```
GTCGCTGACG TGGAGGCGTC CGAAGGGCAG CAGGGTGTGT CGGGGCTCGG -31

ATTAAGACAT CGGAGTCGGA GACCTGAGAG ATGTTAACCA AATTCGAGAC 20
                                 MetLeuThrL ysPheGluTh 7


CAAGAGCGCG CGGGTCAAAG GGCTCAGCTT TCACCCCAAA AGACCTTGGA 70
rLysSerAla ArgValLysG lyLeuSerPh eHisProLys ArgProTrpI 24


TCCTGACTAG TTTACATAAT GGGGTCATCC AGTTATGGGA CTATCGGATG 120
leLeuThrSe rLeuHisAsn GlyValIleG lnLeuTrpAs pTyrArgMet 40


TGCACTCTCA TTGACAAGTT TGATGAACAT GATGGTCCAG TGCGAGGCAT 170      p.G56S
CysThrLeuI leAspLysPh eAspGluHis AspGlyProV alArgGlyIl 57


TGACTTCCAT AAGCAGCAGC CACTGTTCGT CTCTGGAGGA GATGACTATA 220
eAspPheHis LysGlnGlnP roLeuPheVa lSerGlyGly AspAspTyrL 74


AGATTAAGGT TTGGAATTAC AAGCTTCGGC GCTGTCTTTT CACATTGCTT 270
ysIleLysVa lTrpAsnTyr LysLeuArgA rgCysLeuPh eThrLeuLeu 90


GGGCACTTAG ATTATATTCG CACCACGTTT TTTCATCATG AAATATCCTG 320
GlyHisLeuA spTyrIleAr gThrThrPhe PheHisHisG luTyrProTr 107
```

```
GATTCTGAGT GCCTCCGATG ATCAGACCAT CCGAGTGTGG AACTGGCAAT 370
pIleLeuSer AlaSerAspA spGlnThrIl eArgValTrp AsnTrpGlnS 124


CTAGAACCTG TGTTTGTGTG TTAACAGGGC ACAACCATTA TGTGATGTGT 420
erArgThrCy sValCysVal LeuThrGlyH isAsnHisTy rValMetCys 140


GCTCAGTTCC ACCCCACAGA AGACTTGGTA GTATCAGCCA GCCTGGACCA 470
AlaGlnPheH isProThrGl uAspLeuVal ValSerAlaS erLeuAspGl 157


GACTGTGCGC GTTTGGGATA TTTCTGGTCT GAGGAAAAAA AACCTGTCCC 520
nThrValArg ValTrpAspI leSerGlyLe uArgLysLys AsnLeuSerP 174


CTGGTGCGGT GGAATCGGAT GTGAGAGGAA TAACTGGGGT TGATCTATTT 570
roGlyAlaVa lGluSerAsp ValArgGlyI leThrGlyVa lAspLeuPhe 190


GGAACTACAG ATGCAGTGGT GAAGCATGTA CTAGAGGGTC ACGATCGTGG 620   p.H199R  L201P
GlyThrThrA spAlaValVa lLysHisVal LeuGluGlyH isAspArgGl 207


AGTAAACTGG GCTGCCTTCC ACCCCACTAT GCCCCTTATT GTATCTGGGG 670   p.A211V
yValAsnTrp AlaAlaPheH isProThrMe tProLeuIle ValSerGlyA 224


CAGATGATCG TCAAGTGAAG ATCTGGCGCA TGAATGAATC AAAGGCATGG 720   p.R227C  p.K230N  R233C  p.R233H  p.R233L  p.K238E  p.A239P  p.W240R  p.W240L  p.W240S
laAspAspAr gGlnValLys IleTrpArgM etAsnGluSe rLysAlaTrp 240


GAGGTTGATA CCTGCCGGGG CCATTACAAC AATGTATCTT GTGCCGTCTT 770   p.E241K  p.E241A  E241D  p.V242G  p.D243N  p.D243G  G247D
GluValAspT hrCysArgGl yHisTyrAsn AsnValSerC ysAlaValPh 257


CCACCCTCGC CAAGAGTTGA TCCTCAGCAA TTCTGAGGAC AAGAGTATTC 820   H258R  L263S
```

```
eHisProArg GlnGluLeuI leLeuSerAs nSerGluAsp LysSerIleA 274


GAGTCTGGGA TATGTCTAAG CGGACTGGGG TTCAGACTTT CCGCAGAGAC 870    p.R281W  p.Q285H
rgValTrpAs pMetSerLys ArgThrGlyV alGlnThrPh eArgArgAsp 290


CATGATCGTT TCTGGGTCCT AGCTGCTCAC CCTAACCTTA ACCTCTTTGC 920
HisAspArgP heTrpValLe uAlaAlaHis ProAsnLeuA snLeuPheAl 307


AGCAGGCCAT GATGGTGGTA TGATTGTGTT TAAGCTGGAA CGGGAACGGC 970
aAlaGlyHis AspGlyGlyM etIleValPh eLysLeuGlu ArgGluArgP 324


CAGCCTATGC TGTTCATGGC AATATGCTAC ACTATGTCAA GGACCGATTC 1020   p.R339Q
roAlaTyrAl aValHisGly AsnMetLeuH isTyrValLy sAspArgPhe 340


TTACGACAGC TGGATTTCAA CAGCTCCAAA GATGTAGCTG TGATGCAGTT 1070
LeuArgGlnL euAspPheAs nSerSerLys AspValAlaV alMetGlnLe 357


GCGGAGTGGT TCCAAGTTTC CAGTATTCAA TATGTCATAC AATCCAGCAG 1120   R358W  p.Y370S
uArgSerGly SerLysPheP roValPheAs nMetSerTyr AsnProAlaG 374


AAAATGCAGT CCTGCTTTGT ACAAGAGCTA GCAATCTAGA GAATAGTACC 1170
luAsnAlaVa lLeuLeuCys ThrArgAlaS erAsnLeuGl uAsnSerThr 390


TATGACCTGT ACACCATCCC TAAAGATGCT GACTCCCAGA ATCCTGATGC 1220
TyrAspLeuT yrThrIlePr oLysAspAla AspSerGlnA snProAspAl 407


GCCTGAAGGG AAACGATCCT CAGGCCTGAC AGCCGTTTGG GTCGCTCGAA 1270   p.P408S
aProGluGly LysArgSerS erGlyLeuTh rAlaValTrp ValAlaArgA 424
```

```
ATCGGTTTGC TGTCCTAGAT CGGATGCATT CGCTTCTGAT CAAGAATCTG 1320
snArgPheAl aValLeuAsp ArgMetHisS erLeuLeuIl eLysAsnLeu 440


AAGAATGAGA TCACCAAAAA GGTACAGGTG CCCAACTGTG ATGAGATCTT 1370
LysAsnGluI leThrLysLy sValGlnVal ProAsnCysA spGluIlePh 457


CTATGCTGGC ACAGGCAATC TCCTGCTTCG AGATGCGGAC TCTATCACAC 1420
eTyrAlaGly ThrGlyAsnL euLeuLeuAr gAspAlaAsp SerIleThrL 474


TCTTTGACGT ACAGCAGAAG CGGACTCTGG CATCTGTGAA GATTTCTAAA 1470
euPheAspVa lGlnGlnLys ArgThrLeuA laSerValLy sIleSerLys 490


GTGAAATACG TTATCTGGTC AGCAGACATG TCACATGTAG CACTACTAGC 1520
ValLysTyrV alIleTrpSe rAlaAspMet SerHisValA laLeuLeuAl 507


CAAACACGCC ATTGTGATCT GTAACCGCAA ACTGGATGCT TTATGTAACA 1570    p.D519Y
aLysHisAla IleValIleC ysAsnArgLy sLeuAspAla LeuCysAsnI 524


TTCATGAGAA CATTCGTGTC AAGAGTGGGG CCTGGGATGA GAGTGGGGTA 1620    S532I
leHisGluAs nIleArgVal LysSerGlyA laTrpAspGl uSerGlyVal 540


TTTATCTATA CCACAAGCAA CCACATCAAA TATGCTGTCA CCACTGGGGA 1670
PheIleTyrT hrThrSerAs nHisIleLys TyrAlaValT hrThrGlyAs 557


CCACGGGATC ATTCGAACTC TGGATTTACC CATCTATGTC ACACGGGTGA 1720
pHisGlyIle IleArgThrL euAspLeuPr oIleTyrVal ThrArgValL 574
```

```
AGGGCAACAA TGTATACTGC CTAGACAGGG AGTGTCGTCC CCGGGTACTC 1770
ysGlyAsnAs nValTyrCys LeuAspArgG luCysArgPr oArgValLeu 590


ACCATTGATC CCACTGAGTT CAAATTCAAG CTGGCCCTGA TCAACAGAAA 1820
ThrIleAspP roThrGluPh eLysPheLys LeuAlaLeuI leAsnArgLy 607


ATATGATGAG GTACTGCACA TGGTGAGGAA TGCCAAACTA GTTGGCCAGT 1870
sTyrAspGlu ValLeuHisM etValArgAs nAlaLysLeu ValGlyGlnS 624


CTATTATTGC TTATCTCCAG AAGAAGGGCT ATCCTGAAGT GGCACTGCAT 1920
erIleIleAl aTyrLeuGln LysLysGlyT yrProGluVa lAlaLeuHis 640


TTTGTCAAGG ATGAGAAAAC TCGCTTTAGT CTGGCACTGG AGTGTGGAAA 1970
PheValLysA spGluLysTh rArgPheSer LeuAlaLeuG luCysGlyAs 657


CATTGAGATT GCTCTGGAAG CAGCCAAAGC ACTGGATGAC AAGAACTGCT 2020
nIleGluIle AlaLeuGluA laAlaLysAl aLeuAspAsp LysAsnCysT 674


GGGAAAAGCT GGGAGAAGTG GCCCTGCTGC AGGGGAACCA CCAGATTGTG 2070    G678E
rpGluLysLe uGlyGluVal AlaLeuLeuG lnGlyAsnHi sGlnIleVal 690


GAAATGTGCT ATCAGCGTAC CAAAAACTTT GACAAACTTT CCTTCCTGTA 2120    Y694C  T697I
GluMetCysT yrGlnArgTh rLysAsnPhe AspLysLeuS erPheLeuTy 707


TCTTATCACT GGCAACTTAG AAAAACTTCG CAAGATGATG AAGATTGCTG 2170
rLeuIleThr GlyAsnLeuG luLysLeuAr gLysMetMet LysIleAlaG 724


AGATCAGAAA GGACATGAGT GGCCACTATC AGAATGCCCT ATACCTGGGT 2220
```

luIleArgLy sAspMetSer GlyHisTyrG lnAsnAlaLe uTyrLeuGly 740

GATGTGTCAG AGCGTGTGCG GATCCTGAAG AACTGTGGAC AGAAGTCCCT 2270 R745H
AspValSerG luArgValAr gIleLeuLys AsnCysGlyG lnLysSerLe 757

GGCCTATCTC ACAGCTGCTA CCCATGGCTT AGATGAAGAA GCTGAGAGCC 2320
uAlaTyrLeu ThrAlaAlaT hrHisGlyLe uAspGluGlu AlaGluSerL 774

TAAAGGAGAC ATTTGACCCA GAGAAGGAGA CAATCCCAGA CATTGACCCT 2370 T777I
euLysGluTh rPheAspPro GluLysGluT hrIleProAs pIleAspPro 790

AATGCCAAGC TGCTCCAGCC ACCTGCACCT ATCATGCCAT TGGATACCAA 2420 p.Asn807Thr
AsnAlaLysL euLeuGlnPr oProAlaPro IleMetProL euAspThrAs 807

TTGGCCTTTA TTGACTGTAT CCAAAGGATT TTTTGAAGGC ACCATTGCCA 2470
nTrpProLeu LeuThrValS erLysGlyPh ePheGluGly ThrIleAlaS 824

GCAAAGGGAA GGGAGGAGCA CTGGCTGCTG ACATTGACAT TGACACTGTT 2520 p.G828E
erLysGlyLy sGlyGlyAla LeuAlaAlaA spIleAspIl eAspThrVal 840

GGTACAGAGG GCTGGGGAGA GGATGCAGAG CTGCAGTTGG ATGAAGATGG 2570 G844D
GlyThrGluG lyTrpGlyGl uAspAlaGlu LeuGlnLeuA spGluAspGl 857

GTTTGTGGAG GCTACAGAAG GTTTGGGGGA TGATGCTCTT GGCAAGGGAC 2620
yPheValGlu AlaThrGluG lyLeuGlyAs pAspAlaLeu GlyLysGlyG 874

AGGAAGAAGG AGGTGGCTGG GATGTAGAAG AAGATCTGGA GCTCCCTCCT 2670
lnGluGluGl yGlyGlyTrp AspValGluG luAspLeuGl uLeuProPro 890

```
GAGCTGGATA TATCCCCTGG GGCAGCTGGT GGGGCTGAAG ATGGTTTCTT 2720
GluLeuAspI leSerProGl yAlaAlaGly GlyAlaGluA spGlyPhePh 907


TGTGCCCCCA ACCAAGGGAA CAAGTCCAAC TCAGATCTGG TGTAATAACT 2770
eValProPro ThrLysGlyT hrSerProTh rGlnIleTrp CysAsnAsnS 924


CTCAGCTTCC AGTTGATCAC ATCCTGGCAG GCTCTTTCGA AACAGCCATG 2820
erGlnLeuPr oValAspHis IleLeuAlaG lySerPheGl uThrAlaMet 940


CGGCTCCTTC ATGACCAAGT AGGGGTAATC CAGTTTGGCC CCTACAAGCA 2870
ArgLeuLeuH isAspGlnVa lGlyValIle GlnPheGlyP roTyrLysGl 957


ACTGTTCCTA CAGACATACG CCCGAGGCCG CACAACCTAT CAGGCTCTGC 2920
nLeuPheLeu GlnThrTyrA laArgGlyAr gThrThrTyr GlnAlaLeuP 974


CCTGCCTACC CTCCATGTAT GGCTATCCTA ATCGCAACTG GAAGGATGCA 2970
roCysLeuPr oSerMetTyr GlyTyrProA snArgAsnTr pLysAspAla 990


GGGCTGAAGA ATGGTGTACC AGCTGTGGGC CTGAAGCTTA ATGACCTCAT 3020
GlyLeuLysA snGlyValPr oAlaValGly LeuLysLeuA snAspLeuIl 1007


CCAACGGTTG CAGCTGTGCT ACCAGCTCAC CACAGTTGGC AAATTTGAGG 3070    p.C1013S
eGlnArgLeu GlnLeuCysT yrGlnLeuTh rThrValGly LysPheGluG 1024


AGGCTGTGGA AAAATTCCGT TCCATCCTTC TCAGTGTGCC ACTTCTTGTT 3120
luAlaValGl uLysPheArg SerIleLeuL euSerValPr oLeuLeuVal 1040
```

GTGGACAATA AACAAGAGAT TGCAGAGGCC CAGCAGCTCA TCACCATTTG 3170
ValAspAsnL ysGlnGluIl eAlaGluAla GlnGlnLeuI leThrIleCy 1057

CCGTGAGTAC ATTGTGGGTT TGTCCGTGGA GACAGAAAGG AAGAAGCTGC 3220   p.R1058C
sArgGluTyr IleValGlyL euSerValGl uThrGluArg LysLysLeuP 1074

CCAAAGAGAC TCTAGAACAG CAGAAGCGCA TCTGTGAGAT GGCAGCCTAT 3270
roLysGluTh rLeuGluGln GlnLysArgI leCysGluMe tAlaAlaTyr 1090

TTCACCCACT CAAACCTGCA GCCTGTGCAC ATGATCCTGG TGCTGCGTAC 3320
PheThrHisS erAsnLeuGl nProValHis MetIleLeuV alLeuArgTh 1107

AGCCCTCAAT CTGTTCTTCA AGCTCAAGAA CTTCAAGACA GCTGCCACCT 3370
rAlaLeuAsn LeuPhePheL ysLeuLysAs nPheLysThr AlaAlaThrP 1124

TTGCTCGGCG CCTACTAGAA CTCGGGCCCA AGCCTGAGGT GGCCCAACAG 3420
heAlaArgAr gLeuLeuGlu LeuGlyProL ysProGluVa lAlaGlnGln 1140

ACCCGAAAAA TCCTGTCTGC CTGTGAGAAG AATCCCACAG ATGCCTACCA 3470   p.R1142X
ThrArgLysI leLeuSerAl aCysGluLys AsnProThrA spAlaTyrGl 1157

GCTCAATTAT GACATGCACA ACCCCTTTGA CATTTGTGCT GCATCATATC 3520
nLeuAsnTyr AspMetHisA snProPheAs pIleCysAla AlaSerTyrA 1174

GGCCCATCTA CCGTGGAAAG CCAGTAGAAA AGTGTCCACT CAGTGGGGCC 3570
rgProIleTy rArgGlyLys ProValGluL ysCysProLe uSerGlyAla 1190

TGCTATTCCC CTGAGTTCAA AGGTCAAATC TGCAGGGTCA CCACAGTGAC 3620

```
CysTyrSerP roGluPheLy sGlyGlnIle CysArgValT hrThrValTh 1207


AGAGATTGGC AAAGATGTGA TTGGTTTAAG GATCAGTCCT CTGCAGTTTC 3670

rGluIleGly LysAspValI leGlyLeuAr gIleSerPro LeuGlnPheA 1224


GCTAAGGCCC CCTTTGTGTG CATGGGTCAG TCACCATATG TTCCCCCCAG *45

rgStop


AGAATGTGTC TATATCCTCC TTCTAACAGC ACCTTCCCCC TGCAGCTACT *95

CTTCAGATCT GGCTCTCTGT ACCCTAAAAC CTAGTATCTT TTTCTCTTCT *145

ATGGAAAATC CGAAGGTCTA AACTTGACTT TTTTGAGGTC TTCTCAACTT *195

GACTACAGTT GTGCTCATAA TTGTCCTTGC CTTTCCAGCT TAATTATTTT *245

AAGGAACAAA TGAAAACTCT GGGCTGGGTG GAGTGGCTCA TACCTGTAAT *295

CCCAGCACTT TGGGAGGCTA CGGTGGGCAG ATCATCTGAG GCCAGGAGTT *345

CGAGACCTGC CTGGCCAACA TGGCAACACC CCGTCTCTAA TAAAAATATA *395

AAAATTAGCC TGGCATGGTA GCATGCGCCT ATAGTCCCAG CTGCTCAGGA *445

GGCTGAGGCA TGAGAATCGC TTGAACCTAG GAGGTGGAGG TTGCATTCAA *495

CTGAGATCAT ACCACTTCAT TCCAGCCTGG GTGACAGAGC AAGACTCTGT *545

CTCAAAAAAA AAAAAAAGGA AAACTCTGTG ATGGACATTT GTTTAGTAAA *595

TCCCTTCAGT ATTTATCCCT CCTTTCCCCA CAGCAGCTTT CTTTCCTGTC *645

AACTAGAAAG GAGCAGGATG TAATAAATAC ATTTTGGTGT GACTAGGCCA *695

CACCAACTCT TAATCATCTC CCATTTTCCT TAGACATTTA AATTTCAAGG *745

CAGGTACCCT CTGTGTACTC AGAAATTTGA AGAAGTTATT TGGTTTTCCA *795

AAATGCACAC TGCGGGTTAT TGATTTGTTC TTTACAACTA TTGTTCTCAT *845

ATTTCTCACA CTAAATAAAT CTCTATGAGA GCTTCTTGAC TTGGCCATTT *895

ATTTCTTGGA CACTCTCATG TTCTTGTTCA CCCATGCAGG CACCCCACCA *945

AAGTACATAT CTTCCTTCCA GTAATAATTT TTAATTACAA AATAAACATC *995

CACTATTGGA AAAAAAAAA AAAAGCTAGC CGGGCATGGT GGTGGGTGCC *1045
```

```
TGTAATCCCA GCTACTCTGG AGGCTGAGGC AGAGGATTGC TTGAACCCGG *1095

GAGGCGGAGG TTGCAGTAAG CTGAGATCGC GCCACCGCAC TCCAGCCTGG *1145

GCGACAGAGT GAGACTCCAT CTCAAAAAAA AAGAAAGAAA AAAAGAAGCA *1195

CATGTTTTTC ATAGGGTATA TATGAGGACC TAAACTGCTG TGAAAATGAT *1245

AGAAAGCAAG TAGCTCCCTT ATTCTGTTTT TGATTGCAGC CTTTTATCTT *1295

TTGCTAATTA TAGCAATATT TATTGAGCAC CTGCCATGTG ACTGTCACTG *1345

TTCTAGATAT TTTACATGTA ATATACAGAT AAAAGAATAG TACTTTATAT *1395

ATATTACAAT GATACAATGA TTACATTAAC AATACAATAT TTTGCTTGTC *1445

ATATGCTAAG AATAATTGGG TAGAGTGACA TTACTGTGCC TTCGATTAAA *1495

ATAAGTACTT TTTTGCGTGT TAAATTCATG TTTTCAATAA ATAATAAATG *1545

CATATAGTTG AAAAATCA
```