

Nucleotide sequence of human coronavirus, CoV-SARS. Sequence begins near the 5' end of the pol 1b region and extends though all structural genes to the very 3' poly A region. Sequence determined at the Centers for Disease Control, Atlanta, GA.

Length: 17554 Fri, Apr 11, 2003 2:03 PM Check: 1561 ..

```
1 AGTTGGAAAA GATGGCAGAT CAGGCTATGA CCCAAATGTA CAAACAGGCA
51 AGATCTGAGG ACAAGAGGGC AAAAGTAACT AGTGCTATGC AAACAATGCT
101 CTTCACTATG CTTAGGAAGC TTGATAATGA TGCACCTAAC AACATTATCA
151 ACAATGCGCG TGATGGTTGT GTTCCACTCA ACATCATAACC ATTGACTACA
201 GCAGCCAAAC TCATGGTTGT TGTCCCTGAT TATGGTACCT ACAAGAACAC
251 TTGTGATGGT AACACCTTTA CATATGCATC TGCACCTCTGG GAAATCCAGC
301 AAGTTGTTGA TGCGGATAGC AAGATTGTTT AACTTAGTGA AATTAACATG
351 GACAATTCAC CAAATTTGGC TTGGCCTCTT ATTGTTACAG CTCTAAGAGC
401 CAACTCAGCT GTTAAACTAC AGAATAATGA ACTGAGTCCA GTAGCACTAC
451 GACAGATGTC CTGTGCGGCT GGTACCACAC AAACAGCTTG TACTGATGAC
501 AATGCACTTG CCTACTATAA CAATTCGAAG GGAGGTAGGT TTGTGCTGGC
551 ATTACTATCA GACCACCAAG ATCTCAAATG GGCTAGATTC CCTAAGAGTG
601 ATGGTACAGG TACAATTTAC ACAGAACTGG AACCACCTTG TAGGTTTGT
651 ACAGACACAC CAAAAGGGCC TAAAGTGAAA TACTTGTACT TCATCAAAGG
701 CTTAAACAAC CTAAATAGAG GTATGGTGCT GGGCAGTTTA GCTGCTACAG
751 TACGTCTTCA GGCTGGAAAT GCTACAGAAG TACCTGCCAA TTCAACTGTG
801 CTTTCCTTCT GTGCTTTTGC AGTAGACCCT GCTAAAGCAT ATAAGGATTA
851 CCTAGCAAGT GGAGGACAAC CAATCACCAA CTGTGTGAAG ATGTTGTGTA
901 CACACACTGG TACAGGACAG GCAATTACTG TAACACCAGA AGCTAACATG
951 GACCAAGAGT CCTTTGGTGG TGCTTCATGT TGTCTGTATT GTAGATGCCA
1001 CATTGACCAT CCAAATCCTA AAGGATTCTG TGAATTGAAA GGTAAGTACG
1051 TCCAAATACC TACCATTGT GCTAATGACC CAGTGGGTTT TACACTTAGA
1101 AACACAGTCT GTACCGTCTG CGGAATGTGG AAAGGTTATG GCTGTAGTTG
1151 TGACCAACTC CGCGAACCCCT TGATGCAGTC TGCGGATGCA TCAACGTTTT
1201 TAAACGGGTT TGCGGTGTAA GTGCAGCCCG TCTACACCG TGCGGCACAG
1251 GCACTAGTAC TGATGTCGTC TACAGGGCTT TTGATATTTA CAACGAAAAA
1301 GTTGCTGGTT TTGCAAAGTT CCTAAAAACT AATTGCTGTC GCTTCCAGGA
1351 GAAGGATGAG GAAGGCAATT TATTAGACTC TTACTTTGTA GTTAAGAGGC
1401 ATACTATGTC TAACTACCAA CATGAAGAGA CTATTTATAA CTTGGTTAAA
1451 GATTGTCCAG CGGTTGCTGT CCATGACTTT TTCAAGTTTA GAGTAGATGG
1501 TGACATGGTA CCACATATAT CACGTCAGCG TCTAACTAAA TACACAATGG
1551 CTGATTTAGT CTATGCTCTA CGTCATTTT ATGAGGGTAA TTGTGATACA
1601 TTTAAAGAAA TACTCGTCAC ATACAATTGC TGTGATGATG ATTATTTCAA
1651 TAAGAAGGAT TGGTATGACT TCGTAGAGAA TCCTGACATC TTACGCGTAT
1701 ATGCTAACTT AGGTGAGCGT GTACGCCAAT CATTATTAAG GACTGTACAA
1751 TTCTGCGATG CTATGCGTGA TGCAGGCATT GTAGGCGTAC TGACATTAGA
1801 TAATCAGGAT CTTAATGGGA ACTGGTACGA TTTCGGTGAT TTCGTACAAG
1851 TAGCACCAGG CTGCGGAGTT CCTATTGTGG ATTCATATTA CTCATTGCTG
1901 ATGCCCATCC TCACTTTGAC TAGGGCATTG GCTGCTGAGT CCCATATGGA
1951 TGCTGATCTC GCAAACCAC TTATTAAGTG GGATTTGCTG AAATATGATT
2001 TTACGGAAGA GAGACTTTGT CTCTTCGACC GTTATTTTAA ATATTGGGAC
2051 CAGACATACC ATCCCAATTG TATTAAGTGT TTGGATGATA GGTGTATCCT
2101 TCATTGTGCA AACTTTAATG TGTTATTTT TACTGTGTTT CCACCTACAA
2151 GTTTTGGACC ACTAGTAAGA AAAATATTTG TAGATGGTGT TCCTTTTGT
2201 GTTTCAACTG GATACCATT TCGTGAGTTA GGAGTCGTAC ATAATCAGGA
2251 TGTAACCTTA CATAGCTCGC GTCTCAGTTT CAAGGAACCT TTAGTGATAG
2301 CTGCTGATCC AGCTATGCAT GCAGCTTCTG GCAATTTATT GCTAGATAAA
2351 CGCACTACAT GCTTTTCAGT AGCTGCACTA ACAAACAATG TTGCTTTTCA
```

2401 AACTGTCAAA CCCGGTAATT TTAATAAAGA CTTTTATGAC TTTGCTGTGT  
2451 CTAAGGTTTT CTTAAGGAA GGAAGTTCTG TTGAACTAAA AACTTCTTC  
2501 TTTGCTCAGG ATGGCAACGC TGCTATCAGT GATTATGACT ATTATCGTTA  
2551 TAATCTGCCA ACAATGTGTG ATATCAGACA ACTCCTATTC GTAGTTGAAG  
2601 TTGTTGATAA ATACTTTGAT TGTTACGATG GTGGCTGTAT TAATGCCAAC  
2651 CAAGTAATCG TTAACAATCT GGATAAATCA GCTGGTTTCC CATTAAATAA  
2701 ATGGGGTAAG GCTAGACTTT ATTATGACTC AATGAGTTAT GAGGATCAAG  
2751 ATGCACTTTT CGCGTATACT AAGCGTAATG TCATCCCTAC TATAACTCAA  
2801 ATGAATCTTA AGTATGCCAT TAGTGCAAAG AATAGAGCTC GCACCGTAGC  
2851 TGGTGTCTCT ATCTGTAGTA CTATGACAAA TAGACAGTTT CATCAGAAAT  
2901 TATTGAAGTC AATAGCCGCC ACTAGAGGAG CTA CTGTGGT AATTGGAACA  
2951 AGCAAGTTTT ACGGTGGCTG GCATAATATG TTA AAAACTG TTTACAGTGA  
3001 TGTAGAAACT CCACACCTTA TGGGTTGGGA TTATCCAAAA TGTGACAGAG  
3051 CCATGCCTAA CATGCTTAGG ATAATGGCCT CTCTTGTCT TGCTCGCAA  
3101 CATAACACTT GCTGTAATT ATCACACCGT TTCTACAGGT TAGTAAACGA  
3151 GTGTGCGCAA GTATTAAGTG AGATGGTCAT GTGTGGCGGC TCACTATAG  
3201 TTAACCAGG TGGAAATCA TCCGGTGATG CTACA ACTGC TTATGCTAAT  
3251 AGTGTCTTTA ACATTTGTCA AGCTGTTACA GCCAATGTAA ATGCACTTCT  
3301 TTCAACTGAT GGTAATAAGA TAGCTGACAA GTATGTCCGC AATCTACAAC  
3351 ACAGGCTCTA TGAGTGTCTC TATAGAAATA GGGATGTTGA TCATGAATTC  
3401 GTGGATGAGT TTTACGCTTA CCTGCGTAAA CATTCTCCA TGATGATTCT  
3451 TTCTGATGAT GCCGTTGTGT GCTATAACAG TAACTATGCG GCTCAAGGTT  
3501 TAGTAGCTAG CATTAAAGAAC TTTAAGGCAG TTCTTTATTA TCAAAATAAT  
3551 GTGTTTATGT CTGAGGCAA ATGTTGGACT GAGACTGACC TTAATAAAGG  
3601 ACCTCACGAA TTTTGCTCAC AGCATAAAT GCTAGTTAAA CAAGGAGATG  
3651 ATTACGTGTA CCTGCCTTAC CCAGATCCAT CAAGAATATT AGGCGCAGGC  
3701 TGTTTTGTCT ATGATATTGT CAAAACAGAT GGTACACTTA TGATTGAAAG  
3751 GTTCGTGTCA CTGGCTATTG ATGCTTACCC ACTTACAAAA CATCCTAATC  
3801 AGGAGTATGC TGATGTCTTT CACTGTATT TACAATACAT TAGAAAGTTA  
3851 CATGATGAGC TTA CTGGCCA CATGTTGGAC ATGTATTCCG TAATGCTAAC  
3901 TAATGATAAC ACCTCACGGT ACTGGGAACC TGAGTTTTAT GAGGCTATGT  
3951 ACACACCACA TACAGTCTTG CAGGCTGTAG GTGCTTGTGT ATTGTGCAAT  
4001 TCACAGACTT CACTTCGTTG CGGTGCCTGT ATTAGGAGAC CATTCTATG  
4051 TTGCAAGTGC TGCTATGACC ATGTCATTTT AACATCACAC AAATTAGTGT  
4101 TGTCTGTAA TCCCTATGTT TGCAATGCCC CAGGTTGTGA TGCTACTGAT  
4151 GTGACACAAC TGTATCTAGG AGGTATGAGC TATTATTGCA AGTCACATAA  
4201 GCCTCCCATT AGTTTTCCAT TATGTGCTAA TGGTCAGGTT TTTGGTTTAT  
4251 AAAAAACAC ATGTGTAGGC AGTGACAATG TCACTGACTT CAATGCGATA  
4301 GCAACATGTG ATGGACTAA TGCTGGCGAT TACATACTTG CCAACTGTTG  
4351 TACTGAGAGA CTCAAGCTTT TCGCAGCAGA AACGCTCAA GCCACTGAGG  
4401 AAACATTTAA GCTGTCATAT GGTATTGCTA CTGTACGCGA AGTACTCTCT  
4451 GACAGAGAAT TGCATCTTTC ATGGGAGGTT GGAAAACCTA GACCACCATT  
4501 GAACAGAAAC TATGTCTTTA CTGGTTACCG TGTA ACTAAA AATAGTAAAG  
4551 TACAGATTGG AGAGTACACC TTTGAAAAAG GTGACTATGG TGATGCTGTT  
4601 GTGTACAGAG GTACTACGAC ATACAAGTTG AATGTTGGTG ATTACTTTGT  
4651 GTTGACATCT CACTGTAA TGCCACTTAG TGCACCTACT CTAGTGCCAC  
4701 AAGAGCACTA TGTGAGAATT ACTGGCTTGT ACCCAACT CAACATCTCA  
4751 GATGAGTTTT CTAGCAATGT TGCAAATTAT CAAAAGGTCG GCATGCAAAA  
4801 GTACTCTACA CTCCAAGGAC CACCTGGTAC TGGTAAGAGT CATTTTGCCA  
4851 TCGGACTTGC TCTCTATTAC CCATCTGCTC GCATAGTGTA TACGGCATGC  
4901 TCTCATGCAG CTGTTGATGC CCTATGTGAA AAGGCATTA AATATTTGCC  
4951 CATAGATAAA TGTAGTAGAA TCATACCTGC GCGTGC GCGC GTAGAGTGT  
5001 TTGATAAATT CAAAGTGAAT TCAACTAG AACAGTATGT TTTCTGCACT  
5051 GAAATGCAT TGCCAGAAAC AACTGCTGAC ATTGAGTCT TTGATGAAAT  
5101 CTCTATGGCT ACTAATTATG ACTTGAGTGT TGTC AATGCT AGACTTCGTG  
5151 CAAAACACTA CGTCTATATT GCGATCCTG CTCAATTACC AGCCCCCGC

5201 ACATTGCTGA CTAAAGGCAC ACTAGAACCA GAATATTTTA ATTCAGTGTG  
5251 CAGACTTATG AAAACAATAG GTCCAGACAT GTTCCTTGGG ACTTGTGCGC  
5301 GTTGTCTGCG TGAATTGTT GACACTGTGA GTGCTTTAGT TTATGACAAT  
5351 AAGCTAAAAG CACACAAGGA TAAGTCAGCT CAATGCTTCA AAATGTTCTA  
5401 CAAAGGTGTT ATTACACATG ATGTTTCATC TGCAATCAAC AGACCTCAA  
5451 TAGGCGTTGT AAGAGAATTT CTTACACGCA ATCCTGCTTG GAGAAAAGCT  
5501 GTTTTTATCT CACCTTATAA TTCACAGAAC GCTGTAGCTT CAAAAATCTT  
5551 AGGATTGCCT ACGCAGACTG TTGATTCATC ACAGGGTCTT GAATATGACT  
5601 ATGTCATATT CACACAACT ACTGAAACAG CACTCTTG TAATGTCAAC  
5651 CGCTTCAATG TGGCTATCAC AAGGGCAAAA ATTGGCATT TGTGCATAAT  
5701 GTCTGATAGA GATCTTTATG ACAAAGTCA ATTTACAAGT CTAGAAATAC  
5751 CACGTCGCAA TGTGGCTACA TTACAAGCAG AAAATGTAAC TGGACTTTTT  
5801 AAGGACTGTA GTAAGATCAT TACTGGTCTT CATCCTACAC AGGCACCTAC  
5851 ACACCTCAGC GTTGATATAA AGTTCAAGAC TGAAGGATTA TGTGTTGACA  
5901 TACCAGGCAT ACCAAAGGAC ATGACCTACC GTAGACTCAT CTCTATGATG  
5951 GGTTTCAAAA TGAATTACCA AGTCAATGGT TACCCTAATA TGTTTATCAC  
6001 CCGCGAAGAA GCTATTCGTC ACGTTCGTGC GTGGATTGGC TTTGATGTAG  
6051 AGGGCTGTCA TGCAACTAGA GATGCTGTGG GTACTAACCT ACCTCTCCAG  
6101 CTAGGATTTT CTACAGGTGT TAACTTAGTA GCTGTACCGA CTGGTTATGT  
6151 TGACACTGAA AATAACACAG AATTCACCAG AGTTAATGCA AAACCTCCAC  
6201 CAGGTGACCA GTTTAAACAT CTTATAACCAC TCATGTATAA AGGCTTGCCC  
6251 TGGAATGTAG TCGTATTAA GATAGTACAA ATGCTCAGTG ATACACTGAA  
6301 AGGATTGTCA GACAGAGTCG TGTTTCGCTT TTGGGCGCAT GGCTTTGAGC  
6351 TTACATCAAT GAAGTACTTT GTCAAGATTG GACCTGAAAG AACGTGTTGT  
6401 CTGTGTGACA AACGTGCAAC TTGCTTTTCT ACTTCATCAG ATACTTATGC  
6451 CTGCTGGAAT CATTCTGTGG GTTTTGACTA TGTCTATAAC CCATTTATGA  
6501 TTGATGTTCA GCAGTGGGGC TTTACGGGTA ACCTTCAGAG TAACCATGAC  
6551 CAACATTGCC AGGTACATGG AAATGCACAT GTGGCTAGTT GTGATGCTAT  
6601 CATGACTAGA TGTTTAGCAG TCCATGAGTG CTTTGTTAAG CGCGTTGATT  
6651 GGTCTGTTGA ATACCCTATT ATAGGAGATG AACTGAGGGT TAATTCGTCT  
6701 TGCAGAAAAG TACAACACAT GGTTGTGAAG TCTGCATTGC TTGCTGATAA  
6751 GTTTCCAGTT CTTTCATGACA TTGGAAATCC AAAGGCTATC AAGTGTGTGC  
6801 CTCAGGCTGA AGTAGAATGG AAGTTCACG ATGCTCAGCC ATGTAGTGAC  
6851 AAAGCTTACA AAATAGAGGA GCTCTTCTAT TCTTATGCTA CACATCACGA  
6901 TAAATTCATG GATGGTGTTT GTTTGTGTTG GAATTGTAAC GTTGATCGTT  
6951 ACCCAGCCAA TGCAATTGTG TGTAGGTTTG ACACAAGAGC CTTGTCAAAC  
7001 TTGAACTTAC CAGGCTGTGA TGGTGGTAGT TTGTATGTGA ATAAGCATGC  
7051 ATCCACACT CCAGCTTTCG ATAAAAGTGC ATTTACTAAT TAAAGCAAT  
7101 TGCCTTCTT TACTATTCT GATAGTCTT GTGAGTCTCA TGGCAAACAA  
7151 GTAGTGTGCG ATATTGATTA TGTCCACTC AAATCTGCTA CGTGATTAC  
7201 ACGATGCAAT TTAGGTGGTG CTGTTTGCAG ACACCATGCA AATGAGTACC  
7251 GACAGTACTT GGATGCATAT AATATGATGA TTTCTGCTGG ATTTAGCCTA  
7301 TGGATTTACA AACAATTTGA TACTTATAAC CTGTGGAATA CATTTACCAG  
7351 GTTACAGAGT TTAGAAAATG TGGCTTATAA TGTTGTTAAT AAAGGACACT  
7401 TTGATGGACA CGCCGGCGAA GCACCTGTTT CCATCATTAA TAATGCTGTT  
7451 TACACAAAGG TAGATGGTAT TGATGTGGAG ATCTTTGAAA ATAAGACAAC  
7501 ACTTCCTGTT AATGTTGCAT TTGAGCTTTG GGCTAAGCGT AACATTAAC  
7551 CAGTGCCAGA GATTAAGATA CTCAATAATT TGGGTGTTGA TATCGCTGCT  
7601 AATACTGTAA TCTGGGACTA CAAAAGAGAA GCCCAGCAC ATGTATCTAC  
7651 AATAGGTGTC TGCACAATGA CTGACATTGC CAAGAAACCT ACTGAGAGTG  
7701 CTTGTTCTTC ACTTACTGTC TTGTTTGATG GTAGAGTGGA AGGACAGGTA  
7751 GACCTTTTTA GAAACGCCC TAATGGTGT TTAATAACAG AAGGTTCACT  
7801 CAAAGGTCTA ACACCTTCAA AGGGACCAGC ACAAGCTAGC GTCAATGGAG  
7851 TCACATTAAT TGGAGAATCA GTAAAAACAC AGTTTAACTA CTTTAAAGAA  
7901 GTAGACGGCA TTATTCAACA GTTGCCTGAA ACCTACTTTA CTCAGAGCAG  
7951 AGACTTAGAG GATTTTAAAGC CCAGATCACA AATGGAACT GACTTTCTCG

8001 AGCTCGCTAT GGATGAATTC ATACAGCGAT ATAAGCTCGA GGGCTATGCC  
8051 TTCGAACACA TCGTTTATGG AGATTTTCAGT CATGGACAAC TTGGCGGTCT  
8101 TCATTTAATG ATAGGCTTAG CCAAGCGCTC ACAAGATTCA CCACTTAAAT  
8151 TAGAGGATTT TATCCCTATG GACAGCACAG TGAAAAATTA CTTCATAACA  
8201 GATGCGCAAA CAGGTTCATC AAAATGTGTG TGTTCTGTGA TTGATCTTTT  
8251 ACTTGATGAC TTTGTCGAGA TAATAAAGTC ACAAGATTTG TCAGTGATTT  
8301 CAAAAGTGGT CAAGGTTACA ATTGACTATG CTGAAATTTT ATTCATGCTT  
8351 TGGTGTAAGG ATGGACATGT TGAAACCTTC TACCCHAAAAC TACAAGCAAG  
8401 TCAAGCGTGG CAACCAGGTG TTGCGATGCC TAACTTGTAC AAGATGCAAAA  
8451 GAATGCTTCT TGAAAAGTGT GACCTTCAGA ATTATGGTGA AAATGCTGTT  
8501 ATACCAAAAAG GAATAATGAT GAATGTCGCA AAGTATACTC AACTGTGTCA  
8551 ATACTTAAAT ACACTTACTT TAGCTGTACC CTACAACATG AGAGTTATTC  
8601 ACTTTGGTGC TGGCTCTGAT AAAGGAGTTG CACCAGGTAC AGCTGTGCTC  
8651 AGACAATGGT TGCCAACCTG CACACTACTT GTCGATTGAG ATCTTAATGA  
8701 CTTCGTCTCC GACGCAGATT CTACTTTAAT TGGAGACTGT GCAACAGTAC  
8751 ATACGGCTAA TAAATGGGAC CTTATTATTA GCGATATGTA TGACCCTAGG  
8801 ACCAAACATG TGACAAAAGA GAATGACTCT AAAGAAGGGT TTTTCACTTA  
8851 TCTGTGTGGA TTTATAAAGC AAAAAGTAGC CCTGGGTGGT TCTATAGCTG  
8901 TAAAGATAAC AGAGCATTCT TGGAATGCTG ACCTTTACAA GCTTATGGGC  
8951 CATTCTCAT GGTGGACAGC TTTTGTGTTACA AATGTAAATG CATCATCATC  
9001 GGAAGCATT TTAATTGGGG CTAAGTATCT TGGCAAGCCG AAGGAACAAA  
9051 TTGATGGCTA TACCATGCAT GCTAACTACA TTTTCTGGAG GAACACAAAT  
9101 CCTATCCAGT TGTCTTCTA TTCACTCTTT GACATGAGCA AATTTCTCT  
9151 TAAATTAAGA GGAAGTCTG TAATGTCTCT TAAGGAGAAT CAAATCAATG  
9201 ATATGATTTA TTCTTCTG GAAAAAGGTA GGCTTATCAT TAGAGAAAAC  
9251 AACAGAGTTG TGGTTTCAAG TGATATTCTT GTTAACAAC AAACGAACAT  
9301 GTTTATTTT TATTATTTT TACTCTCAC TAGTGGTAGT GACCTTGACC  
9351 GGTGCACCAC TTTTGATGAT GTTCAAGCTC CTAATTACAC TCAACATACT  
9401 TCATCTATGA GGGGGGTTA CTATCCTGAT GAAATTTTA GATCAGACAC  
9451 TCTTTATTTA ACTCAGGATT TATTTCTTCC ATTTTATTCT AATGTTACAG  
9501 GGTTTCATAC TATTAATCAT ACGTTTGGCA ACCCTGTCAT ACCTTTAAG  
9551 GATGGTATTT ATTTGCTGC CACAGAGAAA TCAAATGTTG TCCGTGGTTG  
9601 GGTTTTTGGT TCTACCATGA ACAACAAGTC ACAGTCGGTG ATTATTATTA  
9651 ACAATTCTAC TAATGTTGTT ATACGAGCAT GTAACCTTGA ATTGTGTGAC  
9701 AACCTTTCT TTGCTGTTT TAAACCATG GGTACACAGA CACATACTAT  
9751 GATATTCGAT AATGCATTTA ATTGCATTTT CGAGTACATA TCTGATGCCT  
9801 TTTGCTTGA TGTTTCAGAA AAGTCAGGTA ATTTTAAACA CTTACGAGAG  
9851 TTTGTGTTA AAAATAAAGA TGGGTTTCTC TATGTTTATA AGGGCTATCA  
9901 ACCTATAGAT GTAGTTCGTG ATCTACCTTC TGGTTTTAAC ACTTTGAAAC  
9951 CTATTTTAA GTTGCCTCTT GGTATTAACA TTACAAATTT TAGAGCCATT  
10001 CTACAGCCT TTTACCTGC TCAAGACATT TGGGGCACGT CAGCTGCAGC  
10051 CTATTTTGTT GGCTATTTAA AGCCAACACT ATTTATGCTC AAGTATGATG  
10101 AAAATGGTAC AATCACAGAT GCTGTTGATT GTTCTCAAAA TCCACTTGCT  
10151 GAACTCAAAT GCTCTGTAA GAGCTTTGAG ATTGACAAAAG GAATTTACCA  
10201 GACCTCTAAT TTCAGGGTTG TTCCCTCAGG AGATGTTGTG AGATTCCCTA  
10251 ATATTACAAA CTTGTGTCCT TTTGGAGAGG TTTTAAATGC TACTAAATC  
10301 CCTTCTGTCT ATGCATGGGA GAGAAAAAAA ATTTCTAATT GTGTTGCTGA  
10351 TFACTCTGTG CTCTACAAC CAACATTTTT TTCAACCTTT AAGTGCTATG  
10401 GCGTTTCTGC CACTAAGTTG AATGATCTTT GCTTCTCAA TGTCTATGCA  
10451 GATCTTTTGT TAGTCAAGGG AGATGATGTA AGACAAATAG CGCCAGGACA  
10501 AACTGGTGTT ATTGCTGATT ATAATTATAA ATTGCCAGAT GATTTTCATGG  
10551 GTTGTGTCCT TGCTTGAAT ACTAGGAACA TTGATGCTAC TTCAACTGGT  
10601 AATTATAATT ATAAATATAG GTATCTTAGA CATGGCAAGC TTAGGCCCTT  
10651 TGAGAGAGAC ATATCTAATG TGCTTTCTC CCCTGATGGC AAACCTTGCA  
10701 CCCACCTGC TCTAATTGT TATTGGCCAT TAAATGATTA TGGTTTTTAC  
10751 ACCACTACTG GCATTGGCTA CCAACCTTAC AGAGTTGTAG TACTTTCTT

10801 TGAAC TTTTA AATGCACCGG CCACGGTTTG TGGACCAAAA TTATCCACTG  
10851 ACCTTATTAA GAACCAGTGT GTCAATTTTA ATTTAATGG ACTCACTGGT  
10901 ACTGGTGTGT TAACCTCTT TCCAAAGAGA TTTCAACCAT TTCAACAATT  
10951 TGGCCGTGAT GTTTCTGATT TCACTGATTC CGTTCGAGAT CCTAAAACAT  
11001 CTGAAATATT AGACATTTCA CCTTGCTCTT TTGGGGGTGT AAGTGT AATT  
11051 ACACCTGGAA CAAATGCTTC ATCTGAAGTT GCTGTTCTAT ATCAAGATGT  
11101 TAACTGCACT GATGTTTCTA CAGCAATTCA TGCAGATCAA CTCACACCAG  
11151 CTTGGCGCAT ATATTCTACT GGAAACAATG TATTCCAGAC TCAAGCAGGC  
11201 TGTCTTATAG GAGCTGAGCA TGTCGACACT TCTTATGAGT GCGACATTCC  
11251 TATTGGAGCT GGCATTTGTG CTAGTTACCA TACAGTTTCT TTATTACGTA  
11301 GTACTAGCCA AAAATCTATT GTGGCTTATA CTATGTCTTT AGGTGCTGAT  
11351 AGTTCAATTG CTTACTCTAA TAACACCATT GCTATACCTA CTAAC TTTTC  
11401 AATTAGCATT ACTACAGAAG TAATGCCTGT TTCTATGGCT AAAACCTCCG  
11451 TAGATTGTAA TATGTACATC TGC GGAGATT C TACTGAATG TGCTAATTTG  
11501 CTTCTCCAAT ATGGTAGCTT TTGCACACAA C TAAATCGTG CACTCTCAGG  
11551 TATTGCTGCT GAACAGGATC GCAACACACG TGAAGTGTTC GCTCAAGTCA  
11601 AACAAATGTA CAAAACCCCA ACTTTGAAAT ATTTTGGTGG TTTAATTTT  
11651 TCACAAATAT TACCTGACCC TCTAAAGCCA ACTAAGAGGT CTTTATTGA  
11701 GGACTTGCTC TTTAATAAGG TGACACTCGC TGATGCTGGC TTCATGAAGC  
11751 AATATGGCGA ATGCCTAGGT GATATTAATG CTAGAGATCT CATTGTGCG  
11801 CAGAAGTTCA ATGGACTTAC AGTGTTGCCA CCTCTGCTCA CTGATGATAT  
11851 GATTGCTGCC TACTGCTG CTCTAGTTAG TGGTACTGCC ACTGCTGGAT  
11901 GGACATTTGG TGCTGGCGCT GCTCTTCAA TAC TTTTGC TATGCAAATG  
11951 GCATATAGGT TCAATGGCAT TGGAGTTACC CAAATGTTC TCTATGAGAA  
12001 CCAAAAACAA ATCGCCAACC AATTTAACAA GCGGATTAGT CAAATCAAG  
12051 AATCACTTAC AACACATCA ACTGCATTGG GCAAGCTGCA AGACGTTGTT  
12101 AACCAGAATG C TCAAGCATT AAACACACTT GTTAAACAAC TTAGCTCTAA  
12151 TTTTGGTGCA ATTTCAAGTG TGCTAAATGA TATCCTTTCG CGACTTGATA  
12201 AAGTCGAGGC GGAGGTACAA ATTGACAGGT TAATTACAGG CAGACTTCAA  
12251 AGCCTTCAA CCTATGTAAC ACAACAATA ATCAGGGCTG CTGAAATCAG  
12301 GGCTTCTGCT AATCTTGCTG C TACTAAAAT GTCTGAGTGT GTTCTTGGAC  
12351 AATCAAAAAG AGTTGACTTT TGTGGAAAGG GCTACCACCT TATGTCCTTC  
12401 CCACAAGCAG CCCC GCATGG TGTGTCTTC CTACATGTCA CGTATGTGCC  
12451 ATCCCAGGAG AGGAACTTCA CCACAGCGCC AGCAATTTGT CATGAAGGCA  
12501 AAGCATACTT CCTCGTGAA GGTGTTTTTG TGTTAATGG CACTTCTTGG  
12551 TTTATTACAC AGAGGAACTT CTTTCTCCA CAAATAATTA CTACAGACAA  
12601 TACATTTGTC TCAGGAAATT GTGATGTCGT TATTGGCATC ATTAACAACA  
12651 CAGTTTATGA TCCTCTGCAA CCTGAGCTCG ACTCATTCAA AGAAGAGCTG  
12701 GACAAGTACT TCAAAAATCA TACATCACCA GATGTTGATC TTGGCGCAT  
12751 TTCAGGCACT AACGCTTCTG TCGTCAACAT TCAAAAAGAA ATGACCGCC  
12801 TCAATGAGGT CGTAAAAAT TAAATGAAT CACTCATTGA CCTCAAGAA  
12851 TTGGGAAAAT ATGAGCAATA TATTAAATGG CTTGGTATG TTTGGCTCGG  
12901 CTTCATTGCT GGACTAATTG CCATCGTCAT GGTTACAATC TTGCTTTGTT  
12951 GCATGACTAG TTGTTGCAGT TGCCTCAAGG GTGCATGCTC TTGTGGTTCT  
13001 TGCTGCAAGT TTGATGAGGA TGACTCTGAG CCAGTTCTCA AGGGTGTCAA  
13051 ATTACATTAC ACATAAACGA ACTTATGGAT TTGTTTATGA GATTTTTTAC  
13101 TCTTGGATCA ATTACTGCAC AGCCAGTAAA AATTGACAAT GCTTCTCCTG  
13151 CAAGTACTGT TCATGCTACA GCAACGATAC CGCTACAAGC CTCACTCCCT  
13201 TTCGGATGGC TTGTTATTGG CGTTGCATTT CTGCTGTTT TTCAGAGCGC  
13251 TACCAAAATA ATTGCGCTCA ATAAAAGATG GCAGCTAGCC CTTTATAAGG  
13301 GCTTCCAGTT CATTGCAAT T TACTGCTGC TATTTGTTAC CATCTATTCA  
13351 CATCTTTTGC TTGTCGCTGC AGGTATGGAG GCGCAATTTT TGTACCTCTA  
13401 TGCCTTGATA TATTTTCTAC AATGCATCAA CGCATGTAGA ATTATTATGA  
13451 GATGTTGGCT TTGTTGGAAG TGCAAATCCA AGAACCATT ACTTTATGAT  
13501 GCCAACTACT TTGTTGCTG GCACACACAT AACTATGACT ACTGTATACC  
13551 ATATAACAGT GTCACAGATA CAATTGTCGT TACTGAAGGT GACGGCATT

13601 CAACACCAAA ACTCAAAGAA GACTACCAAA TTGGTGGTTA TTCTGAGGAT  
13651 AGGCACTCAG GTGTTAAAGA CTATGTCGTT GTACATGGCT ATTCACCGA  
13701 AGTTTACTAC CAGCTTGAGT CTACACAAAT TACTACAGAC ACTGGTATTG  
13751 AAAATGCTAC ATTCTTCATC TTAAACAAGC TTGTTAAAGA CCCACCGAAT  
13801 GTGCAAATAC ACACAATCGA CGGCTCTTCA GGAGTTGCTA ATCCAGCAAT  
13851 GGATCCAATT TATGATGAGC CGACGACGAC TACTAGCGTG CTTTGTAAAG  
13901 CACAAGAAAG TGAGTACGAA CTTATGTACT CATTTCGTTT GGAAGAAACA  
13951 GGTACGTTAA TAGTTAATAG CGTACTTCTT TTTCTTGCTT TCGTGGTATT  
14001 CTTGCTAGTC AACTAGCCA TCCTTACTGC GCTTCGATTG TGTGCGTACT  
14051 GCTGCAATAT TGTTAACGTG AGTTTAGTAA AACCAACGGT TTACGTCTAC  
14101 TCGCGTGTTA AAAATCTGAA CTCTTCTGAA GGAGTTCTG ATCTTCTGGT  
14151 CTAACGAAC TAACTATTAT TATTATTCTG TTTGGAAC TTAAACATTGCT  
14201 TATCATGGCA GACAACGGTA CTATTACCGT TGAGGAGCTT AAACAACCTC  
14251 TGGAACAATG GAACCTAGTA ATAGGTTTCC TATTCTAGC CTGGATTATG  
14301 TTACTCAAT TTGCCTATC TAATCGGAAC AGGTTTTTGT ACATAATAAA  
14351 GCTTGTCTTCT TGTGGCTCT TGTGGCCAGT AACACTTGCT TGTTTGTGCT  
14401 TTGCTGCTGT CTACAGAATT AATTGGGTGA CTGGCGGGAT TGGCATTGCA  
14451 ATGGCTTGTA TTGTAGGCTT GATGTGGCTT AGCTACTTCG TTGCTTCTT  
14501 CAGGCTGTTT GCTCGTACCC GCTCAATGTG GTCATTCAAC CCAGAAACAA  
14551 ACATTCTTCT CAATGTGCCT CTCCGGGGGA CAATTGTGAC CAGACCGCTC  
14601 ATGGAAAGTG AACTTGTCAT TGGTGCTGTG ATCATTCTGT GTCATTGCG  
14651 AATGGCCGGA CACCCCTAG GGCCTGTGA CATTAAAGGAC CTGCCAAAAG  
14701 AGTCACTGT GGCTACATCA CGAACGCTTT CTTATTACAA ATTAGGAGCG  
14751 TCGCAGCGTG TAGGCACTGA TTCAGGTTTT GCTGCATACA ACCGCTACCG  
14801 TATTGAAAC TATAAATTA ATACAGACCA CGCCGGTAGC AACGACAATA  
14851 TTGCTTTGCT AGTACAGTAA GTGACAACAG ATGTTTCATC TTGTTGACTT  
14901 CCAGTTACA ATAGCAGAGA TATTGATTAT CATTATGAGG ACTTTCAGGA  
14951 TTGCTATTTG GAATCTTGAC GTTATAATAA GTTCAATAGT GAGACAATTA  
15001 TTAAGCCTC TAACTAAGAA GAATTATTCG GAGTTAGATG ATGAAGAACC  
15051 TATGGAGTTA GATTATCCAT AAAACGAACA TGAAAATTAT TCTCTCCTG  
15101 ACATTGATTG TATTTACATC TTGCGAGCTA TACTACTATC AGGAGTGTGT  
15151 TAGAGGTACG ACTGTACTAC TAAAAGAACC TTGCCATCA GGAACATACG  
15201 AGGGCAATC ACCATTTAC CCTCTTGCTG ACAATAAAT TGCCTAACT  
15251 TGCCTAGCA CACTTTGC TTTTGCTTG GCTGACGGTA CTCGACATAC  
15301 CTATCAGCTG CGTGCAAGAT CAGTTTACC AAAACTTTT ATCAGACAAG  
15351 AGGAGGTTCA ACAAGAGCTC TACTCGCCAC TTTTCTCAT TGTTGCTGCT  
15401 CTAGTATTTT TAATACTTTG CTTACCAATT AAGAGAAAGA CAGAATGAAT  
15451 GAGCTCACTT TAATTGACTT CTATTTGTGC TTTTAGCCT TTCTGCTATT  
15501 CCTTGTTTTA ATAATGCTTA TTATATTTT GTTTTCACT GAAATCCAGG  
15551 ATCTAGAAGA ACCTGTACC AAGTCTAAA CGAACATGAA ACTTCTCATT  
15601 GTTTGACTT GTATTTCTCT ATGCAGTTGC ATATGCACTG TAGTACAGCG  
15651 CTGTGCATCT AATAAACCTC ATGTGCTTGA AGATCCTTGT AAGGTACAAC  
15701 ACTAGGGGTA AACTTATAG CACTGCTTGG CTTTGTGCTC TAGGAAAGGT  
15751 TTTACCTTTT CATAGATGGC AACTATGGT TCAAACATGC ACACCTAATG  
15801 TTAATATCAA CTGTCAAGAT CCAGCTGGTG GTGCGCTTAT AGCTAGGTGT  
15851 TGGTACCTT ATGAAGGTCA CAAACTGCT GCATTTAGAG ACGTACTTGT  
15901 TGTTTTAAAT AAACGAACAA ATTAATAATGT CTGATAATGG ACCCAATCA  
15951 AACCAACGTA GTGCCCCCG CATTACATTT GGTGGACCCA CAGATTCAAC  
16001 TGACAATAAC CAGAATGGAG GACGCAATGG GGCAAGGCCA AACAGCGCC  
16051 GACCCCAAGG TTTACCCAAT AATACTGCGT CTTGGTTTAC AGCTCTCACT  
16101 CAGCATGGCA AGGAGGAACT TAGATTCCT CGAGGCCAGG GCGTTCCAAT  
16151 CAACACCAAT AGTGGTCCAG ATGACCAAAT TGGCTACTAC CGAAGAGCTA  
16201 CCCGACGAGT TCGTGGTGGT GACGGCAAAA TGAAAGAGCT CAGCCCAGA  
16251 TGGTACTTCT ATTACCTAGG AACTGGCCA GAAGCTTAC TTCCCTACGG  
16301 CGTAACAAA GAAGGCATCG TATGGGTTGC AACTGAGGGA GCCTTGAATA  
16351 CACCAAAGA CCACATTGGC ACCCGCAATC CTAATAACAA TGCTGCCACC

16401 GTGCTACAAC TTCCTCAAGG AACAAACATTG CCAAAGGCT TCTACGCAGA  
16451 GGAAGCAGA GCGGCAGTC AAGCCTCTC TCGCTCCTCA TCACGTAGTC  
16501 GCGGTAATTC AAGAAATTCA ACTCCTGGCA GCAGTAGGGG AAATTCTCCT  
16551 GCTCGAATGG CTAGCGGAGG TGGTAAACT GCCCTCGCGC TATTGCTGCT  
16601 AGACAGATTG AACAGCTTG AGAGCAAAGT TTCTGGTAAA GGCCAACAAC  
16651 AACAAAGCCA AACTGTCACT AAGAAATCTG CTGCTGAGGC ATCTAAAAAG  
16701 CCTCGCCAAA AACGTAAGT CACAAAACAG TACAACGTCA CTCAAGCATT  
16751 TGGGAGACGT GGTCCAGAAC AAACCAAGG AAATTTCTGGG GACCAAGACC  
16801 TAATCAGACA AGGAACTGAT TACAAACATT GGCCGCAAAT TGCACAATTT  
16851 GCTCCAAGTG CCTCTGCATT CTTTGGAAATG TCACGCATTG GCATGGAAGT  
16901 CACACCTTCG GGAACATGGC TGACTTATCA TGGAGCCATT AAATTGGATG  
16951 ACAAAGATCC ACAATTCAAA GACAACGTCA TACTGCTGAA CAAGCACATT  
17001 GACGCATACA AACATTCCC ACCAACAGAG CCTAAAAAGG AAAAAAGAA  
17051 AAAGACTGAT GAAGCTCAGC CTTTGCCGCA GAGACAAAAG AAGCAGCCCA  
17101 CTGTGACTCT TCTTCCTGCG GCTGACATGG ATGATTTCTC CAGACAACCT  
17151 CAAAATTCCA TGAGTGGAGC TTCTGCTGAT TCAACTCAGG CATAAACACT  
17201 CATGATGACC ACACAAGGCA GATGGGCTAT GTAAACGTTT TCGCAATTCC  
17251 GTTACGATA CATAGTCTAC TCTGTGCAG AATGAATTCT CGTAACTAAA  
17301 CAGCACAAGT AGGTTTAGTT AACTTTAATC TCACATAGCA ATCTTTAATC  
17351 AATGTGTAAC ATTAGGGAGG ACTTGAAAGA GCCACCACAT TTTCATCGAG  
17401 GCCACGCGGA GTACGATCGA GGGTACAGTG AATAATGCTA GGGAGAGCTG  
17451 CCTATATGGA AGAGCCCTAA TGTGTAATAA TAATTTTAGT AGTGCTATCC  
17501 CCATGTGATT TTAATAGCTT CTTAGGAGAA TGACAAAAAA AAAAAAAAAA  
17551 AAAAA