```
AGTTCCTCTC CCGCCGCGCC GGGCCGCGCT GCCGCTCGCT CCCCGGCCGT -225

GGCGCCTCCG GGCCAGACGC GCTGCAGCCT CCAGCCCGCG GCAAGCGGGG -175

CGGCCGCGCC ACCCCCGGCC CCGCGCCAGC AGCCCCTCGC CGCGCGTCCA -125

GCGTTCCCGG CCAGCAGCCT CCCCATACGC AGGTCCTGCT GGGCCGCCCC -75

GTCGCGCCCC CCACTCTGAA CTCAAGTCAC CGTGGAGCTC CGCCGCCCCG -25

AAACTTTCAC CGCGAGCGGG AAATATGGGA TGTATAAAAT CAAAAGGGAA 26

AGACAGCTTG AGTGACGATG GAGTAGATTT GAAGACTCAA CCAGTACGTA 76

ATACTGAAAG AACTATTTAT GTGAGAGATC CAACGTCCAA TAAACAGCAA 126

AGGCCAGTTC CAGAATCTCA GCTTTTACCT GGACAGAGGT TTCAAACTAA 176

AGATCCAGAG GAACAAGGAG ACATTGTGGT AGCCTTGTAC CCCTATGATG 226

GCATCCACCC GGACGACTTG TCTTTCAAGA AAGGAGAGAA GATGAAAGTC 276

CTGGAGGAGC ATGGAGAATG GTGGAAAGCA AAGTCCCTTT TAACAAAAAA 326

AGAAGGCTTC ATCCCCAGCA ACTATGTGGC CAAACTCAAC ACCTTAGAAA 376

CAGAAGAGTG GTTTTTCAAG GATATAACCA GGAAGGACGC AGAAAGGCAG 426

CTTTTGGCAC CAGGAAATAG CGCTGGAGCT TTCCTTATTA GAGAAAGTGA 476

AACATTAAAA GGAAGCTTCT CTCTGTCTGT CAGAGACTTT GACCCTGTGC 526

ATGGTGATGT TATTAAGCAC TACAAAATTA GAAGTCTGGA TAATGGGGGC 576

TATTACATCT CTCCACGAAT CACTTTTCCC TGTATCAGCG ACATGATTAA 626

ACATTACCAA AAGCAGGCAG ATGGCTTGTG CAGAAGATTG GAGAAGGCTT 676

GTATTAGTCC CAAGCCACAG AAGCCATGGG ATAAAGATGC CTGGGAGATC 726

CCCCGGGAGT CCATCAAGTT GGTGAAAAGG CTTGGCGCTG GGCAGTTTGG 776

GGAAGTCTGG ATGGGTTACT ATAACAACAG TACCAAGGTG GCTGTGAAAA 826
```

CCCTGAAGCC AGGAACTATG TCTGTGCAAG CCTTCCTGGA AGAAGCCAAC 876

CTCATGAAGA CCCTGCAGCA TGACAAGCTC GTGAGGCTCT ACGCTGTGGT 926

CACCAGGGAG GAGCCCATTT ACATCATCAC CGAGTACATG GCCAAGGGCA 976

GTTTGCTGGA TTTCCTGAAG AGCGATGAAG GTGGCAAAGT GCTGCTTCCA 1026

AAGCTCATTG ACTTTTCTGC TCAGATTGCA GAGGGAATGG CATACATCGA 1076

GCGGAAGAAC TACATTCACC GGGACCTGCG AGCAGCTAAT GTTCTGGTCT 1126

CCGAGTCACT CATGTGCAAA ATTGCAGATT TTGGCCTTGC TAGAGTAATT 1176

GAAGATAATG AGTACACAGC AAGGGAAGGT GCTAAGTTCC CTATTAAGTG 1226

GACGGCTCCA GAAGCAATCA ACTTTGGATG TTTCACTATT AAGTCTGATG 1276

TGTGGTCCTT TGGAATCCTC CTATACGAAA TTGTCACCTA TGGGAAAATT 1326

CCCTACCCAG GGAGAACTAA TGCCGACGTG ATGACCGCCC TGTCCCAGGG 1376

CTACAGGATG CCCCGTGTGG AGAACTGCCC AGATGAGCTC TATGACATTA 1426

TGAAAATGTG CTGGAAAGAA AAGGCAGAAG AGAGACCAAC GTTTGACTAC 1476

TTACAGAGCG TCCTGGATGA TTTCTACACA GCCACGGAAG GGCAATACCA 1526   Q507* Y508H Y508F Y508*

GCAGCAGCCT TAGAGCACAG GGAGACCCGT CCATTTGGCA GGGGTGGCTG *37

CCTCATTTAG AGAGGAAAAG TAACCATCAC TGGTTGCACT TATGATTTCA *87

TGTGCGGGGA TCATCTGCCG TGCCTGGATC CTGAAATAGA GGCTAAATTA *137

CTCAGGAAGA ACACCCTCTA AATGGGAAAG TATTCTGTAC TCTTAGATGG *187

ATTCTCCACT CAGTTGCAAC TTGGACTTGT CCTCAGCAGC TGGTAATCTT *237

GCTCTGCTTG ACAACATCTG AGTGCAGCCG TTTGAGAAGA AAACATCTAT *287

TCTCTCCAAA AATGCACCCA ACTAGCTCTA TGTTTACAAA TGGACATAGG *337

ACTCAAAGTT TCAGAGACCA TTGCAATGAA TCCCCAATAA TTGCAGAACT *387

AAACTCATTT ATAAAGCTAA AATAACCGGA TATATACATA GCATGACATT *437

TCTTTGTGCT TTGGCTTACT TGTTTAAAAA AAAAAAAAAA CTAATCCAAC *487

CTGTTAGATT TTGCAGGTGA AGTCAGCAGC TTAAAAATGT CTTTCCCAGA *537

TTTCAATGAT TTTTTTCCCC CTACCTCCCA AAATCTGAGA CTGTTAAAAC *587

ATTTTTCTTC TATGAACACT GCTCAGACCT GCTAGACATG CCATAGGAGT *637

GGCGTGCACA TCTCTCTCTC TTCCAGCAGG AGGAGCCCGT GAGCACGCAC *687

```
AGCTGCCCTG TCTGCTCACC CGAAGGCACC GGGCTCACCT GGACCTCCCA *737

GGAAAGGGAG AAGAGCCTCA GAAACTGCTC TGTGTTTAGA AGGAATATTT *787

TTAAGAGTCC AGCTTTTTCA TTTCCACAAT TTCCTATATC CAGATTTGTT *837

TTGACAATGT AGTTTGGAAG AACTAAGATT CTAATCTCTG AAGAACCTTA *887

TAGGGCCTTC TAAAACATAA GAGTTTCCTT TGTTGCTTCA AATATTTGAA *937

CATTATGTTA AAGATCAAGT ATTAATTTTA GTTGTACTCT AGAAAGCTAA *987

AGTGCCACAT TCGGGGCTAT TTTTATGATT CAGCAATCTT TTCTAAATTG *1037

TGTAGCATGT GTATGAGACT ATTTATACCC AAGGATATGA AGGAACATAA *1087

GTGACTACAA GGCTCTAATA AGCCACGGTG GCAGGAGGTT CAAGCGGTTC *1137

TGTTCACTAA ATTTTTCTCC TGTAAGCTTT GAATGGAAAC TTCTGTATCA *1187

CATGATGTGT TTCACTTATG CTGTTGTGTA TATACCTAAT ATTTCTATTT *1237

TTGATTTTAT TTTAATACAC CTCGTCCAAT AACATCTCAA GCTTTTTATT *1287

TGCATTTACA TTTTCAGCTG TGGTCAGTGT AAAAATTGGT CATCAGCTGG *1337

GGGCGGGGTG GTTAGAAGTG ATTCAACAGA GCTACATGCT TTAAACTTGC *1387

CCAAGTTCTA CCTCCTTCCT TTGAACATTT CAGATTGGAG AACCAAGGAG *1437

TTGATTGCCT GAACACCTGA ACATCCGTTT ATGGGGGCCA GATAGAATTT *1487

GTTTTCAAAT AGGCTTAACA GGCATCATTA AAATTTCATT CTGTGTGTTT *1537

TGTTTAGGCT TGAGGTGCTT AGAAGATGGG ATAAAATATT CTACTTTTTT *1587

CTAAATTTTA ACTTTGTTTC CTATGTGATT TTTTTAAATG TCCTTTCTAA *1637

AATATTCTAA AATTATTGAT TCACAAGTGC CATGTTCAGA ACTATAGAAT *1687

ATTACTGTTA CATAATGTCT GCACAGCTGG TCCCTTGATT CAGTGGTAAG *1737

GTTTTTGTGT ACACCCCCCT GCTTGCATTT TATTTCAGAA CCACAAGTAT *1787

TACCCAATAT GTTACATGGA GAGGAACTAT AAAGAATCCC TAAGGCAAAA *1837

AGAAGTCTCT AGAAAATGAC TAGAGGTTTT TTTTTTAGCA TAACAAATTT *1887

ATTTAAAGAA AATTATTAAA TTTATCTTCG CCTTGTTTTG CTTCTCCCAG *1937

TTCCTCCTCT TCTTGCCATT TTCCACTTGT CTTTCCCTCC CAATCAAGCC *1987

TGTGATCCTT ACCTCCATGT GGGCCCTTCA CCAGCTTGGG CCTCATCTCT *2037

GGTGTCCAGC ATGTGTGGAA GTCACACGTT CCCTTGATGA ACAGCACACA *2087
```

```
CAGTCTCCTT ACTTAGCTAT AGGTTTCCAG CCTCCCTGTG ACAGACAGGC *2137

ATAATGAGGG GCTGAATAGG TGTTTGTAGC ATTTTCGGGT ATCCAGTGGT *2187

GTGCAAAATG GCTCATGTCA TCACACCTCA GGTTATTGTA GAGAACTGGA *2237

AAGACAGAAT CCATACTCCC TACCGCCAAG ATTCTGACTT AGCTGTTGTG *2287

CAGCGGGAGA TGTATGTCAG TCTATTTTAA AAGCTTCTCC AGTCAGCTAG *2337

ACACAGTGGC TCATGCCTGT AATCTCAGCA CTTTGGGAGG CTAAGGCAGG *2387

AGGATCGCTT GAGCTCAGGA GTTAGACACC AGCCTGAGCA ATATAGCAAG *2437

TCTCTACAAA AAATTTTAAA AATTAGCTAA GTGTGGTGGC ACATACCTAT *2487

GGTCCCAGCT ACTCAGGAGG CTGAGGTGGG AGGATTGCTT GAGCCCAGGA *2537

AGTCAAGGCT GCAGTGAGCT ATGATCGTGC CACTGCACTC CAGCCTGGGC *2587

AAGAGTGAGA CCCTGTCTCA AAAAAAAAAA CCAAAAAAAC AAAAAAACCC *2637

TTCCCTGGTG AATATAATGT ATGCTCAGTC TTGAGAGTCA CCATTTGGTT *2687

TGACCCATAT TACAAAGCCC AGCCTTCTTT TTCCTCAGAT GGTTTCATTT *2737

CTCAAAGGGG AGTGGAAGAG AAGGAAGGAG GAAAAAAACG TGCCTTCCTT *2787

CCTGGTGAAC ATGAAGCTCA CTGAACCTAA AGGAGACTCA CAGGCTAGAA *2837

ACCAGTGTCC TTAGTAGTAG GGGCCACCTG GGGCTTCCGT GGCTTCTCCA *2887

GGCATACCTC CTGAGGTTCC ACAGAATCCC CTTGCAGAAA GTACTTCTGA *2937

TGAGCCACTG GGAACGTGTA CTGTCCAAGC CCCCTTCTAG GGGAAATTCA *2987

CTGCCACAGA AGCCAGCAGC CCTGGCTGTC CCAAGGGTGC ATAAAAGGAT *3037

AATAACTAAA AGGAGGGAAT TAATTTTGTT GCCTACACCA GGGATAGATC *3087

GAGGGCACTG AAAGCATCAT AAGTATAAAC TCATCTCCAC GTCAGCTTGT *3137

GGAGTCTGCT TTCTCGCTCT CTCTTTTTTT TTTTTTTTG ACAGAGTCTC *3187

ACTCTGTCAC GCAGGCTAGA GTTGCAGTGG TGCGATCTCA GCTCATTGCA *3237

ACCTCCCCCT CTCTGGTTGG AGCGATTCTT CTCCTGCCTC AGCCTCCCTA *3287

GTAGCTGGAA CTACAGGCAC ACACCATCTC GCCTGGCTAA TTTTTGTATT *3337

TTTAGTAGAG ATGGGGTTTC ACCATGTTAG CCAGGCTGGT CTCGAACTCC *3387

TGACCTCAGG CGATCCGCCT GCCTCGGCTT CCCAAAGTGC TGGGATTACA *3437

AGTGTGAGCC ACTGTGCCCG GTCGGGGTCT GCTTTCTCTT TATTCCCCGT *3487
```

```
CCTCTCACAA ACTACATCCA GGGGCGGTAG CTGCTAAAAG GACTCAATTC *3537

TGCAACCTCA GCTTCTGACA CCTTCAAATA AATTAAATGC TGGTCCTAGG *3587

GCTCGCTCCC CAGCTCTTCT GATGTCACTG CCAGGTTCCA CAGCCAGTTG *3637

TCGGACATCT CCTGCTTCCT CAGCCAAGAA CACCTCTCAG ATGAGATCAA *3687

ACCGCCCAGC TTTCCTCTTA GCATCTTCCT TTATTTTCCA CTTCTGGCTG *3737

GTGGAAGAGA GATCGTAGTT CTCTATTCCC TACATTTTAA CTCCACGAAT *3787

ATTTTAGATT AAGTGAAGTC TTGCTGTTCT TCCTGATGGA AATGGTAATT *3837

ATAAACATAC AAAGATACCC ATACACACAT TCTTTACTCA GGGCCCTCTT *3887

GTGCTTTGGC AAATGATAAA AATTATCTCT AACTCACACA ACGACCCTGG *3937

AAGATCTGGA CAATGCTATC ACTATTTCAC AGCTGAGAGA TCGGGAGCAT *3987

TTAAAGCCAC AGAAGGAATA GATTGGCTTC GCTTCATATT GTTTCCTTGT *4037

GAAATAAAAC CAGTGAAACA CC
```