

# SEQUENCE LISTING

<110> F. Hoffmann - La Roche AG

<120> Prediction of Genotoxicity

<130> 24699WO

<160> 9

<170> PatentIn version 3.5

<210> 1

<211> 2328

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (3)..(3)

<223> n is a, c, g, t or u

<400> 1

```

rgngcgggaa gcaggggcg ggccctctggt ggcggtcggg aactcgggtg gaggcggcaa      60
cattgtttca agttggccaa attgacaaga gcgagaggta tactgcgttc catcccgacc      120
cgggggccacg gtactgggcc ctgtttcccc ctccctcgcc cccgagagcc aggggtccgcc      180
ttctgcaggg ttcccaggcc cccgctccag ggccgggctg acccgactcg ctggcgcttc      240
atggagaact tccaaaaggt ggaaaagatc ggagagggca cgtacggagt tgtgtacaaa      300
gccagaaaca agttgacggg agaggtggtg gcgcttaaga aaatccgcct ggacactgag      360
actgagggtg tgcccagtac tgccatccga gagatctctc tgcttaagga gcttaaccat      420
cctaataattg tcaagctgct ggatgtcatt cacacagaaa ataaactcta cctgggttttt      480
gaattttctgc accaagatct caagaaattc atggatgcct ctgctctcac tggcattcct      540
cttccccctca tcaagagcta tctgttccag ctgctccagg gcctagcttt ctgccattct      600
catcgggtcc tccaccgaga ccttaaacct cagaatctgc ttattaacac agagggggcc      660
atcaagctag cagactttgg actagccaga gcttttggag tccctgttcg tacttacacc      720
catgaggtgg tgaccctgtg gtaccgagct cctgaaatcc tcctgggctg caaatattat      780
tccacagctg tggacatctg gagcctgggc tgcattcttg ctgagatggt gactcgccgg      840
gccctattcc ctggagattc tgagattgac cagctcttcc ggatctttcg gactctgggg      900
accccagatg aggtgggtg gccaggagtt acttctatgc ctgattacaa gccaagtttc      960
cccaagtggg ccgggcaaga ttttagtaaa gttgtacctc ccctggatga agatggacgg      1020
agcttggttat cgcaaagtgt gcactacgac cctaacaagc ggatttcggc caaggcagcc      1080
ctggctcacc ctttcttcca ggatgtgacc aagccagtac cccatcttcg actctgatag      1140
ccttcttgaa gccccagcc ctaatctcac cctctcctcc agtgtgggct tgaccaggct      1200

```

tggccttggg	ctatttggac	tcaggtgggc	cctctgaact	tgccttaaac	actcaccttc	1260
tagtcttggc	cagccaactc	tgggaataca	ggggtgaaag	gggggaacca	gtgaaaatga	1320
aaggaagttt	cagtattaga	tgcacttaag	ttagcctcca	ccaccctttc	ccccttctct	1380
tagttattgc	tgaagaggg	tgggtataaaa	ataattttta	aaaagccttc	ctacacgtta	1440
gatttgccgt	accaatctct	gaatgcccc	taattattat	ttccagtgtt	tgggatgacc	1500
aggatcccaa	gcctcctgct	gccacaatgt	ttataaaggc	caaagtatag	cgggggctaa	1560
gttggtgctt	ttgagaacca	agtaaaacaa	aaccactggg	aggagtctat	tttaaagaat	1620
tcggttgaaa	aaatagatcc	aatcagttta	taccctagtt	agtgttttgc	ctcaccta	1680
aggctgggag	actgaagact	cagcccgggt	ggggctgcag	aaaaatgatt	ggccccagtc	1740
cccttgtttg	tcccttctac	aggcatgagg	aatctgggag	gccctgagac	agggattgtg	1800
cttcattcca	atctattgct	tcaccatggc	cttatgaggc	aggtgagaga	tgtttgaatt	1860
tttctcttcc	ttttagtatt	cttagttgtt	cagttgcaa	ggatccctga	tcccattttc	1920
ctctgacgtc	cacctcctac	cccataggag	ttagaagtta	gggttttaggc	atcattttga	1980
gaatgctgac	actttttcag	ggctgtgatt	gagtgagggc	atgggtaaaa	atatttcttt	2040
aaaagaagga	tgaacaatta	tatttatatt	tcaggttata	tccaatagta	gagttggctt	2100
tttttttttt	tttttggcca	tagtgggtgg	atgtgttgc	atgtgcacct	tggggttttg	2160
taatgacagt	gctaaaaaaa	aaaagcattt	tttttttatg	atgtgtctct	gtcacccctg	2220
tccttgagtg	ctcttgctat	taacgttatt	tgtaatttag	ttttagctc	attaaaaaaa	2280
tgtgcctagt	tttataaaaa	aaaaaaaaaa	aaacaaaaaa	aaaaaaaaaa		2328

<210> 2  
 <211> 1339  
 <212> RNA  
 <213> Homo sapiens

<400> 2	
ggcgagacga	gccgacgcag cagcgagcc cagggggcgc gacgcagcag cgcggaacccg 60
gagcgacaag	cgccccgacg gagcgccagg acgagagaca ccaaagagaa cacgccgagg 120
gagacaagga	gggaaggaaa aggaggagca gcagcagcaa aaagaaggaa gagacacaag 180
caggcccagg	agaacaagcg cgcaaaacaa caccaaaagg gaagccaagg aaagcaggca 240
aaagagaaaag	aacaagcgac gcacagagag acagaaagac acaccaagga ggaaccggac 300
acgccaaaaga	gaccagaaaag ccggacagaa ccaagagcaa gccggagaag ggaagaagag 360
aaaaagcaaa	cacaggacac cacagaccac acgcgcacag ggaagagcac cgaaggaaaa 420
gaaccaggag	gagaggagag aggagggcac cgacgcagag ggagacgaca aggcaagaag 480
aaaggaacag	ggaaggagcg gaaaagggga ggcacgacaa aagcgggagg agacagagca 540

gaaaaaagaa aaaggggaaga acggaagcgc cgccagaaaa caagcggaac acgaaacaac 600  
 agaccccaac agacccgcgg ccagagggaa gggagcacag gcacagcagg aacagggaca 660  
 gacacgacca aaagaaaagg caccacgacg gacaacagaa agaggcaaca gaagcaagcg 720  
 gaagcacaga aaaggaccac acagacaaag ccgaaaacac aggcagcgac acacagaggc 780  
 gaaacccaaa aaaaacggag aacgcaccaa aaaccagaaa aaggagacgg aggcaacaag 840  
 agacgaacac acagacagga cacaagacaa agagcaccga agaagcccag ggggccaac 900  
 caggagcggg gcaaggagca cagaaacacg ggaccgacca acacacgaag aaggagcaag 960  
 caagaggaaa ggacggaccc accaaaacaa gaacagaaaa ccaggaaacg aaaacaccac 1020  
 gacgaagacg ggagaacaca gcgccggcag aagcaagacg cgaaacccga aggaaagccc 1080  
 aagaggaaca gagcgccgac ccacagaaaa gggagagacc agccaaaaga aaccagaga 1140  
 agccaaagca cccgaccga agaaaagaaa gacgaaggac agccccgaag agacacagac 1200  
 gacagcaaaa aaagagacag cgaaacaaca aagccagagg ggaaggcaaa gacaagcaag 1260  
 gaagaacacc agaagaaagg aacagaaaaa ggcagaagga agacgaacag gcaagaccac 1320  
 ccacaaaaag aaggacaag 1339

<210> 3  
 <211> 629  
 <212> PRT  
 <213> Homo sapiens

<400> 3

Met Ala Val Pro Pro Gly His Gly Pro Phe Ser Gly Phe Pro Gly Pro  
 1 5 10 15

Gln Glu His Thr Gln Val Leu Pro Asp Val Arg Leu Leu Pro Arg Arg  
 20 25 30

Leu Pro Leu Ala Phe Arg Asp Ala Thr Ser Ala Pro Leu Arg Lys Leu  
 35 40 45

Ser Val Asp Leu Ile Lys Thr Tyr Lys His Ile Asn Glu Val Tyr Tyr  
 50 55 60

Ala Lys Lys Lys Arg Arg Ala Gln Gln Ala Pro Pro Gln Asp Ser Ser  
 65 70 75 80

Asn Lys Lys Glu Lys Lys Val Leu Asn His Gly Tyr Asp Asp Asp Asn  
 85 90 95

His Asp Tyr Ile Val Arg Ser Gly Glu Arg Trp Leu Glu Arg Tyr Glu  
 100 105 110

Ile Asp Ser Leu Ile Gly Lys Gly Ser Phe Gly Gln Val Val Lys Ala  
115 120 125

Tyr Asp His Gln Thr Gln Glu Leu Val Ala Ile Lys Ile Ile Lys Asn  
130 135 140

Lys Lys Ala Phe Leu Asn Gln Ala Gln Ile Glu Leu Arg Leu Leu Glu  
145 150 155 160

Leu Met Asn Gln His Asp Thr Glu Met Lys Tyr Tyr Ile Val His Leu  
165 170 175

Lys Arg His Phe Met Phe Arg Asn His Leu Cys Leu Val Phe Glu Leu  
180 185 190

Leu Ser Tyr Asn Leu Tyr Asp Leu Leu Arg Asn Thr His Phe Arg Gly  
195 200 205

Val Ser Leu Asn Leu Thr Arg Lys Leu Ala Gln Gln Leu Cys Thr Ala  
210 215 220

Leu Leu Phe Leu Ala Thr Pro Glu Leu Ser Ile Ile His Cys Asp Leu  
225 230 235 240

Lys Pro Glu Asn Ile Leu Leu Cys Asn Pro Lys Arg Ser Ala Ile Lys  
245 250 255

Ile Val Asp Phe Gly Ser Ser Cys Gln Leu Gly Gln Arg Ile Tyr Gln  
260 265 270

Tyr Ile Gln Ser Arg Phe Tyr Arg Ser Pro Glu Val Leu Leu Gly Thr  
275 280 285

Pro Tyr Asp Leu Ala Ile Asp Met Trp Ser Leu Gly Cys Ile Leu Val  
290 295 300

Glu Met His Thr Gly Glu Pro Leu Phe Ser Gly Ser Asn Glu Val Asp  
305 310 315 320

Gln Met Asn Arg Ile Val Glu Val Leu Gly Ile Pro Pro Ala Ala Met  
325 330 335

Leu Asp Gln Ala Pro Lys Ala Arg Lys Tyr Phe Glu Arg Leu Pro Gly  
340 345 350

Gly Gly Trp Thr Leu Arg Arg Thr Lys Glu Leu Arg Lys Asp Tyr Gln

355

360

365

Gly Pro Gly Thr Arg Arg Leu Gln Glu Val Leu Gly Val Gln Thr Gly  
370 375 380

Gly Pro Gly Gly Arg Arg Ala Gly Glu Pro Gly His Ser Pro Ala Asp  
385 390 395 400

Tyr Leu Arg Phe Gln Asp Leu Val Leu Arg Met Leu Glu Tyr Glu Pro  
405 410 415

Ala Ala Arg Ile Ser Pro Leu Gly Ala Leu Gln His Gly Phe Phe Arg  
420 425 430

Arg Thr Ala Asp Glu Ala Thr Asn Thr Gly Pro Ala Gly Ser Ser Ala  
435 440 445

Ser Thr Ser Pro Ala Pro Leu Asp Thr Cys Pro Ser Ser Ser Thr Ala  
450 455 460

Ser Ser Ile Ser Ser Ser Gly Gly Ser Ser Gly Ser Ser Ser Asp Asn  
465 470 475 480

Arg Thr Tyr Arg Tyr Ser Asn Arg Tyr Cys Gly Gly Pro Gly Pro Pro  
485 490 495

Ile Thr Asp Cys Glu Met Asn Ser Pro Gln Val Pro Pro Ser Gln Pro  
500 505 510

Leu Arg Pro Trp Ala Gly Gly Asp Val Pro His Lys Thr His Gln Ala  
515 520 525

Pro Ala Ser Ala Ser Ser Leu Pro Gly Thr Gly Ala Gln Leu Pro Pro  
530 535 540

Gln Pro Arg Tyr Leu Gly Arg Pro Pro Ser Pro Thr Ser Pro Pro Pro  
545 550 555 560

Pro Glu Leu Met Asp Val Ser Leu Val Gly Gly Pro Ala Asp Cys Ser  
565 570 575

Pro Pro His Pro Ala Pro Ala Pro Gln His Pro Ala Ala Ser Ala Leu  
580 585 590

Arg Thr Arg Met Thr Gly Gly Arg Pro Pro Leu Pro Pro Pro Asp Asp  
595 600 605

Pro Ala Thr Leu Gly Pro His Leu Gly Leu Arg Gly Val Pro Gln Ser  
610 615 620

Thr Ala Ala Ser Ser  
625

<210> 4  
<211> 544  
<212> PRT  
<213> Homo sapiens

<400> 4

Met Cys Thr Val Val Asp Pro Arg Ile Val Arg Arg Tyr Leu Leu Arg  
1 5 10 15

Arg Gln Leu Gly Gln Gly Ala Tyr Gly Ile Val Trp Lys Ala Val Asp  
20 25 30

Arg Arg Thr Gly Glu Val Val Ala Ile Lys Lys Ile Phe Asp Ala Phe  
35 40 45

Arg Asp Lys Thr Asp Ala Gln Arg Thr Phe Arg Glu Ile Thr Leu Leu  
50 55 60

Gln Glu Phe Gly Asp His Pro Asn Ile Ile Ser Leu Leu Asp Val Ile  
65 70 75 80

Arg Ala Glu Asn Asp Arg Asp Ile Tyr Leu Val Phe Glu Phe Met Asp  
85 90 95

Thr Asp Leu Asn Ala Val Ile Arg Lys Gly Gly Leu Leu Gln Asp Val  
100 105 110

His Val Arg Ser Ile Phe Tyr Gln Leu Leu Arg Ala Thr Arg Phe Leu  
115 120 125

His Ser Gly His Val Val His Arg Asp Gln Lys Pro Ser Asn Val Leu  
130 135 140

Leu Asp Ala Asn Cys Thr Val Lys Leu Cys Asp Phe Gly Leu Ala Arg  
145 150 155 160

Ser Leu Gly Asp Leu Pro Glu Gly Pro Glu Asp Gln Ala Val Thr Glu  
165 170 175

Tyr Val Ala Thr Arg Trp Tyr Arg Ala Pro Glu Val Leu Leu Ser Ser  
180 185 190

His Arg Tyr Thr Leu Gly Val Asp Met Trp Ser Leu Gly Cys Ile Leu  
195 200 205

Gly Glu Met Leu Arg Gly Arg Pro Leu Phe Pro Gly Thr Ser Thr Leu  
210 215 220

His Gln Leu Glu Leu Ile Leu Glu Thr Ile Pro Pro Pro Ser Glu Glu  
225 230 235 240

Asp Leu Leu Ala Leu Gly Ser Gly Cys Arg Ala Ser Val Leu His Gln  
245 250 255

Leu Gly Ser Arg Pro Arg Gln Thr Leu Asp Ala Leu Leu Pro Pro Asp  
260 265 270

Thr Ser Pro Glu Ala Leu Asp Leu Leu Arg Arg Leu Leu Val Phe Ala  
275 280 285

Pro Asp Lys Arg Leu Ser Ala Thr Gln Ala Leu Gln His Pro Tyr Val  
290 295 300

Gln Arg Phe His Cys Pro Ser Asp Glu Trp Ala Arg Glu Ala Asp Val  
305 310 315 320

Arg Pro Arg Ala His Glu Gly Val Gln Leu Ser Val Pro Glu Tyr Arg  
325 330 335

Ser Arg Val Tyr Gln Met Ile Leu Glu Cys Gly Gly Ser Ser Gly Thr  
340 345 350

Ser Arg Glu Lys Gly Pro Glu Gly Val Ser Pro Ser Gln Ala His Leu  
355 360 365

His Lys Pro Arg Ala Asp Pro Gln Leu Pro Ser Arg Thr Pro Val Gln  
370 375 380

Gly Pro Arg Pro Arg Pro Gln Ser Ser Pro Gly His Asp Pro Ala Glu  
385 390 395 400

His Glu Ser Pro Arg Ala Ala Lys Asn Val Pro Arg Gln Asn Ser Ala  
405 410 415

Pro Leu Leu Gln Thr Ala Leu Leu Gly Asn Gly Glu Arg Pro Pro Gly  
420 425 430

Ala Lys Glu Ala Pro Pro Leu Thr Leu Ser Leu Val Lys Pro Ser Gly  
435 440 445

Arg Gly Ala Ala Pro Ser Leu Thr Ser Gln Ala Ala Ala Gln Val Ala  
 450 455 460

Asn Gln Ala Leu Ile Arg Gly Asp Trp Asn Arg Gly Gly Gly Val Arg  
 465 470 475 480

Val Ala Ser Val Gln Gln Val Pro Pro Arg Leu Pro Pro Glu Ala Arg  
 485 490 495

Pro Gly Arg Arg Met Phe Ser Thr Ser Ala Leu Gln Gly Ala Gln Gly  
 500 505 510

Gly Ala Arg Ala Leu Leu Gly Gly Tyr Ser Gln Ala Tyr Gly Thr Val  
 515 520 525

Cys His Ser Ala Leu Gly His Leu Pro Leu Leu Glu Gly His His Val  
 530 535 540

<210> 5  
 <211> 1789  
 <212> RNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> (3)..(3)  
 <223> n is a, c, g, or u

<400> 5  
 rgncccaagc cagagcggcg cggccggaag aggccagggc ccgggggagg cggcggcagc 60  
 ggcggcggcg gggcagcccg ggcagcccga gccccgcagc cgggccggcc ggcgccagag 120  
 cggcggcggg cccgggaggc ggccccgggg ccgggcaggc cgcggacagc cgcgcggagc 180  
 ccggcggcgg aggcggagga ggcggcggcg gccccggagg ccggcccccg cccaggcggc 240  
 accggcggcg gaaaggcacg cggggccagg ggggggcgcg gggccccgag ccggggggga 300  
 cccggcggca gcggcggagg aggcagcggg ggccccggcg caggcacagc cccgccgccc 360  
 gggggaagcg ggccggacag cgggaaggga ccacagcgag ccaccaggcc aaggcccaga 420  
 gcgccccaa aagggcacac ggacacaaag gaggcaaggc caggggcgga ccaggcacgg 480  
 cggcagagac cagggaacag cgccacaaga aggccagga caagaggcaa gaaccgagag 540  
 cgcagacagc gaagcggacc acgcaaagga ggcgagaacc acccagggcg agaagaaaga 600  
 cgagcaccaa acgggcggaa agggccgaga caggaccggg ggccccgccac caccaaggcc 660  
 aaggaccacc caccagcaa gggacagacc agccccgcag cggccacacc accccagggc 720  
 gggcaccgcg acacaagccc cagaaccgcg gggaccgcg acgcgccccaa gccgcgaggc 780



```

aggcaaagca gggccgaggg gagcccaagc ccacacgccg cacaccgggc cccagagcca      840
cggagccacg aacaccccacc acgagggcag cggcggacgg cagagccccg ggccagccca      900
ccccgggggac aggggggggac cagcgggggag acacaagggc gggaacacca acccgggaac      960
aaaccgagag agaaccccaa cacacggagc aagccccaga aaagccaccc cggacaaagg     1020
gcaaaccgaa cgccgccaga ggccacgcgc cgccagccgc ggagacaccc cacccaaggc     1080
cccccacaga ggccggcgca cagccgagaa cgcgagcggg aaccacgcgc caacaaccgc     1140
ccaccccccc ccaaccaggc gggaacccca ccaaccgccc aacgccacca cccccacga      1200
ggccccagcg ggcacaccac cccaccccg cccacaagcaa cgagaccga ccagccagac      1260
ggcagcgacc gagccacacc accccacaac ccccgagggc cccaccaagc accccaccc      1320
acgggagccc caagaggggc gggaaggggg gccaaagccca caagcccgcc cggcggggcc      1380
cagacagagg gcagaggaaa gagcccggc caccacagcc cccaccag cccaccccg      1440
ggggcaagag gaaacggggg ggaggggaaga gaaggacagg ggggggggag aggacccac      1500
ccccggcccc ccccccccc agaccccacc cccagaccc ccccccccg cccgaaaaga      1560
accagcccag cccgcccccc ccccgcccc cggggaaaag agaaacaaag aaaacgcgac      1620
gcaccgcaa ccggccccgc cccccacagc gaaccccccc gcccgcccc aaggcacccc      1680
cccacccac ccggaggggc aggggagggg gagagcccga gcagccacag aagggccgga      1740
cagacccgc aaaaaaggca gaaaaccgaa aaaaaaaaaa aaaaaaaaaa      1789

```

```

<210> 6
<211> 433
<212> PRT
<213> Homo sapiens

```

```

<400> 6

```

```

Met Ser Gly Arg Pro Arg Thr Thr Ser Phe Ala Glu Ser Cys Lys Pro
1           5           10           15

```

```

Val Gln Gln Pro Ser Ala Phe Gly Ser Met Lys Val Ser Arg Asp Lys
          20           25           30

```

```

Asp Gly Ser Lys Val Thr Thr Val Val Ala Thr Pro Gly Gln Gly Pro
          35           40           45

```

```

Asp Arg Pro Gln Glu Val Ser Tyr Thr Asp Thr Lys Val Ile Gly Asn
          50           55           60

```

```

Gly Ser Phe Gly Val Val Tyr Gln Ala Lys Leu Cys Asp Ser Gly Glu
65           70           75           80

```

```

Leu Val Ala Ile Lys Lys Val Leu Gln Asp Lys Arg Phe Lys Asn Arg

```

85																90																95															
Glu	Leu	Gln	Ile	Met	Arg	Lys	Leu	Asp	His	Cys	Asn	Ile	Val	Arg	Leu																																
100								105								110																															
Arg	Tyr	Phe	Phe	Tyr	Ser	Ser	Gly	Glu	Lys	Lys	Asp	Glu	Val	Tyr	Leu																																
115								120								125																															
Asn	Leu	Val	Leu	Asp	Tyr	Val	Pro	Glu	Thr	Val	Tyr	Arg	Val	Ala	Arg																																
130								135								140																															
His	Tyr	Ser	Arg	Ala	Lys	Gln	Thr	Leu	Pro	Val	Ile	Tyr	Val	Lys	Leu																																
145				150								155								160																											
Tyr	Met	Tyr	Gln	Leu	Phe	Arg	Ser	Leu	Ala	Tyr	Ile	His	Ser	Phe	Gly																																
				165								170								175																											
Ile	Cys	His	Arg	Asp	Ile	Lys	Pro	Gln	Asn	Leu	Leu	Leu	Asp	Pro	Asp																																
				180								185								190																											
Thr	Ala	Val	Leu	Lys	Leu	Cys	Asp	Phe	Gly	Ser	Ala	Lys	Gln	Leu	Val																																
				195				200								205																															
Arg	Gly	Glu	Pro	Asn	Val	Ser	Tyr	Ile	Cys	Ser	Arg	Tyr	Tyr	Arg	Ala																																
210								215								220																															
Pro	Glu	Leu	Ile	Phe	Gly	Ala	Thr	Asp	Tyr	Thr	Ser	Ser	Ile	Asp	Val																																
225				230								235								240																											
Trp	Ser	Ala	Gly	Cys	Val	Leu	Ala	Glu	Leu	Leu	Leu	Gly	Gln	Pro	Ile																																
				245								250								255																											
Phe	Pro	Gly	Asp	Ser	Gly	Val	Asp	Gln	Leu	Val	Glu	Ile	Ile	Lys	Val																																
				260				265								270																															
Leu	Gly	Thr	Pro	Thr	Arg	Glu	Gln	Ile	Arg	Glu	Met	Asn	Pro	Asn	Tyr																																
				275				280								285																															
Thr	Glu	Phe	Lys	Phe	Pro	Gln	Ile	Lys	Ala	His	Pro	Trp	Thr	Lys	Asp																																
290								295								300																															
Ser	Ser	Gly	Thr	Gly	His	Phe	Thr	Ser	Gly	Val	Arg	Val	Phe	Arg	Pro																																
305				310								315								320																											
Arg	Thr	Pro	Pro	Glu	Ala	Ile	Ala	Leu	Cys	Ser	Arg	Leu	Leu	Glu	Tyr																																
				325								330								335																											

Thr Pro Thr Ala Arg Leu Thr Pro Leu Glu Ala Cys Ala His Ser Phe  
 340 345 350

Phe Asp Glu Leu Arg Asp Pro Asn Val Lys Leu Pro Asn Gly Arg Asp  
 355 360 365

Thr Pro Ala Leu Phe Asn Phe Thr Thr Gln Glu Leu Ser Ser Asn Pro  
 370 375 380

Pro Leu Ala Thr Ile Leu Ile Pro Pro His Ala Arg Ile Gln Ala Ala  
 385 390 395 400

Ala Ser Thr Pro Thr Asn Ala Thr Ala Ala Ser Asp Ala Asn Thr Gly  
 405 410 415

Asp Arg Gly Gln Thr Asn Asn Ala Ala Ser Ala Ser Ala Ser Asn Ser  
 420 425 430

Thr

<210> 7  
 <211> 2503  
 <212> RNA  
 <213> Homo sapiens

<400> 7  
 gcggagcgcg gcgcgcgcgg cggggggccgg gcgcggcccg ggaccgccgc caccgcgcgc 60  
 agaccgccgg ccgcgagccg ccggagcacc ggacaagccg gcgccaacga gccgggggca 120  
 cgcccgcgag ggccaagcca ggccggcgag cgggacgccc ccccgcccag ccaccgccgc 180  
 cgccgcccgc cccccccagc cggcgggcgg ccggggcccag caaccaggcg aagacacggg 240  
 acggggcgccg cggcgaacag gaggagaagg aggcgcgcgg cccaccgggg ccgccgcccc 300  
 aggcgcgcgc gcgcggggcc cgcgggccgag ggccgcgcgc ccccgccgac gccaggacgg 360  
 agaagaagac aaacggcagc gcaagacacc cgagggggccg aggcaagaca agaccaaggg 420  
 ccccgagcag aaggccggag agaggggggg ggcggcagga cccggagagg ccccccacag 480  
 gcgcccgggg aaccgcgcac gggggcccacc agccgcacca gagaggcacg aggacgaaga 540  
 ggggcgaggg gagaggacca ggccagccac gcccgagag ggagcccag gagaggcgag 600  
 cgcaaccacc cccacgcaag acccagcagg acacaacaag cgccacacac cagcgacacc 660  
 ggcgcggagg gcaccggaga agcgacccca aagccccacg acaagcccc agccgcccgc 720  
 ccgcggcgag cagcagaggc gggaaacgga gaccacaaag cggacaaacg ggcgagggag 780  
 cagccaccgc acaaaggcaa aagcaagcca cagacaaccg ggcaccaagg agacagacgg 840  
 aacagaagag ggggcacccg caccgccacc gggaaggccc gcccaaggacc caaacacgcc 900

aacacgacgc acagacaacc acacggagaa gccccacccg cgagaccgga caaggaccga	960
agcagaccgg agacggggaa cacacaacag cacaacggaa acgccgccag cgccccggcc	1020
ggccacgcca ccggcagaag ggcacaccga gacccaagcc ccagaaccgc cacaacgaga	1080
ggggagagcc aagcggcgac ggccggcccc agccaagcaa cccaacaaaag acaacccaag	1140
agggggacac gggaccggcc cccgacaccg cgggccacgg acaccaccca gagacagggg	1200
gggggcgcac cagagaggcc acaggccgcc cccccgggcc cacggggagg aacagcacac	1260
cacccgacag gaacccaac gaggagacgg gccaggcacc gccaacgagg agcaagacaa	1320
caacacccca agaccgagcc gagggccgag ccacgcaccc cgacgaagcg acggggccga	1380
ccccaccaa gcggcaggag ggcgaaacgg acccgcacag gagccagaaa caccacccca	1440
gcgggggagc ggaccacaaa cccgacacac ccaagcacia aggagacagc acaaaaggag	1500
gccagcccgg ccgagagccg accaggcagg ccagcccgcg ggggacaccg agcaagccac	1560
agaccgaggc cccagcaggc agcggcggag ggagccacac cccacaggg cagcccccaa	1620
cacaccccg ccccgccacc gccgagccag caccgcccac gccccgcgcc gcccaaacac	1680
cccaccaggc cgcaaccac ccaggccgcg cggggcaaca aagcccacac cccacggcg	1740
gccgaggagg ccggggacag caggccggcc agccccaca cgaggccagg cccccccac	1800
aaccagcccc caggaccaca cccacggcc agccaggggc cagagcagcc caggcgggga	1860
ccgaccagac aagagggaca agccgagcga ggcacccgcc gcccgcgcgc ccaccgccc	1920
agggggccgg cacagaaaaa agagagaaag caaacgaac agacggcggg cccaccccc	1980
gcggggcccc ccacagaggc aagaagcccg cccagccca gagacaggga cacagcccca	2040
ggaacccgac acaccagacc cgggaggcag ggaaagcagc cacagccacc gccccacccc	2100
cggccgcac cccaggcagg gggacgggga caccagagac cgggaaagga caaggggggg	2160
gccccacccc ccgagcccg agcggggccc ccccagggc cccagccca gccccgccc	2220
accacccggg cggggagggg cccgccagga acgaccagcc agcgaggagc caaaggcaag	2280
gcacaagcag gggggggagg ggggggaggg gggcccaggg gccccacccc ggggagggga	2340
ggcagggcag ggacagccca gggcagcccg gagggggacc cccccccac ccaagcccgg	2400
ggcccgaag ggggggaggg caggggggga gccccaggg gggggggggg ccgaagcacc	2460
aaacgcgaga aaaaaaagc aaaggaaaaa aaaaaaaaaa aaa	2503

<210> 8  
 <211> 304  
 <212> RNA  
 <213> Homo sapiens

<220>

<221> misc\_feature  
<222> (16)..(16)  
<223> n is a, c, g, or u

<220>  
<221> misc\_feature  
<222> (21)..(21)  
<223> n is a, c, g, or u

<220>  
<221> misc\_feature  
<222> (24)..(24)  
<223> n is a, c, g, or u

<220>  
<221> misc\_feature  
<222> (53)..(53)  
<223> n is a, c, g, or u

<220>  
<221> misc\_feature  
<222> (60)..(60)  
<223> n is a, c, g, or u

<220>  
<221> misc\_feature  
<222> (81)..(81)  
<223> n is a, c, g, or u

<220>  
<221> misc\_feature  
<222> (89)..(89)  
<223> n is a, c, g, or u

<220>  
<221> misc\_feature  
<222> (100)..(100)  
<223> n is a, c, g, or u

<220>  
<221> misc\_feature  
<222> (131)..(131)  
<223> n is a, c, g, or u

<220>  
<221> misc\_feature  
<222> (149)..(149)  
<223> n is a, c, g, or u

<220>  
<221> misc\_feature  
<222> (170)..(170)  
<223> n is a, c, g, or u

<220>  
<221> misc\_feature  
<222> (175)..(175)  
<223> n is a, c, g, or u

<220>  
<221> misc\_feature  
<222> (193)..(194)  
<223> n is a, c, g, or u

<220>  
 <221> misc\_feature  
 <222> (210)..(210)  
 <223> n is a, c, g, or u

<220>  
 <221> misc\_feature  
 <222> (247)..(247)  
 <223> n is a, c, g, or u

<220>  
 <221> misc\_feature  
 <222> (249)..(249)  
 <223> n is a, c, g, or u

<220>  
 <221> misc\_feature  
 <222> (251)..(251)  
 <223> n is a, c, g, or u

<220>  
 <221> misc\_feature  
 <222> (255)..(255)  
 <223> n is a, c, g, or u

<220>  
 <221> misc\_feature  
 <222> (292)..(292)  
 <223> n is a, c, g, or u

<220>  
 <221> misc\_feature  
 <222> (300)..(300)  
 <223> n is a, c, g, or u

<400> 8  
 mkkkrrrsrgs dssamnsskd nvkngrshsm hshygskkrr hsvggsgsma mrngsrdrvhn 60  
 kmgsdgsdas gssdvsgvcr nrhrrsmdnk rsadrdgykn sdmsrrsrra ssggkmykkg 120  
 gyavykgrsk nvakrhgaca rvskdkhanv hdvhdksvyd kdkymddcgn msmhmvkyrg 180  
 aychrrkvhr dknnkgkadg araksvkysn vvwyrdvgss ysdmwigvgcm asgrgsvdhr 240  
 gswgssnkny nkyknhards gkyskkrvsa amkhvyrsg rhasvsskkdg rnssyghgkn 300  
 rrsm 304

<210> 9  
 <211> 1318  
 <212> RNA  
 <213> Homo sapiens

<400> 9  
 ccccgccccg agccaggcag ccggcaggca gaggcccacc cgggccgggc ccagaccggg 60  
 ggggccccgc aggagggggc gagcccgcac gcccgaaaca cgcagccgcg cggcgcgagg 120  
 gcggggagcg ggggcgcccc aagagggcg ggggaggaag ggcccgagc cgcgggccga 180  
 ggccgaaggc gccgacccca ccccgaaagc gcgacggagc agggccgcca cacaagcaga 240

ccgcggggccc ccagcgcacc cggaagagcc caccaccccg acgagcccg agcccagacc	300
gccgaagagg agacgagaca agggccacgc gcgggcgccc ggggaacgca cagacaaaag	360
cgcaccccac cagaaacggg ggagggggggc agcagggacc aggggaaggga cagaggacac	420
cacgcccga gccaaccggc acgagcagca ggaccgggag gaggcccagc gagaagccga	480
cagcacgccc caacacccca acacccgccc gggcacgcga gggaacgggg gcaagcagag	540
gccggcgcgcc accaccaaga gaggacgcgg gaagagaaga aaggcgaagg acaaaggcaa	600
cccgaccgag gacaaaccgg cgcgcggggga cgcagaggcc gaggccacag ccaagggagc	660
ccacagagac gaagcccacc aaaagcggag agagggggcag ccagaaggac gggccagaac	720
aagcagcacc agggagggcc ccgccaggcc gacccgcagg acgggcagcc cagcgggcac	780
cacccaccga gccccagagc cccggcagag cacggcacga gagcggacga gcggcccagg	840
cgcggcaagc cagaggggga agggccagac agggccaaaa ggggacaggg gcccgcgga	900
gaaccaacca gcaccacaa agccccaggc accagcagcg gcagcccga ccgagagacc	960
gggacccgca cagcgcccac accccccca gcagcggagg cgcgcagccc ccagcccggc	1020
caacaacacc caaacgaaaa agcagcagga gaagaggccc cgcccggaag gagggccacc	1080
ccaggaacac caccaccac caggaccac acgggggagc ggggcaggac aacaccagcc	1140
gcacccgccc cccaagagc aaaaccgggc aaggggacac gagggggggg ggggggggga	1200
aaagggaac gggggaagga acaggccgag caggacggag ccacaaggcg acccaaacgg	1260
gagcaggaga aggaaacaag aaaaaggga gcagggggag aaaaaaaaaa aaaaaaaaaa	1318