ctcttatggc tgcaaagatc gccgaccgtg agaccggcaa agcactgggc	1140
ccaaatcaag tcggtgaatt gtgtattaag ggccctatgg tctctaaagg ctacgtgaac	1200
aatgtggagg ccactaaaga agccattgat gatgatggct ggctccatag cggcgacttc	1260
ggttactatg atgaggacga acacttctat gtggtcgatc gctacaaaga attgattaag	1320
tacaaaggct ctcaagtcgc cccagccgaa ctggaagaaa ttttgctgaa gaacccttgt	1380
atccgcgacg tggccgtcgt gggtatccca gacttggaag ctggtgagtt gcctagcgcc	1440
tttgtggtga aacaacctgg aaaggagatc actgctaagg aggtctacga ctatttggcc	1500
gagcgcgtgt ctcacaccaa atatctgcgt ggcggcgtcc gcttcgtcga ttccatccca	1560
cgcaacgtga ccggtaagat cactcgtaaa gaattgctga agcaactcct cgaaaaagct	1620
adcadc	1626
<210> SEQ ID NO 7 <211> LENGTH: 1626 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase	
<400> SEQUENCE: 7	
atgatgaaac gcgaaaagaa cgtgatctac ggcccagaac cactgcatcc actggaagac	60
ctcaccgctg gtgagatgct cttccgagca ctgcgtaaac atagtcacct ccctcaagca	120
ctcgtggacg tcgtgggaga cgagagcctc tcctacaaag aatttttcga agctactgtg	180
ctgttggccc aaagcctcca taattgtggg tacaaaatga acgatgtggt gagcatttgt	240
gctgagaata acactcgctt ctttattcct gtaatcgctg cttggtacat cggcatgatt	300
gtcgcccctg tgaatgaatc ttacatccca gatgagctgt gtaaggttat gggtattagc	360
aaacctcaaa tcgtctttac taccaaaaac atcttgaata aggtcttgga agtccagtct	420
cgtactaact tcatcaaacg catcattatt ctggataccg tcgaaaacat ccacggctgt	480
gagageetee etaaetteat etetegttae agegatggta atategetaa ttteaageee	540
ttgcattttg atccagtcga gcaagtggcc gctattttgt gctcctccgg caccactggt	600
ttgcctaaag gtgtcatgca gactcaccag aatatctgtg tgcgtttgat ccacgctctc	660
gaccetegtg tgggtaetea attgateeet ggegtgaetg tgetggtgta tetgeettte	720
tttcacgcct ttggtttctc tattaccctg ggctatttca tggtcggctt gcgtgtcatc	780
atgtttcgtc gcttcgacca agaagccttc ttgaaggcta ttcaagacta cgaggtgcgt	840
tccgtgatca acgtcccttc agtcattttg ttcctgagca aatctccttt ggttgacaag	900
tatgatctga gcagcttgcg tgagctgtgc tgtggcgctg ctcctttggc caaagaagtg	960
gccgaggtcg ctgctaagcg tctgaacctc cctggtatcc gctgcggttt tggtttgact	1000
	1020

cgcgtgactc ctcttatggc tgcaaagatc gccgaccgtg agaccggcaa agcactgggc 1140

				-0011011	iueu	
ccaaatcaag	tcggtgaatt	gtgtattaag	ggccctatgg	tctctaaagg	ctacgtgaac	1200
aatgtggagg	ccactaaaga	agccattgat	gatgatggct	ggctccatag	cggcgacttc	1260
ggttactatg	atgaggacga	acacttctat	gtggtcgatc	gctacaaaga	attgattaag	1320
tacaaaggct	ctcaagtcgc	accagccgaa	ctggaagaaa	ttttgctgaa	gaacccttgt	1380
atccgcgacg	tggccgtcgt	gggtatccca	gacttggaag	ctggcgagtt	gcctagcgcc	1440
tttgtggtga	aacaacccgg	caaggagatc	actgctaagg	aggtctacga	ctatttggcc	1500
gagcgcgtgt	ctcacaccaa	atatctgcgt	ggcggcgtcc	gcttcgtcga	ttctattcca	1560
cgcaacgtta	ccggtaagat	cactcgtaaa	gagttgctga	agcaactcct	cgaaaaagct	1620
ggcggc						1626
<220> FEATU	TH: 1626 : DNA NISM: Artific	_		netic lucife	erase	
<400> SEQUE	ENCE: 8					
atgatgaaac	gcgaaaagaa	cgtgatctac	ggcccagaac	cactgcatcc	actggaagac	60
ctcaccgctg	gtgagatgct	cttccgagca	ctgcgtaaac	atagtcacct	ccctcaagca	120
ctcgtggacg	tcgtgggaga	cgagaacctc	tcctacaaag	aatttttcga	agctactgtg	180
ctgttggccc	aaagcctcca	taattgtggg	tacaaaatga	acgatgtggt	gagcatttgt	240
gctgagaata	acactcgctt	ctttattcct	gtaatcgctg	cttggtacat	cggcatgatt	300
gtcgcccctg	tgaatgaatc	ttacatccca	gatgagctgt	gtaaggttat	gggtattagc	360
aaacctcaaa	tcgtctttac	taccaaaaac	atcttgaata	aggtcttgga	agtccagtct	420
cgtactaact	tcatcaaacg	catcattatt	ctggataccg	tcgaaaacat	ccacggctgt	480
gagageetee	ctaacttcat	ctctcgttac	agcgatggta	atatcgctaa	tttcaagccc	540
ttgcattttg	atccagtcga	gcaagtggcc	gctattttgt	gctcctccgg	caccactggt	600
ttgcctaaag	gtgtcatgca	gactcaccag	aatatctgtg	tgcgtttgat	ccacgctctc	660
gaccctcgtg	tgggtactca	attgatctct	ggcgtgactg	tgctggtgta	tctgcctttc	720
tttcacgcct	ttggtttctc	tattaccctg	ggctatttca	tggtcggctt	gcgtgtcatc	780
atgtttcgtc	gcttcgacca	agaagccttc	ttgaaggcta	ttcaagacta	cgaggtgcgt	840
tccgtgatca	acgtcccttc	agtcattttg	ttcctgagca	aatctccttt	ggttgacaag	900
tatgatctga	gcagcttgcg	tgagctgtgc	tgtggcgctg	ctcctttggc	caaagaagtg	960
gccgaggtcg	ctgctaagcg	tctgaacctc	cctggtatcc	gctgcggttt	tggtttgact	1020
gagagcactt	ctgctaacat	ccatagcttg	cgagacgagt	ttaagtctgg	tagcctgggt	1080
cgcgtgactc	ctcttatggc	tgcaaagatc	gccgaccgtg	agaccggcaa	agcactgggc	1140
ccaaatcaag	tcggtgaatt	gtgtattaag	ggccctatgg	tctctaaagg	ctacgtgaac	1200
aatgtggagg	ccactaaaga	agccattgat	gatgatggct	ggctccatag	cggcgacttc	1260
ggttactatg	atgaggacga	acacttctat	gtggtcgatc	gctacaaaga	attgattaag	1320
tacaaaggct	ctcaagtcgc	accagccgaa	ctggaagaaa	ttttgctgaa	gaacccttgt	1380

atccgcgacg tggccgtcgt gggtatccca gacttggaag ctggcgagtt gcctagcgcc 1440

tttgtggtga aacaacccgg caaggagatc actgctaagg aggtctacga ctatttggcc	1500
gagcgcgtgt ctcacaccaa atatctgcgt ggcggcgtcc gcttcgtcga ttctattcca	1560
cgcaacgtta ccggtaagat cactcgtaaa gagttgctga agcaactcct cgaaaaagct	1620
ggcggc	1626
<210> SEQ ID NO 9 <211> LENGTH: 1626 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase	
<400> SEQUENCE: 9	
atgatgaaac gcgaaaagaa cgtgatctac ggcccagaac cactgcatcc actggaagac	60
ctcaccgctg gtgagatgct cttccgagca ctgcgtaaac atagtcacct ccctcaagca	120
ctcgtggacg tcgtgggaga cgagagcctc tcctacaaag aatttttcga agctactgtg	180
ctgttggccc aaagcctcca taattgtggg tacaaaatga acgatgtggt gagcatttgt	240
gctgagaata acactcgctt ctttattcct gtaatcgctg cttggtacat cggcatgatt	300
gtcgcccctg tgaatgaatc ttacatccca gatgagctgt gtaaggttat gggtattagc	360
aaacctcaaa tcgtctttac taccaaaaac atcttgaata aggtcttgga agtccagtct	420
cgtactaact tcatcaaacg catcattatt ctggataccg tcgaaaacat ccacggctgt	480
gagageetee etaaetteat etetegttae agegatggta atategetaa ttteaageee	540
ttgcattttg atccagtcga gcaagtggcc gctattttgt gctcctccgg caccactggt	600
ttgcctaaag gtgtcatgca gactcaccag aatatctgtg tgcgtttgat ccacgctctc	660
gaccctcgtg tgggtactca attgatccct ggcgtgactg tgctggtgta tctgcctttc	720
tttcacgcct ttggtttctc tattaccctg ggctatttca tggtcggctt gcgtgtcatc	780
atgtttcgtc gcttcgacca agaagccttc ttgaaggcta ttcaagacta cgaggtgcgt	840
teegtgatea aegteeette agteattttg tteetgagea aateteettt ggttgacaag	900
tatgatctga gcagcttgcg tgagctgtgc tgtggcgctg ctcctttggc caaagaagtg	960
gccgaggtcg ctgctaagcg tctgaacctc cctggtatcc gctgcggttt tggtttgact	1020
gagagcactt ctgctaacat ccatagcttg cgagacgagt ttaagtctgg tagcctgggt	1080
cgcgtgactc ctcttatggc tgcaaagatc gccgaccgtg agaccggcaa agcactgggc	1140
ccaaatcaag tcggtgaatt gtgtattaag ggccctatgg tctctaaagg ctacgtgaac	1200
aatgtggagg ccactaaaga agccattgat gatgatggct ggctccatag cggcgacttc	1260
ggttactatg atgaggacga acacttctat gtggtcgatc gctacaaaga attgattaag	1320
tacaaaggct ctcaagtcgc accagccgaa ctggaagaaa ttttgctgaa gaacccttgt	1380
atccgcgacg tggccgtcgt gggtatccca gacttggaag ctggcgagtt gcctagcgcc	1440
tttgtggtga aacaacccgg caaggagatc actgctaagg aggtctacga ctatttggcc	1500
gagcgcgtgt ctcacaccaa atatctgcgt ggcggcgtcc gcttcgtcga ttctattcca	1560
cgcaacgtta ccggtaagat cactcgtaaa gagttgctga agcaactcct cgaaaaagct	1620
ggcggc	1626

<211> LENGTH: 1626

-continued

```
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Sequence of a synthetic luciferase
<400> SEQUENCE: 10
atgatgaagc gtgagaaaaa tgtgatttat ggtcctgaac cattgcatcc tctggaggat
                                                                      60
ttgactgctg gcgaaatgct gtttcgcgcc ttgcgcaagc acagccatct gccacaggct
                                                                     120
ttggtcgacg tggtcggtga tgagtctctg agctacaaag aattctttga ggccaccgtg
                                                                     180
ttgctggctc aaagcttgca caactgtggc tataagatga atgacgtcgt gtctatctgc
                                                                     240
gccgaaaaca atactcgttt ctttattcct gtcatcgctg cctggtatat tggtatgatc
                                                                     300
gtggctccag tcaacgagag ctacattcct gatgaactgt gtaaagtgat gggcatctct
                                                                     360
aagccacaga ttgtcttcac cactaaaaat atcttgaaca aggtgctgga ggtccaaagc
                                                                      420
cgcaccaatt ttattaaacg tatcattatc ttggacactg tggaaaacat tcatggttgc
                                                                      480
gagtctctgc ctaatttcat cagccgctac tctgatggca acattgccaa ttttaaacca
                                                                     540
ttgcacttcg accctgtcga acaggtggct gccatcctgt gtagctctgg taccactggc
                                                                      600
ttgccaaagg gtgtcatgca aacccatcag aacatttgcg tgcgtctgat ccacgctctc
gatcctcgct acggcactca actgattcca ggtgtcaccg tgttggtcta tctgcctttt
                                                                     720
ttccatgctt ttggcttcca catcactttg ggttacttta tggtgggcct gcgtgtcatt
atgttccgcc gttttgacca ggaggccttc ttgaaagcta tccaagatta tgaagtgcgc
                                                                     900
tctqtcatta atqtqccaaq cqtcatcctq tttttqtcta aqaqccctct qqtqqacaaa
tacgatttgt ctagcctgcg tgagttgtgt tgcggtgccg ctccactggc caaggaagtc
                                                                     960
gctgaggtgg ccgctaaacg cttgaacctg cctggcattc gttgtggttt cggcttgacc
                                                                    1020
                                                                    1080
gaatctacta gcgccattat ccaatctctg cgcgacgagt ttaagagcgg ttctttgggc
cgtgtcaccc cactgatggc tgccaaaatt gctgatcgcg aaactggtaa ggccttgggc
                                                                    1140
cctaaccagg tgggtgagct gtgcatcaaa ggcccaatgg tcagcaaggg ttatgtgaat
                                                                    1200
                                                                    1260
aacgtcgaag ctaccaaaga ggccattgac gatgacggct ggttgcattc tggtgatttc
ggctactatg acgaagatga gcacttttac gtggtcgacc gttataagga actgatcaaa
                                                                    1320
tacaagggta gccaagtggc tcctgccgaa ttggaggaaa ttctgttgaa aaatccatgt
                                                                    1380
atcogcgatg togctgtggt cggcattcct gacctggagg ccggtgaatt gccatctgct
                                                                    1440
ttcgtggtca agcagcctgg caaagagatc actgccaagg aagtgtatga ttacctggct
                                                                    1500
gagcgtgtca gccataccaa atatttgcgc ggtggcgtgc gttttgtcga ctctattcca
                                                                    1560
cgtaacgtga ctggtaagat cacccgcaaa gaactgttga agcaactgtt ggagaaagcc
                                                                    1620
ggcggt
                                                                    1626
<210> SEQ ID NO 11
<211> LENGTH: 1626
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<223> OTHER INFORMATION: Sequence of a synthetic luciferase
```

atgatgaagc gtgagaaaaa tgtgatttat ggtcctgaac cattgcatcc tctggaggat

ttgactgccg gcgaaatgct gtttcgcgcc ttgcgcaagc acagccatct gccacaagct 12	0
ttggtggacg tggtcggtga tgaatctctg agctacaaag agttctttga ggcaaccgtg 18	0
ttgctggctc agagcttgca caactgtggc tataagatga atgacgtcgt gtctatctgc 24	0
gccgaaaaca atactcgttt ctttattcct gtcatcgctg cctggtatat tggtatgatc 30	0
gtggctccag tcaacgagag ctacattcct gatgaactgt gtaaagtgat gggcatctct 36	0
aagccacaga ttgtcttcac cactaaaaat atcttgaaca aagtgctgga ggtccaaagc 42	0
cgcaccaatt ttattaaacg tatcattatc ttggacactg tggaaaacat tcatggttgc 48	0
gaatctctgc ctaatttcat cagccgctac tctgatggca acattgccaa ttttaaacca 54	0
ttgcacttcg accetgtcga acaggtggct gccatcetgt gtagetctgg tactaetgge 60	0
ttgccaaagg gtgtcatgca aacccatcag aacatttgcg tgcgtctgat ccacgctctc 66	0
gatecteget aeggeaceca aetgatteet ggtgteaeeg tgttggteta tetgeetttt 72	0
ttccatgctt ttggcttcca catcactttg ggttacttta tggtgggcct gcgtgtcatt 78	0
atgttccgcc gttttgacca ggaggctttc ttgaaagcta tccaagatta tgaagtgcgc 84	0
tctgtcatta atgtgccaag cgtcatcctg tttttgtcta agagccctct ggtggacaaa 90	0
tacgatttgt cttctctgcg tgagttgtgt tgcggtgccg ctccactggc caaggaagtc 96	0
gctgaggtgg ccgctaaacg cttgaacctg cctggcattc gttgtggttt cggcttgacc 102	0
gaatctacta gcgccattat ccaatctctg cgcgacgaat ttaagagcgg ttctttgggc 108	0
cgtgtcaccc cactgatggc tgccaaaatt gctgatcgcg aaactggtaa ggccttgggc 114	0
cctaaccagg tgggtgagct gtgcatcaaa ggcccaatgg tcagcaaggg ttatgtgaat 120	0
aacgtcgaag ctaccaaaga ggccatcgac gatgacggct ggttgcattc tggtgatttc 126	0
ggctactatg acgaagatga gcacttttac gtggtggacc gttataagga actgatcaaa 132	0
tacaagggta gccaagtggc tcctgccgaa ttggaggaga ttctgttgaa aaatccatgt 138	0
atccgcgatg tcgctgtggt cggcattcct gacctggagg ccggtgaatt gccatctgct 144	0
ttcgtggtca agcagcctgg taaagagatc actgccaagg aagtgtatga ttacctggct 150	0
gaacgtgtca gccataccaa atatttgcgc ggtggcgtgc gttttgtgga ctctattcca 156	0
cgtaacgtga ctggtaagat cacccgcaaa gaactgttga agcaactgtt ggagaaagcc 162	0
ggcggt 162	6
<210> SEQ ID NO 12 <211> LENGTH: 1626 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase	
<400> SEQUENCE: 12	
atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctttgcaccc tttggaggat 6	0
ttgactgccg gcgaaatgct gtttcgcgct ttgcgtaagc actctcattt gcctcaagcc 12	0
ttggtcgatg tggtcggcga tgaatctttg agctataagg agttttttga ggcaaccgtc 18	0
ttgctggctc agtctttgca taattgcggc tacaagatga acgacgtcgt ctctatttgt 24	0

gccgaaaaca atacccgttt cttcattcca gtcatcgccg cctggtatat cggtatgatc gtggctccag tcaacgagag ctacattcct gacgaactgt gtaaagtcat gggtatctct

-continued	
aagccacaga ttgtgttcac cactaagaat attttgaaca aagtgctgga agtccaaagc	420
cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcatggttgc	480
gaatctctgc ctaatttcat tagccgctat tctgacggca acatcgccaa ctttaaacct	540
ttgcatttcg accetgtgga acaagtgget getateetgt gtageagegg tactaetgge	600
ctcccaaagg gcgtcatgca gacccatcaa aacatttgcg tgcgtctgat ccatgctctc	660
gatccacgct acggcactca gctgattcct ggtgtcaccg tcttggtcta cctgcctttc	720
ttccatgctt tcggcttcca cattactttg ggttacttta tggtcggtct gcgtgtcatt	780
atgttccgcc gttttgatca ggaggctttt ttgaaagcca tccaagatta tgaagtccgc	840
agcgtcatta acgtgcctag cgtgatcctg tttttgtcta agagcccact cgtggacaag	900
tacgacttgt cttccctgcg tgagttgtgt tgcggtgccg ccccactggc taaggaggtc	960
gctgaagtgg ccgccaaacg cttgaatctg ccaggcattc gttgtggctt cggcctcacc	1020
gaatctacca gcgctattat tcaatctctc cgcgatgagt ttaagagcgg ctctttgggc	1080
cgtgtcactc cactcatggc tgctaaaatc gctgatcgcg aaactggtaa ggctttgggc	1140
cctaaccaag tgggcgagct gtgtatcaaa ggccctatgg tgagcaaggg ttatgtcaat	1200
aacgtcgaag ctaccaagga ggccatcgac gacgacggct ggctgcattc tggtgatttt	1260
ggctactacg acgaagatga gcatttttac gtcgtggatc gttacaagga gctgatcaaa	1320
tacaagggta gccaggtggc tccagccgag ttggaggaga ttctgttgaa aaatccatgc	1380
atccgtgatg tcgctgtggt cggcattcct gatctggagg ccggtgaact gccttctgct	1440
ttcgtcgtca agcagcctgg taaagaaatc accgccaaag aagtgtatga ttacctggct	1500
gaacgtgtga gccataccaa gtacttgcgt ggcggcgtgc gttttgtgga cagcattcca	1560
cgtaatgtga ctggtaaaat tacccgcaag gaactgttga agcaattgtt ggagaaggcc	1620
ggcggt	1626
<210> SEQ ID NO 13 <211> LENGTH: 1626 <212> TYPE: DNA 213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase	
<400> SEQUENCE: 13	60
atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctttgcatcc tttggaggat	60
ttgactgccg gcgaaatgct gtttcgtgct ttgcgtaaac actotcattt gcctcaagcc	120
ttggtcgatg tggtcggcga tgaatctttg agctacaagg agttttttga ggcaaccgtc	180
ttgctggctc agtccttgca taattgtggc tacaagatga acgacgtcgt ctccatttgt	240
gcagaaaaca ataccegttt etteatteea gteategeeg eatggtatat eggtatgate	300
gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct	360
aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc	420
cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc	480
gaatctttgc ctaattttat tagccgctat tcagacggaa acatcgccaa ctttaagcct	540
ctccatttcg accctgtgga acaagttgct gcaatcctgt gtagcagcgg tactactgga	600

ctcccaaagg gagtcatgca gacccatcaa aacatttgcg tgcgtctgat ccatgctctc

720
780
840
900
960
1020
1080
1140
1200
1260
1320
1380
1440
1500
1560
1620
1626
1626
60
1626 60 120
1626 60 120 180
1626 60 120 180 240
1626 60 120 180 240 300
1626 60 120 180 240 300 360
1626 60 120 180 240 300 360 420
1626 60 120 180 240 300 360 420 480
1626 60 120 180 240 300 360 420 480 540
60 120 180 240 300 360 420 480 540
60 120 180 240 300 360 420 480 540 600

agtgtcatca acgtgcctag cgtgatcctg tttttgtcta agagcccact cgtggacaag tacgacttgt cttcactgcg tgaattgtgt tgcggtgccg ctccactggc taaggaggtc

gctgaagtgg ccgccaaacg cttgaatctt ccagggattc gttgtggctt cggcctcacc 1020 gaatctacca gcgctattat tcagtctctc cgcgatgagt ttaagagcgg ctctttgggc 1080 cgtgtcactc cactcatggc tgctaagatc gctgatcgcg aaactggtaa ggctttgggc 1140 cctaaccaag tgggcgagct gtgtatcaaa ggccctatgg tgagcaaggg ttatgtcaat 1200 aacgtcgaag ctaccaagga ggccatcgac gacgacggct ggttgcattc tggtgatttt 1260 ggatattacg acgaagatga gcattttac gtcgtggatc gttacaagga gctgatcaaa 1320 tacaagggta gccaggttgc tccagctgag ttggaggaga ttctgttgaa aaatccatgc 1380 attcgcgatg tcgctgtggt cggcattcct gatctggagg ccggcgaact gccttctgct 1440 ttcgttgtca agcagcctgg taaagaaatt accgccaaag aagtgtatga ttacctggct 1500 gaacgtgtga gccatactaa gtacttgcgt ggcggcgtgc gttttgttga ctccatccct 1560 cgtaacgtaa caggcaaaat tacccgcaag gagctgttga aacaattgtt ggagaaggcc 1620 ggcggt
cgtgtcactc cactcatggc tgctaagatc gctgatcgcg aaactggtaa ggctttgggc 1140 cctaaccaag tgggcgagct gtgtatcaaa ggccctatgg tgagcaaggg ttatgtcaat 1200 aacgtcgaag ctaccaagga ggccatcgac gacgacggct ggttgcattc tggtgatttt 1260 ggatattacg acgaagatga gcattttac gtcgtggatc gttacaagga gctgatcaaa 1320 atcaagggta gccaggttgc tccagctgag ttggaggaga ttctgttgaa aaatccatgc 1380 attcgcgatg tcgctgtggt cggcattcct gatctggagg ccggcgaact gccttctgct 1440 ttcgttgtca agcagctgg taaagaaatt accgccaaag aagtgtatga ttacctggct 1500 gaacgtgtga gccatactaa gtacttgcgt ggcggcgtgc gttttgttga ctccatccct 1560 cgtaacgtaa caggcaaaat tacccgcaag gagctgttga aacaattgtt ggagaaggcc 1620
cctaaccaag tgggcgagct gtgtatcaaa ggccctatgg tgagcaaggg ttatgtcaat 1200 aacgtcgaag ctaccaagga ggccatcgac gacgacggct ggttgcattc tggtgatttt 1260 ggatattacg acgaagatga gcattttac gtcgtggatc gttacaagga gctgatcaaa 1320 tacaagggta gccaggttgc tccagctgag ttggaggaga ttctgttgaa aaatccatgc 1380 attcgcgatg tcgctgtggt cggcattcct gatctggagg ccggcgaact gccttctgct 1440 ttcgttgtca agcagctgg taaagaaatt accgccaaag aagtgtatga ttacctggct 1500 gaacgtgtga gccatactaa gtacttgcgt ggcggcgtgc gttttgttga ctccatccct 1560 cgtaacgtaa caggcaaaat tacccgcaag gagctgttga aacaattgtt ggagaaggcc 1620
aacgtcgaag ctaccaagga ggccatcgac gacgacggct ggttgcattc tggtgatttt 1260 ggatattacg acgaagatga gcattttac gtcgtggatc gttacaagga gctgatcaaa 1320 tacaagggta gccaggttgc tccagctgag ttggaggaga ttctgttgaa aaatccatgc 1380 attcgcgatg tcgctgtggt cggcattcct gatctggagg ccggcgaact gccttctgct 1440 ttcgttgtca agcagctgg taaagaaatt accgccaaag aagtgtatga ttacctggct 1500 gaacgtgtga gccatactaa gtacttgcgt ggcggcgtgc gttttgttga ctccatccct 1560 cgtaacgtaa caggcaaaat tacccgcaag gagctgttga aacaattgtt ggagaaggcc 1620
ggatattacg acgaagatga gcattttac gtcgtggatc gttacaagga gctgatcaaa 1320 tacaagggta gccaggttgc tccagctgag ttggaggaga ttctgttgaa aaatccatgc 1380 attcgcgatg tcgctgtggt cggcattcct gatctggagg ccggcgaact gccttctgct 1440 ttcgttgtca agcagcctgg taaagaaatt accgccaaag aagtgtatga ttacctggct 1500 gaacgtgtga gccatactaa gtacttgcgt ggcggcgtgc gttttgttga ctccatccct 1560 cgtaacgtaa caggcaaaat tacccgcaag gagctgttga aacaattgtt ggagaaggcc 1620
tacaagggta gccaggttgc tccagctgag ttggaggaga ttctgttgaa aaatccatgc 1380 attcgcgatg tcgctgtggt cggcattcct gatctggagg ccggcgaact gccttctgct 1440 ttcgttgtca agcagcctgg taaagaaatt accgccaaag aagtgtatga ttacctggct 1500 gaacgtgtga gccatactaa gtacttgcgt ggcggcgtgc gttttgttga ctccatccct 1560 cgtaacgtaa caggcaaaat tacccgcaag gagctgttga aacaattgtt ggagaaggcc 1620
attegegatg tegetgtggt eggeatteet gatetggagg eeggegaact geettetget 1440 ttegttgtea ageageetgg taaagaaatt acegecaaag aagtgtatga ttacetgget 1500 gaaegtgtga geeatactaa gtacttgegt ggeggegtge gttttgttga etceateeet 1560 egtaaegtaa eaggeaaaat taceegeaag gagetgttga aacaattgtt ggagaaggee 1620
ttcgttgtca agcagcctgg taaagaaatt accgccaaag aagtgtatga ttacctggct 1500 gaacgtgtga gccatactaa gtacttgcgt ggcggcgtgc gttttgttga ctccatccct 1560 cgtaacgtaa caggcaaaat tacccgcaag gagctgttga aacaattgtt ggagaaggcc 1620
gaacgtgtga gccatactaa gtacttgcgt ggcggcgtgc gttttgttga ctccatccct 1560 cgtaacgtaa caggcaaaat tacccgcaag gagctgttga aacaattgtt ggagaaggcc 1620
cgtaacgtaa caggcaaaat tacccgcaag gagctgttga aacaattgtt ggagaaggcc 1620
ggcggt 1626
<pre><210> SEQ ID NO 15 <211> LENGTH: 1626 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase</pre>
22235 OTHER INFORMATION: Sequence of a synthetic futificase
<400> SEQUENCE: 15
<400> SEQUENCE: 15
<pre><400> SEQUENCE: 15 atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat 60</pre>
<pre><400> SEQUENCE: 15 atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actcttattt gcctcaagcc 120</pre>
<pre><400> SEQUENCE: 15 atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat 60 ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actcttattt gcctcaagcc 120 ttggtcgatg tggtcggcga tgaatctttg agctacaagg agtttttga ggcaaccgtc 180</pre>
<pre><400> SEQUENCE: 15 atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat 60 ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actcttattt gcctcaagcc 120 ttggtcgatg tggtcggcga tgaatctttg agctacaagg agttttttga ggcaaccgtc 180 ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt 240</pre>
<pre><400> SEQUENCE: 15 atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat 60 ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actcttattt gcctcaagcc 120 ttggtcgatg tggtcggcga tgaatctttg agctacaagg agttttttga ggcaaccgtc 180 ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt 240 gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc 300</pre>
<400> SEQUENCE: 15 atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat 60 ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actcttattt gcctcaagcc 120 ttggtcgatg tggtcggcga tgaatctttg agctacaagg agttttttga ggcaaccgtc 180 ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt 240 gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc 300 gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct 360
<pre><400> SEQUENCE: 15 atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat 60 ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actcttattt gcctcaagcc 120 ttggtcgatg tggtcggcga tgaatctttg agctacaagg agtttttga ggcaaccgtc 180 ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt 240 gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc 300 gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct 360 aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc 420</pre>
<400> SEQUENCE: 15 atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat 60 ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actcttattt gcctcaagcc 120 ttggtcgatg tggtcggcga tgaatctttg agctacaagg agttttttga ggcaaccgtc 180 ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc 300 gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct 360 aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc 420 cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc 480
atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat 60 cttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actcttattt gcctcaagcc 120 cttggtcgatg tggtcggcga tgaatctttg agctacaagg agtttttga ggcaaccgtc ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt 240 gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc 300 gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct 360 aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc 420 cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc gaatctttgc ctaatttcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca 540 gaatctttgc ctaattcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca 540 gaatctttgc ctaattcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca 540 gaatctttgc ctaattcatcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca 540 gaatctttgc ctaattcatcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca 540 gaatctttgc ctaattcatcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca 540 gaatctttgc ctaattcatcatcatcatcatcatcatcatcatcatcatc
<pre><400> SEQUENCE: 15 atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat 60 ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actcttattt gcctcaagcc 120 ttggtcgatg tggtcggcga tgaatctttg agctacaagg agttttttga ggcaaccgtc 180 ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt 240 gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc 300 gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct 360 aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc 420 cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc 480 gaatctttgc ctaatttcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca 540 ctccacttcg accctgtgga acaagttgca gccattctgt gtagcagcgg tactactgga 600</pre>
atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat 60 ctcgactgagc gcgaaatgct gtttcgtgct ctccgcaagc actcttatt gcctcaagcc 120 cttggtcgatg tggtcggcga tgaatctttg agctacaagg agtttttga ggcaaccgtc 180 ctgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt 240 gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc 300 gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct 360 aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc 420 cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc gaatctttgc ctaatttcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca 540 ctccacttcg accctgtgga acaagttgca gccattctgt gtagcagcgg tactactgga 600 ctcccaaagg gagtcatgca gacccatcaa aacatttgcg tgcgtctgat ccatgctctc 660 ctcccaaagg gagtcatgca gacccatcaa aacatttgcg tgcgtctgat ccatgctctc
atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat 600 ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actcttattt gcctcaagcc 1200 ttggtcgatg tggtcggcga tgaatctttg agctacaagg agttttttga ggcaaccgtc 1800 ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt 2400 gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc 3000 gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct 3600 aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc 4200 cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc 4800 gaatctttgc ctaatttcat ctctcgctat tcagacggca acatcgcaaa ctttaaaacca 5400 ctcccaaagg gagtcatgca gacccatcaa aacatttgcg tgcgtctgat ccatgctctc 6600 gatccacgct acggcactca gctgattcct ggtgtcaccg tcttggtcta cttgcctttc 7200 gatccacgct acggcactca gctgattcct ggtgtcaccg tcttggtcta cttgcctttc 7200 gatccacgct acggcactca gctgattcct ggtgtcaccg tcttggtcta cttgcctttc
atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat 60 ctctgactgcg gcgaaatgct gtttcgtgct ctccgcaagc actcttattt gcctcaagcc 120 cttggtcgatg tggtcggcga tgaatctttg agctacaagg agttttttga ggcaaccgtc 180 cttgctggctc agtccctcca caattgtggc tacaaggatga acgacgtcgt tagtatctgt gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc 300 gtggctccag tcaaccgaga ctacattccc gacgaactgt gtaaagtcat gggtatctct aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc 420 cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc gaatctttgc ctaattcct ctccgctat tcagacggca acatcgcaaa ctttaaacca 540 ctccacatcg accctgtgga acaagttgca gccattctgt gtagcagcgg tactactgga 600 ctcccaaagg gagtcatgca gacccatcaa aacatttgcg tgggtctgat ccatgctctc gatccacagct acggcactca gctgattcct ggtgtcaccg tcttggtcta cttgccttc 720 ctccatgct tcggctttca tattactttg ggttacttta tggtcggtct ccgcgtgatt 780 ctccatgctt tcggctttca tattactttg ggttacttta tggtcggtct ccgcgtgatt 780 ctccatgctt tcggctttca tattactttg ggttacttta tggtcggtct ccgcgtgatt
atgatgaage gtgagaaaaa tgtcatctat ggccctgage ctctccatce tttggaggat 600 ttgactgccg gcgaaatgct gtttcgtgct ctccgcaage actcttattt gcctcaagec 1200 ttggtcgatg tggtcggcga tgaatctttg agctacaagg agttttttga ggcaaccgtc 1800 ttgctggctc agtccctcca caattgtggc tacaaggatga acgacgtcgt tagtatctgt 2400 gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc 3000 gtggctccag tcaacgaggag ctacattccc gacgaactgt gtaaagtcat gggtatctct 3600 aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc 4200 cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc 4800 gaatctttgc ctaattcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca 5400 ctccacattcg accctgtgga acaagttgca gccattctgt gtagcagcgg tactactgga 6000 ctcccaaagg gagtcatgca gacccatcaa aacatttgcg tgcgtctgat ccatgctctc gatccacgct acggcactca gctgattcct ggtgtcaccg tcttggtcta cttgccttc 7200 gatccatgct tcggctttca tattactttg ggttacttta tggtcggtct ccgcgtgatt 7800 atgttccatgct tcggctttca tattactttg ggttacttta tggtcggtct ccgcgtgatt 7800 atgttcccc gttttgatca ggaggctttc ttgaaagcca tccaagatta tgaagtccgc 8400 atgttcccc gttttgatca ggaggctttc ttgaaagcca tccaagatta tgaagtccgc
atgatgaage gtgagaaaaa tgtcatctat ggccctgage ctctccatcc tttggaggat 600 ttgactgccg gcgaaatgct gtttcgtgct ctccgcaage actcttattt gcctcaagcc 1200 ttggtcgatg tggtcggcga tgaatctttg agctacaagg agttttttga ggcaaccgtc 1800 ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt 2400 gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc 3000 gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct 3600 aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc 4200 cgcaccaact ttattaagcg tacaatcact ttggacactg tggagaatat tcacggttgc 4800 gaatctttge ctaattcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca 5400 ctccaacttcg accctgtgga acaagttgca gccattctgt gtagcagcgg tactactgga 6000 ctcccaaagg gagtcatgca gacccatcaa aacatttgcg tgcgtctgat ccatgctctc 6600 gatccacgct acggcactca gctgattcct ggtgtcaccg tcttggtcta cttgccttcc 7200 ttccatgctt tcggctttca tattactttg ggttacttta tggtcggtct ccgcgtgatt 7800 atggtccacca gctgtttgatca ggaggctttc ttgaaagcca tccaagatta tgaagtccgc 8400 agtgtcatca acgtgcctag cgtgatcctg ttttgtcta agagcccact cgtggacaag 9000 agtgtcatca acgtgcctag cgtgatcctg tttttgtcta agagcccact cgtggacaag 9000 agtgtcatca acgtgcctag cgtgatcctg tttttgtcta agagcccact cgtggacaag
atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat 60 ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actcttattt gcctcaagcc 120 ttggtcgatg tggtcggcga tgaatctttg agctacaagg agttttttga ggcaaccgtc 180 ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt 240 gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc 300 gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct 360 aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc 420 cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc 480 gaatctttgc ctaatttcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca 540 ctccacttcg accctgtgga acaagttgca gccattctgt gtagcagcgg tactactgga 600 ctcccaaagg gagtcatgca gacccatcaa aacatttgcg tgcgtctgat ccatgctctc 660 gatccacgct acggcactca gctgattcct ggtgtcaccg tcttggtcta cttgccttc 720 ttccatgctt tcggctttca tattactttg ggttacttta tggtcggtct ccgcgtgatt 780 atgttccgc gttttgatca ggaggcttc ttgaaagcca tccaagatta tgaagtccgc aggtgtcatca acgtgcctag cgtgatcctg ttttgtcta agagccact cgtggacaag 900 tacgacttgt cttcactgc tgaattgtg tgcggtccg ctccactggc taaggaggtc 960 tacgacttgt cttcactgc tgaattgtg tgcggtccc ccgcgtgatca 260 gatcactca acgtgcctag cgtgatcctg tttttgtcta agagcccact cgtggacaag 900 tacgacttgt cttcactgc tgaattgtgt tgcggtccc ctccactgc taaggaggtc 960 tacgacttgt cttcactgc tgaattgtgt tgcggtccc ctccactgc taaggaggtc
atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat 60 ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actcttattt gcctcaagcc 120 ttggtcgatg tggtcggcga tgaatctttg agctacaagg agttttttga ggcaaccgtc 180 ttgctggctc agtccctcca caattgtggc tacaaggtag acgacgtcgt tagtatctgt 240 gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc 300 gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc 420 cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc gaatctttgc ctaatttcg ctcaattcatc ttggacactg tggagaatat tcacggttgc 480 ctccacatcg accctgtgga acaagttgca gccattctgt gtagcagcgg tactactga 600 ctcccacagg aggccatca gctgattcct ggtgtcaccg tcttggtcta cttgccttc 720 gatccacgct acggcactca gctgattcct ggtgtcaccg tcttggtcta cttgccttc 720 ttccatgctt tcggctttca tattacttg ggttacttta tggtcggtct ccgcgtgatt 780 atgttccgc gttttgatca ggaggcttc ttgaaagcca tccaagatta tgaagtccgc aggtgtcatca acgtgcctag cgtgatcctg tttttgtcta agagcccact cgtggacaag 900 tacgacttgt cttcactgcg tgaattgtgt tgcggtgcc ctccactggc taaggaggtc 960 gctgaagtgg ccgccaaacg cttgaatctt ccagggattc gttgtggct cggcctcacc 1020 gctgaagtgg ccgccaaacg cttgaatctt ccagggattc gttgtggct cggcctcacc

aacgttgaag ctaccaagga ggccatcgac gacgacggct ggttgcattc tggtgatttt 1260

-continued	
ggatattacg acgaagatga gcatttttac gtcgtggatc gttacaagga gctgatcaaa	1320
tacaagggta gccaggttgc tccagctgag ttggaggaga ttctgttgaa aaatccatgc	1380
attcgcgatg tcgctgtggt cggcattcct gatctggagg ccggcgaact gccttctgct	1440
ttcgttgtca agcagcctgg taaagaaatt accgccaaag aagtgtatga ttacctggct	1500
gaacgtgtga gccatactaa gtacttgcgt ggcggcgtgc gttttgttga ctccatccct	1560
cgtaacgtaa caggcaaaat tacccgcaag gagctgttga aacaattgtt ggagaaggcc	1620
ggcggt	1626
<210> SEQ ID NO 16 <211> LENGTH: 1626 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase	
<400> SEQUENCE: 16	
atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat	60
ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actctcattt gcctcaagcc	120
ttggtcgatg tggtcggcga tgaatctttg agctacaagg agttttttga ggcaaccgtc	180
ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt	240
gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc	300
gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct	360
aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc	420
cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc	480
gaatctttgc ctaatttcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca	540
ctccacttcg accctgtgga acaagttgca gccattctgt gtagcagcgg tactactgga	600
ctcccaaagg gagtcatgca gacccatcaa aacatttgcg tgcgtctgat ccatgctctc	660
gatccacgct acggcactca gctgattcct ggtgtcaccg tcttggtcta cttgcctttc	720
ttccatgctt tcggctttca tattactttg ggttacttta tggtcggtct ccgcgtgatt	780
atgttccgcc gttttgatca ggaggctttc ttgaaagcca tccaagatta tgaagtccgc	840
agtgtcatca acgtgcctag cgtgatcctg tttttgtcta agagcccact cgtggacaag	900
tacgacttgt cttcactgcg tgaattgtgt tgcggtgccg ctccactggc taaggaggtc	960
gctgaagtgg ccgccaaacg cttgaatctt ccagggattc gttgtggctt cggcctcacc	1020
gaatctacca gcgctattat tcagtctctc cgcgatgagt ttaagagcgg ctctttgggc	1080
cgtgtcactc cactcatggc tgctaagatc gctgatcgcg aaactggtaa ggctttgggc	1140
ccgaaccaag tgggcgagct gtgtatcaaa ggccctatgg tgagcaaggg ttatgtcaat	1200
aacgttgaag ctaccaagga ggccatcgac gacgacggct ggttgcattc tggtgatttt	1260
ggatattacg acgaagatga gcatttttac gtcgtggatc gttacaagga gctgatcaaa	1320
tacaagggta gccaggttgc tccagctgag ttggaggaga ttctgttgaa aaatccatgc	1380
attogogatg togotgtggt oggoattoot gatotggagg coggogaact goottotgot	1440
ttcgttgtca agcagcctgg taaagaaatt accgccaaag aagtgtatga ttacctggct	1500

gaacgtgtga gccatactaa gtacttgcgt ggcggcgtgc gttttgttga ctccatccct 1560

cgtaacgtaa caggcaaaat tacccgcaag gagctgttga aacaattgtt ggagaaggcc	1620
ggcggt	1626
<210> SEQ ID NO 17 <211> LENGTH: 1626 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase	
<400> SEQUENCE: 17	
atgatgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat	60
ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actctcattt gcctcaagcc	120
ttggtcgatg tggtcggcga tgaatctttg agctacaagg agttttttga ggcaaccgtc	180
ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt	240
gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc	300
gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct	360
aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc	420
cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc	480
gaatctttgc ctaatttcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca	540
ctccacttcg accctgtgga acaagttgca gccattctgt gtagcagcgg tactactgga	600
ctcccaaagg gagtcatgca gacccatcaa aacatttgcg tgcgtctgat ccatgctctc	660
gatccacgct acggcactca gctgattcct ggtgtcaccg tcttggtcta cttgcctttc	720
ttccatgctt tcggctttca tattactttg ggttacttta tggtcggtct ccgcgtgatt	780
atgttccgcc gttttgatca ggaggctttc ttgaaagcca tccaagatta tgaagtccgc	840
agtgtcatca acgtgcctag cgtgatcctg tttttgtcta agagcccact cgtggacaag	900
tacgacttgt cttcactgcg tgaattgtgt tgcggtgccg ctccactggc taaggaggtc	960
gctgaagtgg ccgccaaacg cttgaatctt ccagggattc gttgtggctt cggcctcacc	1020
gaatctacca gcgctattat tcagtctctc ggggatgagt ttaagagcgg ctctttgggc	1080
cgtgtcactc cactcatggc tgctaagatc gctgatcgcg aaactggtaa ggctttgggc	1140
ccgaaccaag tgggcgagct gtgtatcaaa ggccctatgg tgagcaaggg ttatgtcaat	1200
aacgttgaag ctaccaagga ggccatcgac gacgacggct ggttgcattc tggtgatttt	1260
ggatattacg acgaagatga gcatttttac gtcgtggatc gttacaagga gctgatcaaa	1320
tacaagggta gccaggttgc tccagctgag ttggaggaga ttctgttgaa aaatccatgc	1380
attogogatg togotgtggt oggoattoot gatotggagg coggogaact goottotgot	1440
ttcgttgtca agcagcctgg taaagaaatt accgccaaag aagtgtatga ttacctggct	1500
gaacgtgtga gccatactaa gtacttgcgt ggcggcgtgc gttttgttga ctccatccct	1560
cgtaacgtaa caggcaaaat tacccgcaag gagctgttga aacaattgtt ggagaaggcc	1620
ggcggt	1626

<210> SEQ ID NO 18 <211> LENGTH: 1626 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE:

<223> OTHER INFORMATION: Sequence of a synthetic luciferase
<400> SEQUENCE: 18
atgataaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat 60
ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actctcattt gcctcaagcc 120
ttggtcgatg tggtcggcga tgaatctttg agctacaagg agttttttga ggcaaccgtc 180
ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt 240
gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc 300
gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct 360
aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc 420
cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc 480
gaatetttge etaattteat etetegetat teagaeggea acategeaaa etttaaacea 540
ctccacttcg accctgtgga acaagttgca gccattctgt gtagcagcgg tactactgga 600
ctcccaaagg gagtcatgca gacccatcaa aacatttgcg tgcgtctgat ccatgctctc 660
gatccacgct acggcactca gctgattcct ggtgtcaccg tcttggtcta cttgcctttc 720
ttccatgctt tcggctttca tattactttg ggttacttta tggtcggtct ccgcgtgatt 780
atgttccgcc gttttgatca ggaggctttc ttgaaagcca tccaagatta tgaagtccgc 840
agtgtcatca acgtgcctag cgtgatcctg tttttgtcta agagcccact cgtggacaag 900
tacgacttgt cttcactgcg tgaattgtgt tgcggtgccg ctccactggc taaggaggtc 960
gctgaagtgg ccgccaaacg cttgaatctt ccagggattc gttgtggctt cggcctcacc 1020
gaatctacca gtgcgattat ccagactctc ggggatgagt ttaagagcgg ctctttgggc 1080
cgtgtcactc cactcatggc tgctaagatc gctgatcgcg aaactggtaa ggctttgggc 1140
ccgaaccaag tgggcgagct gtgtatcaaa ggccctatgg tgagcaaggg ttatgtcaat 1200
aacgttgaag ctaccaagga ggccatcgac gacgacggct ggttgcattc tggtgatttt 1260
ggatattacg acgaagatga gcatttttac gtcgtggatc gttacaagga gctgatcaaa 1320
tacaagggta gccaggttgc tccagctgag ttggaggaga ttctgttgaa aaatccatgc 1380
attogogatg togotgtggt oggoattoot gatotggagg coggogaact goottotgot 1440
ttcgttgtca agcagcctgg tacagaaatt accgccaaag aagtgtatga ttacctggct 1500
gaacgtgtga gccatactaa gtacttgcgt ggcggcgtgc gttttgttga ctccatccct 1560
cgtaacgtaa caggcaaaat tacccgcaag gagctgttga aacaattgtt ggtgaaggcc 1620
ggcggt 1626
<210> SEQ ID NO 19 <211> LENGTH: 933 <212> TYPE: DNA <213> ORGANISM: Renilla reniformis
<400> SEQUENCE: 19
atgacttcga aagtttatga tccagaacaa aggaaacgga tgataactgg tccgcagtgg 60
tgggccagat gtaaacaaat gaatgttctt gattcattta ttaattatta tgattcagaa 120
aaacatgcag aaaatgctgt tatttttta catggtaacg cggcctcttc ttatttatgg 180
cgacatgttg tgccacatat tgagccagta gcgcggtgta ttataccaga tcttattggt 240

atgggcaaat caggcaaatc tggtaatggt tcttataggt tacttgatca ttacaaatat	300
cttactgcat ggtttgaact tcttaattta ccaaagaaga tcatttttgt cggccatgat	360
tggggtgctt gtttggcatt tcattatagc tatgagcatc aagataagat	420
gttcacgctg aaagtgtagt agatgtgatt gaatcatggg atgaatggcc tgatattgaa	480
gaagatattg cgttgatcaa atctgaagaa ggagaaaaaa tggttttgga gaataacttc	540
ttcgtggaaa ccatgttgcc atcaaaaatc atgagaaagt tagaaccaga agaatttgca	600
gcatatcttg aaccattcaa agagaaaggt gaagttcgtc gtccaacatt atcatggcct	660
cgtgaaatcc cgttagtaaa aggtggtaaa cctgacgttg tacaaattgt taggaattat	720
aatgcttatc tacgtgcaag tgatgattta ccaaaaatgt ttattgaatc ggatccagga	780
ttcttttcca atgctattgt tgaaggcgcc aagaagtttc ctaatactga atttgtcaaa	840
gtaaaaggtc ttcattttc gcaagaagat gcacctgatg aaatgggaaa atatatcaaa	900
tcgttcgttg agcgagttct caaaaatgaa caa	933
<pre><210> SEQ ID NO 20 <211> LENGTH: 933 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase</pre>	
<400> SEQUENCE: 20	
atggetteca aggtgtaega eecegageag egeaagegea tgateaeegg eeeteagtgg	60
tgggcccgct gcaagcagat gaacgtgctg gactccttca tcaactacta cgacagcgag	120
aagcacgccg agaacgccgt gatcttcctg cacggcaacg ccgcctccag ctacctgtgg	180
aggeaegtgg tgeeteaeat egageeegtg geeegetgea teateeetga eetgategge	240
atgggcaagt ccggcaagag cggcaacggc tcctaccgcc tgctggacca ctacaagtac	300
atgggcaagt ccggcaagag cggcaacggc tcctaccgcc tgctggacca ctacaagtac	300 360
	300 360 420
ctgaccgcct ggttcgagct gctgaacctg cccaagaaga tcatcttcgt gggccacgac	300 360
ctgaccgcct ggttcgagct gctgaacctg cccaagaaga tcatcttcgt gggccacgac tgggggagcct gcctggcctt ccactactcc tacgagcacc aggacaagat caaggccatc	300 360 420 480 540
ctgaccgcct ggttcgagct gctgaacctg cccaagaaga tcatcttcgt gggccacgac tggggagcct gcctggcctt ccactactcc tacgagcacc aggacaagat caaggccatc gtgcacgccg agagcgtggt ggacgtgatc gagtcctggg acgagtggcc tgacatcgag gaggacatcg ccctgatcaa gagcgaggag ggcgagaaga tggtgctgga gaacaacttc ttcgtggaga ccatgctgcc cagcaagatc atgcgcaagc tggagcctga ggagttcgcc	300 360 420 480 540
ctgaccgcct ggttcgagct gctgaacctg cccaagaaga tcatcttcgt gggccacgac tggggagcct gcctggcctt ccactactcc tacgagcacc aggacaagat caaggccatc gtgcacgccg agagcgtggt ggacgtgatc gagtcctggg acgagtggcc tgacatcgag gaggacatcg ccctgatcaa gagcgaggag ggcgagaaga tggtgctgga gaacaacttc	300 360 420 480 540 600
ctgaccgcct ggttcgagct gctgaacctg cccaagaaga tcatcttcgt gggccacgac tggggagcct gcctggcctt ccactactcc tacgagcacc aggacaagat caaggccatc gtgcacgccg agagcgtggt ggacgtgatc gagtcctggg acgagtggcc tgacatcgag gaggacatcg ccctgatcaa gagcgaggag ggcgagaaga tggtgctgga gaacaacttc ttcgtggaga ccatgctgcc cagcaagatc atgcgcaagc tggagcctga ggagttcgcc	300 360 420 480 540 600 660
ctgaccgcct ggttcgagct gctgaacctg cccaagaaga tcatcttcgt gggccacgac tggggagcct gcctggcctt ccactactcc tacgagcacc aggacaagat caaggccatc gtgcacgccg agagcgtggt ggacgtgatc gagtcctggg acgagtggcc tgacatcgag gaggacatcg ccctgatcaa gagcgaggag ggcgagaaga tggtgctgga gaacaacttc ttcgtggaga ccatgctgcc cagcaagatc atgcgcaagc tggagcctga ggagttcgcc gcctacctg agccctcaa ggagaagggc gaggtgcgcc gccctaccct gtcctggccc	300 360 420 480 540 600
ctgaccgcct ggttcgagct gctgaacctg cccaagaaga tcatcttcgt gggccacgac tggggagcct gcctggcctt ccactactcc tacgagcacc aggacaagat caaggccatc gtgcacgccg agagcgtggt ggacgtgatc gagtcctggg acgagtggcc tgacatcgag gaggacatcg ccctgatcaa gagcgaggag ggcgagaaga tggtgctgga gaacaacttc ttcgtggaga ccatgctgcc cagcaagatc atgcgcaagc tggagcctga ggagttcgcc gcctacctg agccttcaa ggagaagggc gaggtgcgcc gccctaccct gtcctggccc cgcgagatcc ctctggtgaa gggcggcaag cccgacgtgg tgcagatcgt gcgcaactac	300 360 420 480 540 600 660
ctgaccgcct ggttcgagct gctgaacctg cccaagaaga tcatcttcgt gggccacgac tggggagcct gcctggcctt ccactactcc tacgagcacc aggacaagat caaggccatc gtgcacgccg agagcgtggt ggacgtgatc gagtcctggg acgagtggcc tgacatcgag gaggacatcg ccctgatcaa gagcgaggag ggcgagaaga tggtgctgga gaacaacttc ttcgtggaga ccatgctgcc cagcaagatc atgcgcaagc tggagcctga ggagttcgcc gcctacctgg agcccttcaa ggagaagggc gaggtgcgcc gccctaccct gtcctggccc cgcgagatcc ctctggtgaa gggcggcaag cccgacgtgg tgcagatcgt gcgcaactac aacgcctacc tgcgcgcca cgacgacctg cctaagatgt tcatcgagtc cgaccctggc	300 360 420 480 540 600 660 720

SEQ ID O 21

LENGTH 933

<212> TYPE: DNA <213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Sequence of a synthetic luciferase

agcttcgtgg agcgcgtgct gaagaacgag cag

-continued

-continued	
<400> SEQUENCE: 21	
atggcttcca aggtgtacga ccccgagcaa cgcaaacgca tgatcactgg gcctcagtgg	60
tgggctcgct gcaagcaaat gaacgtgctg gactccttca tcaactacta tgattccgag	120
aagcacgccg agaacgccgt gatttttctg catggtaacg ctgcctccag ctacctgtgg	180
aggcacgtcg tgcctcacat cgagcccgtg gctcgctgca tcatccctga tctgatcgga	240
atgggtaagt ccggcaagag cgggaatggc tcatatcgcc tcctggatca ctacaagtac	300
ctcaccgctt ggttcgagct gctgaacctt ccaaagaaaa tcatctttgt gggccacgac	360
tggggggctt gtctggcctt tcactactcc tacgagcacc aagacaagat caaggccatc	420
gtccatgctg agagtgtcgt ggacgtgatc gagtcctggg acgagtggcc tgacatcgag	480
gaggatatcg ccctgatcaa gagcgaagag ggcgagaaaa tggtgcttga gaataacttc	540
ttcgtcgaga ccatgctccc aagcaagatc atgcggaaac tggagcctga ggagttcgct	600
gcctacctgg agcccttcaa ggagaagggc gaggttagac ggcctaccct ctcctggcct	660
cgcgagatcc ctctcgttaa gggaggcaag cccgacgtcg tccagattgt ccgcaactac	720
aacgcctacc ttcgggccag cgacgatctg cctaagatgt tcatcgagtc cgaccctggg	780
ttcttttcca acgctattgt cgagggagct aagaagttcc ctaacaccga gttcgtgaag	840
gtgaagggcc tccacttcag ccaggaggac gctccagatg aaatgggtaa gtacatcaag	900
agcttcgtgg agcgcgtgct gaagaacgag cag	933
<210> SEQ ID NO 22 <211> LENGTH: 933 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase	
<400> SEQUENCE: 22	
atggcttcca aggtgtacga ccccgagcaa cgcaaacgca tgatcactgg gcctcagtgg	60
tgggctcgct gcaagcaaat gaacgtgctg gactccttca tcaactacta tgattccgag	120
aagcacgccg agaacgccgt gatttttctg catggtaacg ctgcctccag ctacctgtgg	180
aggcacgtcg tgcctcacat cgagcccgtg gctagatgca tcatccctga tctgatcgga	240
atgggtaagt ccggcaagag cgggaatggc tcatatcgcc tcctggatca ctacaagtac	300
ctcaccgctt ggttcgagct gctgaacctt ccaaagaaaa tcatctttgt gggccacgac	360
tggggggett gtctggcctt tcactactcc tacgagcacc aagacaagat caaggccatc	420
gtccatgctg agagtgtcgt ggacgtgatc gagtcctggg acgagtggcc tgacatcgag	480
gaggatatog cootgatoaa gagogaagag ggogagaaaa tggtgottga gaataactto	540
ttcgtcgaga ccatgctccc aagcaagatc atgcggaaac tggagcctga ggagttcgct	600
gcctacctgg agccattcaa ggagaagggc gaggttagac ggcctaccct ctcctggcct	660
cgcgagatcc ctctcgttaa gggaggcaag cccgacgtcg tccagattgt ccgcaactac	720
aacgcctacc ttcgggccag cgacgatctg cctaagatgt tcatcgagtc cgaccctggg	780
	700
ttcttttcca acgctattgt cgagggagct aagaagttcc ctaacaccga gttcgtgaag	840

933

<211	> LE	Q ID NGTH	: 54												
		GANI		_	phor	us p	lagi	.opht	hala	mus					
		QUEN			Tara	7 an	77-1	т1 о	Птть	C1	Dro	C lu	Dro	Tou	u; a
1	Met	Lys	ALG	5	пуь	Abii	vai	116	10	GIY	PLO	Giu	PLU	15	пть
Pro	Leu	Glu	Asp 20	Leu	Thr	Ala	Gly	Glu 25	Met	Leu	Phe	Arg	Ala 30	Leu	Arg
Lys	His	Ser 35	His	Leu	Pro	Gln	Ala 40	Leu	Val	Asp	Val	Phe 45	Gly	Asp	Glu
Ser	Leu 50	Ser	Tyr	Lys	Glu	Phe 55	Phe	Glu	Ala	Thr	C y s 60	Leu	Leu	Ala	Gln
Ser 65	Leu	His	Asn	Суѕ	Gl y 70	Tyr	Lys	Met	Asn	Asp 75	Val	Val	Ser	Ile	C y s 80
Ala	Glu	Asn	Asn	L y s 85	Arg	Phe	Phe	Ile	Pro 90	Ile	Ile	Ala	Ala	Trp 95	Tyr
Ile	Gly	Met	Ile 100	Val	Ala	Pro	Val	Asn 105	Glu	Ser	Tyr	Ile	Pro 110	Asp	Glu
Leu	Cys	Lys 115	Val	Met	Gly	Ile	Ser 120	Lys	Pro	Gln	Ile	Val 125	Phe	Cys	Thr
Lys	Asn 130	Ile	Leu	Asn	Lys	Val 135	Leu	Glu	Val	Gln	Ser 140	Arg	Thr	Asn	Phe
Ile 145	Lys	Arg	Ile	Ile	Ile 150	Leu	Asp	Thr	Val	Glu 155	Asn	Ile	His	Gly	Cys 160
Glu	Ser	Leu	Pro	Asn 165	Phe	Ile	Ser	Arg	Ty r 170	Ser	Asp	Gly	Asn	Ile 175	Ala
Asn	Phe	Lys	Pro 180	Leu	His	Tyr	Asp	Pro 185	Val	Glu	Gln	Val	Ala 190	Ala	Ile
Leu	Сув	Ser 195	Ser	Gly	Thr	Thr	Gly 200	Leu	Pro	Lys	Gly	Val 205	Met	Gln	Thr
His	Gln 210	Asn	Ile	Сув	Val	A rg 215	Leu	Ile	His	Ala	Leu 220	Asp	Pro	Arg	Ala
Gl y 225	Thr	Gln	Leu	Ile	Pro 230	Gly	Val	Thr	Val	Leu 235	Val	Tyr	Leu	Pro	Phe 240
Phe	His	Ala	Phe	Gly 245	Phe	Ser	Ile	Asn	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly
Leu	Arg	Val	Ile 260	Met	Leu	Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys
Ala	Ile	Gln 275	Asp	Tyr	Glu	Val	Arg 280	Ser	Val	Ile	Asn	Val 285	Pro	Ala	Ile
Ile	Leu 290	Phe	Leu	Ser	Lys	Ser 295	Pro	Leu	Val	Asp	L y s 300	Tyr	Asp	Leu	Ser
Ser 305	Leu	Arg	Glu	Leu	Cys 310	Cys	Gly	Ala	Ala	Pro 315	Leu	Ala	Lys	Glu	Val 320
Ala	Glu	Val	Ala	Val 325	Lys	Arg	Leu	Asn	Leu 330	Pro	Gly	Ile	Arg	Cys 335	Gly
Phe	Gly	Leu	Thr 340	Glu	Ser	Thr	Ser	Ala 345	Asn	Ile	His	Ser	Leu 350	Gly	Asp
Glu	Phe	L y s 355	Ser	Gly	Ser	Leu	Gly 360	Arg	Val	Thr	Pro	Leu 365	Met	Ala	Ala

Lys Ile Ala Asp Arg Glu Thr Gly Lys Ala Leu Gly Pro Asn Gln Val Gly Glu Leu Cys Val Lys Gly Pro Met Val Ser Lys Gly Tyr Val Asn Asn Val Glu Ala Thr Lys Glu Ala Ile Asp Asp Asp Gly Trp Leu His Ser Gly Asp Phe Gly Tyr Tyr Asp Glu Asp Glu His Phe Tyr Val Val Asp Arg Tyr Lys Glu Leu Ile Lys Tyr Lys Gly Ser Gln Val Ala Pro $435 \ \ \, 440 \ \ \, 445$ Ala Glu Leu Glu Glu Ile Leu Leu Lys Asn Pro Cys Ile Arg Asp Val Ala Val Val Gly Ile Pro Asp Leu Glu Ala Gly Glu Leu Pro Ser Ala 465 470470475475 Phe Val Val Lys Gln Pro Gly Lys Glu Ile Thr Ala Lys Glu Val Tyr Asp Tyr Leu Ala Glu Arg Val Ser His Thr Lys Tyr Leu Arg Gly Gly Arg Lys Glu Leu Leu Lys Gln Leu Leu Glu Lys Ser Ser Lys Leu $530 \hspace{1.5cm} 535 \hspace{1.5cm} 540 \hspace{1.5cm}$ <210> SEQ ID NO 24 <211> LENGTH: 542 <212> TYPE: PRT <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of clone YG#81-6G01 <400> SEQUENCE: 24 Met Met Lys Arg Glu Lys Asn Val Ile Tyr Gly Pro Glu Pro Leu His Pro Leu Glu Asp Leu Thr Ala Gly Glu Met Leu Phe Arg Ala Leu Arg 20 25 30Lys His Ser His Leu Pro Gln Ala Leu Val Asp Val Val Gly Asp Glu 40 Ser Leu Ser Tyr Lys Glu Phe Phe Glu Ala Thr Val Leu Leu Ala Gln Ser Leu His Asn Cys Gly Tyr Lys Met Asn Asp Val Val Ser Ile Cys Ala Glu Asn Asn Thr Arg Phe Phe Ile Pro Val Ile Ala Ala Trp Tyr Ile Gly Met Ile Val Ala Pro Val Asn Glu Ser Tyr Ile Pro Asp Glu Leu Cys Lys Val Met Gly Ile Ser Lys Pro Gln Ile Val Phe Thr Thr Lys Asn Ile Leu Asn Lys Val Leu Glu Val Gln Ser Arg Thr Asn Phe

Asn Phe Lys Pro Leu His Phe Asp Pro Val Glu Gln Val Ala Ala Ile 185 Leu Cys Ser Ser Gly Thr Thr Gly Leu Pro Lys Gly Val Met Gln Thr $195 \hspace{1.5cm} 200 \hspace{1.5cm} 205 \hspace{1.5cm}$ His Gln Asn Ile Cys Val Arg Leu Ile His Ala Leu Asp Pro Arg Ala Gly Thr Gln Leu Ile Pro Gly Val Thr Val Leu Val Tyr Leu Pro Phe 235 Phe His Ala Phe Gly Phe Ser Ile Thr Leu Gly Tyr Phe Met Val Gly 250 Leu Arg Val Ile Met Phe Arg Arg Phe Asp Gln Glu Ala Phe Leu Lys Ala Ile Gln Asp Tyr Glu Val Arg Ser Val Ile Asn Val Pro Ser Val Ile Leu Phe Leu Ser Lys Ser Pro Leu Val Asp Lys Tyr Asp Leu Ser 295 Ser Leu Arg Glu Leu Cys Cys Gly Ala Ala Pro Leu Ala Lys Glu Val Ala Glu Val Ala Ala Lys Arg Leu Asn Leu Pro Gly Ile Arg Cys Gly $325 \hspace{1.5cm} 330 \hspace{1.5cm} 335$ Phe Gly Leu Thr Glu Ser Thr Ser Ala Asn Ile His Ser Leu Arg Asp Glu Phe Lys Ser Gly Ser Leu Gly Arg Val Thr Pro Leu Met Ala Ala 355 360 Lys Ile Ala Asp Arg Glu Thr Gly Lys Ala Leu Gly Pro Asn Gln Val\$370\$Gly Glu Leu Cys Ile Lys Gly Pro Met Val Ser Lys Gly Tyr Val Asn 385 390395395 Asn Val Glu Ala Thr Lys Glu Ala Ile Asp Asp Asp Gly Trp Leu His 410 Ser Gly Asp Phe Gly Tyr Tyr Asp Glu Asp Glu His Phe Tyr Val Val 425 Asp Arg Tyr Lys Glu Leu Ile Lys Tyr Lys Gly Ser Gln Val Ala Pro $435 \ \ \, 440 \ \ \, 445$ Ala Glu Leu Glu Glu Ile Leu Leu Lys Asn Pro Cys Ile Arg Asp Val Ala Val Val Gly Ile Pro Asp Leu Glu Ala Gly Glu Leu Pro Ser Ala 465 470470475475 Phe Val Val Lys Gln Pro Gly Lys Glu Ile Thr Ala Lys Glu Val Tyr 490 Asp Tyr Leu Ala Glu Arg Val Ser His Thr Lys Tyr Leu Arg Gly Gly Val Arg Phe Val Asp Ser Ile Pro Arg Asn Val Thr Gly Lys Ile Thr Arg Lys Glu Leu Leu Lys Gln Leu Leu Glu Lys Ala Gly Gly

<210> SEQ ID NO 25

<211> LENGTH: 542

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> 0	THER	INFO	ORMAT	CION:	Sec	quenc	e of	as	synth	netio	c luc	cifer	case	
<400> S	EQUEI	NCE:	25											
Met Met 1	Lys	Arg	Glu 5	Lys	Asn	Val	Ile	Ty r 10	Gly	Pro	Glu	Pro	Leu 15	His
Pro Leu	Glu	Asp 20	Leu	Thr	Ala	Gly	Glu 25	Met	Leu	Phe	Arg	Ala 30	Leu	Arg
Lys His	Ser 35	His	Leu	Pro	Gln	Ala 40	Leu	Val	Asp	Val	Val 45	Gly	Asp	Glu
Ser Leu 50	Ser	Tyr	Lys	Glu	Phe 55	Phe	Glu	Ala	Thr	Val 60	Leu	Leu	Ala	Gln
Ser Leu 65	His	Asn	Cys	Gly 70	Tyr	Lys	Met	Asn	Asp 75	Val	Val	Ser	Ile	Cys 80
Ala Glu	Asn	Asn	Thr 85	Arg	Phe	Phe	Ile	Pro 90	Val	Ile	Ala	Ala	Trp 95	Tyr
Ile Gly	Met	Ile 100	Val	Ala	Pro	Val	Asn 105	Glu	Ser	Tyr	Ile	Pro 110	Asp	Glu
Leu Cys	L y s 115	Val	Met	Gly	Ile	Ser 120	Lys	Pro	Gln	Ile	Val 125	Phe	Thr	Thr
Lys Asn 130		Leu	Asn	Lys	Val 135	Leu	Glu	Val	Gln	Ser 140	Arg	Thr	Asn	Phe
Ile Lys 145	Arg	Ile	Ile	Ile 150	Leu	Asp	Thr	Val	Glu 155	Asn	Ile	His	Gly	Cys 160
Glu Ser	Leu	Pro	Asn 165	Phe	Ile	Ser	Arg	Ty r 170	Ser	Asp	Gly	Asn	Ile 175	Ala
Asn Phe	L y s	Pro 180	Leu	His	Phe	Asp	Pro 185	Val	Glu	Gln	Val	Ala 190	Ala	Ile
Leu Cys	Ser 195	Ser	Gly	Thr	Thr	Gly 200	Leu	Pro	Lys	Gly	Val 205	Met	Gln	Thr
His Glr 210		Ile	Cys	Val	Arg 215	Leu	Ile	His	Ala	Leu 220	Asp	Pro	Arg	Val
Gly Thr 225	Gln	Leu	Ile	Pro 230	Gly	Val	Thr	Val	Leu 235	Val	Tyr	Leu	Pro	Phe 240
Phe His	Ala	Phe	Gly 245	Phe	Ser	Ile	Thr	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly
Leu Arg	Val	Ile 260	Met	Phe	Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys
Ala Ile	Gln 275	Asp	Tyr	Glu	Val	Arg 280	Ser	Val	Ile	Asn	Val 285	Pro	Ser	Val
Ile Leu 290		Leu	Ser	Lys	Ser 295	Pro	Leu	Val	Asp	Lys 300	Tyr	Asp	Leu	Ser
Ser Leu 305	Arg	Glu	Leu	Cys 310	Cys	Gly	Ala	Ala	Pro 315	Leu	Ala	Lys	Glu	Val 320
Ala Glu	Val	Ala	Ala 325	Lys	Arg	Leu	Asn	Leu 330	Pro	Gly	Ile	Arg	Cys 335	Gly
Phe Gly	Leu	Thr 340	Glu	Ser	Thr	Ser	Ala 345	Asn	Ile	His	Ser	Leu 350	Arg	Asp
Glu Phe	L y s 355	Ser	Gly	Ser	Leu	Gly 360	Arg	Val	Thr	Pro	Leu 365	Met	Ala	Ala
Lys Ile		Asp	Arg	Glu	Thr 375	Gly	Lys	Ala	Leu	Gly 380	Pro	Asn	Gln	Val

Gly 385															
	Glu	Leu	Cys	Ile	L y s 390	Gly	Pro	Met	Val	Ser 395	Lys	Gly	Tyr	Val	Asn 400
Asn	Val	Glu	Ala	Thr 405	Lys	Glu	Ala	Ile	Asp 410	Asp	Asp	Gly	Trp	Leu 415	His
Ser	Gly	Asp	Phe 420	Gly	Tyr	Tyr	Asp	Glu 425	Asp	Glu	His	Phe	Ty r 430	Val	Val
Asp	Arg	Ty r 435	Lys	Glu	Leu	Ile	Lys 440	Tyr	Lys	Gly	Ser	Gln 445	Val	Ala	Pro
Ala	Glu 450	Leu	Glu	Glu	Ile	Leu 455	Leu	Lys	Asn	Pro	Cys 460	Ile	Arg	Asp	Val
Ala 465	Val	Val	Gly	Ile	Pro 470	Asp	Leu	Glu	Ala	Gl y 475	Glu	Leu	Pro	Ser	Ala 480
Phe	Val	Val	Lys	Gln 485	Pro	Gly	Lys	Glu	Ile 490	Thr	Ala	Lys	Glu	Val 495	Tyr
Asp	Tyr	Leu	Ala 500	Glu	Arg	Val	Ser	His 505	Thr	Lys	Tyr	Leu	Arg 510	Gly	Gly
Val	Arg	Phe 515	Val	Asp	Ser	Ile	Pro 520	Arg	Asn	Val	Thr	Gly 525	Lys	Ile	Thr
Arg	Lys 530	Glu	Leu	Leu	Lys	Gln 535	Leu	Leu	Glu	Lys	Ala 540	Gly	Gly		
<220 <220	3> OF 0> FE 3> OT 0> SE	ATUF HER	RE: INFO	RMAT			-		as	ynth	netio	: luc	cifer	ase	
		-GOEL	ICE:	26											
Met 1	Met	-			Lys	Asn	Val	Ile	Ty r	Gly	Pro	Glu	Pro	Leu 15	His
1		Lys	Arg	Glu 5	-				10	-				15	
1 Pro	Met	L y s Glu	Arg Asp 20	Glu 5 Leu	Thr	Ala	Gly	Glu 25	10 Met	Leu	Phe	Arg	Ala 30	15 Leu	Arg
1 Pro Lys	Met Leu	Lys Glu Ser 35	Arg Asp 20 His	Glu 5 Leu Leu	Thr Pro	Ala Gln	Gly Ala 40	Glu 25 Leu	10 Met Val	Leu Asp	Phe Val	Arg Val 45	Ala 30 Gly	15 Leu Asp	Arg Glu
1 Pro Lys Ser	Met Leu His	Lys Glu Ser 35 Ser	Arg Asp 20 His	Glu 5 Leu Leu	Thr Pro Glu	Ala Gln Phe 55	Gly Ala 40 Phe	Glu 25 Leu Glu	10 Met Val Ala	Leu Asp Thr	Phe Val Val 60	Arg Val 45 Leu	Ala 30 Gly Leu	15 Leu Asp Ala	Arg Glu Gln
1 Pro Lys Ser Ser 65	Met Leu His Leu 50	Lys Glu Ser 35 Ser	Arg Asp 20 His Tyr	Glu 5 Leu Leu Lys	Thr Pro Glu Gly 70	Ala Gln Phe 55	Gly Ala 40 Phe	Glu 25 Leu Glu Met	10 Met Val Ala Asn	Leu Asp Thr Asp 75	Phe Val Val 60 Val	Arg Val 45 Leu Val	Ala 30 Gly Leu Ser	15 Leu Asp Ala Ile	Arg Glu Gln Cys 80
1 Pro Lys Ser Ser 65 Ala	Met Leu His Leu 50	Lys Glu Ser 35 Ser His	Arg Asp 20 His Tyr Asn	Glu 5 Leu Leu Lys Cys Thr 85	Thr Pro Glu Gly 70 Arg	Ala Gln Phe 55 Tyr	Gly Ala 40 Phe Lys	Glu 25 Leu Glu Met	10 Met Val Ala Asn Pro	Leu Asp Thr Asp 75	Phe Val Val 60 Val	Arg Val 45 Leu Val	Ala 30 Gly Leu Ser	15 Leu Asp Ala Ile Trp 95	Arg Glu Gln Cys 80 Tyr
1 Pro Lys Ser Ser 65 Ala	Met Leu His Leu 50 Leu	Lys Glu Ser 35 Ser His Asn	Arg Asp 20 His Tyr Asn Ile 100	Glu 5 Leu Leu Lys Cys Thr 85 Val	Thr Pro Glu Gly 70 Arg	Ala Gln Phe 55 Tyr Phe	Gly Ala 40 Phe Lys Val	Glu 25 Leu Glu Met Ile Asn 105	10 Met Val Ala Asn Pro 90 Glu	Leu Asp Thr Asp 75 Val	Phe Val Val 60 Val Ile	Arg Val 45 Leu Val Ala	Ala 30 Gly Leu Ser Ala Pro	Leu Asp Ala Ile Trp 95 Asp	Arg Glu Gln Cys 80 Tyr
1 Pro Lys Ser 65 Ala Ile	Met Leu His Leu 50 Leu Glu	Lys Glu Ser 35 Ser His Asn Met Lys 115	Arg Asp 20 His Tyr Asn Asn Ile 100 Val	Glu 5 Leu Lys Cys Thr 85 Val	Thr Pro Glu Gly 70 Arg Ala	Ala Gln Phe 55 Tyr Phe Ile	Gly Ala 40 Phe Lys Phe Val Ser 120	Glu 25 Leu Glu Met Ile Asn 105 Lys	Met Val Ala Asn Pro 90 Glu Pro	Leu Asp Thr Asp 75 Val Ser	Phe Val Val 60 Val Ile Tyr	Arg Val 45 Leu Val Ala Ile Val 125	Ala 30 Gly Leu Ser Ala Pro 110	15 Leu Asp Ala Ile Trp 95 Asp	Arg Glu Gln Cys 80 Tyr Glu Thr
1 Pro Lys Ser Ser 65 Ala Ile Leu	Met Leu His 50 Leu Glu Gly Cys Asn	Lys Glu Ser 35 Ser His Asn Met Lys 115	Arg Asp 20 His Tyr Asn Ile 100 Val	Glu 5 Leu Lys Cys Thr 85 Val Met Asn	Thr Pro Glu Gly 70 Arg Ala Gly Lys	Ala Gln Phe 55 Tyr Phe Ule Val 135	Gly Ala 40 Phe Lys Phe Val Ser 120 Leu	Glu 25 Leu Glu Met Ile Asn 105 Lys Glu	Met Val Ala Asn Pro 90 Glu Pro Val	Leu Asp Thr Asp 75 Val Ser Gln	Phe Val Val 60 Val Ile Tyr Ile Ser 140	Arg Val 45 Leu Val Ala Ile Val 125 Arg	Ala 30 Gly Leu Ser Ala Pro 110 Phe	15 Leu Asp Ala Ile Trp 95 Asp	Arg Glu Gln Cys 80 Tyr Glu Thr
1 Pro Lys Ser Ser Ala Ile Leu Lys Ile 145	Met Leu His Leu 50 Leu Glu Gly Cys Asn 130	Lys Glu Ser 35 Ser His Asn Met Lys 115 Ile Arg	Arg Asp 20 His Tyr Asn Asn Ule 100 Ual	Glu 5 Leu Lys Cys Thr 85 Val Met Asn	Thr Pro Glu Gly 70 Arg Ala Gly Lys Ile 150	Ala Gln Phe 55 Tyr Phe Val 135 Leu	Gly Ala 40 Phe Lys Phe Val Ser 120 Leu Asp	Glu 25 Leu Glu Met Ile Asn 105 Lys Glu	Met Val Ala Asn Pro 90 Glu Pro Val	Leu Asp Thr Asp 75 Val Ser Gln Gln Glu 155	Phe Val Val 60 Val Ile Tyr Ile Ser 140 Asn	Arg Val 45 Leu Val Ala Ile Val 125 Arg	Ala 30 Gly Leu Ser Ala Pro 110 Phe Thr	15 Leu Asp Ala Ile Trp 95 Asp Thr Asn	Arg Glu Gln Cys 80 Tyr Glu Thr Phe Cys 160

<400> SEQUENCE: 27

Leu Cys Ser Se 195	er Gly Thr	Thr Gly 1	Leu Pro L y s	Gly Val 205	Met Gln Thr
His Gln Asn I	e Cys Val.	Arg Leu : 215	Ile His Ala	Leu Asp 220	Pro Arg Val
Gly Thr Gln Le	eu Ile Pro 230	Gly Val	Thr Val Leu 235	Val Tyr	Leu Pro Phe 240
Phe His Ala Pl	e Gly Phe 245	Ser Ile	Thr Leu Gly 250	Tyr Phe	Met Val Gly 255
Leu Arg Val I			Phe Asp Gln 265	Glu Ala	Phe Leu L y s 270
Ala Ile Gln A 275	sp Tyr Glu	Val Arg s	Ser Val Ile	Asn Val 285	Pro Ser Val
Ile Leu Phe Le 290	eu Ser Lys	Ser Pro 1 295	Leu Val Asp	Lys Tyr 300	Asp Leu Ser
Ser Leu Arg G	u Leu Cys 310	Cys Gly A	Ala Ala Pro 315	Leu Ala	Lys Glu Val 320
Ala Glu Val A	a Ala Lys 325	Arg Leu A	Asn Leu Pro 330	Gly Ile	Arg Cys Gly 335
Phe Gly Leu Tl	ır Glu Ser 10		Ala Asn Ile 345	His Ser	Leu Arg Asp 350
Glu Phe Lys So	er Gly Ser	Leu Gly 2	Arg Val Thr	Pro Leu 365	Met Ala Ala
Lys Ile Ala As	sp Arg Glu	Thr Gly 1	Lys Ala Leu	Gly Pro 380	Asn Gln Val
Gly Glu Leu Cy 385	rs Ile Lys 390	Gly Pro I	Met Val Ser 395	Lys Gly	Tyr Val Asn 400
Asn Val Glu A	a Thr Lys 405	Glu Ala :	Ile Asp Asp 410	Asp Gly	Trp Leu His 415
Ser Gly Asp Pl			Glu Asp Glu 425	His Phe	Tyr Val Val 430
Asp Arg Tyr Ly 435	rs Glu Leu	Ile Lys 7	Tyr Lys Gly	Ser Gln 445	Val Ala Pro
Ala Glu Leu G 450	u Glu Ile	Leu Leu 1 455	L y s Asn Pro	Cys Ile 460	Arg Asp Val
Ala Val Val G 465	y Ile Pro 470	Asp Leu (Glu Ala Gly 475	Glu Leu	Pro Ser Ala 480
Phe Val Val L	s Gln Pro 485	Gly Lys	Glu Ile Thr 490	Ala Lys	Glu Val Ty r 495
Asp Tyr Leu A	-		His Thr Lys 505	Tyr Leu	Arg Gly Gly 510
Val Arg Phe Va 515	ıl Asp Ser	Ile Pro 2 520	Arg Asn Val	Thr Gly 525	Lys Ile Thr
Arg Lys Glu Lo	eu Leu Lys	Gln Leu 1 535	Leu Glu Lys	Ala Gly 540	Gly
<210> SEQ ID N <211> LENGTH: <212> TYPE: PH <213> ORGANISN <220> FEATURE: <223> OTHER IN	542 T : Artificia	_		etic luc	iferase

Met 1	Met	Lys	Arg	Glu 5	Lys	Asn	Val	Ile	Tyr 10	Gly	Pro	Glu	Pro	Leu 15	His
Pro	Leu	Glu	Asp 20	Leu	Thr	Ala	Gly	Glu 25	Met	Leu	Phe	Arg	Ala 30	Leu	Arg
Lys	His	Ser 35	His	Leu	Pro	Gln	Ala 40	Leu	Val	Asp	Val	Val 45	Gly	Asp	Glu
Ser	Leu 50	Ser	Tyr	Lys	Glu	Phe 55	Phe	Glu	Ala	Thr	Val 60	Leu	Leu	Ala	Gln
Ser 65	Leu	His	Asn	Cys	Gl y 70	Tyr	Lys	Met	Asn	Asp 75	Val	Val	Ser	Ile	C y s 80
Ala	Glu	Asn	Asn	Thr 85	Arg	Phe	Phe	Ile	Pro 90	Val	Ile	Ala	Ala	Trp 95	Tyr
Ile	Gly	Met	Ile 100	Val	Ala	Pro	Val	Asn 105	Glu	Ser	Tyr	Ile	Pro 110	Asp	Glu
Leu	Суѕ	Lys 115	Val	Met	Gly	Ile	Ser 120	Lys	Pro	Gln	Ile	Val 125	Phe	Thr	Thr
Lys	Asn 130	Ile	Leu	Asn	Lys	Val 135	Leu	Glu	Val	Gln	Ser 140	Arg	Thr	Asn	Phe
Ile 145	Lys	Arg	Ile	Ile	Ile 150	Leu	Asp	Thr	Val	Glu 155	Asn	Ile	His	Gly	Cys 160
Glu	Ser	Leu	Pro	Asn 165	Phe	Ile	Ser	Arg	Ty r 170	Ser	Asp	Gly	Asn	Ile 175	Ala
Asn	Phe	Lys	Pro 180	Leu	His	Phe	Asp	Pro 185	Val	Glu	Gln	Val	Ala 190	Ala	Ile
Leu	Cys	Ser 195	Ser	Gly	Thr	Thr	Gly 200	Leu	Pro	Lys	Gly	Val 205	Met	Gln	Thr
His	Gln 210	Asn	Ile	Cys	Val	Arg 215	Leu	Ile	His	Ala	Leu 220	Asp	Pro	Arg	Val
Gl y 225	Thr	Gln	Leu	Ile	Pro 230	Gly	Val	Thr	Val	Leu 235	Val	Tyr	Leu	Pro	Phe 240
Phe	His	Ala	Phe	Gly 245	Phe	Ser	Ile	Thr	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly
Leu	Arg	Val	Ile 260	Met	Phe	Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys
Ala	Ile	Gln 275	Asp	Tyr	Glu	Val	Arg 280	Ser	Val	Ile	Asn	Val 285	Pro	Ser	Val
	Leu 290		Leu	Ser		Ser 295		Leu	Val		L y s 300	Tyr	Asp	Leu	Ser
Ser 305	Leu	Arg	Glu	Leu	Cys 310	Cys	Gly	Ala	Ala	Pro 315	Leu	Ala	Lys	Glu	Val 320
Ala	Glu	Val	Ala	Ala 325	Lys	Arg	Leu	Asn	Leu 330	Pro	Gly	Ile	Arg	Cys 335	Gly
Phe	Gly	Leu	Thr 340	Glu	Ser	Thr	Ser	Ala 345	Asn	Ile	His	Ser	Leu 350	Arg	Asp
Glu	Phe	Lys 355	Ser	Gly	Ser	Leu	Gl y 360	Arg	Val	Thr	Pro	Leu 365	Met	Ala	Ala
Lys	Ile 370	Ala	Asp	Arg	Glu	Thr 375	Gly	Lys	Ala	Leu	Gly 380	Pro	Asn	Gln	Val
Gly 385	Glu	Leu	Сув	Ile	L y s 390	Gly	Pro	Met	Val	Ser 395	Lys	Gly	Tyr	Val	Asn 400
Asn	Val	Glu	Ala	Thr	Lys	Glu	Ala	Ile	Asp	Asp	Asp	Gly	Trp	Leu	His

405

-continued

415

410

				405					410					413	
Ser	Gly	Asp	Phe 420	Gly	Tyr	Tyr	Asp	Glu 425	Asp	Glu	His	Phe	Tyr 430	Val	Val
Asp	Arg	Tyr 435	Lys	Glu	Leu	Ile	Lys 440	Tyr	Lys	Gly	Ser	Gln 445	Val	Ala	Pro
Ala	Glu 450	Leu	Glu	Glu	Ile	Leu 455	Leu	Lys	Asn	Pro	Cys 460	Ile	Arg	Asp	Val
Ala 465	Val	Val	Gly	Ile	Pro 470	Asp	Leu	Glu	Ala	Gly 475	Glu	Leu	Pro	Ser	Ala 480
Phe	Val	Val	Lys	Gln 485	Pro	Gly	Lys	Glu	Ile 490	Thr	Ala	Lys	Glu	Val 495	Tyr
Asp	Tyr	Leu	Ala 500	Glu	Arg	Val	Ser	His 505	Thr	Lys	Tyr	Leu	Arg 510	Gly	Gly
Val	Arg	Phe 515	Val	Asp	Ser	Ile	Pro 520	Arg	Asn	Val	Thr	Gly 525	Lys	Ile	Thr
Arg	Lys 530	Glu	Leu	Leu	Lys	Gln 535	Leu	Leu	Glu	Lys	Ala 540	Gly	Gly		
<211 <212 <213 <220)> SE .> LE ?> TY 3> OR 0> FE 3> OT	NGTH PE: GANI ATUR	: 54 PRT SM:	2 Arti					i a s	ynth	ıetic	: luc	ifer	ase	
<400	> SE	QUEN	CE:	28											
Met 1	Met	Lys	Arg	Glu 5	Lys	Asn	Val	Ile	Ty r 10	Gly	Pro	Glu	Pro	Leu 15	His
Pro	Leu	Glu	Asp 20	Leu	Thr	Ala	Gly	Glu 25	Met	Leu	Phe	Arg	Ala 30	Leu	Arg
Lys	His	Ser 35	His	Leu	Pro	Gln	Ala 40	Leu	Val	Asp	Val	Val 45	Gly	Asp	Glu
Ser	Leu 50	Ser	Tyr	Lys	Glu	Phe 55	Phe	Glu	Ala	Thr	Val 60	Leu	Leu	Ala	Gln
Ser 65	Leu	His	Asn	Суѕ	Gly 70	Tyr	Lys	Met	Asn	Asp 75	Val	Val	Ser	Ile	C y s 80
Ala	Glu	Asn	Asn	Thr 85	Arg	Phe	Phe	Ile	Pro 90	Val	Ile	Ala	Ala	Trp 95	Tyr
Ile	Gly	Met	Ile 100	Val	Ala	Pro	Val	Asn 105	Glu	Ser	Tyr	Ile	Pro 110	Asp	Glu
Leu	Суѕ	L y s 115	Val	Met	Gly	Ile	Ser 120	Lys	Pro	Gln	Ile	Val 125	Phe	Thr	Thr
Lys	Asn 130	Ile	Leu	Asn	Lys	Val 135	Leu	Glu	Val	Gln	Ser 140	Arg	Thr	Asn	Phe
Ile 145	Lys	Arg	Ile	Ile	Ile 150	Leu	Asp	Thr	Val	Glu 155	Asn	Ile	His	Gly	C y s 160
Glu	Ser	Leu	Pro	Asn 165	Phe	Ile	Ser	Arg	Ty r 170	Ser	Asp	Gly	Asn	Ile 175	Ala
Asn	Phe	Lys	Pro 180	Leu	His	Phe	Asp	Pro 185	Val	Glu	Gln	Val	Ala 190	Ala	Ile
Leu	Cys	Ser 195	Ser	Gly	Thr	Thr	Gly 200	Leu	Pro	Lys	Gly	Val 205	Met	Gln	Thr
His	Gln	Asn	Ile	Cys	Val	Arg	Leu	Ile	His	Ala	Leu	Asp	Pro	Arg	Val

215

210

-continued

220

Gly 225	Thr	Gln	Leu	Ile	Pro 230	Gly	Val	Thr	Val	Leu 235	Val	Tyr	Leu	Pro	Phe 240
Phe	His	Ala	Phe	Gly 245	Phe	Ser	Ile	Thr	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly
Leu	Arg	Val	Ile 260	Met	Phe	Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys
Ala	Ile	Gln 275	Asp	Tyr	Glu	Val	Arg 280	Ser	Val	Ile	Asn	Val 285	Pro	Ser	Val
Ile	Leu 290	Phe	Leu	Ser	Lys	Ser 295	Pro	Leu	Val	Asp	L y s 300	Tyr	Asp	Leu	Ser
Ser 305	Leu	Arg	Glu	Leu	C y s 310	Сув	Gly	Ala	Ala	Pro 315	Leu	Ala	Lys	Glu	Val 320
Ala	Glu	Val	Ala	Ala 325	Lys	Arg	Leu	Asn	Leu 330	Pro	Gly	Ile	Arg	Cys 335	Gly
Phe	Gly	Leu	Thr 340	Glu	Ser	Thr	Ser	Ala 345	Asn	Ile	His	Ser	Leu 350	Arg	Asp
Glu	Phe	Ly s 355	Ser	Gly	Ser	Leu	Gl y 360	Arg	Val	Thr	Pro	Leu 365	Met	Ala	Ala
Lys	Ile 370	Ala	Asp	Arg	Glu	Thr 375	Gly	Lys	Ala	Leu	Gly 380	Pro	Asn	Gln	Val
Gly 385	Glu	Leu	Cys	Ile	L y s 390	Gly	Pro	Met	Val	Ser 395	Lys	Gly	Tyr	Val	Asn 400
Asn	Val	Glu	Ala	Thr 405	Lys	Glu	Ala	Ile	Asp 410	Asp	Asp	Gly	Trp	Leu 415	His
Ser	Gly	Asp	Phe 420	Gly	Tyr	Tyr	Asp	Glu 425	Asp	Glu	His	Phe	Ty r 430	Val	Val
Asp	Arg	Tyr 435	Lys	Glu	Leu	Ile	Lys 440	Tyr	Lys	Gly	Ser	Gln 445	Val	Ala	Pro
Ala	Glu 450	Leu	Glu	Glu	Ile	Leu 455	Leu	Lys	Asn	Pro	Cys 460	Ile	Arg	Asp	Val
Ala 465	Val	Val	Gly	Ile	Pro 470	Asp	Leu	Glu	Ala	Gly 475	Glu	Leu	Pro	Ser	Ala 480
Phe	Val	Val	Lys	Gln 485	Pro	Gly	Lys	Glu	Ile 490	Thr	Ala	Lys	Glu	Val 495	Tyr
Asp	Tyr	Leu	Ala 500	Glu	Arg	Val	Ser	His 505	Thr	Lys	Tyr	Leu	Arg 510	Gly	Gly
Val	Arg	Phe 515		Asp	Ser		Pro 520		Asn	Val	Thr	Gly 525		Ile	Thr
Arg	L y s 530	Glu	Leu	Leu	Lys	Gln 535	Leu	Leu	Glu	Lys	Ala 540	Gly	Gly		
<211 <212 <213 <220)> SE l> LE 2> TY 3> OF 0> FE 3> OT	NGTH PE: RGANI ATUF	: 54 PRT SM: RE:	l2 Arti			_		i a s	ynth	netio	: luc	cifer	ase	
)> SE				T	7	T7 c 7	т1 -	m	G1	D	G1	D	T	TT 2 -
Met 1	met	ьys	Arg	GIu 5	ьуѕ	Asn	val	11e	Tyr 10	GIY	Pro	GLU	Pro	Leu 15	Hls

Pro Leu Glu Asp Leu Thr Ala Gly Glu Met Leu Phe Arg Ala Leu Arg

			20					25					30		
Lys	His	Ser 35	His	Leu	Pro	Gln	Ala 40	Leu	Val	Asp	Val	Val 45	Gly	Asp	Glu
Ser	Leu 50	Ser	Tyr	Lys	Glu	Phe 55	Phe	Glu	Ala	Thr	Val 60	Leu	Leu	Ala	Gln
Ser 65	Leu	His	Asn	Cys	Gly 70	Tyr	Lys	Met	Asn	Asp 75	Val	Val	Ser	Ile	Cys 80
Ala	Glu	Asn	Asn	Thr 85	Arg	Phe	Phe	Ile	Pro 90	Val	Ile	Ala	Ala	Trp 95	Tyr
Ile	Gly	Met	Ile 100	Val	Ala	Pro	Val	Asn 105	Glu	Ser	Tyr	Ile	Pro 110	Asp	Glu
Leu	Cys	L y s 115	Val	Met	Gly	Ile	Ser 120	Lys	Pro	Gln	Ile	Val 125	Phe	Thr	Thr
Lys	Asn 130	Ile	Leu	Asn	Lys	Val 135	Leu	Glu	Val	Gln	Ser 140	Arg	Thr	Asn	Phe
Ile 145	Lys	Arg	Ile	Ile	Ile 150	Leu	Asp	Thr	Val	Glu 155	Asn	Ile	His	Gly	Cys 160
Glu	Ser	Leu	Pro	Asn 165	Phe	Ile	Ser	Arg	Ty r 170	Ser	Asp	Gly	Asn	Ile 175	Ala
Asn	Phe	Lys	Pro 180	Leu	His	Phe	Asp	Pro 185	Val	Glu	Gln	Val	Ala 190	Ala	Ile
Leu	Cys	Ser 195	Ser	Gly	Thr	Thr	Gly 200	Leu	Pro	Lys	Gly	Val 205	Met	Gln	Thr
His	Gln 210	Asn	Ile	Cys	Val	Arg 215	Leu	Ile	His	Ala	Leu 220	Asp	Pro	Arg	Val
Gl y 225	Thr	Gln	Leu	Ile	Pro 230	Gly	Val	Thr	Val	Leu 235	Val	Tyr	Leu	Pro	Phe 240
Phe	His	Ala	Phe	Gly 245	Phe	Ser	Ile	Thr	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly
Leu	Arg	Val	Ile 260	Met	Phe	Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys
Ala	Ile	Gln 275	Ąsp	Tyr	Glu	Val	Arg 280	Ser	Val	Ile	Asn	Val 285	Pro	Ser	Val
Ile	Leu 290	Phe	Leu	Ser	Lys	Ser 295	Pro	Leu	Val	Asp	L y s 300	Tyr	Asp	Leu	Ser
Ser 305	Leu	Arg	Glu	Leu	Cys 310	Сув	Gly	Ala	Ala	Pro 315	Leu	Ala	Lys	Glu	Val 320
Ala	Glu	Val	Ala	Ala 325	Lys	Arg	Leu	Asn	Leu 330	Pro	Gly	Ile	Arg	С у в 335	Gly
Phe	Gly	Leu	Thr 340	Glu	Ser	Thr	Ser	Ala 345	Asn	Ile	His	Ser	Leu 350	Arg	Asp
Glu	Phe	L y s 355	Ser	Gly	Ser	Leu	Gly 360	Arg	Val	Thr	Pro	Leu 365	Met	Ala	Ala
Lys	Ile 370	Ala	Asp	Arg	Glu	Thr 375	Gly	Lys	Ala	Leu	Gly 380	Pro	Asn	Gln	Val
Gl y 385	Glu	Leu	Cys	Ile	L y s 390	Gly	Pro	Met	Val	Ser 395	Lys	Gly	Tyr	Val	Asn 400
Asn	Val	Glu	Ala	Thr 405	Lys	Glu	Ala	Ile	Asp 410	Asp	Asp	Gly	Trp	Leu 415	His
Ser	Gly	Asp	Phe 420	Gly	Tyr	Tyr	Asp	Glu 425	Asp	Glu	His	Phe	Tyr 430	Val	Val

Asp Arg Tyr Lys Glu Leu Ile Lys Tyr Lys Gly Ser Gln Val Ala Pro $435 \ \ \, 440 \ \ \, 445$ Ala Glu Leu Glu Glu Ile Leu Leu Lys Asn Pro Cys Ile Arg Asp Val 455 Ala Val Val Gly Ile Pro Asp Leu Glu Ala Gly Glu Leu Pro Ser Ala 465 470470475475 Phe Val Val Lys Gln Pro Gly Lys Glu Ile Thr Ala Lys Glu Val Tyr Asp Tyr Leu Ala Glu Arg Val Ser His Thr Lys Tyr Leu Arg Gly Gly 505 <210> SEQ ID NO 30 <211> LENGTH: 542 <212> TYPE: PRT <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 30 Met Met Lys Arg Glu Lys Asn Val Ile Tyr Gly Pro Glu Pro Leu His 1 $$ 10 $$ 15 Pro Leu Glu Asp Leu Thr Ala Gly Glu Met Leu Phe Arg Ala Leu Arg 20 25 30Lys His Ser His Leu Pro Gln Ala Leu Val Asp Val Val Gly Asp Glu 35 40 45 Ser Leu His Asn Cys Gly Tyr Lys Met Asn Asp Val Val Ser Ile Cys 65 70 75 80 Ile Gly Met Ile Val Ala Pro Val Asn Glu Ser Tyr Ile Pro Asp Glu 105 Leu Cys Lys Val Met Gly Ile Ser Lys Pro Gln Ile Val Phe Thr Thr 120 Lys Asn Ile Leu Asn Lys Val Leu Glu Val Gln Ser Arg Thr Asn Phe Ile Lys Arg Ile Ile Ile Leu Asp Thr Val Glu Asn Ile His Gly Cys Glu Ser Leu Pro Asn Phe Ile Ser Arg Tyr Ser Asp Gly Asn Ile Ala Asn Phe Lys Pro Leu His Phe Asp Pro Val Glu Gln Val Ala Ala Ile Leu Cys Ser Ser Gly Thr Thr Gly Leu Pro Lys Gly Val Met Gln Thr 195 $$ 200 $$ 205 Gly Thr Gln Leu Ile Ser Gly Val Thr Val Leu Val Tyr Leu Pro Phe

Leu Arg Val Ile Met Phe Arg Arg Phe Asp Gln Glu Ala Phe Leu Lys 265 Ala Ile Gln Asp Tyr Glu Val Arg Ser Val Ile Asn Val Pro Ser Val Ile Leu Phe Leu Ser Lys Ser Pro Leu Val Asp Lys Tyr Asp Leu Ser 290 295 300 Ser Leu Arg Glu Leu Cys Cys Gly Ala Ala Pro Leu Ala Lys Glu Val 305 310310315 Ala Glu Val Ala Ala Lys Arg Leu Asn Leu Pro Gly Ile Arg Cys Gly Phe Gly Leu Thr Glu Ser Thr Ser Ala Asn Ile His Ser Leu Arg Asp Glu Phe Lys Ser Gly Ser Leu Gly Arg Val Thr Pro Leu Met Ala Ala Lys Ile Ala Asp Arg Glu Thr Gly Lys Ala Leu Gly Pro Asn Gln Val Gly Glu Leu Cys Ile Lys Gly Pro Met Val Ser Lys Gly Tyr Val Asn 385 390 395 400 Asn Val Glu Ala Thr Lys Glu Ala Ile Asp Asp Asp Gly Trp Leu His $405 \hspace{1.5cm} 410 \hspace{1.5cm} 415 \hspace{1.5cm}$ Ser Gly Asp Phe Gly Tyr Tyr Asp Glu Asp Glu His Phe Tyr Val Val 420 \$425\$Asp Arg Tyr Lys Glu Leu Ile Lys Tyr Lys Gly Ser Gln Val Ala Pro $435 \ \ \, 440 \ \ \, 445$ Ala Glu Leu Glu Glu Ile Leu Leu Lys Asn Pro Cys Ile Arg Asp Val $450 \ \ 460$ Ala Val Val Gly Ile Pro Asp Leu Glu Ala Gly Glu Leu Pro Ser Ala Phe Val Val Lys Gln Pro Gly Lys Glu Ile Thr Ala Lys Glu Val Tyr 490 Asp Tyr Leu Ala Glu Arg Val Ser His Thr Lys Tyr Leu Arg Gly Gly 500 505 510505 Arg Lys Glu Leu Leu Lys Gln Leu Leu Glu Lys Ala Gly Gly 530 540 <210> SEQ ID NO 31 <211> LENGTH: 542 <212> TYPE: PRT <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 31 Met Met Lys Arg Glu Lys Asn Val Ile Tyr Gly Pro Glu Pro Leu His Pro Leu Glu Asp Leu Thr Ala Gly Glu Met Leu Phe Arg Ala Leu Arg 20 25 30Lys His Ser His Leu Pro Gln Ala Leu Val Asp Val Val Gly Asp Glu $35 \ \ \,$ 40 $\ \ \,$ 45

Phe His Ala Phe Gly Phe Ser Ile Thr Leu Gly Tyr Phe Met Val Gly

250

Ser	Leu 50	Ser	Tyr	Lys	Glu	Phe 55	Phe	Glu	Ala	Thr	Val 60	Leu	Leu	Ala	Gln
Ser 65	Leu	His	Asn	Cys	Gly 70	Tyr	Lys	Met	Asn	Asp 75	Val	Val	Ser	Ile	Cys 80
Ala	Glu	Asn	Asn	Thr 85	Arg	Phe	Phe	Ile	Pro 90	Val	Ile	Ala	Ala	Trp 95	Tyr
Ile	Gly	Met	Ile 100	Val	Ala	Pro	Val	Asn 105	Glu	Ser	Tyr	Ile	Pro 110	Asp	Glu
Leu	Cys	L y s 115	Val	Met	Gly	Ile	Ser 120	Lys	Pro	Gln	Ile	Val 125	Phe	Thr	Thr
Lys	Asn 130	Ile	Leu	Asn	Lys	Val 135	Leu	Glu	Val	Gln	Ser 140	Arg	Thr	Asn	Phe
Ile 145	Lys	Arg	Ile	Ile	Ile 150	Leu	Asp	Thr	Val	Glu 155	Asn	Ile	His	Gly	Cys 160
Glu	Ser	Leu	Pro	Asn 165	Phe	Ile	Ser	Arg	Ty r 170	Ser	Asp	Gly	Asn	Ile 175	Ala
Asn	Phe	Lys	Pro 180	Leu	His	Phe	Asp	Pro 185	Val	Glu	Gln	Val	Ala 190	Ala	Ile
Leu	Cys	Ser 195	Ser	Gly	Thr	Thr	Gly 200	Leu	Pro	Lys	Gly	Val 205	Met	Gln	Thr
His	Gln 210	Asn	Ile	Cys	Val	Arg 215	Leu	Ile	His	Ala	Leu 220	Asp	Pro	Arg	Val
Gl y 225	Thr	Gln	Leu	Ile	Pro 230	Gly	Val	Thr	Val	Leu 235	Val	Tyr	Leu	Pro	Phe 240
Phe	His	Ala	Phe	Gl y 245	Phe	Ser	Ile	Thr	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly
Leu	Arg	Val	Ile 260	Met	Phe	Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys
Ala	Ile	Gln 275	Asp	Tyr	Glu	Val	Arg 280	Ser	Val	Ile	Asn	Val 285	Pro	Ser	Val
Ile	Leu 290	Phe	Leu	Ser	Lys	Ser 295	Pro	Leu	Val	Asp	L y s 300	Tyr	Asp	Leu	Ser
Ser 305	Leu	Arg	Glu	Leu	Cys 310	Cys	Gly	Ala	Ala	Pro 315	Leu	Ala	Lys	Glu	Val 320
Ala	Glu	Val	Ala	Ala 325	Lys	Arg	Leu	Asn	Leu 330	Pro	Gly	Ile	Arg	Cys 335	Gly
Phe	Gly	Leu	Thr 340	Glu	Ser	Thr	Ser	Ala 345	Asn	Ile	His	Ser	Leu 350	Arg	Asp
Glu	Phe	L y s 355	Ser	Gly	Ser	Leu	Gl y 360	Arg	Val	Thr	Pro	Leu 365	Met	Ala	Ala
Lys	Ile 370	Ala	Asp	Arg	Glu	Thr 375	Gly	Lys	Ala	Leu	Gly 380	Pro	Asn	Gln	Val
Gl y 385	Glu	Leu	Сув	Ile	Lys 390	Gly	Pro	Met	Val	Ser 395	Lys	Gly	Tyr	Val	Asn 400
Asn	Val	Glu	Ala	Thr 405	Lys	Glu	Ala	Ile	Asp 410	Asp	Asp	Gly	Trp	Leu 415	His
Ser	Gly	Asp	Phe 420	Gly	Tyr	Tyr	Asp	Glu 425	Asp	Glu	His	Phe	Tyr 430	Val	Val
Asp	Arg	Ty r 435	Lys	Glu	Leu	Ile	L y s 440	Tyr	Lys	Gly	Ser	Gln 445	Val	Ala	Pro

Ala	Glu 450	Leu	Glu	Glu	Ile	Leu 455	Leu	Lys	Asn	Pro	Cys 460	Ile	Arg	Asp	Val
Ala 465	Val	Val	Gly	Ile	Pro 470	Asp	Leu	Glu	Ala	Gl y 475	Glu	Leu	Pro	Ser	Ala 480
Phe	Val	Val	Lys	Gln 485	Pro	Gly	Lys	Glu	Ile 490	Thr	Ala	Lys	Glu	Val 495	Tyr
Asp	Tyr	Leu	Ala 500	Glu	Arg	Val	Ser	His 505	Thr	Lys	Tyr	Leu	Arg 510	Gly	Gly
Val	Arg	Phe 515	Val	Asp	Ser	Ile	Pro 520	Arg	Asn	Val	Thr	Gl y 525	Lys	Ile	Thr
Arg	Lys 530	Glu	Leu	Leu	Lys	Gln 535	Leu	Leu	Glu	Lys	Ala 540	Gly	Gly		
<211 <212 <213 <220	.> LE !> TY !> OR !> FE	ATUF	: 54 PRT SM: E:	2 Arti			_	ice e of	as	ynth	etic	: luc	ifer	ase	
<400	> SE	QUEN	ICE:	32											
Met 1	Met	Lys	Arg	Glu 5	Lys	Asn	Val	Ile	Ty r 10	Gly	Pro	Glu	Pro	Leu 15	His
Pro	Leu	Glu	Asp 20	Leu	Thr	Ala	Gly	Glu 25	Met	Leu	Phe	Arg	Ala 30	Leu	Arg
Lys	His	Ser 35	His	Leu	Pro	Gln	Ala 40	Leu	Val	qaA	Val	Val 45	Gly	Asp	Glu
Ser	Leu 50	Ser	Tyr	Lys	Glu	Phe 55	Phe	Glu	Ala	Thr	Val 60	Leu	Leu	Ala	Gln
Ser 65	Leu	His	Asn	Cys	Gly 70	Tyr	Lys	Met	Asn	Asp 75	Val	Val	Ser	Ile	Cys 80
Ala	Glu	Asn	Asn	Thr 85	Arg	Phe	Phe	Ile	Pro 90	Val	Ile	Ala	Ala	Trp 95	Tyr
Ile	Gly	Met	Ile 100	Val	Ala	Pro	Val	Asn 105	Glu	Ser	Tyr	Ile	Pro 110	Asp	Glu
Leu	Сув	Lys 115	Val	Met	Gly	Ile	Ser 120	Lys	Pro	Gln	Ile	Val 125	Phe	Thr	Thr
Lys	Asn 130	Ile	Leu	Asn	Lys	Val 135	Leu	Glu	Val	Gln	Ser 140	Arg	Thr	Asn	Phe
Ile 145	Lys	Arg	Ile	Ile	Ile 150	Leu	Asp	Thr	Val	Glu 155	Asn	Ile	His	Gly	Cys 160
Glu	Ser	Leu	Pro	Asn 165	Phe	Ile	Ser	Arg	Ty r 170	Ser	Asp	Gly	Asn	Ile 175	Ala
Asn	Phe	Lys	Pro 180	Leu	His	Phe	Asp	Pro 185	Val	Glu	Gln	Val	Ala 190	Ala	Ile
Leu	Cys	Ser 195	Ser	Gly	Thr	Thr	Gl y 200	Leu	Pro	Lys	Gly	Val 205	Met	Gln	Thr
His	Gln 210	Asn	Ile	Cys	Val	Arg 215	Leu	Ile	His	Ala	Leu 220	Asp	Pro	Arg	Tyr
Gl y 225	Thr	Gln	Leu	Ile	Pro 230	Gly	Val	Thr	Val	Leu 235	Val	Tyr	Leu	Pro	Phe 240
Phe	His	Ala	Phe	Gl y 245	Phe	His	Ile	Thr	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly

														<u> </u>	
Leu	Arg	Val	Ile 260	Met	Phe	Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys
Ala	Ile	Gln 275	Asp	Tyr	Glu	Val	Arg 280	Ser	Val	Ile	Asn	Val 285	Pro	Ser	Val
Ile	Leu 290	Phe	Leu	Ser	Lys	Ser 295	Pro	Leu	Val	Asp	L y s 300	Tyr	Asp	Leu	Ser
Ser 305	Leu	Arg	Glu	Leu	Cys 310	Сув	Gly	Ala	Ala	Pro 315	Leu	Ala	Lys	Glu	Val 320
Ala	Glu	Val	Ala	Ala 325	Lys	Arg	Leu	Asn	Leu 330	Pro	Gly	Ile	Arg	Cys 335	Gly
Phe	Gly	Leu	Thr 340	Glu	Ser	Thr	Ser	Ala 345	Ile	Ile	Gln	Ser	Leu 350	Arg	Asp
Glu	Phe	Lys 355	Ser	Gly	Ser	Leu	Gly 360	Arg	Val	Thr	Pro	Leu 365	Met	Ala	Ala
Lys	Ile 370	Ala	Asp	Arg	Glu	Thr 375	Gly	Lys	Ala	Leu	Gl y 380	Pro	Asn	Gln	Val
Gly 385	Glu	Leu	Cys	Ile	L y s 390	Gly	Pro	Met	Val	Ser 395	Lys	Gly	Tyr	Val	Asn 400
Asn	Val	Glu	Ala	Thr 405	Lys	Glu	Ala	Ile	Asp 410	Asp	Asp	Gly	Trp	Leu 415	His
Ser	Gly	Asp	Phe 420	Gly	Tyr	Tyr	Asp	Glu 425	Asp	Glu	His	Phe	Tyr 430	Val	Val
Asp	Arg	Ty r 435	Lys	Glu	Leu	Ile	Lys 440	Tyr	Lys	Gly	Ser	Gln 445	Val	Ala	Pro
Ala	Glu 450	Leu	Glu	Glu	Ile	Leu 455	Leu	Lys	Asn	Pro	Cys 460	Ile	Arg	Asp	Val
Ala 465	Val	Val	Gly	Ile	Pro 470	Asp	Leu	Glu	Ala	Gly 475	Glu	Leu	Pro	Ser	Ala 480
Phe	Val	Val	Lys	Gln 485	Pro	Gly	Lys	Glu	Ile 490	Thr	Ala	Lys	Glu	Val 495	Tyr
Asp	Tyr	Leu	Ala 500	Glu	Arg	Val	Ser	His 505	Thr	Lys	Tyr	Leu	Arg 510	Gly	Gly
Val	Arg	Phe 515	Val	Asp	Ser	Ile	Pro 520	Arg	Asn	Val	Thr	Gly 525	Lys	Ile	Thr
Arg	L y s 530	Glu	Leu	Leu	Lys	Gln 535	Leu	Leu	Glu	Lys	Ala 540	Gly	Gly		
<210> SEQ ID NO 33 <211> LENGTH: 542															
	2> TY 3> OF			Art	ficia	al Se	equer	ıce							
<220)> FE	EATUF	RE:				_		as	ynth	netio	c luc	cifer	ase	
<400)> SE	QUEN	ICE:	33											
Met 1	Met	Lys	Arg	Glu 5	Lys	Asn	Val	Ile	Ty r 10	Gly	Pro	Glu	Pro	Leu 15	His
Pro	Leu	Glu	Asp 20	Leu	Thr	Ala	Gly	Glu 25	Met	Leu	Phe	Arg	Ala 30	Leu	Arg
Lys	His	Ser 35	His	Leu	Pro	Gln	Ala 40	Leu	Val	Asp	Val	Val 45	Gly	Asp	Glu
Ser	Leu 50	Ser	Tyr	Lys	Glu	Phe 55	Phe	Glu	Ala	Thr	Val 60	Leu	Leu	Ala	Gln

Ser 65	Leu	His	Asn	Cys	Gly 70	Tyr	Lys	Met	Asn	Asp 75	Val	Val	Ser	Ile	Cys 80
Ala	Glu	Asn	Asn	Thr 85	Arg	Phe	Phe	Ile	Pro 90	Val	Ile	Ala	Ala	Trp 95	Tyr
Ile	Gly	Met	Ile 100	Val	Ala	Pro	Val	Asn 105	Glu	Ser	Tyr	Ile	Pro 110	Asp	Glu
Leu	Cys	Lys 115	Val	Met	Gly	Ile	Ser 120	Lys	Pro	Gln	Ile	Val 125	Phe	Thr	Thr
Lys	Asn 130	Ile	Leu	Asn	Lys	Val 135	Leu	Glu	Val	Gln	Ser 140	Arg	Thr	Asn	Phe
Ile 145	Lys	Arg	Ile	Ile	Ile 150	Leu	Asp	Thr	Val	Glu 155	Asn	Ile	His	Gly	C y s 160
Glu	Ser	Leu	Pro	Asn 165	Phe	Ile	Ser	Arg	Ty r 170	Ser	Asp	Gly	Asn	Ile 175	Ala
Asn	Phe	Lys	Pro 180	Leu	His	Phe	Asp	Pro 185	Val	Glu	Gln	Val	Ala 190	Ala	Ile
Leu	Cys	Ser 195	Ser	Gly	Thr	Thr	Gly 200	Leu	Pro	Lys	Gly	Val 205	Met	Gln	Thr
His	Gln 210	Asn	Ile	Суѕ	Val	Arg 215	Leu	Ile	His	Ala	Leu 220	Asp	Pro	Arg	Tyr
Gl y 225	Thr	Gln	Leu	Ile	Pro 230	Gly	Val	Thr	Val	Leu 235	Val	Tyr	Leu	Pro	Phe 240
Phe	His	Ala	Phe	Gl y 245	Phe	His	Ile	Thr	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly
Leu	Arg	Val	Ile 260	Met	Phe	Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys
Ala	Ile	Gln 275	Asp	Tyr	Glu	Val	Arg 280	Ser	Val	Ile	Asn	Val 285	Pro	Ser	Val
Ile	Leu 290	Phe	Leu	Ser	Lys	Ser 295	Pro	Leu	Val	Asp	L y s 300	Tyr	Asp	Leu	Ser
Ser 305	Leu	Arg	Glu	Leu	Cys 310	Cys	Gly	Ala	Ala	Pro 315	Leu	Ala	Lys	Glu	Val 320
Ala	Glu	Val	Ala	Ala 325	Lys	Arg	Leu	Asn	Leu 330	Pro	Gly	Ile	Arg	Cys 335	Gly
Phe	Gly	Leu	Thr 340	Glu	Ser	Thr	Ser	Ala 345	Ile	Ile	Gln	Ser	Leu 350	Arg	Asp
Glu	Phe	Lys 355	Ser	Gly	Ser	Leu	Gl y 360	Arg	Val	Thr	Pro	Leu 365	Met	Ala	Ala
Lys	Ile 370	Ala	Asp	Arg	Glu	Thr 375	Gly	Lys	Ala	Leu	Gly 380	Pro	Asn	Gln	Val
Gly 385	Glu	Leu	Cys	Ile	L y s 390	Gly	Pro	Met	Val	Ser 395	Lys	Gly	Tyr	Val	Asn 400
Asn	Val	Glu	Ala	Thr 405	Lys	Glu	Ala	Ile	Asp 410	Asp	Asp	Gly	Trp	Leu 415	His
Ser	Gly	Asp	Phe 420	Gly	Tyr	Tyr	Asp	Glu 425	Asp	Glu	His	Phe	Ty r 430	Val	Val
Asp	Arg	Ty r 435	Lys	Glu	Leu	Ile	Lys 440	Tyr	Lys	Gly	Ser	Gln 445	Val	Ala	Pro
Ala	Glu 450	Leu	Glu	Glu	Ile	Leu 455	Leu	Lys	Asn	Pro	Cys 460	Ile	Arg	Asp	Val
Ala	Val	Val	Gly	Ile	Pro	Asp	Leu	Glu	Ala	Gly	Glu	Leu	Pro	Ser	Ala

465					470					475					480
Phe	Val	Val	Lys	Gln 485	Pro	Gly	Lys	Glu	Ile 490	Thr	Ala	Lys	Glu	Val 495	Tyr
Asp	Tyr	Leu	Ala 500	Glu	Arg	Val	Ser	His 505	Thr	Lys	Tyr	Leu	Arg 510	Gly	Gly
Val	Arg	Phe 515	Val	Asp	Ser	Ile	Pro 520	Arg	Asn	Val	Thr	Gl y 525	Lys	Ile	Thr
Arg	Lys 530	Glu	Leu	Leu	Lys	Gln 535	Leu	Leu	Glu	Lys	Ala 540	Gly	Gly		
<211 <212 <213 <220 <223	L> LE 2> TY 3> OF 0> FE 3> OT	ENGTH PE: RGANI EATUR THER	SM: RE: INFO	Arti ORMAI			_		ā a s	s y ntl	netio	: luc	cifer	rase	
			ICE:		_	_					_		_	_	
1				5					Tyr 10					15	
Pro	Leu	Glu	Asp 20	Leu	Thr	Ala	Gly	Glu 25	Met	Leu	Phe	Arg	Ala 30	Leu	Arg
Lys	His	Ser 35	His	Leu	Pro	Gln	Ala 40	Leu	Val	Asp	Val	Val 45	Gly	Asp	Glu
Ser	Leu 50	Ser	Tyr	Lys	Glu	Phe 55	Phe	Glu	Ala	Thr	Val 60	Leu	Leu	Ala	Gln
Ser 65	Leu	His	Asn	Cys	Gly 70	Tyr	Lys	Met	Asn	Asp 75	Val	Val	Ser	Ile	Cys 80
Ala	Glu	Asn	Asn	Thr 85	Arg	Phe	Phe	Ile	Pro 90	Val	Ile	Ala	Ala	T rp 95	Tyr
Ile	Gly	Met	Ile 100	Val	Ala	Pro	Val	Asn 105	Glu	Ser	Tyr	Ile	Pro 110	Asp	Glu
Leu	Cys	Lys 115	Val	Met	Gly	Ile	Ser 120	Lys	Pro	Gln	Ile	Val 125	Phe	Thr	Thr
Lys	Asn 130	Ile	Leu	Asn	Lys	Val 135	Leu	Glu	Val	Gln	Ser 140	Arg	Thr	Asn	Phe
Ile 145	Lys	Arg	Ile	Ile	Ile 150	Leu	Asp	Thr	Val	Glu 155	Asn	Ile	His	Gly	Cys 160
Glu	Ser	Leu	Pro	Asn 165	Phe	Ile	Ser	Arg	Ty r 170	Ser	Asp	Gly	Asn	Ile 175	Ala
Asn	Phe	Lys	Pro 180	Leu	His	Phe	Asp	Pro 185	Val	Glu	Gln	Val	Ala 190	Ala	Ile
Leu	Сув	Ser 195	Ser	Gly	Thr	Thr	Gly 200	Leu	Pro	Lys	Gly	Val 205	Met	Gln	Thr
His	Gln 210	Asn	Ile	Сув	Val	Arg 215	Leu	Ile	His	Ala	Leu 220	Asp	Pro	Arg	Tyr
Gl y 225	Thr	Gln	Leu	Ile	Pro 230	Gly	Val	Thr	Val	Leu 235	Val	Tyr	Leu	Pro	Phe 240
Phe	His	Ala	Phe	Gly 245	Phe	His	Ile	Thr	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly
Leu	Arg	Val	Ile 260	Met	Phe	Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys
Ala	Ile	Gln	Asp	Tyr	Glu	Val	Arg	Ser	Val	Ile	Asn	Val	Pro	Ser	Val

275 280 285						
Ile Leu Phe Leu Ser Lys Ser Pro Leu Val Asp Lys Tyr Asp Leu Ser 290 295 300						
Ser Leu Arg Glu Leu Cys Cys Gly Ala Ala Pro Leu Ala Lys Glu Val 305 310 315 320						
Ala Glu Val Ala Ala Lys Arg Leu Asn Leu Pro Gly Ile Arg Cys Gly 325 330 335						
Phe Gly Leu Thr Glu Ser Thr Ser Ala Ile Ile Gln Ser Leu Arg Asp 340 345 350						
Glu Phe Lys Ser Gly Ser Leu Gly Arg Val Thr Pro Leu Met Ala Ala 355 360 365						
Lys Ile Ala Asp Arg Glu Thr Gly Lys Ala Leu Gly Pro Asn Gln Val 370 375 380						
Gly Glu Leu Cys Ile Lys Gly Pro Met Val Ser Lys Gly Tyr Val Asn 385 390 395 400						
Asn Val Glu Ala Thr Lys Glu Ala Ile Asp Asp Asp Gly Trp Leu His 405 410 415						
Ser Gly Asp Phe Gly Tyr Tyr Asp Glu Asp Glu His Phe Tyr Val Val 420 425 430						
Asp Arg Tyr Lys Glu Leu Ile Lys Tyr Lys Gly Ser Gln Val Ala Pro 435 440 445						
Ala Glu Leu Glu Glu Ile Leu Leu Lys Asn Pro Cys Ile Arg Asp Val 450 455 460						
Ala Val Val Gly Ile Pro Asp Leu Glu Ala Gly Glu Leu Pro Ser Ala 465 470 475 480						
Phe Val Val Lys Gln Pro Gly Lys Glu Ile Thr Ala Lys Glu Val Tyr 485 490 495						
Asp Tyr Leu Ala Glu Arg Val Ser His Thr Lys Tyr Leu Arg Gly Gly 500 505 510						
Val Arg Phe Val Asp Ser Ile Pro Arg Asn Val Thr Gly Lys Ile Thr 515 520 525						
Arg Lys Glu Leu Leu Lys Gln Leu Leu Glu Lys Ala Gly Gly 530 535 540						
<210> SEQ ID NO 35 <211> LENGTH: 29 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide						
<400> SEQUENCE: 35						
acgccagccc aagcttaggc ctgagtggc	29					
<210> SEQ ID NO 36 <211> LENGTH: 44 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide						
<400> SEQUENCE: 36						
cttaattctc cccatccccc tgttgacaat taatcatcgg ctcg	44					

<210> SEQ ID NO 37

	-continued
<211> LENGTH: 40	
<212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE:	
<pre><223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 37	
tataatgtga ggaattgcga gcggataaca atttcacaca	40
<210> SEQ ID NO 38	
<211> LENGTH: 40	
<212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
12237 OTHER INFORMATION. AIR OTTIGORACTED CIAC	
<400> SEQUENCE: 38	
atgggatgtt acctagacca atatgaaata tttggtaaat	40
<210> SEQ ID NO 39	
<211> LENGTH: 40	
<pre><212> TYPE: DNA <213> ORGANISM: Artificial Sequence</pre>	
<220> FEATURE:	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 39	
aaatgottaa tgaatttoaa aaaaaaaaaa aaaggaatto	40
<210> SEQ ID NO 40	
<211> LENGTH: 40	
<212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 40	
gatatcaagc ttatcgatac cgtcgacctc gaggattata	40
<210> SEQ ID NO 41 <211> LENGTH: 37	
<212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 41	
tagaaaaagg cctcggcggc cgctagttca gtcagtt	37
<210> SEQ ID NO 42	
<211> LENGTH: 17	
<212> TYPE: DNA <213> ORGANISM: Artificial Sequence	
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 42	
aactgactga actagcg	17
010, GEO. TD. NO. 42	
<210> SEQ ID NO 43 <211> LENGTH: 40	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE:	

	nicinaea
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 43	
gccgccgagg cctttttcta tataatcctc gaggtcgacg	40
<210> SEQ ID NO 44 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 44	
gtatcgataa gcttgatatc gaattccttt ttttttttt	40
<210> SEQ ID NO 45 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide <400> SEQUENCE: 45	
agcttgatat cgaattcctt ttttttttt tttgaaattc	40
<210> SEQ ID NO 46 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 46	
ttgaaattca ttaagcattt atttaccaaa tatttcatat	40
<pre><210> SEQ ID NO 47 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide <400> SEQUENCE: 47</pre>	
tggtctaggt aacatcccat cactagcttt ttttctata	40
<210> SEQ ID NO 48 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 48	
tcgcaattcc tcacattata cgagccgatg attaattgtc	40
<210> SEQ ID NO 49 <211> LENGTH: 53 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide <400> SEQUENCE: 49	

aacaggggga tggggagaat taaggccact caggcctaag cttggg	ctgg cgt 53
<210> SEQ ID NO 50 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 50	
ggaaacagga tcccatgatg aaacgcgaaa agaacgtgat	40
<210> SEQ ID NO 51 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 51	
ctacggccca gaaccactgc atccactgga agacctcacc	40
<210> SEQ ID NO 52 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 52	
gctggtgaga tgctcttccg agcactgcgt aaacatagtc	40
<210> SEQ ID NO 53 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 53	
acctccctca agcactcgtg gacgtcgtgg gagacgagag	40
<210> SEQ ID NO 54 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 54	
cctctcctac aaagaatttt tcgaagctac tgtgctgttg	40
<210> SEQ ID NO 55 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 55	
gcccaaagcc tccataattg tgggtacaaa atgaacgatg	40
<210> SEQ ID NO 56	

	-continued
<211> LENGTH: 40	
<pre><212> TYPE: DNA <213> ORGANISM: Artificial Sequence</pre>	
<220> FEATURE:	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 56	
tggtgagcat ttgtgctgag aataacactc gcttctttat	40
<210> SEQ ID NO 57	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 57	
	40
tcctgtaatc gctgcttggt acatcggcat gattgtcgcc	9
<210> SEQ ID NO 58	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 58	
cctgtgaatg aatcttacat cccagatgag ctgtgtaagg	40
<210> SEQ ID NO 59	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 59	
ttatgggtat tagcaaacct caaatcgtct ttactaccaa	40
<210> SEQ ID NO 60	
<211> LENGTH: 40 <212> TYPE: DNA	
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 60	
aaacatcttg aataaggtct tggaagtcca gtctcgtact	40
and a second and a second seco	4 V
<210> SEQ ID NO 61	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 61	
aacttcatca aacgcatcat tattctggat accgtcgaaa	40
and the second s	
<210> SEQ ID NO 62	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE:	

40
40
40
40
4 0
40

tgtatctgcc tttctttcac gcctttggtt tctctattac	40
<210> SEQ ID NO 69 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 69	
cctgggctat ttcatggtcg gcttgcgtgt catcatgttt	40
<210> SEQ ID NO 70 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 70	
cgtcgcttcg accaagaagc cttcttgaag gctattcaag	40
<pre><210> SEQ ID NO 71 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 71	
actacgaggt gcgttccgtg atcaacgtcc cttcagtcat	40
<pre><210> SEQ ID NO 72 <211> LENGTH: 43 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 72	ta 43
tttgttcctg agcaaatctc ctttggttga caagtatgat c	Lý 43
<210> SEQ ID NO 73 <211> LENGTH: 37 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 73	
agcagettge gtgagetgtg etgtggeget geteett	37
<210> SEQ ID NO 74 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 74	
tggccaaaga agtggccgag gtcgctgcta agcgtctgaa	40
<210> SEQ ID NO 75	

	-continued
<211> LENGTH: 40	
<212> TYPE: DNA <213> ORGANISM: Artificial Sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 75	
cctccctggt atccgctgcg gttttggttt gactgagagc	40
<210> SEQ ID NO 76	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 76	
acttotgota acatocatag ottgogagao gagtttaagt	40
<210> SEQ ID NO 77	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 77	
	40
ctggtagcct gggtcgcgtg actcctctta tggctgcaaa	4 U
<210> SEQ ID NO 78	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 78	
gategeegae egtgagaeeg geaaageaet gggeeeaaat	40
<210> SEQ ID NO 79	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 79	
caagtoggtg aattgtgtat taagggooot atggtotota	40
<210> SEQ ID NO 80	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 80	
aaggctacgt gaacaatgtg gaggccacta aagaagccat	40
augustages gaygeeacca augustgeeac	
<210> SEQ ID NO 81	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE:	

	-continued
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 81	
tgatgatgat ggctggctcc atagcggcga cttcggttac	40
<210> SEQ ID NO 82 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 82	
tatgatgagg acgaacactt ctatgtggtc gatcgctaca	40
<210> SEQ ID NO 83 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 83	
aagaattgat taagtacaaa ggctctcaag tcgcaccagc	40
<210> SEQ ID NO 84 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 84	
cgaactggaa gaaattttgc tgaagaaccc ttgtatccgc	40
<210> SEQ ID NO 85 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 85	
gacgtggccg tcgtgggtat cccagacttg gaagctggcg	40
<210> SEQ ID NO 86 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 86	
agttgcctag cgcctttgtg gtgaaacaac ccggcaagga	40
<210> SEQ ID NO 87 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide <400> SEQUENCE: 87	

gatcactgct aaggaggtct acgactattt ggccgagcgc	 40
<pre><210> SEQ ID NO 88 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 88	40
gtgtctcaca ccaaatatct gcgtggcggc gtccgcttcg	40
<210> SEQ ID NO 89 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 89	
tcgattctat tccacgcaac gttaccggta agatcactcg	40
<210> SEQ ID NO 90 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 90	
taaagagttg ctgaagcaac tcctcgaaaa agctggcggc	40
<210> SEQ ID NO 91 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 91	
tagtaaagtc ttcatgatta tatagaaaaa aaagctagtg	40
<210> SEQ ID NO 92 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 92	
taatcatgaa gactttacta gccgccagct ttttcgagga	40
<pre><210> SEQ ID NO 93 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 93	
gttgcttcag caactcttta cgagtgatct taccggtaac	40
<210> SEQ ID NO 94	

-continued	
<211> LENGTH: 39	
<212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 94	
gttgcgtgga atagaatcga cgaagcggac gccgccacg	39
<210> SEQ ID NO 95	
<211> LENGTH: 41	
<212> TYPE: DNA	
<pre><213> ORGANISM: Artificial Sequence</pre>	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
2225 STILL INTORPATION. AN OTIGORACTE STAGE	
<400> SEQUENCE: 95	
cagatatttg gtgtgagaca cgcgctcggc caaatagtcg t	41
<210> SEQ ID NO 96	
<211> LENGTH: 40	
<212> TYPE: DNA <213> ORGANISM: Artificial Sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 96	
agacctcctt agcagtgatc tccttgccgg gttgtttcac	40
<210> SEQ ID NO 97	
<211> LENGTH: 40	
<212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 97	
	40
cacaaaggcg ctaggcaact cgccagcttc caagtctggg	40
<210> SEQ ID NO 98	
<211> LENGTH: 40	
<212> TYPE: DNA	
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 98	
atacccacga cggccacgtc gcggatacaa gggttcttca	40
<210> SEQ ID NO 99	
<211> LENGTH: 40	
<212> TYPE: DNA	
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 99	
gcaaaatttc ttccagttcg gctggtgcga cttgagagcc	40
<210> SEQ ID NO 100	
<211> LENGTH: 40	
<212> TYPE: DNA	
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>	
NAZOV FERIURE:	

<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 100	
tttgtactta atcaattctt tgtagcgatc gaccacatag	40
<210> SEQ ID NO 101 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 101	
aagtgttcgt cctcatcata gtaaccgaag tcgccgctat	40
<pre><210> SEQ ID NO 102 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide <400> SEQUENCE: 102</pre>	
ggagccagcc atcatcatca atggcttctt tagtggcctc	40
ggagecagee accateatea attggettett tagtggette	***
<210> SEQ ID NO 103 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 103	
cacattgttc acgtagcctt tagagaccat agggccctta	40
<210> SEQ ID NO 104 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 104	
atacacaatt caccgacttg atttgggccc agtgctttgc	40
<pre><210> SEQ ID NO 105 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 105	
cggtctcacg gtcggcgatc tttgcagcca taagaggagt	40
<210> SEQ ID NO 106 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide <400> SEQUENCE: 106	

	-continued
cacgcgaccc aggctaccag acttaaactc gtctcgcaag	40
<210> SEQ ID NO 107 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 107	
ctatggatgt tagcagaagt gctctcagtc aaaccaaaac	40
<pre><210> SEQ ID NO 108 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 108	
cgcagcggat accagggagg ttcagacgct tagcagcgac	4 0
<210> SEQ ID NO 109 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 109	
ctcggccact tctttggcca aaggagcagc gccacagcac	40
<pre><210> SEQ ID NO 110 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 110	
agetcacgca agetgetcag atcatacttg tcaaccaaag	40
<210> SEQ ID NO 111 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 111	
gagatttgct caggaacaaa atgactgaag ggacgttgat	40
<210> SEQ ID NO 112 <211> LENGTH: 36 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 112	26
cacggaacgc acctcgtagt cttgaatagc cttcaa	36
<210> SEQ ID NO 113	

	-continued	
<211> LENGTH: 44		
<212> TYPE: DNA		
<213> ORGANISM: Artificial Sequence		
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>		
<400> SEQUENCE: 113		
gaaggettet tggtegaage gaegaaacat gatgaeaege	aagc	44
<210> SEQ ID NO 114		
<211> LENGTH: 40		
<212> TYPE: DNA		
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>		
<pre><220> FEATORE: <223> OTHER INFORMATION: An oligonucleotide</pre>		
<400> SEQUENCE: 114		
cgaccatgaa atagcccagg gtaatagaga aaccaaaggc		40
<210> SEQ ID NO 115		
<211> LENGTH: 40		
<212> TYPE: DNA		
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>		
<223> OTHER INFORMATION: An oligonucleotide		
<400> SEQUENCE: 115		
gtgaaagaaa ggcagataca ccagcacagt cacgccaggg		40
<210> SEQ ID NO 116		
<211> LENGTH: 40		
<212> TYPE: DNA		
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>		
<pre><223> OTHER INFORMATION: An oligonucleotide</pre>		
<400> SEQUENCE: 116		
atcaattgag tacccacacg agggtcgaga gcgtggatca		40
<210> SEQ ID NO 117		
<211> SEQ 1D NO 117 <211> LENGTH: 40		
<212> TYPE: DNA		
<213> ORGANISM: Artificial Sequence		
<220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide		
<400> SEQUENCE: 117		
aacgcacaca gatattctgg tgagtctgca tgacaccttt		40
010 270 77 110		
<210> SEQ ID NO 118 <211> LENGTH: 40		
<211> HENGIN: 40 <212> TYPE: DNA		
<213> ORGANISM: Artificial Sequence		
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>		
<400> SEQUENCE: 118		
aggcaaacca gtggtgccgg aggagcacaa aatagcggcc		40
<210> SEQ ID NO 119		
<211> LENGTH: 40		
<212> TYPE: DNA		
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>		

	continued
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 119	
acttgctcga ctggatcaaa atgcaagggc ttgaaattag	40
<210> SEQ ID NO 120 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 120	
cgatattacc atcgctgtaa cgagagatga agttagggag	40
<210> SEQ ID NO 121 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 121	
gctctcacag ccgtggatgt tttcgacggt atccagaata	40
<210> SEQ ID NO 122 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 122	
atgatgcgtt tgatgaagtt agtacgagac tggacttcca	40
<210> SEQ ID NO 123 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 123	
agaccttatt caagatgttt ttggtagtaa agacgatttg	40
<210> SEQ ID NO 124 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 124	
aggtttgcta atacccataa ccttacacag ctcatctggg	40
<210> SEQ ID NO 125 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide <400> SEQUENCE: 125	

atgtaagatt cattcacagg ggcgacaatc atgccgatgt	40
<210> SEQ ID NO 126 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 126	
accaagcagc gattacagga ataaagaagc gagtgttatt	40
<210> SEQ ID NO 127 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 127	
ctcagcacaa atgctcacca catcgttcat tttgtaccca	40
<210> SEQ ID NO 128 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 128	
caattatgga ggctttgggc caacagcaca gtagcttcga	40
<210> SEQ ID NO 129 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 129	
aaaattettt gtaggagagg etetegtete eeaegaegte	40
<210> SEQ ID NO 130 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 130	
cacgagtgct tgagggaggt gactatgttt acgcagtgct	40
<210> SEQ ID NO 131 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 131	
cggaagagca tctcaccagc ggtgaggtct tccagtggat	40
<210> SEQ ID NO 132	

	-continued
<211> LENGTH: 40	
<212> TYPE: DNA <213> ORGANISM: Artificial Sequence	
<220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 132	
gcagtggttc tgggccgtag atcacgttct tttcgcgttt	40
010 070 70 00 100	
<210> SEQ ID NO 133 <211> LENGTH: 40	
<212> TYPE: DNA <213> ORGANISM: Artificial Sequence	
<220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 133	
catcatggga tcctgtttcc tgtgtgaaat tgttatccgc	40
210. SEC TO NO 134	
<210> SEQ ID NO 134 <211> LENGTH: 40	
<212> TYPE: DNA <213> ORGANISM: Artificial Sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 134	
ggaaacagga tcccatgatg aagcgtgaga aaaatgtcat	40
<210> SEQ ID NO 135	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence <220> FEATURE:	
<pre><220> FEATORE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 135	
ctatggccct gagcctctcc atcctttgga ggatttgact	4 0
<210> SEQ ID NO 136	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
-	
<400> SEQUENCE: 136	
gccggcgaaa tgctgtttcg tgctctccgc aagcactctc	40
<210> SEQ ID NO 137	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 137	
atttgcctca agccttggtc gatgtggtcg gcgatgaatc	40
<210> SEQ ID NO 138	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE:	

	-continued
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 138	
tttgagctac aaggagtttt ttgaggcaac cgtcttgctg	40
<210> SEQ ID NO 139 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 139	
gctcagtccc tccacaattg tggctacaag atgaacgacg	40
<210> SEQ ID NO 140 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 140	
tcgttagtat ctgtgctgaa aacaataccc gtttcttcat	40
<210> SEQ ID NO 141 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 141	
tccagtcatc gccgcatggt atatcggtat gatcgtggct	40
<210> SEQ ID NO 142 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 142	
ccagtcaacg agagctacat tcccgacgaa ctgtgtaaag	40
<210> SEQ ID NO 143 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 143	
tcatgggtat ctctaagcca cagattgtct tcaccactaa	40
<210> SEQ ID NO 144 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide <400> SEQUENCE: 144	

gaatattotg aacaaagtoo tggaagtooa aagoogcaco	40
<210> SEQ ID NO 145 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 145	
aactttatta agcgtatcat catcttggac actgtggaga	40
<210> SEQ ID NO 146 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 146	
atattcacgg ttgcgaatct ttgcctaatt tcatctctcg	40
<210> SEQ ID NO 147 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 147	
ctattcagac ggcaacatcg caaactttaa accactccac	40
<210> SEQ ID NO 148 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 148	
ttcgaccctg tggaacaagt tgcagccatt ctgtgtagca	40
<210> SEQ ID NO 149 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 149	
gcggtactac tggactccca aagggagtca tgcagaccca	40
<pre><210> SEQ ID NO 150 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 150	
tcaaaacatt tgcgtgcgtc tgatccatgc tctcgatcca	40
<210> SEQ ID NO 151	

	-continued
<211> LENGTH: 40	
<212> TYPE: DNA <213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 151	
cgctacggca ctcagctgat tcctggtgtc accgtcttgg	4 0
<210> SEQ ID NO 152	
<211> LENGTH: 40	
<212> TYPE: DNA <213> ORGANISM: Artificial Sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 152	
totacttgcc tttcttccat gctttcggct ttcatattac	40
<210> SEQ ID NO 153 <211> LENGTH: 40	
<212> TYPE: DNA	
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 153	
tttgggttac tttatggtcg gtctccgcgt gattatgttc	40
<210> SEQ ID NO 154	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence <220> FEATURE:	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 154	
cgccgttttg atcaggaggc tttcttgaaa gccatccaag	4 0
<210> SEQ ID NO 155	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 155	
	40
attatgaagt ccgcagtgtc atcaacgtgc ctagcgtgat	40
<210> SEQ ID NO 156	
<211> LENGTH: 40 <212> TYPE: DNA	
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>	
<pre><220> FEATORE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 156	
cctgtttttg tctaagagcc cactcgtgga caagtacgac	40
<210> SEQ ID NO 157	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE:	

	-concinued
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 157	
ttgtcttcac tgcgtgaatt gtgttgcggt gccgctccac	40
<210> SEQ ID NO 158 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 158	
tggctaagga ggtcgctgaa gtggccgcca aacgcttgaa	40
<pre><210> SEQ ID NO 159 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide <400> SEQUENCE: 159</pre>	
tcttccaggg attcgttgtg gcttcggcct caccgaatct	40
<210> SEQ ID NO 160 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 160	
accagogota ttattcagto totoogogat gagtttaaga	40
<pre><210> SEQ ID NO 161 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide <400> SEQUENCE: 161</pre>	
gcggctcttt gggccgtgtc actccactca tggctgctaa	40
<210> SEQ ID NO 162 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 162	
gatcgctgat cgcgaaactg gtaaggcttt gggccctaac	40
<210> SEQ ID NO 163 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide <400> SEQUENCE: 163	

	-conclined
caagtgggcg agctgtgtat caaaggccct atggtgagca	40
<210> SEQ ID NO 164 <211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 164	
agggttatgt caataacgtc gaagctacca aggaggccat	40
<210> SEQ ID NO 165 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE:	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 165	
cgacgacgac ggctggttgc attctggtga ttttggatat	40
<210> SEQ ID NO 166 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE:	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 166	
tacgacgaag atgagcattt ttacgtcgtg gatcgttaca	40
<210> SEQ ID NO 167 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 167	
aggagctgat caaatacaag ggtagccagg ttgctccagc	40
<210> SEQ ID NO 168 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 168	
tgagttggag gagattctgt tgaaaaatcc atgcattcgc	40
<pre><210> SEQ ID NO 169 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 169	
gatgtcgctg tggtcggcat tcctgatctg gaggccggcg	40
<210> SEQ ID NO 170	

	-continued
<211> LENGTH: 40	
<212> TYPE: DNA <213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 170	
aactgccttc tgctttcgtt gtcaagcagc ctggtaaaga	40
<210> SEQ ID NO 171 <211> LENGTH: 40	
<212> TYPE: DNA <213> ORGANISM: Artificial Sequence	
<220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 171	
aattaccgcc aaagaagtgt atgattacct ggctgaacgt	40
<210> SEQ ID NO 172 <211> LENGTH: 40	
<212> TYPE: DNA	
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 172	
gtgagccata ctaagtactt gcgtggcggc gtgcgttttg	40
<210> SEQ ID NO 173 <211> LENGTH: 40	
<212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence <220> FEATURE:	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 173	
ttgactccat ccctcgtaac gtaacaggca aaattacccg	40
<210> SEQ ID NO 174	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence <220> FEATURE:	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 174	
caaggagctg ttgaaacaat tgttggagaa ggccggcggt	40
<210> SEQ ID NO 175 <211> LENGTH: 40	
<212> TYPE: DNA	
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 175	
tagtaaagtc ttcatgatta tatagaaaaa aaagctagtg	4 0
<210> SEQ ID NO 176	
<211> LENGTH: 40 <212> TYPE: DNA	
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>	
THE TAXABLE VALUE	

<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 176	
taatcatgaa gactttacta accgccggcc ttctccaaca	40
<210> SEQ ID NO 177 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 177	
attgtttcaa cagctccttg cgggtaattt tgcctgttac	40
<pre><210> SEQ ID NO 178 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide <400> SEQUENCE: 178</pre>	
gttacgaggg atggagtcaa caaaacgcac gccgccacgc	40
<210> SEQ ID NO 179 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 179	
aagtacttag tatggctcac acgttcagcc aggtaatcat	40
<210> SEQ ID NO 180 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 180	
acacttettt ggeggtaatt tetttaecag getgettgae	40
<210> SEQ ID NO 181 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 181	
aacgaaagca gaaggcagtt cgccggcctc cagatcagga	40
<210> SEQ ID NO 182 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	

	40
atgccgacca cagcgacatc gcgaatgcat ggatttttca	40
<210> SEQ ID NO 183 <211> LENGTH: 40	
<212> TYPE: DNA	
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 183	
acagaatoto otocaactoa gotggagoaa ootggotaco	40
<210> SEQ ID NO 184 <211> LENGTH: 40	
<212> TYPE: DNA	
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 184	
cttgtatttg atcagctcct tgtaacgatc cacgacgtaa	40
<210> SEQ ID NO 185 <211> LENGTH: 40	
<212> TYPE: DNA	
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 185	
aaatgotoat ottogtogta atatooaaaa toacoagaat	40
<210> SEQ ID NO 186 <211> LENGTH: 40	
<212> TYPE: DNA	
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 186	
gcaaccagec gtegtegteg atggeeteet tggtagette	40
<210> SEQ ID NO 187 <211> LENGTH: 40	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence <220> FEATURE:	
<pre><220> FEATORE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 187	
gacgttattg acataaccct tgctcaccat agggcctttg	4 0
<210> SEQ ID NO 188	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 188	
atacacaget egeceacttg gttagggeec aaageettac	40
<210> SEQ ID NO 189	

<pre><211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence</pre>	
<220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 189	
cagtttcgcg atcagcgatc ttagcagcca tgagtggagt	40
<210> SEQ ID NO 190 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 190	
gacacggccc aaagagccgc tcttaaactc atcgcggaga	40
<pre><210> SEQ ID NO 191 <211> LENGTH: 37 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 191	
gactgaataa tagcgctggt agattcggtg aggccga	37
<pre><210> SEQ ID NO 192 <211> LENGTH: 43 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 192	
agccacaacg aatccctgga agattcaagc gtttggcggc c	43
<210> SEQ ID NO 193 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 193	40
ttcagcgacc tccttagcca gtggagcggc accgcaacac	40
<210> SEQ ID NO 194 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 194	
aattcacgca gtgaagacaa gtcgtacttg tccacgagtg	40
<210> SEQ ID NO 195 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE:	

40
40
40
40
40
40

gacgcacgca aatgttttga tgggtctgca tgactccctt	40
<210> SEQ ID NO 202 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 202	
tgggagtcca gtagtaccgc tgctacacag aatggctgca	40
<210> SEQ ID NO 203 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 203	
acttgttcca cagggtcgaa gtggagtggt ttaaagtttg	40
<pre><210> SEQ ID NO 204 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 204	
cgatgttgcc gtctgaatag cgagagatga aattaggcaa	40
<pre><210> SEQ ID NO 205 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 205	
agattogcaa cogtgaatat totocacagt gtocaagatg	40
<210> SEQ ID NO 206 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 206	
atgatacgct taataaagtt ggtgcggctt tggacttcca	40
<210> SEQ ID NO 207 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 207	
ggactttgtt cagaatattc ttagtggtga agacaatctg	40
<210> SEQ ID NO 208	

	-continued
<211> LENGTH: 40	
<212> TYPE: DNA <213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 208	
tggcttagag atacccatga ctttacacag ttcgtcggga	40
<210> SEQ ID NO 209	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence <220> FEATURE:	
<pre><223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 209	
atgtagctct cgttgactgg agccacgatc ataccgatat	40
<210> SEQ ID NO 210	
<211> LENGTH: 40 <212> TYPE: DNA	
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE:</pre>	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 210	
accatgcggc gatgactgga atgaagaaac gggtattgtt	4 0
<210> SEQ ID NO 211	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence <220> FEATURE:	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 211	
ttcagcacag atactaacga cgtcgttcat cttgtagcca	4 0
<210> SEQ ID NO 212	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 212	
caattgtgga gggactgagc cagcaagacg gttgcctcaa	40
<210> SEQ ID NO 213 <211> LENGTH: 40	
<212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence <220> FEATURE:	
<223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 213	
aaaactcctt gtagctcaaa gattcatcgc cgaccacatc	40
<210> SEQ ID NO 214	
<211> LENGTH: 40 <212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE:	

				-concr			
<223> OTHER II	NFORMATION: Ar	n oligonucleo	tide				
<400> SEQUENCI	E: 214						
gaccaaggct tg	aggcaaat gagaq	gtgctt gcgga	gagca			40	
<210> SEQ ID I <211> LENGTH: <212> TYPE: DI <213> ORGANISI <220> FEATURE: <223> OTHER II	40 NA M: Artificial S :	_	otide				
<400> SEQUENCI	E: 215						
cgaaacagca tt	tegeegge agtea	aaatcc tccaaa	aggat			40	
<210> SEQ ID I <211> LENGTH: <212> TYPE: DI <213> ORGANISI <220> FEATURE: <223> OTHER II	40 NA 4: Artificial S : NFORMATION: Ar		otide				
<400> SEQUENCI							
ggagaggete ag	ggccatag atgad	catttt tctca	egett			40	
<210> SEQ ID I <211> LENGTH: <212> TYPE: DI <213> ORGANISI <220> FEATURE: <223> OTHER II <4400> SEQUENCI	40 NA M: Artificial S : NFORMATION: Ar	_	otide				
catcatggga to		tgaaat tgtta	cege			40	
<210> SEQ ID I <211> LENGTH: <212> TYPE: PI <213> ORGANISI <220> FEATURE <223> OTHER II <4400> SEQUENCI	NO 218 542 RT 1: Artificial S INFORMATION: Se	Sequence		c lucif	erase		
Met Met Lys A		n Val Ile Tvi	Gl v Pro	Glu Pr	o Leu His		
1	5	10	, -14	-	15		
Pro Leu Glu A		a Gly Glu Met 25	Leu Phe	e Arg Al	a Leu Arg		
L y s His Ser H. 35	is Leu Pro Glr	n Ala Leu Val 40	l Asp Val	. Val Gl: 45	y Asp Glu		
Ser Leu Ser T	yr Lys Glu Phe 55	e Phe Glu Ala	a Thr Val	. Leu Le	u Ala Gln		
Ser Leu His A	sn Cys Gly Tyr 70	r Lys Met Ası	n Asp Val 75	. Val Se	r Ile Cys 80		
Ala Glu Asn A	sn Thr Arg Phe	e Phe Ile Pro 90	Val Ile	e Ala Al	a Trp Tyr 95		
Ile Gly Met I	le Val Ala Pro 00	o Val Asn Glu 105	ı Ser Tyr	Ile Pr			
Leu Cys Lys V	al Met Gly Ile	e Ser Lys Pro	Gln Ile	Val Ph	e Thr Thr		

Lys	Asn 130	Ile	Leu	Asn	Lys	Val 135	Leu	Glu	Val	Gln	Ser 140	Arg	Thr	Asn	Phe
Ile 145	Lys	Arg	Ile	Ile	Ile 150	Leu	Asp	Thr	Val	Glu 155	Asn	Ile	His	Gly	C y s 160
Glu	Ser	Leu	Pro	Asn 165	Phe	Ile	Ser	Arg	Ty r 170	Ser	Asp	Gly	Asn	Ile 175	Ala
Asn	Phe	Lys	Pro 180	Leu	His	Phe	Asp	Pro 185	Val	Glu	Gln	Val	Ala 190	Ala	Ile
Leu	Cys	Ser 195	Ser	Gly	Thr	Thr	Gly 200	Leu	Pro	Lys	Gly	Val 205	Met	Gln	Thr
His	Gln 210	Asn	Ile	Cys	Val	Arg 215	Leu	Ile	His	Ala	Leu 220	Asp	Pro	Arg	Tyr
Gly 225	Thr	Gln	Leu	Ile	Pro 230	Gly	Val	Thr	Val	Leu 235	Val	Tyr	Leu	Pro	Phe 240
Phe	His	Ala	Phe	Gly 245	Phe	His	Ile	Thr	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly
Leu	Arg	Val	Ile 260	Met	Phe	Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys
Ala	Ile	Gln 275	Asp	Tyr	Glu	Val	Arg 280	Ser	Val	Ile	Asn	Val 285	Pro	Ser	Val
Ile	Leu 290	Phe	Leu	Ser	Lys	Ser 295	Pro	Leu	Val	Asp	L y s 300	Tyr	Asp	Leu	Ser
Ser 305	Leu	Arg	Glu	Leu	Cys 310	Сув	Gly	Ala	Ala	Pro 315	Leu	Ala	Lys	Glu	Val 320
Ala	Glu	Val	Ala	Ala 325	Lys	Arg	Leu	Asn	Leu 330	Pro	Gly	Ile	Arg	Cys 335	Gly
Phe	Gly	Leu	Thr 340	Glu	Ser	Thr	Ser	Ala 345	Ile	Ile	Gln	Ser	Leu 350	Arg	Asp
Glu	Phe	L y s 355	Ser	Gly	Ser	Leu	Gly 360	Arg	Val	Thr	Pro	Leu 365	Met	Ala	Ala
Lys	Ile 370	Ala	Asp	Arg	Glu	Thr 375	Gly	Lys	Ala	Leu	Gl y 380	Pro	Asn	Gln	Val
Gl y 385	Glu	Leu	Cys	Ile	L y s 390	Gly	Pro	Met	Val	Ser 395	Lys	Gly	Tyr	Val	Asn 400
Asn	Val	Glu	Ala	Thr 405	Lys	Glu	Ala	Ile	Asp 410	Asp	Asp	Gly	Trp	Leu 415	His
Ser	Gly		Phe 420		Tyr		Asp			Glu	His		Ty r 430		Val
Asp	Arg	Ty r 435	Lys	Glu	Leu	Ile	Lys 440	Tyr	Lys	Gly	Ser	Gln 445	Val	Ala	Pro
Ala	Glu 450	Leu	Glu	Glu	Ile	Leu 455	Leu	Lys	Asn	Pro	Cys 460	Ile	Arg	Asp	Val
Ala 465	Val	Val	Gly	Ile	Pro 470	Asp	Leu	Glu	Ala	Gl y 475	Glu	Leu	Pro	Ser	Ala 480
Phe	Val	Val	Lys	Gln 485	Pro	Gly	Lys	Glu	Ile 490	Thr	Ala	Lys	Glu	Val 495	Tyr
Asp	Tyr	Leu	Ala 500	Glu	Arg	Val	Ser	His 505	Thr	Lys	Tyr	Leu	Arg 510	Gly	Gly
Val	Arg	Phe 515	Val	Asp	Ser	Ile	Pro 520	Arg	Asn	Val	Thr	Gl y 525	Lys	Ile	Thr
Arg	Lys	Glu	Leu	Leu	Lys	Gln	Leu	Leu	Glu	Lys	Ala	Gly	Gly		

	530					535					540				
<211 <212 <213 <220)> SE L> LE 2> TY 3> OF)> FE 3> OT	NGTH PE: RGANI ATUF	I: 54 PRT SM: RE:	12 Arti			_		: a s	s y nth	netio	: luc	cifer	ase	
<400)> SE	QUEN	ICE:	219											
Met 1	Met	Lys	Arg	Glu 5	Lys	Asn	Val	Ile	Ty r 10	Gly	Pro	Glu	Pro	Leu 15	His
Pro	Leu	Glu	Asp 20	Leu	Thr	Ala	Gly	Glu 25	Met	Leu	Phe	Arg	Ala 30	Leu	Arg
Lys	His	Ser 35	His	Leu	Pro	Gln	Ala 40	Leu	Val	Asp	Val	Val 45	Gly	Asp	Glu
Ser	Leu 50	Ser	Tyr	Lys	Glu	Phe 55	Phe	Glu	Ala	Thr	Val 60	Leu	Leu	Ala	Gln
Ser 65	Leu	His	Asn	Суѕ	Gly 70	Tyr	Lys	Met	Asn	Asp 75	Val	Val	Ser	Ile	C y s 80
Ala	Glu	Asn	Asn	Thr 85	Arg	Phe	Phe	Ile	Pro 90	Val	Ile	Ala	Ala	T rp 95	Tyr
Ile	Gly	Met	Ile 100	Val	Ala	Pro	Val	Asn 105	Glu	Ser	Tyr	Ile	Pro 110	Asp	Glu
Leu	Cys	L y s 115	Val	Met	Gly	Ile	Ser 120	Lys	Pro	Gln	Ile	Val 125	Phe	Thr	Thr
Lys	Asn 130	Ile	Leu	Asn	Lys	Val 135	Leu	Glu	Val	Gln	Ser 140	Arg	Thr	Asn	Phe
Ile 145	Lys	Arg	Ile	Ile	Ile 150	Leu	Asp	Thr	Val	Glu 155	Asn	Ile	His	Gly	Cys 160
Glu	Ser	Leu	Pro	Asn 165	Phe	Ile	Ser	Arg	Ty r 170	Ser	Asp	Gly	Asn	Ile 175	Ala
Asn	Phe	Lys	Pro 180	Leu	His	Phe	Asp	Pro 185	Val	Glu	Gln	Val	Ala 190	Ala	Ile
Leu	Cys	Ser 195	Ser	Gly	Thr	Thr	Gl y 200	Leu	Pro	Lys	Gly	Val 205	Met	Gln	Thr
His	Gln 210	Asn	Ile	Cys	Val	Arg 215	Leu	Ile	His	Ala	Leu 220	Asp	Pro	Arg	Tyr
Gl y 225		Gln			Pro 230		Val		Val			Tyr	Leu	Pro	Phe 240
Phe	His	Ala	Phe	Gly 245	Phe	His	Ile	Thr	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly
Leu	Arg	Val	Ile 260	Met	Phe	Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys
Ala	Ile	Gln 275	Asp	Tyr	Glu	Val	Arg 280	Ser	Val	Ile	Asn	Val 285	Pro	Ser	Val
Ile	Leu 290	Phe	Leu	Ser	Lys	Ser 295	Pro	Leu	Val	Asp	L y s 300	Tyr	Asp	Leu	Ser
Ser 305	Leu	Arg	Glu	Leu	Cys 310	Cys	Gly	Ala	Ala	Pro 315	Leu	Ala	Lys	Glu	Val 320
Ala	Glu	Val	Ala	Ala 325	Lys	Arg	Leu	Asn	Leu 330	Pro	Gly	Ile	Arg	Cys 335	Gly
Phe	Gly	Leu	Thr	Glu	Ser	Thr	Ser	Ala	Ile	Ile	Gln	Ser	Leu	Arg	Asp

340

-continued

			340					343					350		
Glu	Phe	Lys 355	Ser	Gly	Ser	Leu	Gl y 360	Arg	Val	Thr	Pro	Leu 365	Met	Ala	Ala
Lys	Ile 370	Ala	Asp	Arg	Glu	Thr 375	Gly	Lys	Ala	Leu	Gl y 380	Pro	Asn	Gln	Val
Gly 385	Glu	Leu	Cys	Ile	Lys 390	Gly	Pro	Met	Val	Ser 395	Lys	Gly	Tyr	Val	Asn 400
Asn	Val	Glu	Ala	Thr 405	Lys	Glu	Ala	Ile	Asp 410	qaA	Ąsp	Gly	Trp	Leu 415	His
Ser	Gly	Asp	Phe 420	Gly	Tyr	Tyr	Asp	Glu 425	Asp	Glu	His	Phe	Tyr 430	Val	Val
Asp	Arg	Ty r 435	Lys	Glu	Leu	Ile	Lys 440	Tyr	Lys	Gly	Ser	Gln 445	Val	Ala	Pro
Ala	Glu 450	Leu	Glu	Glu	Ile	Leu 455	Leu	Lys	Asn	Pro	C y s 460	Ile	Arg	Asp	Val
Ala 465	Val	Val	Gly	Ile	Pro 470	Asp	Leu	Glu	Ala	Gl y 475	Glu	Leu	Pro	Ser	Ala 480
Phe	Val	Val	Lys	Gln 485	Pro	Gly	Lys	Glu	Ile 490	Thr	Ala	Lys	Glu	Val 495	Tyr
Asp	Tyr	Leu	Ala 500	Glu	Arg	Val	Ser	His 505	Thr	Lys	Tyr	Leu	Arg 510	Gly	Gly
Val	Arg	Phe 515	Val	Asp	Ser	Ile	Pro 520	Arg	Asn	Val	Thr	Gly 525	Lys	Ile	Thr
Arg	Lys 530	Glu	Leu	Leu	Lys	Gln 535	Leu	Leu	Glu	Lys	Ala 540	Gly	Gly		
<211 <212 <213 <220)> FE	NGTH PE: GANI ATUF	: 54 PRT SM:				_		i a s	ynth	etic	: luc	:ifer	ase	
<211 <212 <213 <220 <223	l> LE ?> TY 3> OR)> FE	NGTH PE: GANI ATUR HER	FRT SM: SM: INFO	Arti ORMAT			_		i a s	ynth	etic	: luc	:ifer	case	
<211 <212 <213 <223 <223 <400	l> LE 2> TY 3> OR 0> FE 3> OT 0> SE	NGTH PE: GANI ATUF HER	FE SA	Arti ORMAT	'ION:	Sec	quenc	e of							His
<211 <212 <213 <220 <223 <400 Met	L> LE 2> TY 3> OF 0> FE 3> OT 0> SE	HER CQUEN	PRT SM: INFO	Arti ORMAT 220 Glu	ION:	Sec Asn	quenc	e of	Ty r 10	Gly	Pro	Glu	Pro	Leu 15	
<211 <212 <213 <220 <223 <400 Met 1 Pro	l> LE 2> TY 3> OR 0> FE 3> OT 0> SE Met Leu	GANI GANI HER LYS	PRT SM: SM: EE: INFO ICE: Arg	Arti DRMAT 220 Glu 5	Lys Thr	Asn Ala	Val Gly	e of Ile Glu 25	Tyr 10 Met	Gly Leu	Pro Phe	Glu Arg	Pro Ala 30	Leu 15 Leu	Arg
<211 <212 <213 <223 <400 Met 1 Pro	l> LE 2> TY 3> OF 3> OT 3> OT 0> SE Met Leu	CATUF CANICATUR CHER CQUEN Lys Glu Ser 35	I: 54 PRT SM: E: INFO ICE: Arg Asp 20 Tyr	Arti DRMAT 220 Glu 5 Leu	Lys Thr	Asn Ala Gln	Val Gly Ala 40	e of Ile Glu 25 Leu	Tyr 10 Met Val	Gly Leu Asp	Pro Phe Val	Glu Arg Val 45	Pro Ala 30 Gly	Leu 15 Leu Asp	Arg Glu
<211 <212 <213 <220 <223 <400 Met 1 Pro	L> LE 2> TYY 3> OR 3> OT 3> OT 1> SE Met Leu His	NGTH PE: GANI ATUF HER QUEN Lys Glu Ser 35 Ser	I: 54 PRT SM: SM: INFO ICE: Arg Asp 20 Tyr	Arti Arti 220 Glu 5 Leu	Lys Thr Pro Glu	Asn Ala Gln Phe 55	Val Gly Ala 40	e of Ile Glu 25 Leu Glu	Tyr 10 Met Val	Gly Leu Asp	Pro Phe Val Val 60	Glu Arg Val 45 Leu	Pro Ala 30 Gly Leu	Leu 15 Leu Asp	Arg Glu Gln
<211 <212 <213 <220 <223 <400 Met 1 Pro Lys Ser 65	L> LE NOTE NOTE NOTE NOTE NOTE NOTE NOTE NOTE	NGTH PE: GANI ATUF HER QUEN Lys Glu Ser 35 Ser His	I: 54 PRT SM: SM: INFO ICE: Arg Asp 20 Tyr Tyr Asn	Arti RMAT 220 Glu 5 Leu Leu Lys	Lys Thr Pro Glu Gly 70	Asn Ala Gln Phe 55	Val Gly Ala 40 Phe	Ile Glu 25 Leu Glu	Tyr 10 Met Val Ala	Gly Leu Asp Thr Asp	Pro Phe Val Val 60 Val	Glu Arg Val 45 Leu Val	Pro Ala 30 Gly Leu Ser	Leu 15 Leu Asp Ala	Arg Glu Gln Cys 80
<211 <212 <213 <220 <223 <400 Met 1 Pro Lys Ser 65 Ala	L> LE TY S> OR S> FE S> OI Net Leu His Leu 50 Leu Glu	NGTH PE: GANI ATUF HER QUEN Lys Glu Ser 35 Ser His	I: 54 PRT SM: RE: INFO GCE: Arg Asp 20 Tyr Tyr Asn Asn	Arti RMAT 220 Glu 5 Leu Leu Lys Cys	Lys Thr Pro Glu Gly 70 Arg	Asn Ala Gln Phe 55 Tyr	Val Gly Ala 40 Phe Lys	Ile Glu 25 Leu Glu Met	Tyr 10 Met Val Ala Asn Pro	Gly Leu Asp Thr Asp 75	Pro Phe Val Val 60 Val	Glu Arg Val 45 Leu Val	Pro Ala 30 Gly Leu Ser	Leu 15 Leu Asp Ala Ile Trp 95	Arg Glu Gln Cys 80 Tyr
<211 <212 <213 <220 <223 <400 Met 1 Pro Lys Ser 65 Ala Ile	L> LE TY SO OR SO OR SO OR HO Leu His Leu Glu Gly	NGTH PE: GGANI HER QUEN Lys Glu Ser 35 Ser His Asn	I: 54 PRT ISM: ISM: ISM: INFO ISM: INFO INFO INFO INFO INFO INFO INFO INFO	Arti ORMAT 220 Glu 5 Leu Leu Lys Cys Thr 85	Lys Thr Pro Glu Gly 70 Arg	Asn Ala Gln Phe 55 Tyr Phe	Val Gly Ala 40 Phe Lys Phe Val	Ile Glu 25 Leu Glu Met Ile Asn 105	Tyr 10 Met Val Ala Asn Pro 90 Glu	Gly Leu Asp Thr Asp 75 Val	Pro Phe Val Val 60 Val Ile	Glu Arg Val 45 Leu Val Ala	Pro Ala 30 Gly Leu Ser Ala Pro 110	Leu 15 Leu Asp Ala Ile Trp 95	Arg Glu Gln Cys 80 Tyr
<211 <212 <213 <220 <223 <400 Met 1 Pro Lys Ser Ser 65 Ala Ile	L> LE TY S FE S FE S OT Met Leu His Leu Glu Gly Cys	NGTH PE: GGANI ATUF HER QUEN Lys Glu Ser 35 Ser His Asn Met Lys	I: 54 PRT: ISM: ISM: ISM: INFO ICE: Arg Asp 20 Tyr Tyr Asn Asn Ile 100 Val	Arti ORMAI 220 Glu 5 Leu Leu Lys Cys Thr 85 Val	Lys Thr Pro Glu Gly 70 Arg Ala Gly	Asn Ala Gln Phe 55 Tyr Phe Ile	Val Gly Ala 40 Phe Lys Phe Val Ser 120	Glu 25 Leu Glu Met Ile Asn 105 Lys	Tyr 10 Met Val Ala Asn Pro 90 Glu	Gly Leu Asp Thr Asp 75 Val Ser Gln	Pro Phe Val Val 60 Val Ile Tyr	Glu Arg Val 45 Leu Val Ala Ile Val 125	Pro Ala 30 Gly Leu Ser Ala Pro 110	Leu 15 Leu Asp Ala Ile Trp 95 Asp Thr	Arg Glu Gln Cys 80 Tyr Glu Thr

345

145					150					155					160
Glu	Ser	Leu	Pro	Asn 165	Phe	Ile	Ser	Arg	Ty r 170	Ser	Asp	Gly	Asn	Ile 175	Ala
Asn	Phe	Lys	Pro 180	Leu	His	Phe	Asp	Pro 185	Val	Glu	Gln	Val	Ala 190	Ala	Ile
Leu	Cys	Ser 195	Ser	Gly	Thr	Thr	Gly 200	Leu	Pro	Lys	Gly	Val 205	Met	Gln	Thr
His	Gln 210	Asn	Ile	Cys	Val	Arg 215	Leu	Ile	His	Ala	Leu 220	Asp	Pro	Arg	Tyr
Gl y 225	Thr	Gln	Leu	Ile	Pro 230	Gly	Val	Thr	Val	Leu 235	Val	Tyr	Leu	Pro	Phe 240
Phe	His	Ala	Phe	Gly 245	Phe	His	Ile	Thr	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly
Leu	Arg	Val	Ile 260	Met	Phe	Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys
Ala	Ile	Gln 275	Asp	Tyr	Glu	Val	Arg 280	Ser	Val	Ile	Asn	Val 285	Pro	Ser	Val
Ile	Leu 290	Phe	Leu	Ser	Lys	Ser 295	Pro	Leu	Val	Asp	Lys 300	Tyr	Asp	Leu	Ser
Ser 305	Leu	Arg	Glu	Leu	Cys 310	Суѕ	Gly	Ala	Ala	Pro 315	Leu	Ala	Lys	Glu	Val 320
Ala	Glu	Val	Ala	Ala 325	Lys	Arg	Leu	Asn	Leu 330	Pro	Gly	Ile	Arg	Cys 335	Gly
Phe	Gly	Leu	Thr 340	Glu	Ser	Thr	Ser	Ala 345	Ile	Ile	Gln	Ser	Leu 350	Arg	Asp
Glu	Phe	L y s 355	Ser	Gly	Ser	Leu	Gly 360	Arg	Val	Thr	Pro	Leu 365	Met	Ala	Ala
Lys	Ile 370	Ala	Asp	Arg	Glu	Thr 375	Gly	Lys	Ala	Leu	Gly 380	Pro	Asn	Gln	Val
Gly 385	Glu	Leu	Cys	Ile	Lys 390	Gly	Pro	Met	Val	Ser 395	Lys	Gly	Tyr	Val	Asn 400
Asn	Val	Glu	Ala	Thr 405	Lys	Glu	Ala	Ile	Asp 410	Asp	Asp	Gly	Trp	Leu 415	His
Ser	Gly	Asp	Phe 420	Gly	Tyr	Tyr	Asp	Glu 425	Asp	Glu	His	Phe	Tyr 430	Val	Val
Asp	Arg	Ty r 435	Lys	Glu	Leu	Ile	Lys 440	Tyr	Lys	Gly	Ser	Gln 445	Val	Ala	Pro
Ala	Glu 450	Leu	Glu	Glu	Ile	Leu 455	Leu	Lys	Asn	Pro	Cys 460	Ile	Arg	Asp	Val
Ala 465	Val	Val	Gly	Ile	Pro 470	Asp	Leu	Glu	Ala	Gly 475	Glu	Leu	Pro	Ser	Ala 480
Phe	Val	Val	Lys	Gln 485	Pro	Gly	Lys	Glu	Ile 490	Thr	Ala	Lys	Glu	Val 495	Tyr
Asp	Tyr	Leu	Ala 500	Glu	Arg	Val	Ser	His 505	Thr	Lys	Tyr	Leu	Arg 510	Gly	Gly
Val	Arg	Phe 515	Val	Asp	Ser	Ile	Pro 520	Arg	Asn	Val	Thr	Gly 525	Lys	Ile	Thr
Arg	Lys 530	Glu	Leu	Leu	Lys	Gln 535	Leu	Leu	Glu	Lys	Ala 540	Gly	Gly		

<212 <213 <220	2> TY 3> OF 0> FE	PE: GANI ATUF	SM: RE:	Arti			equen		: a s	yntl	netio	: luc	cifer	ase	
<400)> SE	QUEN	ICE:	221											
Met 1	Met	Lys	Arg	Glu 5	Lys	Asn	Val	Ile	Ty r 10	Gly	Pro	Glu	Pro	Leu 15	His
Pro	Leu	Glu	Asp 20	Leu	Thr	Ala	Gly	Glu 25	Met	Leu	Phe	Arg	Ala 30	Leu	Arg
Lys	His	Ser 35	His	Leu	Pro	Gln	Ala 40	Leu	Val	Asp	Val	Val 45	Gly	Asp	Glu
Ser	Leu 50	Ser	Tyr	Lys	Glu	Phe 55	Phe	Glu	Ala	Thr	Val 60	Leu	Leu	Ala	Gln
Ser 65	Leu	His	Asn	Суѕ	Gly 70	Tyr	Lys	Met	Asn	Asp 75	Val	Val	Ser	Ile	C y s 80
Ala	Glu	Asn	Asn	Thr 85	Arg	Phe	Phe	Ile	Pro 90	Val	Ile	Ala	Ala	Trp 95	Tyr
Ile	Gly	Met	Ile 100	Val	Ala	Pro	Val	Asn 105	Glu	Ser	Tyr	Ile	Pro 110	Asp	Glu
Leu	Cys	Lys 115	Val	Met	Gly	Ile	Ser 120	Lys	Pro	Gln	Ile	Val 125	Phe	Thr	Thr
Lys	Asn 130	Ile	Leu	Asn	Lys	Val 135	Leu	Glu	Val	Gln	Ser 140	Arg	Thr	Asn	Phe
Ile 145	Lys	Arg	Ile	Ile	Ile 150	Leu	Asp	Thr	Val	Glu 155	Asn	Ile	His	Gly	Cys 160
Glu	Ser	Leu	Pro	Asn 165	Phe	Ile	Ser	Arg	Ty r 170	Ser	Asp	Gly	Asn	Ile 175	Ala
Asn	Phe	Lys	Pro 180	Leu	His	Phe	Asp	Pro 185	Val	Glu	Gln	Val	Ala 190	Ala	Ile
Leu	Суѕ	Ser 195	Ser	Gly	Thr	Thr	Gl y 200	Leu	Pro	Lys	Gly	Val 205	Met	Gln	Thr
His	Gln 210	Asn	Ile	Cys	Val	Arg 215	Leu	Ile	His	Ala	Leu 220	Asp	Pro	Arg	Tyr
Gl y 225	Thr	Gln	Leu	Ile	Pro 230	Gly	Val	Thr	Val	Leu 235	Val	Tyr	Leu	Pro	Phe 240
Phe	His	Ala	Phe	Gly 245	Phe	His	Ile	Thr	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly
Leu	Arg	Val	Ile 260	Met	Phe	Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys
Ala	Ile	Gln 275	Asp	Tyr	Glu	Val	Arg 280	Ser	Val	Ile	Asn	Val 285	Pro	Ser	Val
Ile	Leu 290	Phe	Leu	Ser	Lys	Ser 295	Pro	Leu	Val	Asp	L y s 300	Tyr	Asp	Leu	Ser
Ser 305	Leu	Arg	Glu	Leu	Cys 310	Cys	Gly	Ala	Ala	Pro 315	Leu	Ala	Lys	Glu	Val 320
Ala	Glu	Val	Ala	Ala 325	Lys	Arg	Leu	Asn	Leu 330	Pro	Gly	Ile	Arg	Cys 335	Gly
Phe	Gly	Leu	Thr 340	Glu	Ser	Thr	Ser	Ala 345	Ile	Ile	Gln	Ser	Leu 350	Arg	Asp
Glu	Phe	Lys 355	Ser	Gly	Ser	Leu	Gly 360	Arg	Val	Thr	Pro	Leu 365	Met	Ala	Ala

Gly Glu Leu Cys Ile Lys Gly Pro Met Val Ser Lys Gly Tyr Val Asn Asn Val Glu Ala Thr Lys Glu Ala Ile Asp Asp Asp Gly Trp Leu His Ser Gly Asp Phe Gly Tyr Tyr Asp Glu Asp Glu His Phe Tyr Val Val Asp Arg Tyr Lys Glu Leu Ile Lys Tyr Lys Gly Ser Gln Val Ala Pro $435 \ \ \, 440 \ \ \, 445$ Ala Glu Leu Glu Glu Ile Leu Leu Lys Asn Pro Cys Ile Arg Asp Val Ala Val Val Gly Ile Pro Asp Leu Glu Ala Gly Glu Leu Pro Ser Ala 465 470470475475 Phe Val Val Lys Gln Pro Gly Lys Glu Ile Thr Ala Lys Glu Val Tyr Asp Tyr Leu Ala Glu Arg Val Ser His Thr Lys Tyr Leu Arg Gly Gly Arg Lys Glu Leu Leu Lys Gln Leu Leu Glu Lys Ala Gly Gly 530 535 540<210> SEQ ID NO 222 <211> LENGTH: 542 <212> TYPE: PRT <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 222 Met Met Lys Arg Glu Lys Asn Val Ile Tyr Gly Pro Glu Pro Leu His Pro Leu Glu Asp Leu Thr Ala Gly Glu Met Leu Phe Arg Ala Leu Arg 20 25 30Lys His Ser His Leu Pro Gln Ala Leu Val Asp Val Val Gly Asp Glu 40 Ser Leu Ser Tyr Lys Glu Phe Phe Glu Ala Thr Val Leu Leu Ala Gln Ser Leu His Asn Cys Gly Tyr Lys Met Asn Asp Val Val Ser Ile Cys Ala Glu Asn Asn Thr Arg Phe Phe Ile Pro Val Ile Ala Ala Trp Tyr Ile Gly Met Ile Val Ala Pro Val Asn Glu Ser Tyr Ile Pro Asp Glu Leu Cys Lys Val Met Gly Ile Ser Lys Pro Gln Ile Val Phe Thr Thr Lys Asn Ile Leu Asn Lys Val Leu Glu Val Gln Ser Arg Thr Asn Phe

Lys Ile Ala Asp Arg Glu Thr Gly Lys Ala Leu Gly Pro Asn Gln Val

Asn Phe Lys Pro Leu His Phe Asp Pro Val Glu Gln Val Ala Ala Ile 185 Leu Cys Ser Ser Gly Thr Thr Gly Leu Pro Lys Gly Val Met Gln Thr $195 \hspace{1.5cm} 200 \hspace{1.5cm} 205 \hspace{1.5cm}$ His Gln Asn Ile Cys Val Arg Leu Ile His Ala Leu Asp Pro Arg Tyr Gly Thr Gln Leu Ile Pro Gly Val Thr Val Leu Val Tyr Leu Pro Phe 235 Phe His Ala Phe Gly Phe His Ile Thr Leu Gly Tyr Phe Met Val Gly 250 Leu Arg Val Ile Met Phe Arg Arg Phe Asp Gln Glu Ala Phe Leu Lys Ala Ile Gln Asp Tyr Glu Val Arg Ser Val Ile Asn Val Pro Ser Val Ile Leu Phe Leu Ser Lys Ser Pro Leu Val Asp Lys Tyr Asp Leu Ser 295 Ser Leu Arg Glu Leu Cys Cys Gly Ala Ala Pro Leu Ala Lys Glu Val Ala Glu Val Ala Ala Lys Arg Leu Asn Leu Pro Gly Ile Arg Cys Gly $325 \hspace{1.5cm} 330 \hspace{1.5cm} 335$ Phe Gly Leu Thr Glu Ser Thr Ser Ala Ile Ile Gln Ser Leu Gly Asp Glu Phe Lys Ser Gly Ser Leu Gly Arg Val Thr Pro Leu Met Ala Ala 355 360 Lys Ile Ala Asp Arg Glu Thr Gly Lys Ala Leu Gly Pro Asn Gln Val\$370\$Gly Glu Leu Cys Ile Lys Gly Pro Met Val Ser Lys Gly Tyr Val Asn 385 390395395 Asn Val Glu Ala Thr Lys Glu Ala Ile Asp Asp Asp Gly Trp Leu His 410 Ser Gly Asp Phe Gly Tyr Tyr Asp Glu Asp Glu His Phe Tyr Val Val 425 Asp Arg Tyr Lys Glu Leu Ile Lys Tyr Lys Gly Ser Gln Val Ala Pro $435 \ \ \, 440 \ \ \, 445$ Ala Glu Leu Glu Glu Ile Leu Leu Lys Asn Pro Cys Ile Arg Asp Val Ala Val Val Gly Ile Pro Asp Leu Glu Ala Gly Glu Leu Pro Ser Ala 465 470470475 Phe Val Val Lys Gln Pro Gly Lys Glu Ile Thr Ala Lys Glu Val Tyr 490 Asp Tyr Leu Ala Glu Arg Val Ser His Thr Lys Tyr Leu Arg Gly Gly Val Arg Phe Val Asp Ser Ile Pro Arg Asn Val Thr Gly Lys Ile Thr Arg Lys Glu Leu Leu Lys Gln Leu Leu Glu Lys Ala Gly Gly

<210> SEQ ID NO 223

<211> LENGTH: 542

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER	R INFORMA	TION: Se	quenc	e of	as	synth	netio	: luc	cifer	ase	
<400> SEQUE	NCE: 223										
Met Ile Ly:	arg Glu 5	Lys Asn	Val	Ile	Ty r 10	Gly	Pro	Glu	Pro	Leu 15	His
Pro Leu Glu	Asp Leu 20	Thr Ala	Gly	Glu 25	Met	Leu	Phe	Arg	Ala 30	Leu	Arg
Lys His Ser 35	r His Leu	Pro Gln	Ala 40	Leu	Val	Asp	Val	Val 45	Gly	Asp	Glu
Ser Leu Ser 50	Tyr Lys	Glu Phe 55	Phe	Glu	Ala	Thr	Val 60	Leu	Leu	Ala	Gln
Ser Leu Hi: 65	s Asn Cys	Gly Tyr 70	Lys	Met	Asn	Asp 75	Val	Val	Ser	Ile	Cys 80
Ala Glu Ası	Asn Thr 85	Arg Phe	Phe	Ile	Pro 90	Val	Ile	Ala	Ala	T rp 95	Tyr
Ile Gly Me	Ile Val	Ala Pro	Val	Asn 105	Glu	Ser	Tyr	Ile	Pro 110	Asp	Glu
Leu Cys Lys		Gly Ile	Ser 120	Lys	Pro	Gln	Ile	Val 125	Phe	Thr	Thr
Lys Asn Ile 130	e Leu Asn	Lys Val	Leu	Glu	Val	Gln	Ser 140	Arg	Thr	Asn	Phe
Ile Lys Are	g Ile Ile	Ile Leu 150	Asp	Thr	Val	Glu 155	Asn	Ile	His	Gly	Cys 160
Glu Ser Le	ı Pro Asn 165	Phe Ile	Ser	Arg	Ty r 170	Ser	Asp	Gly	Asn	Ile 175	Ala
Asn Phe Lys	Pro Leu 180	His Phe	Asp	Pro 185	Val	Glu	Gln	Val	Ala 190	Ala	Ile
Leu Cys Ser 19		Thr Thr	Gl y 200	Leu	Pro	Lys	Gly	Val 205	Met	Gln	Thr
His Gln Ası 210	n Ile Cys	Val Arg 215	Leu	Ile	His	Ala	Leu 220	Asp	Pro	Arg	Tyr
Gly Thr Gli 225	n Leu Ile	Pro Gly 230	Val	Thr	Val	Leu 235	Val	Tyr	Leu	Pro	Phe 240
Phe His Ala	a Phe Gly 245		Ile	Thr	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly
Leu Arg Va	l Ile Met 260	Phe Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys
Ala Ile Gli 27		Glu Val	Arg 280	Ser	Val	Ile	Asn	Val 285	Pro	Ser	Val
Ile Leu Pho 290	e Leu Ser	Lys Ser 295		Leu	Val	Asp	L y s 300	Tyr	Asp	Leu	Ser
Ser Leu Arg	g Glu Leu	Cys Cys 310	Gly	Ala	Ala	Pro 315	Leu	Ala	Lys	Glu	Val 320
Ala Glu Va	l Ala Ala 325		Leu	Asn	Leu 330	Pro	Gly	Ile	Arg	Cys 335	Gly
Phe Gly Let	Thr Glu 340	Ser Thr	Ser	Ala 345	Ile	Ile	Gln	Thr	Leu 350	Gly	Asp
Glu Phe Lys	_	Ser Leu	Gl y 360	Arg	Val	Thr	Pro	Leu 365	Met	Ala	Ala
Lys Ile Ala 370	a Asp Arg	Glu Thr 375		Lys	Ala	Leu	Gly 380	Pro	Asn	Gln	Val

Gly 385	Glu	Leu	Cys	Ile	L y s 390	Gly	Pro	Met	Val	Ser 395	Lys	Gly	Tyr	Val	Asn 400
Asn	Val	Glu	Ala	Thr 405	Lys	Glu	Ala	Ile	Asp 410	Asp	Asp	Gly	Trp	Leu 415	His
Ser	Gly	Asp	Phe 420	Gly	Tyr	Tyr	Asp	Glu 425	Asp	Glu	His	Phe	Tyr 430	Val	Val
Asp	Arg	Tyr 435	Lys	Glu	Leu	Ile	Lys 440	Tyr	Lys	Gly	Ser	Gln 445	Val	Ala	Pro
Ala	Glu 450	Leu	Glu	Glu	Ile	Leu 455	Leu	Lys	Asn	Pro	Cys 460	Ile	Arg	Asp	Val
Ala 465	Val	Val	Gly	Ile	Pro 470	Asp	Leu	Glu	Ala	Gly 475	Glu	Leu	Pro	Ser	Ala 480
Phe	Val	Val	Lys	Gln 485	Pro	Gly	Thr	Glu	Ile 490	Thr	Ala	Lys	Glu	Val 495	Tyr
Asp	Tyr	Leu	Ala 500	Glu	Arg	Val	Ser	His 505	Thr	Lys	Tyr	Leu	Arg 510	Gly	Gly
Val	Arg	Phe 515	Val	Asp	Ser	Ile	Pro 520	Arg	Asn	Val	Thr	Gly 525	Lys	Ile	Thr
Arg	L y s 530	Glu	Leu	Leu	Lys	Gln 535	Leu	Leu	Val	Lys	Ala 540	Gly	Gly		
<213 <213 <213	0> SE 1> LE 2> TY 3> OF	NGTH PE:	: 31 PRT SM:	l1 Reni	illa	reni	Lform	nis							
<400	O> SE	QUEN	ICE:	224											
Met 1	Thr	Ser	Lys	Val 5	Tyr	Asp	Pro	Glu	Gln 10	Arg	Lys	Arg	Met	Ile 15	Thr
Gly	Pro	Gln	Trp 20	Trp	Ala	Arg	Cys	Lys 25	Gln	Met	Asn	Val	Leu 30	Asp	Ser
Phe	Ile	Asn 35	Tyr	Tyr	Asp	Ser	Glu 40	Lys	His	Ala	Glu	Asn 45	Ala	Val	Ile
Phe	Leu 50	His	Gly	Asn	Ala	Ala 55	Ser	Ser	Tyr	Leu	Trp 60	Arg	His	Val	Val
Pro 65	His	Ile	Glu	Pro	Val 70	Ala	Arg	Сув	Ile	Ile 75	Pro	Asp	Leu	Ile	Gly 80
Met	Gly	Lys	Ser	Gly 85	Lys	Ser	Gly	Asn	Gly 90	Ser	Tyr	Arg	Leu	Leu 95	Asp
His	Tyr	Lys	Ty r 100	Leu	Thr	Ala	Trp	Phe 105	Glu	Leu	Leu	Asn	Leu 110	Pro	Lys
Lys	Ile	Ile 115	Phe	Val	Gly	His	Asp 120	Trp	Gly	Ala	Cys	Leu 125	Ala	Phe	His
Tyr	Ser 130	Tyr	Glu	His	Gln	Asp 135	Lys	Ile	Lys	Ala	Ile 140	Val	His	Ala	Glu
Ser 145	Val	Val	Asp	Val	Ile 150	Glu	Ser	Trp	Asp	Glu 155	Trp	Pro	Asp	Ile	Glu 160
Glu	Asp	Ile	Ala	Leu 165	Ile	Lys	Ser	Glu	Glu 170	Gly	Glu	Lys	Met	Val 175	Leu
Glu	Asn	Asn	Phe 180	Phe	Val	Glu	Thr	Met 185	Leu	Pro	Ser	Lys	Ile 190	Met	Arg
Lys	Leu	Glu 195	Pro	Glu	Glu	Phe	Ala 200	Ala	Tyr	Leu	Glu	Pro 205	Phe	Lys	Glu

Lys	Gl y 210	Glu	Val	Arg	Arg	Pro 215	Thr	Leu	Ser	Trp	Pro 220	Arg	Glu	Ile	Pro
Leu 225	Val	Lys	Gly	Gly	L y s 230	Pro	Asp	Val	Val	Gln 235	Ile	Val	Arg	Asn	Ty r 240
Asn	Ala	Tyr	Leu	Arg 245	Ala	Ser	Asp	Asp	Leu 250	Pro	Lys	Met	Phe	Ile 255	Glu
Ser	Asp	Pro	Gly 260	Phe	Phe	Ser	Asn	Ala 265	Ile	Val	Glu	Gly	Ala 270	Lys	Lys
Phe	Pro	Asn 275	Thr	Glu	Phe	Val	L y s 280	Val	Lys	Gly	Leu	His 285	Phe	Ser	Gln
Glu	Asp 290	Ala	Pro	Asp	Glu	Met 295	Gly	Lys	Tyr	Ile	L y s 300	Ser	Phe	Val	Glu
Arg 305	Val	Leu	Lys	Asn	Glu 310	Gln									
<211 <212 <213 <220 <223)> FE 3> OT	NGTH PE: GANI ATUR HER	PRT SM: SM: INFO	1 Arti RMAT			-	ce e of	a s	ynth	ıetic	: luc	:ifer	ase	
<400)> SE	QUEN	CE:	225											
Met 1	Ala	Ser	Lys	Val 5	Tyr	Asp	Pro	Glu	Gln 10	Arg	Lys	Arg	Met	Ile 15	Thr
Gly	Pro	Gln	Trp 20	Trp	Ala	Arg	Сув	Lys 25	Gln	Met	Asn	Val	Leu 30	Asp	Ser
Phe	Ile	Asn 35	Tyr	Tyr	Asp	Ser	Glu 40	Lys	His	Ala	Glu	Asn 45	Ala	Val	Ile
Phe	Leu 50	His	Gly	Asn	Ala	Ala 55	Ser	Ser	Tyr	Leu	Trp 60	Arg	His	Val	Val
Pro 65	His	Ile	Glu	Pro	Val 70	Ala	Arg	Сув	Ile	Ile 75	Pro	Asp	Leu	Ile	Gly 80
Met	Gly	Lys	Ser	Gly 85	Lys	Ser	Gly	Asn	Gl y 90	Ser	Tyr	Arg	Leu	Leu 95	Asp
His	Tyr	Lys	Ty r 100	Leu	Thr	Ala	Trp	Phe 105	Glu	Leu	Leu	Asn	Leu 110	Pro	Lys
Lys	Ile	Ile 115	Phe	Val	Gly	His	Asp 120	Trp	Gly	Ala	Суѕ	Leu 125	Ala	Phe	His
Tyr	Ser 130	Tyr	Glu	His	Gln	Asp 135	Lys	Ile	Lys	Ala	Ile 140	Val	His	Ala	Glu
Ser 145	Val	Val	Asp	Val	Ile 150	Glu	Ser	Trp	Asp	Glu 155	Trp	Pro	Asp	Ile	Glu 160
Glu	Asp	Ile	Ala	Leu 165	Ile	Lys	Ser	Glu	Glu 170	Gly	Glu	Lys	Met	Val 175	Leu
Glu	Asn	Asn	Phe 180	Phe	Val	Glu	Thr	Met 185	Leu	Pro	Ser	Lys	Ile 190	Met	Arg
Lys	Leu	Glu 195	Pro	Glu	Glu	Phe	Ala 200	Ala	Tyr	Leu	Glu	Pro 205	Phe	Lys	Glu
Lys	Gly 210	Glu	Val	Arg	Arg	Pro 215	Thr	Leu	Ser	Trp	Pro 220	Arg	Glu	Ile	Pro
Leu 225	Val	Lys	Gly	Gly	Lys 230	Pro	Asp	Val	Val	Gln 235	Ile	Val	Arg	Asn	Ty r 240

Asn	Ala	Tyr	Leu	Arg 245	Ala	Ser	Asp	Asp	Leu 250	Pro	Lys	Met	Phe	Ile 255	Glu
Ser	Asp	Pro	Gly 260	Phe	Phe	Ser	Asn	Ala 265	Ile	Val	Glu	Gly	Ala 270	Lys	Lys
Phe	Pro	A sn 275	Thr	Glu	Phe	Val	Lys 280	Val	Lys	Gly	Leu	His 285	Phe	Ser	Gln
Glu	Asp 290	Ala	Pro	Asp	Glu	Met 295	Gly	Lys	Tyr	Ile	Lys 300	Ser	Phe	Val	Glu
Arg 305	Val	Leu	Lys	Asn	Glu 310	Gln									
<211 <212 <213 <220)> SE .> LE !> TY B> OR 0> FE B> OT	NGTH PE: GANI ATUR	: 31 PRT SM: RE:	1 Arti			_		ac	.vn+h	etic	· luc	ifer	·ASE	
)> SE					200	, u o	0 01		7	.0010	- 140		420	
Met 1	Ala	Ser	Lys	Val 5	Tyr	Asp	Pro	Glu	Gln 10	Arg	Lys	Arg	Met	Ile 15	Thr
Gly	Pro	Gln	Trp 20	Trp	Ala	Arg	Cys	Lys 25	Gln	Met	Asn	Val	Leu 30	Asp	Ser
Phe	Ile	Asn 35	Tyr	Tyr	Asp	Ser	Glu 40	Lys	His	Ala	Glu	Asn 45	Ala	Val	Ile
Phe	Leu 50	His	Gly	Asn	Ala	Ala 55	Ser	Ser	Tyr	Leu	Trp 60	Arg	His	Val	Val
Pro 65	His	Ile	Glu	Pro	Val 70	Ala	Arg	Cys	Ile	Ile 75	Pro	Asp	Leu	Ile	Gly 80
Met	Gly	Lys	Ser	Gly 85	Lys	Ser	Gly	Asn	Gl y 90	Ser	Tyr	Arg	Leu	Leu 95	Asp
His	Tyr	Lys	Tyr 100	Leu	Thr	Ala	Trp	Phe 105	Glu	Leu	Leu	Asn	Leu 110	Pro	Lys
Lys	Ile	Ile 115	Phe	Val	Gly	His	Asp 120	Trp	Gly	Ala	Cys	Leu 125	Ala	Phe	His
Tyr	Ser 130	Tyr	Glu	His	Gln	Asp 135	Lys	Ile	Lys	Ala	Ile 140	Val	His	Ala	Glu
Ser 145	Val	Val	Asp	Val	Ile 150	Glu	Ser	Trp	Asp	Glu 155	Trp	Pro	Asp	Ile	Glu 160
Glu	Asp	Ile	Ala	Leu 165	Ile	Lys	Ser	Glu	Glu 170	Gly	Glu	Lys	Met	Val 175	Leu
Glu	Asn	Asn	Phe 180	Phe	Val	Glu	Thr	Met 185	Leu	Pro	Ser	Lys	Ile 190	Met	Arg
Lys	Leu	Glu 195	Pro	Glu	Glu	Phe	Ala 200	Ala	Tyr	Leu	Glu	Pro 205	Phe	Lys	Glu
Lys	Gly 210	Glu	Val	Arg	Arg	Pro 215	Thr	Leu	Ser	Trp	Pro 220	Arg	Glu	Ile	Pro
Leu 225	Val	Lys	Gly	Gly	L y s 230	Pro	Asp	Val	Val	Gln 235	Ile	Val	Arg	Asn	Ty r 240
Asn	Ala	Tyr	Leu	Arg 245	Ala	Ser	Asp	Asp	Leu 250	Pro	Lys	Met	Phe	Ile 255	Glu
Ser	Asp	Pro	Gl y 260	Phe	Phe	Ser	Asn	Ala 265	Ile	Val	Glu	Gly	Ala 270	Lys	Lys

Phe Pro Asn Thr Glu Phe Val Lys Val Lys Gly Leu His Phe Ser Gln 275 280 285
Glu Asp Ala Pro Asp Glu Met Gly Lys Tyr Ile Lys Ser Phe Val Glu 290 295 300
Arg Val Leu Lys Asn Glu Gln 305 310
<pre><210> SEQ ID NO 227 <211> LENGTH: 311 <212> TYPE: PRT <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase</pre>
<400> SEQUENCE: 227
Met Ala Ser Lys Val Tyr Asp Pro Glu Gln Arg Lys Arg Met Ile Thr 1 5 10 15
Gly Pro Gln Trp Trp Ala Arg Cys Lys Gln Met Asn Val Leu Asp Ser 20 25 30
Phe Ile Asn Tyr Tyr Asp Ser Glu Lys His Ala Glu Asn Ala Val Ile 35 40 45
Phe Leu His Gly Asn Ala Ala Ser Ser Tyr Leu Trp Arg His Val Val 50 55 60
Pro His Ile Glu Pro Val Ala Arg Cys Ile Ile Pro Asp Leu Ile Gly 65 70 75 80
Met Gly Lys Ser Gly Lys Ser Gly Asn Gly Ser Tyr Arg Leu Leu Asp 85 90 95
His Tyr Lys Tyr Leu Thr Ala Trp Phe Glu Leu Leu Asn Leu Pro Lys 100 105 110
Lys Ile Ile Phe Val Gly His Asp Trp Gly Ala Cys Leu Ala Phe His 115 120 125
Tyr Ser Tyr Glu His Gln Asp Lys Ile Lys Ala Ile Val His Ala Glu 130 135 140
Ser Val Val Asp Val Ile Glu Ser Trp Asp Glu Trp Pro Asp Ile Glu 145 150 150 160
Glu Asp Ile Ala Leu Ile Lys Ser Glu Glu Gly Glu Lys Met Val Leu 165 170 175
Glu Asn Asn Phe Phe Val Glu Thr Met Leu Pro Ser Lys Ile Met Arg 180 185 190
Lys Leu Glu Pro Glu Glu Phe Ala Ala Tyr Leu Glu Pro Phe Lys Glu 195 200 205
Lys Gly Glu Val Arg Arg Pro Thr Leu Ser Trp Pro Arg Glu Ile Pro 210 215 220
Leu Val Lys Gly Gly Lys Pro Asp Val Val Gln Ile Val Arg Asn Tyr 225 230 230 235 240
Asn Ala Tyr Leu Arg Ala Ser Asp Asp Leu Pro Lys Met Phe Ile Glu 245 250 255
Ser Asp Pro Gly Phe Phe Ser Asn Ala Ile Val Glu Gly Ala Lys Lys 260 265 270
Phe Pro Asn Thr Glu Phe Val Lys Val Lys Gly Leu His Phe Ser Gln 275 280 285
Glu Asp Ala Pro Asp Glu Met Gly Lys Tyr Ile Lys Ser Phe Val Glu 290 295 300

```
Arg Val Leu Lys Asn Glu Gln
<210> SEQ ID NO 228
<211> LENGTH: 14
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A consensus sequence
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)...(14)
<223> OTHER INFORMATION: n = A,T,C or G
<400> SEQUENCE: 228
yggmnnnnng ccaa
                                                                           14
<210> SEQ ID NO 229
<211> LENGTH: 38
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A primer
<400> SEQUENCE: 229
gtactgagac gacgccagcc caagcttagg cctgagtg
                                                                           38
<210> SEQ ID NO 230
<211> LENGTH: 38
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A primer
<400> SEQUENCE: 230
                                                                           38
ggcatgagcg tgaactgact gaactagcgg ccgccgag
<210> SEQ ID NO 231
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A primer
<400> SEQUENCE: 231
ggatcccatg gtgaagcgtg agaa
                                                                           24
<210> SEQ ID NO 232
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A primer
<400> SEQUENCE: 232
ggatcccatg gtgaaacgcg a
                                                                           21
<210> SEQ ID NO 233
<211> LENGTH: 31
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A primer
```

<400> SEQUENCE: 233	
ctagcttttt tttctagata atcatgaaga c	31
<210> SEQ ID NO 234 <211> LENGTH: 54 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: A primer <400> SEQUENCE: 234	
caaaaagctt ggcattccgg tactgttggt aaagccacca tggtgaagcg agag	54
<210> SEQ ID NO 235 <211> LENGTH: 26 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: A primer	
<400> SEQUENCE: 235	
caattgttgt tgttaacttg tttatt	26
<210> SEQ ID NO 236 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: A primer	
<400> SEQUENCE: 236	
aaccatggct tccaaggtgt acgaccccga gcaacgcaaa	40
<210> SEQ ID NO 237 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: A primer	
<400> SEQUENCE: 237	
gctctagaat tactgctcgt tcttcagcac gcgctccacg	40
<210> SEQ ID NO 238 <211> LENGTH: 31 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: A primer	
<400> SEQUENCE: 238	
cgctagccat ggcttcgaaa gtttatgatc c	31
<210> SEQ ID NO 239 <211> LENGTH: 25 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: A primer	
<400> SEQUENCE: 239	
ggccagtaac tctagaatta ttgtt	25

```
<210> SEQ ID NO 240
<211> LENGTH: 5
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: An oligonucleotide
<400> SEQUENCE: 240
tataa
                                                                          5
<210> SEQ ID NO 241
<211> LENGTH: 6
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: An oligonucleotide
<400> SEQUENCE: 241
stratg
                                                                           6
<210> SEQ ID NO 242
<211> LENGTH: 9
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: An oligonucleotide
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)...(9)
OTHER INFORMATION: n = A, T, C r G
<400> SEQUENCE: 242
mttncnnma
<210> SEQ ID NO 243
<211> LENGTH: 5
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: An oligonucleotide
<400> SEQUENCE: 243
tratg
<210> SEQ ID NO 244
<211> LENGTH: 7
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A consensus sequence
<400> SEQUENCE: 244
tgastma
                                                                          7
<210> SEQ ID NO 245
<211> LENGTH: 14
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A consensus sequence
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)...(14)
<223> OTHER INFORMATION: n = A, T, C or G
```

<400> SEQUENCE: 245	
yggmnnnnng ccaa	14
<210> SEQ ID NO 246 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 246	
aaccatggct tccaaggtgt acgaccccga gcaacgcaaa	40
<210> SEQ ID NO 247 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 247	
cgcatgatca ctgggcctca gtggtgggct cgctgcaagc	40
<210> SEQ ID NO 248 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 248	
aaatgaacgt gctggactcc ttcatcaact actatgattc	40
<210> SEQ ID NO 249 <211> LENGTH: 50 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 249	
cgagaagcac gccgagaacg ccgtgatttt tctgcatggt aacgctgcct	50
<210> SEQ ID NO 250 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 250	
ccagctacct gtggaggcac gtcgtgcctc acatcgagcc	40
<210> SEQ ID NO 251 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 251	
cgtggctaga tgcatcatcc ctgatctgat cggaatgggt	40

	SEQ ID NO 252 LENGTH: 40	
<212>	TYPE: DNA	
<213>	ORGANISM: Artificial Sequence	
	FEATURE:	
<223>	OTHER INFORMATION: An oligonucleotide	
	•	
<400>	SEQUENCE: 252	
aagtc	cggca agagcgggaa tggctcatat cgcctcctgg	40
<210>	SEQ ID NO 253	
<211>	LENGTH: 40	
<212>	TYPE: DNA	
<213>	ORGANISM: Artificial Sequence	
	FEATURE:	
<223>	OTHER INFORMATION: An oligonucleotide	
<400>	SEQUENCE: 253	
atcact	tacaa gtacctcacc gcttggttcg agctgctgaa	40
	SEQ ID NO 254	
	LENGTH: 40	
	TYPE: DNA	
	ORGANISM: Artificial Sequence	
	FEATURE:	
<223>	OTHER INFORMATION: An oligonucleotide	
-100-	CECHENCE - 254	
<400>	SEQUENCE: 254	
aa++ a	caaag aaaatcatct ttgtgggcca cgactggggg	40
CCCCC	saaag aaaaccaccc ccgcgggcca cgaccggggg	40
<210×	SEQ ID NO 255	
	LENGTH: 40	
	TYPE: DNA	
	ORGANISM: Artificial Sequence	
	FEATURE:	
	OTHER INFORMATION: An oligonucleotide	
	•	
<400>	SEQUENCE: 255	
gcttg	tctgg cctttcacta ctcctacgag caccaagaca	40
010	GTO TD NO 056	
	SEQ ID NO 256	
	LENGTH: 40	
	TYPE: DNA	
	ORGANISM: Artificial Sequence	
	FEATURE: OTHER INFORMATION: An oligonucleotide	
\2237	OTHER INFORMATION: All Oligonacteotide	
<400>	SEQUENCE: 256	
agatca	aaggc catcgtccat gctgagagtg tcgtggacgt	40
,		
<210>	SEQ ID NO 257	
<211>	LENGTH: 45	
<212>	TYPE: DNA	
<213>	ORGANISM: Artificial Sequence	
	FEATURE:	
<223>	OTHER INFORMATION: An oligonucleotide	
<400>	SEQUENCE: 257	
		45
gatcg	agtcc tgggacgagt ggcctgacat cgaggaggat atcgc	45
-210-	CEO ID NO 250	
	SEQ ID NO 258 LENGTH: 40	
	TYPE: DNA	
~~ + ~ >	TITLE PARTY	

	-concinaea
<pre><213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 258	
cctgatcaag agcgaagagg gcgagaaaat ggtgcttgag	40
<pre><210> SEQ ID NO 259 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 259	
aataacttct tcgtcgagac catgctccca agcaagatca	40
<210> SEQ ID NO 260 <211> LENGTH: 45 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 260	
tgcggaaact ggagcctgag gagttcgctg cctacctgga	gccat 45
<pre><210> SEQ ID NO 261 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide <400> SEQUENCE: 261</pre>	
tcaaggagaa gggcgaggtt agacggccta ccctctcctg	40
<210> SEQ ID NO 262 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 262	
gcctcgcgag atccctctcg ttaagggagg caagcccgac	40
<210> SEQ ID NO 263 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 263	
gtcgtccaga ttgtccgcaa ctacaacgcc taccttcggg	40
<210> SEQ ID NO 264 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	

	-continued	
<400> SEQUENCE: 264		
ccagcgacga tctgcctaag atgttcatcg agtccgaccc		40
<210> SEQ ID NO 265 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide		
<400> SEQUENCE: 265		
tgggttcttt tccaacgcta ttgtcgaggg agctaagaag		40
<pre><210> SEQ ID NO 266 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide <400> SEQUENCE: 266</pre>		
ttccctaaca ccgagttcgt gaaggtgaag ggcctccact		40
<210> SEQ ID NO 267 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide		
<400> SEQUENCE: 267		
tcagccagga ggacgctcca gatgaaatgg gtaagtacat		40
<210> SEQ ID NO 268 <211> LENGTH: 49 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide		
<400> SEQUENCE: 268		
caagagette gtggagegeg tgetgaagaa egageagtaa	ttctagagc	49
<210> SEQ ID NO 269 <211> LENGTH: 29 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide		
<400> SEQUENCE: 269		
gctctagaat tactgctcgt tcttcagca	:	29
<210> SEQ ID NO 270 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide		
<400> SEQUENCE: 270		
cgcgctccac gaagctcttg atgtacttac ccatttcatc		40

<211><212><213><223>	SEQ ID NO 271 LENGTH: 40 TYPE: DNA ORGANISM: Artificial Sequence FEATURE: OTHER INFORMATION: An oligonucleotide SEQUENCE: 271	
	cgtcc tcctggctga agtggaggcc cttcaccttc	40
<211><212><213><220><223><400>	SEQ ID NO 272 LENGTH: 40 TYPE: DNA ORGANISM: Artificial Sequence FEATURE: OTHER INFORMATION: An oligonucleotide SEQUENCE: 272	40
acgaa	stegy tyttagygaa teetetaget teetegataa	40
<211> <212> <213> <220>	SEQ ID NO 273 LENGTH: 40 TYPE: DNA ORGANISM: Artificial Sequence FEATURE: OTHER INFORMATION: An oligonucleotide	
<400>	SEQUENCE: 273	
tagcg	ttgga aaagaaccca gggtcggact cgatgaacat	40
<211><212><213><220><223>	SEQ ID NO 274 LENGTH: 40 TYPE: DNA ORGANISM: Artificial Sequence FEATURE: OTHER INFORMATION: An oligonucleotide SEQUENCE: 274	
cttag	gcaga tegtegetgg eeegaaggta ggegttgtag	40
<211> <212> <213> <220>	SEQ ID NO 275 LENGTH: 40 TYPE: DNA ORGANISM: Artificial Sequence FEATURE: OTHER INFORMATION: An oligonucleotide	
<400>	SEQUENCE: 275	
ttgcg	gacaa tetggacgae gtegggettg eetecettaa	40
<211><212><213><223>	SEQ ID NO 276 LENGTH: 40 TYPE: DNA ORGANISM: Artificial Sequence FEATURE: OTHER INFORMATION: An oligonucleotide SEQUENCE: 276	
	gggat ctcgcgaggc caggagaggg taggccgtct	40
<211>	SEQ ID NO 277 LENGTH: 40 TYPE: DNA	

<pre><213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	9	
<400> SEQUENCE: 277		
aacctcgccc ttctccttga atggctccag gtaggcagc	3	40
<pre><210> SEQ ID NO 278 <211> LENGTH: 45 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	e.	
<400> SEQUENCE: 278		
aactcctcag gctccagttt ccgcatgatc ttgcttggg	a gcatg	45
<210> SEQ ID NO 279 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	à	
<400> SEQUENCE: 279		
gtctcgacga agaagttatt ctcaagcacc attttctcg	2	40
<pre><210> SEQ ID NO 280 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide <400> SEQUENCE: 280</pre>	9	
cetetteget ettgateagg gegatateet eetegatgte		40
<210> SEQ ID NO 281 <211> LENGTH: 43 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	9	
<400> SEQUENCE: 281		
aggccactcg tcccaggact cgatcacgtc cacgacactc	c tca	43
<210> SEQ ID NO 282 <211> LENGTH: 42 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	a	
<400> SEQUENCE: 282		
gcatggacga tggccttgat cttgtcttgg tgctcgtag	g ag	42
<210> SEQ ID NO 283 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	à	

<400> SEQUENCE: 283	
tagtgaaagg ccagacaagc cccccagtcg tggcccacaa	40
<pre><210> SEQ ID NO 284 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 284	
agatgatttt ctttggaagg ttcagcagct cgaaccaagc	4 0
<210> SEQ ID NO 285 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 285	
ggtgaggtac ttgtagtgat ccaggaggcg atatgagcca	4 0
<210> SEQ ID NO 286 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 286	
ttcccgctct tgccggactt acccattccg atcagatcag	4 0
<210> SEQ ID NO 287 <211> LENGTH: 45 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 287	
ggatgatgca tctagccacg ggctcgatgt gaggcacgac	gtgcc 45
<210> SEQ ID NO 288 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide	
<400> SEQUENCE: 288	
tccacaggta gctggaggca gcgttaccat gcagaaaaat	40
<pre><210> SEQ ID NO 289 <211> LENGTH: 45 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: An oligonucleotide</pre>	
<400> SEQUENCE: 289	
cacggcgttc tcggcgtgct tctcggaatc atagtagttg	atgaa 45

<211> <212> <213> <220>	SEQ ID NO 290 LENGTH: 40 TYPE: DNA ORGANISM: Artificial Sequence FEATURE: OTHER INFORMATION: An oligonucleotide	
	SEQUENCE: 290 ccagc acgttcattt gcttgcagcg agcccaccac	40
		10
<211> <212> <213> <220>	SEQ ID NO 291 LENGTH: 40 TYPE: DNA ORGANISM: Artificial Sequence FEATURE: OTHER INFORMATION: An oligonucleotide	
<400>	SEQUENCE: 291	
tgagg	cccag tgatcatgcg tttgcgttgc tcggggtcgt	40
<211> <212> <213> <220>	SEQ ID NO 292 LENGTH: 20 TYPE: DNA ORGANISM: Artificial Sequence FEATURE: OTHER INFORMATION: An oligonucleotide	
	SEQUENCE: 292	
acacc-	ttgga agccatggtt	20
<211><212><213><220><223>	SEQ ID NO 293 LENGTH: 10 TYPE: DNA ORGANISM: Artificial Sequence FEATURE: OTHER INFORMATION: A Kozak sequence	
	SEQUENCE: 293	1.0
aacca ⁻	egget	10
<211> <212> <213> <220>	SEQ ID NO 294 LENGTH: 12 TYPE: DNA ORGANISM: Artificial Sequence FEATURE: OTHER INFORMATION: An oligonucleotide	
<400>	SEQUENCE: 294	
taatt	ctaga gc	12
<211><212><213><220><223>	SEQ ID NO 295 LENGTH: 32 TYPE: DNA ORGANISM: Artificial Sequence FEATURE: OTHER INFORMATION: A primer SEQUENCE: 295	
	gccat ggtaaagcgt gagaaaaatg to	32
<211>	SEQ ID NO 296 LENGTH: 33 TYPE: DNA	

<213> ORGANISM: Artific <220> FEATURE: <223> OTHER INFORMATIO	-			
<400> SEQUENCE: 296				
ccgactctag attactaacc	gccggccttc acc		33	
<210> SEQ ID NO 297 <211> LENGTH: 1626 <212> TYPE: DNA <213> ORGANISM: Artific <220> FEATURE: <223> OTHER INFORMATIO	cial Sequence ON: Sequence of a synth	netic luciferase		
<400> SEQUENCE: 297				
atggtgaaac gcgaaaagaa	cgtgatctac ggcccagaac	cactgcatcc actggaage	ac 60	
ctcaccgctg gtgagatgct	cttccgagca ctgcgtaaac	atagtcacct ccctcaage	ca 120	
ctcgtggacg tcgtgggaga	cgagagcctc tcctacaaag	aatttttcga agctactg	tg 180	
ctgttggccc aaagcctcca	taattgtggg tacaaaatga	acgatgtggt gagcattte	gt 240	
gctgagaata acactcgctt	ctttattcct gtaatcgctg	cttggtacat cggcatga	tt 300	
gtcgcccctg tgaatgaatc	ttacatccca gatgagctgt	gtaaggttat gggtatta	gc 360	
aaacctcaaa tcgtctttac	taccaaaaac atcttgaata	aggtettgga agtecagte	ct 420	
cgtactaact tcatcaaacg	catcattatt ctggataccg	tcgaaaacat ccacggcto	gt 480	
gagageetee etaaetteat	ctctcgttac agcgatggta	atatcgctaa tttcaagc	cc 540	
ttgcattttg atccagtcga	gcaagtggcc gctattttgt	gctcctccgg caccactge	gt 600	
ttgcctaaag gtgtcatgca	gactcaccag aatatctgtg	tgcgtttgat ccacgctc	tc 660	
gaccctcgtg tgggtactca	attgatccct ggcgtgactg	tgctggtgta tctgcctt	tc 720	
tttcacgcct ttggtttctc	tattaccctg ggctatttca	tggtcggctt gcgtgtca	tc 780	
atgtttcgtc gcttcgacca	agaagccttc ttgaaggcta	ttcaagacta cgaggtgc	gt 840	
teegtgatea aegteeette	agtcattttg ttcctgagca	aatctccttt ggttgaca	ag 900	
tatgatctga gcagcttgcg	tgagctgtgc tgtggcgctg	ctcctttggc caaagaag		
gccgaggtcg ctgctaagcg	tctgaacctc cctggtatcc	gctgcggttt tggtttga		
gagagcactt ctgctaacat	ccatagcttg cgagacgagt	ttaagtctgg tagcctgg		
cgcgtgactc ctcttatggc	tgcaaagatc gccgaccgtg	agaccggcaa agcactgg		
	gtgtattaag ggccctatgg			
aatgtggagg ccactaaaga	agccattgat gatgatggct	ggctccatag cggcgact		
ggttactatg atgaggacga	acacttctat gtggtcgatc	gctacaaaga attgatta		
	accagccgaa ctggaagaaa			
atccgcgacg tggccgtcgt	gggtatccca gacttggaag	ctggcgagtt gcctagcg		
	caaggagatc actgctaagg			
	atatetgegt ggeggegtee			
cgcaacgtta ccggtaagat	cactcgtaaa gagttgctga	agcaactcct cgaaaaag		
ggegge			626	

<211> LENGTH: 542 <212> TYPE: PRT <213> ORGANISM: Artificial Sequence															
		ATUR		Artı	ncia	ıı se	equen	.ce							
<223	FO <	HER	INFC	RMAT	'ION:	Sec	uenc	e of	a s	ynth	etic	luc	ifer	ase	
<400> SEQUENCE: 298															
Met 1	Val	Lys	Arg	Glu 5	Lys	Asn	Val	Ile	Ty r 10	Gly	Pro	Glu	Pro	Leu 15	His
Pro	Leu	Glu	Asp 20	Leu	Thr	Ala	Gly	Glu 25	Met	Leu	Phe	Arg	Ala 30	Leu	Arg
Lys	His	Ser 35	His	Leu	Pro	Gln	Ala 40	Leu	Val	Asp	Val	Val 45	Gly	Asp	Glu
Ser	Leu 50	Ser	Tyr	Lys	Glu	Phe 55	Phe	Glu	Ala	Thr	Val 60	Leu	Leu	Ala	Gln
Ser 65	Leu	His	Asn	Суѕ	Gl y 70	Tyr	Lys	Met	Asn	Asp 75	Val	Val	Ser	Ile	Cys 80
Ala	Glu	Asn	Asn	Thr 85	Arg	Phe	Phe	Ile	Pro 90	Val	Ile	Ala	Ala	Trp 95	Tyr
Ile	Gly	Met	Ile 100	Val	Ala	Pro	Val	Asn 105	Glu	Ser	Tyr	Ile	Pro 110	Asp	Glu
Leu	Cys	Lys 115	Val	Met	Gly	Ile	Ser 120	Lys	Pro	Gln	Ile	Val 125	Phe	Thr	Thr
Lys	Asn 130	Ile	Leu	Asn	Lys	Val 135	Leu	Glu	Val	Gln	Ser 140	Arg	Thr	Asn	Phe
Ile 145	Lys	Arg	Ile	Ile	Ile 150	Leu	Asp	Thr	Val	Glu 155	Asn	Ile	His	Gly	C y s 160
Glu	Ser	Leu	Pro	Asn 165	Phe	Ile	Ser	Arg	Ty r 170	Ser	Asp	Gly	Asn	Ile 175	Ala
Asn	Phe	Lys	Pro 180	Leu	His	Phe	Asp	Pro 185	Val	Glu	Gln	Val	Ala 190	Ala	Ile
Leu	Cys	Ser 195	Ser	Gly	Thr	Thr	Gl y 200	Leu	Pro	Lys	Gly	Val 205	Met	Gln	Thr
His	Gln 210	Asn	Ile	Cys	Val	Arg 215	Leu	Ile	His	Ala	Leu 220	Asp	Pro	Arg	Val
Gl y 225	Thr	Gln	Leu	Ile	Pro 230	Gly	Val	Thr	Val	Leu 235	Val	Tyr	Leu	Pro	Phe 240
Phe	His	Ala	Phe	Gly 245	Phe	Ser	Ile	Thr	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly
Leu	Arg	Val	Ile 260	Met	Phe	Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys
Ala	Ile	Gln 275	Asp	Tyr	Glu	Val	Arg 280	Ser	Val	Ile	Asn	Val 285	Pro	Ser	Val
Ile	Leu 290	Phe	Leu	Ser	Lys	Ser 295	Pro	Leu	Val	Asp	L y s 300	Tyr	Asp	Leu	Ser
Ser 305	Leu	Arg	Glu	Leu	Cys 310	Сув	Gly	Ala	Ala	Pro 315	Leu	Ala	Lys	Glu	Val 320
Ala	Glu	Val	Ala	Ala 325	Lys	Arg	Leu	Asn	Leu 330	Pro	Gly	Ile	Arg	Cys 335	Gly
Phe	Gly	Leu	Thr 340	Glu	Ser	Thr	Ser	Ala 345	Asn	Ile	His	Ser	Leu 350	Arg	Asp
Glu	Phe	L y s 355	Ser	Gly	Ser	Leu	Gly 360	Arg	Val	Thr	Pro	Leu 365	Met	Ala	Ala

Lys Ile Ala Asp Arg Glu Thr Gly Lys Ala Leu Gly Pro Asn Gln Val 370 375 380	
Gly Glu Leu Cys Ile Lys Gly Pro Met Val Ser Lys Gly Tyr Val Asn 385 390 395 400	
Asn Val Glu Ala Thr Lys Glu Ala Ile Asp Asp Asp Gly Trp Leu His 405 410 415	
Ser Gly Asp Phe Gly Tyr Tyr Asp Glu Asp Glu His Phe Tyr Val Val 420 425 430	
Asp Arg Tyr Lys Glu Leu Ile Lys Tyr Lys Gly Ser Gln Val Ala Pro 435 440 445	
Ala Glu Leu Glu Glu Ile Leu Leu Lys Asn Pro Cys Ile Arg Asp Val 450 455 460	
Ala Val Val Gly Ile Pro Asp Leu Glu Ala Gly Glu Leu Pro Ser Ala 465 470 475 480	
Phe Val Val Lys Gln Pro Gly Lys Glu Ile Thr Ala Lys Glu Val Tyr 485 490 495	
Asp Tyr Leu Ala Glu Arg Val Ser His Thr Lys Tyr Leu Arg Gly Gly 500 505 510	
Val Arg Phe Val Asp Ser Ile Pro Arg Asn Val Thr Gly Lys Ile Thr 515 520 525	
Arg Lys Glu Leu Leu Lys Gln Leu Leu Glu Lys Ala Gly Gly 530 540	
<211> LENGTH: 1626 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence	
<pre><220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 299</pre>	
<pre><220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase</pre>	60
<220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 299	60 120
<pre><220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 299 atggtgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat</pre>	
<220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 299 atggtgaage gtgagaaaaa tgtcatctat ggccctgage etetecatee tttggaggat ttgactgccg gcgaaatget gtttcgtget etecgcaage acteteattt gcctcaagee	120
<220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 299 atggtgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actctcattt gcctcaagcc ttggtcgatg tggtcggcga tgaatctttg agctacaagg agtttttga ggcaaccgtc	120 180
<pre><220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 299 atggtgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actctcattt gcctcaagcc ttggtcgatg tggtcggcga tgaatctttg agctacaagg agtttttga ggcaaccgtc ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt</pre>	120 180 240
<pre><220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 299 atggtgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actctcattt gcctcaagcc ttggtcgatg tggtcggcga tgaatctttg agctacaagg agtttttga ggcaaccgtc ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc</pre>	120 180 240 300
<pre><220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 299 atggtgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actctcattt gcctcaagcc ttggtcgatg tggtcggcga tgaatctttg agctacaagg agtttttga ggcaaccgtc ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc</pre>	120 180 240 300 360 420 480
<pre><220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 299 atggggaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actctcattt gcctcaagcc ttggtcgatg tggtcggcga tgaatctttg agctacaagg agtttttga ggcaaccgtc ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc gtggctccag tcaaccgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc gaatctttgc ctaatttcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca</pre>	120 180 240 300 360 420 480 540
<pre><220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 299 atggtgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actctcattt gcctcaagcc ttggtcgatg tggtcggcga tgaatctttg agctacaagg agtttttga ggcaaccgtc ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc gaatctttgc ctaatttcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca ctccacttcg accctgtgga acaagttgca gccattctgt gtagcagcgg tactactgga</pre>	120 180 240 300 360 420 480 540
<pre><220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 299 atggggaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actctcattt gcctcaagcc ttggtcgatg tggtcggcga tgaatctttg agctacaagg agtttttga ggcaaccgtc ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc gtggctccag tcaaccgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc cgcaccaact ttattaagcg tatcatcatc ttggacactg tggagaatat tcacggttgc gaatctttgc ctaatttcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca ctccacttcg accctgtgga acaagttgca gccattctgt gtagcagcgg tactactgga ctcccaaaagg gagtcatgca gacccatcaa aacatttgcg tgcgtctgat ccatgctctc</pre>	120 180 240 300 360 420 480 540 600
<pre><220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 299 atggtgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actctcattt gcctcaagcc ttggtcgatg tggtcggcga tgaatctttg agctacaagg agtttttga ggcaaccgtc ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc cgcaccaact ttattaagcg tacatcatc ttggacactg tggagaatat tcacggttgc gaatctttgc ctaatttcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca ctccacttcg accctgtgga acaagttgca gccattctgt gtagcagcgg tactactgga ctcccaaagg gagtcatgca gacccatcaa aacatttgcg tgcgtctgat ccatgctctc gatccacgct acggcactca gctgattcct ggtgtcaccg tcttggtcta cttgcctttc</pre>	120 180 240 300 360 420 480 540
<pre><220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 299 atgggaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actctcattt gcctcaagcc ttggtcgatg tggtcggcga tgaatctttg agctacaagg agtttttga ggcaaccgtc ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc cgcaccaact ttattaagcg tacatcatc ttggacactg tggagaatat tcacggttgc gaatctttgc ctaattcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca ctccactcg accctgtgga acaagttgca gccattctgt gtagcagcgg tactactgga ctcccaaagg gagtcatgca gacccatcaa aacatttgcg tgcgtctgat ccatgctctc gatccacgct acggcactca gctgattcct ggtgtcaccg tcttggtcta cttgccttct ttccatgctt tcggctttca tattactttg ggttacttta tggtcggtct ccgcgtgatt</pre>	120 180 240 300 360 420 480 540 600 720
<pre><220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 299 atggtgaagc gtgagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actctcattt gcctcaagcc ttggtcgatg tggtcggcga tgaatctttg agctacaagg agtttttga ggcaaccgtc ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct aagccacaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc cgcaccaact ttattaagcg tacatcatc ttggacactg tggagaatat tcacggttgc gaatctttgc ctaatttcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca ctccacttcg accctgtgga acaagttgca gccattctgt gtagcagcgg tactactgga ctcccaaagg gagtcatgca gacccatcaa aacatttgcg tgcgtctgat ccatgctctc gatccacgct acggcactca gctgattcct ggtgtcaccg tcttggtcta cttgcctttc</pre>	120 180 240 300 360 420 480 540 600 660 720
<pre><220> FEATURE: <223> OTHER INFORMATION: Sequence of a synthetic luciferase <400> SEQUENCE: 299 atggggaagc gtggagaaaaa tgtcatctat ggccctgagc ctctccatcc tttggaggat ttgactgccg gcgaaatgct gtttcgtgct ctccgcaagc actctcattt gcctcaagcc ttggtcgatg tggtcggcga tgaatctttg agctacaagg agtttttga ggcaaccgtc ttgctggctc agtccctcca caattgtggc tacaagatga acgacgtcgt tagtatctgt gctgaaaaca atacccgttt cttcattcca gtcatcgccg catggtatat cggtatgatc gtggctccag tcaacgagag ctacattccc gacgaactgt gtaaagtcat gggtatctct aagccacaaga ttgtcttcac cactaagaat attctgaaca aagtcctgga agtccaaagc cgcaccaact ttattaagcg tacatcatc ttggacactg tggagaatat tcacggttgc gaatctttgc ctaatttcat ctctcgctat tcagacggca acatcgcaaa ctttaaacca ctccacttcg accctgtgga acaagttgca gccattctgt gtagcagcgg tactactgga ctcccaaagg gagtcatgca gacccatcaa aacatttgcg tgcgtctgat ccatgctctc gatccacgct acggcactca gctgattcct ggtgtcaccg tcttggtcta cttgccttct ttccatgctt tcggctttca tattactttg ggttacttta tggtcggtct ccgcgtgatt atgttccgcc gttttgatca ggaggctttc ttgaaagcca tccaagatta tgaagtccgc</pre>	120 180 240 300 360 420 480 540 600 720 780 840

gctgaagtgg	ccgccaaacg	cttgaatct	t ccagggattc	gttgtggctt	cggcctcacc 020
gaatctacca	gcgctattat	tcagtctct	c cgcgatgagt	ttaagagcgg	ctctttgggc 080
cgtgtcactc	cactcatggc	tgctaagat	c gctgatcgcg	aaactggtaa	ggctttgggc 140
ccgaaccaag	tgggcgagct	gtgtatcaa	a ggccctatgg	tgagcaaggg	ttatgtcaat 200
aacgttgaag	ctaccaagga	ggccatcga	c gacgacggct	ggttgcattc	tggtgatttt 260
ggatattacg	acgaagatga	gcatttta	c gtcgtggatc	gttacaagga	gctgatcaaa 320
tacaagggta	gccaggttgc	tccagctga	g ttggaggaga	ttctgttgaa	aaatccatgc 380
attcgcgatg	tcgctgtggt	cggcattcc	t gatctggagg	ccggcgaact	gccttctgct 440
ttcgttgtca	agcagcctgg	taaagaaat	t accgccaaag	aagtgtatga	ttacctggct 500
gaacgtgtga	gccatactaa	gtacttgcg	t ggcggcgtgc	gttttgttga	ctccatccct 560
cgtaacgtaa	caggcaaaat	tacccgcaa	g gagctgttga	aacaattgtt	ggagaaggcc 620
ggcggt					626
<220> FEAT	TH: 542 : PRT NISM: Artifi URE:	-	nce ce of a syntl	netic lucife	rase
<400> SEQU	ENCE: 300				
Met Val Ly 1	s Arg Glu L 5	ys Asn Val	Ile Tyr Gly 10	Pro Glu Pro	Leu His 15
Pro Leu Gl	u Asp Leu T 20	hr Ala Gly	Glu Met Leu 25	Phe Arg Ala	Leu Arg
Lys His Se 35	r His Leu P	ro Gln Ala 40	Leu Val Asp	Val Val Gly 45	Asp Glu
Ser Leu Se 50	r Tyr Lys G	lu Phe Phe 55	Glu Ala Thr	Val Leu Leu 60	Ala Gln
Ser Leu Hi 65	s Asn Cys G 7		Met Asn Asp 75	Val Val Ser	Ile Cys 80
Ala Glu As	n Asn Thr A 85	rg Phe Phe	Ile Pro Val 90	Ile Ala Ala	Trp Tyr 95
Ile Gly Me	t Ile Val A 100	la Pro Val	Asn Glu Ser 105	Tyr Ile Pro	-
Leu Cys Ly		ly Ile Ser 120	Lys Pro Gln	Ile Val Phe 125	Thr Thr
Lys Asn Il 130	e Leu Asn L	ys Val Leu 135	Glu Val Gln	Ser Arg Thr 140	Asn Phe
Ile Lys Ar 145		le Leu Asp 50	Thr Val Glu 155	Asn Ile His	Gly Cys 160
Glu Ser Le	u Pro Asn P 165	he Ile Ser	Arg Tyr Ser 170	Asp Gly Asn	Ile Ala 175
Asn Phe Ly	s Pro Leu H 180	is Phe Asp	Pro Val Glu 185	Gln Val Ala	
Leu Cys Se 19		hr Thr Gl y 200	Leu Pro Lys	Gly Val Met 205	Gln Thr

His Gln Asn Ile Cys Val Arg Leu Ile His Ala Leu Asp Pro Arg Tyr 210 $$ 215 $$ 220

													<u> </u>	<u> </u>		
Gly 225	Thr	Gln	Leu	Ile	Pro 230	Gly	Val	Thr	Val	Leu 235	Val	Tyr	Leu	Pro	Phe 240	
Phe	His	Ala	Phe	Gly 245	Phe	His	Ile	Thr	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly	
Leu	Arg	Val	Ile 260	Met	Phe	Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys	
Ala	Ile	Gln 275	Asp	Tyr	Glu	Val	Arg 280	Ser	Val	Ile	Asn	Val 285	Pro	Ser	Val	
Ile	Leu 290	Phe	Leu	Ser	Lys	Ser 295	Pro	Leu	Val	Asp	Lys 300	Tyr	Asp	Leu	Ser	
Ser 305	Leu	Arg	Glu	Leu	Cys 310	Cys	Gly	Ala	Ala	Pro 315	Leu	Ala	Lys	Glu	Val 320	
Ala	Glu	Val	Ala	Ala 325	Lys	Arg	Leu	Asn	Leu 330	Pro	Gly	Ile	Arg	Cys 335	Gly	
Phe	Gly	Leu	Thr 340	Glu	Ser	Thr	Ser	Ala 345	Ile	Ile	Gln	Ser	Leu 350	Arg	Asp	
Glu	Phe	Lys 355	Ser	Gly	Ser	Leu	Gly 360	Arg	Val	Thr	Pro	Leu 365	Met	Ala	Ala	
Lys	Ile 370	Ala	Asp	Arg	Glu	Thr 375	Gly	Lys	Ala	Leu	Gly 380	Pro	Asn	Gln	Val	
Gly 385	Glu	Leu	Cys	Ile	L y s 390	Gly	Pro	Met	Val	Ser 395	Lys	Gly	Tyr	Val	Asn 400	
Asn	Val	Glu	Ala	Thr 405	Lys	Glu	Ala	Ile	Asp 410	Asp	Asp	Gly	Trp	Leu 415	His	
Ser	Gly	Asp	Phe 420	Gly	Tyr	Tyr	Asp	Glu 425	Asp	Glu	His	Phe	Tyr 430	Val	Val	
Asp	Arg	Ty r 435	Lys	Glu	Leu	Ile	Lys 440	Tyr	Lys	Gly	Ser	Gln 445	Val	Ala	Pro	
Ala	Glu 450	Leu	Glu	Glu	Ile	Leu 455	Leu	Lys	Asn	Pro	Cys 460	Ile	Arg	Asp	Val	
Ala 465	Val	Val	Gly	Ile	Pro 470	Asp	Leu	Glu	Ala	Gly 475	Glu	Leu	Pro	Ser	Ala 480	
Phe	Val	Val	Lys	Gln 485	Pro	Gly	Lys	Glu	Ile 490	Thr	Ala	Lys	Glu	Val 495	Tyr	
Asp	Tyr	Leu	Ala 500	Glu	Arg	Val	Ser	His 505	Thr	Lys	Tyr	Leu	Arg 510	Gly	Gly	
Val	Arg	Phe 515	Val	Asp	Ser	Ile	Pro 520	Arg	Asn	Val	Thr	Gly 525	Lys	Ile	Thr	
Arg	Lys 530	Glu	Leu	Leu	Lys	Gln 535	Leu	Leu	Glu	Lys	Ala 540	Gly	Gly			
<211 <212 <213 <220 <223	.> LE ?> TY 3> OF 0> FE 3> OT	ENGTH PE: RGANI EATUR THER	SM: RE:	Arti Arti			-		ās	s y ntl	netio	: luc	cifer	rase		
		-			aa to	gtcat	tctat	t ggo	cacto	gagc	ctc	tcca-	taa -	tttg	gaggat	60
ttgá	actgo	ccg (gcga	aatg	ct gi	tttc	gtgci	t cto	ccgca	aagc	acto	ctca	ttt «	gccto	caagcc	120
ttg	gtaga	atg t	tggto	egge	ga to	gaato	cttt	g ago	ctaca	aagg	agt	tttt	tga (ggca	accgtc	180

ttgctggctc	agtccctcca	caattgtggc	tacaagatga	acgacgtcgt	tagtatctgt	240
gctgaaaaca	atacccgttt	cttcattcca	gtcatcgccg	catggtatat	cggtatgatc	300
gtggctccag	tcaacgagag	ctacattccc	gacgaactgt	gtaaagtcat	gggtatctct	360
aagccacaga	ttgtcttcac	cactaagaat	attctgaaca	aagtcctgga	agtccaaagc	420
cgcaccaact	ttattaagcg	tatcatcatc	ttggacactg	tggagaatat	tcacggttgc	480
gaatctttgc	ctaatttcat	ctctcgctat	tcagacggca	acatcgcaaa	ctttaaacca	540
ctccacttcg	accctgtgga	acaagttgca	gccattctgt	gtagcagcgg	tactactgga	600
ctcccaaagg	gagtcatgca	gacccatcaa	aacatttgcg	tgcgtctgat	ccatgctctc	660
gatccacgct	acggcactca	gctgattcct	ggtgtcaccg	tcttggtcta	cttgcctttc	720
ttccatgctt	tcggctttca	tattactttg	ggttacttta	tggtcggtct	ccgcgtgatt	780
atgttccgcc	gttttgatca	ggaggctttc	ttgaaagcca	tccaagatta	tgaagtccgc	840
agtgtcatca	acgtgcctag	cgtgatcctg	tttttgtcta	agagcccact	cgtggacaag	900
tacgacttgt	cttcactgcg	tgaattgtgt	tgcggtgccg	ctccactggc	taaggaggtc	960
gctgaagtgg	ccgccaaacg	cttgaatctt	ccagggattc	gttgtggctt	cggcctcacc	1020
gaatctacca	gtgcgattat	ccagactctc	ggggatgagt	ttaagagcgg	ctctttgggc	1080
cgtgtcactc	cactcatggc	tgctaagatc	gctgatcgcg	aaactggtaa	ggctttgggc	1140
ccgaaccaag	tgggcgagct	gtgtatcaaa	ggccctatgg	tgagcaaggg	ttatgtcaat	1200
aacgttgaag	ctaccaagga	ggccatcgac	gacgacggct	ggttgcattc	tggtgatttt	1260
ggatattacg	acgaagatga	gcatttttac	gtcgtggatc	gttacaagga	gctgatcaaa	1320
tacaagggta	gccaggttgc	tccagctgag	ttggaggaga	ttctgttgaa	aaatccatgc	1380
attcgcgatg	tegetgtggt	cggcattcct	gatctggagg	ccggcgaact	gccttctgct	1440
ttcgttgtca	agcagcctgg	tacagaaatt	accgccaaag	aagtgtatga	ttacctggct	1500
gaacgtgtga	gccatactaa	gtacttgcgt	ggcggcgtgc	gttttgttga	ctccatccct	1560
cgtaacgtaa	caggcaaaat	tacccgcaag	gagctgttga	aacaattgtt	ggtgaaggcc	1620
ggcggt						1626

```
<210> SEQ ID NO 302
```

Met Val Lys Arg Glu Lys Asn Val Ile Tyr Gly Pro Glu Pro Leu His 1 $$ 10 $$ 15

Lys His Ser His Leu Pro Gln Ala Leu Val Asp Val Val Gly Asp Glu $35 \ \ \,$ 40 $\ \ \,$ 45

Ser Leu His Asn Cys Gly Tyr Lys Met Asn Asp Val Val Ser Ile Cys 65 70 75 80

Ala Glu Asn Asn Thr Arg Phe Phe Ile Pro Val Ile Ala Ala Trp Tyr

<211> LENGTH: 542

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Sequence of a synthetic luciferase

<400> SEQUENCE: 302

				85					90					95	
Ile	Gly	Met	Ile 100	Val	Ala	Pro	Val	Asn 105	Glu	Ser	Tyr	Ile	Pro	Asp	Glu
Leu	Cys	L y s 115	Val	Met	Gly	Ile	Ser 120	Lys	Pro	Gln	Ile	Val 125	Phe	Thr	Thr
Lys	Asn 130	Ile	Leu	Asn	Lys	Val 135	Leu	Glu	Val	Gln	Ser 140	Arg	Thr	Asn	Phe
Ile 145	Lys	Arg	Ile	Ile	Ile 150	Leu	Asp	Thr	Val	Glu 155	Asn	Ile	His	Gly	Cys 160
Glu	Ser	Leu	Pro	Asn 165	Phe	Ile	Ser	Arg	Ty r 170	Ser	Asp	Gly	Asn	Ile 175	Ala
Asn	Phe	Lys	Pro 180	Leu	His	Phe	Asp	Pro 185	Val	Glu	Gln	Val	Ala 190	Ala	Ile
Leu	Cys	Ser 195	Ser	Gly	Thr	Thr	Gly 200	Leu	Pro	Lys	Gly	Val 205	Met	Gln	Thr
His	Gln 210	Asn	Ile	Cys	Val	Arg 215	Leu	Ile	His	Ala	Leu 220	Asp	Pro	Arg	Tyr
Gly 225	Thr	Gln	Leu	Ile	Pro 230	Gly	Val	Thr	Val	Leu 235	Val	Tyr	Leu	Pro	Phe 240
Phe	His	Ala	Phe	Gly 245	Phe	His	Ile	Thr	Leu 250	Gly	Tyr	Phe	Met	Val 255	Gly
Leu	Arg	Val	Ile 260	Met	Phe	Arg	Arg	Phe 265	Asp	Gln	Glu	Ala	Phe 270	Leu	Lys
Ala	Ile	Gln 275	Asp	Tyr	Glu	Val	A rg 280	Ser	Val	Ile	Asn	Val 285	Pro	Ser	Val
Ile	Leu 290	Phe	Leu	Ser	Lys	Ser 295	Pro	Leu	Val	Asp	Lys 300	Tyr	Asp	Leu	Ser
Ser 305	Leu	Arg	Glu	Leu	Cys 310	Cys	Gly	Ala	Ala	Pro 315	Leu	Ala	Lys	Glu	Val 320
Ala	Glu	Val	Ala	Ala 325	Lys	Arg	Leu	Asn	Leu 330	Pro	Gly	Ile	Arg	Cys 335	Gly
Phe	Gly	Leu	Thr 340	Glu	Ser	Thr	Ser	Ala 345	Ile	Ile	Gln	Thr	Leu 350	Gly	Asp
Glu	Phe	Lys 355	Ser	Gly	Ser	Leu	Gly 360	Arg	Val	Thr	Pro	Leu 365	Met	Ala	Ala
Lys	Ile 370	Ala	Asp	Arg	Glu	Thr 375	Gly	Lys	Ala	Leu	Gly 380	Pro	Asn	Gln	Val
Gly 385	Glu	Leu	Cys	Ile	Lys	Gly	Pro	Met	Val	Ser	Lys	Gly	Tyr	Val	Asn 400
Asn	Val	Glu	Ala	Thr	Lys	Glu	Ala	Ile	Asp	Asp	Asp	Gly	Trp	Leu 415	His
Ser	Gly	Asp	Phe		Tyr	Tyr	Asp	Glu 425		Glu	His	Phe	Ty r 430		Val
Asp	Arg	Tyr 435		Glu	Leu	Ile	L y s 440		Lys	Gly	Ser	Gln 445		Ala	Pro
Ala	Glu		Glu	Glu	Ile	Leu		Lys	Asn	Pro	Cys		Arg	Asp	Val
														-	

	450					455					460				
Ala 465	Val	Val	Gly	Ile	Pro 470	Asp	Leu	Glu	Ala	Gl y 475	Glu	Leu	Pro	Ser	Ala 480
Phe	Val	Val	Lys	Gln 485	Pro	Gly	Thr	Glu	Ile 490	Thr	Ala	Lys	Glu	Val 495	Tyr
Asp	Tyr	Leu	Ala 500	Glu	Arg	Val	Ser	His 505	Thr	Lys	Tyr	Leu	Arg 510	Gly	Gly
Val	Arg	Phe 515	Val	Asp	Ser	Ile	Pro 520	Arg	Asn	Val	Thr	Gly 525	Lys	Ile	Thr
Arg	Lys 530	Glu	Leu	Leu	Lys	Gln 535	Leu	Leu	Val	Lys	Ala 540	Gly	Gly		

1-47. (canceled)

- **48**. A method to prepare a synthetic nucleic acid molecule comprising an open reading frame, comprising:
 - a) altering a plurality of transcription regulatory sequences in a parent nucleic acid sequence which encodes a polypeptide having at least 100 amino acids to yield a synthetic nucleic acid molecule which has at least 3-fold fewer transcription regulatory sequences relative to the parent nucleic acid sequence, wherein the transcription regulatory sequences are selected from the group consisting of transcription factor binding sequences, intron splice sites, poly(A) addition sites, enhancer sequences and promoter sequences; and
 - b) altering greater than 25% of the codons in the synthetic nucleic acid sequence which has a decreased number of transcription regulatory sequences to yield a further synthetic nucleic acid molecule, wherein the codons which are altered do not result in an increased number of transcription regulatory sequences, wherein the further synthetic nucleic acid molecule encodes a polypeptide with at least 85% amino acid sequence identity to the polypeptide encoded by the parent nucleic acid sequence.
- **49**. A method to prepare a synthetic nucleic acid molecule comprising an open reading frame, comprising:
 - a) altering greater than 25% of the codons in a parent nucleic acid sequence which encodes a polypeptide having at least 100 amino acids to yield a codon-altered synthetic nucleic acid molecule, and
 - b) altering a plurality of transcription regulatory sequences in the codon-altered synthetic nucleic acid molecule to yield a further synthetic nucleic acid molecule which has at least 3-fold fewer transcription regulatory sequences relative to a synthetic nucleic acid molecule with a random selection of codons at the codons which differ, wherein the transcription regulatory sequences are selected from the group consisting of transcription factor binding sequences, intron splice sites, poly(A) addition sites, enhancer sequences and promoter sequences, and wherein the further synthetic nucleic acid molecule encodes a polypeptide with at least 85% amino acid sequence identity to the polypeptide encoded by the parent nucleic acid sequence.
- **50**. The method of claim 48 or 49 wherein the parent nucleic acid sequence encodes a reporter molecule.

- **51**. The method of claim 48 or 49 wherein the parent nucleic acid sequence encodes a luciferase.
- **52**. The method of claim 48 or 49 wherein the synthetic nucleic acid molecule hybridizes under medium stringency hybridization conditions to the parent nucleic acid sequence.
- **53**. The method of claim 48 or 49 wherein the codons which are altered encode the same amino acid as the corresponding codons in the parent nucleic acid sequence.
 - 54. (canceled)
- **55**. A method for preparing at least two synthetic nucleic acid molecules which are codon distinct versions of a parent nucleic acid sequence which encodes a polypeptide, comprising:
 - a) altering a parent nucleic acid sequence to yield a synthetic nucleic acid molecule having an increased number of a first plurality of codons that are employed more frequently in a selected host cell relative to the number of those codons in the parent nucleic acid sequence; and
 - b) altering the parent nucleic acid sequence to yield a further synthetic nucleic acid molecule having an increased number of a second plurality of codons that are employed more frequently in the host cell relative to the number of those codons in the parent nucleic acid sequence, wherein the first plurality of codons is different than the second plurality of codons, and wherein the synthetic and the further synthetic nucleic acid molecules encode the same polypeptide.
- **56**. The method of claim 55 further comprising altering a plurality of transcription regulatory sequences in the synthetic nucleic acid molecule, the further synthetic nucleic acid molecule, or both, to yield at least one yet further synthetic nucleic acid molecule which has at least 3-fold fewer transcription regulatory sequences relative to the synthetic nucleic acid molecule, the further synthetic nucleic acid molecule, or both.
- **57**. The method of claim 55 further comprising altering at least one codon in the first synthetic sequence to yield a first modified synthetic sequence which encodes a polypeptide with at least one amino acid substitution relative to the polypeptide encoded by the first synthetic nucleic acid sequence.
- **58**. The method of claim 56 further comprising altering at least one codon in the second synthetic sequence to yield a second modified synthetic sequence which encodes a

polypeptide with at least one amino acid substitution relative to the polypeptide encoded by the first synthetic nucleic acid sequence.

59. The method of claim 55 wherein the synthetic sequences encode a luciferase.

60-64. (canceled)

65. The method of claim 48 or 49 further comprising altering the further synthetic nucleic acid molecule to

encode a polypeptide having at least one amino acid substitution relative to the polypeptide encoded by the parent nucleic acid sequence.

66. The method of claim 48 or 49 wherein the altering of transcription regulatory sequences does not introduce amino acid substitutions to the polypeptide encoded by the synthetic nucleic acid molecule.

* * * * *