The following mRNA sequences are used with the Seblastian program (<http://seblastian.crg.es>) in Parts II and III of the case study “Computers and Micronutrients: Using Bioinformatic Tools to Uncover Selenoproteins and Mutations” by Winyoo Chowanadisai and Bryant H. Keirns, Department of Nutritional Sciences, Oklahoma State University, Stillwater, OK. This document may be shared with students in case the links are broken, the students have trouble finding the sequences on the webpage, or if there are difficulties accessing the website by internet.

*Sequence 1:* Human Glutathione Peroxidase 1

*Source:* <https://www.ncbi.nlm.nih.gov/nuccore/NM_000581.4>

 1 aaaaggaggc gcctgctggc ctccccttac agtgcttgtt cggggcgctc cgctggcttc

 61 ttggacaatt gcgccatgtg tgctgctcgg ctagcggcgg cggcggcggc ggcccagtcg

 121 gtgtatgcct tctcggcgcg cccgctggcc ggcggggagc ctgtgagcct gggctccctg

 181 cggggcaagg tactacttat cgagaatgtg gcgtccctct gaggcaccac ggtccgggac

 241 tacacccaga tgaacgagct gcagcggcgc ctcggacccc ggggcctggt ggtgctcggc

 301 ttcccgtgca accagtttgg gcatcaggag aacgccaaga acgaagagat tctgaattcc

 361 ctcaagtacg tccggcctgg tggtgggttc gagcccaact tcatgctctt cgagaagtgc

 421 gaggtgaacg gtgcgggggc gcaccctctc ttcgccttcc tgcgggaggc cctgccagct

 481 cccagcgacg acgccaccgc gcttatgacc gaccccaagc tcatcacctg gtctccggtg

 541 tgtcgcaacg atgttgcctg gaactttgag aagttcctgg tgggccctga cggtgtgccc

 601 ctacgcaggt acagccgccg cttccagacc attgacatcg agcctgacat cgaagccctg

 661 ctgtctcaag ggcccagctg tgcctagggc gcccctccta ccccggctgc ttggcagttg

 721 cagtgctgct gtctcggggg ggttttcatc tatgagggtg tttcctctaa acctacgagg

 781 gaggaacacc tgatcttaca gaaaatacca cctcgagatg ggtgctggtc ctgttgatcc

 841 cagtctctgc cagaccaagg cgagtttccc cactaataaa gtgccgggtg tcagcagaa

*Sequence 2:* Mouse Low Density Lipoprotein Receptor

*Source:* <https://www.ncbi.nlm.nih.gov/nuccore/NM_010700.3>

 1 gcagactcct cccccgcctg gaaacctcgc ccctagtact gggaatgact ctgggcgtgc

 61 ggcgtagttt gcagccggga caccgtgagg cttgcgagcc cagattcgca gccgagacac

 121 cgtggggccc gcgatccagt gtttgcagcg ggaacatttc ggggtctgtg atccgagtga

 181 ggacgcaacg cagaagctaa ggatgagcac cgcggatctg atgcgtcgct gggtcatcgc

 241 cctgctcctg gctgctgccg gagttgcagc agaagactca tgcagcagga acgagttcca

 301 gtgtagagac ggaaaatgca tcgctagcaa gtgggtgtgc gatggcagcc ccgagtgccc

 361 ggatggctcc gatgagtccc cagagacatg catgtctgtc acctgtcagt ccaatcaatt

 421 cagctgtgga ggccgtgtca gccgatgcat tcctgactcc tggagatgtg atggacaggt

 481 agactgtgaa aatgactcag acgaacaagg ctgtcccccc aagacgtgct cccaggatga

 541 cttccgatgc caggatggca agtgcatctc cccgcagttt gtgtgtgatg gagaccgaga

 601 ttgcctagat ggctctgatg aggcccactg ccaggccacc acttgtggcc ccgcccactt

 661 ccgctgcaac tcatccatat gcatccccag tctttgggcc tgcgacgggg atgtcgactg

 721 tgttgacggc tccgatgagt ggccacagaa ctgccagggc cgagacacgg cctccaaagg

 781 cgttagcagc ccctgctcct ccctggagtt ccactgtggt agcagtgagt gtatccatcg

 841 cagctgggtc tgtgacggcg aggcagactg caaggacaag tcagatgagg agcactgcgc

 901 ggtggccacc tgccgacctg atgaattcca gtgtgcagat ggctcctgca ttcacggtag

 961 ccgccagtgt gaccgtgaac atgactgcaa ggacatgagc gacgagctcg gctgcgtcaa

 1021 tgtgacacag tgtgatggcc ccaacaagtt caagtgtcac agtggggagt gcatcagctt

 1081 ggacaaggtg tgcgactccg cccgcgactg ccaggactgg tcggatgagc ccatcaagga

 1141 gtgcaagacc aacgagtgtt tggacaacaa tggtggctgt tcccacatct gcaaggacct

 1201 caagattggc tctgagtgcc tgtgtcccag cggcttccgg ttggtggacc tccacaggtg

 1261 tgaagatatt gacgagtgtc aggagccaga cacctgcagc cagctctgtg tgaacctgga

 1321 aggcagctac aagtgtgagt gccaggccgg cttccacatg gacccacaca ccagggtctg

 1381 caaggctgtg ggctccatag gctatctgct cttcaccaac cgccacgagg tccggaagat

 1441 gaccctggac cgcagcgagt acaccagtct gctccccaac ctgaagaatg tggtggctct

 1501 cgacacggag gtgaccaaca atagaatcta ctggtccgac ctgtcccaaa aaaagatcta

 1561 cagcgccctg atggaccagg cccctaactt gtcctacgac accatcatca gtgaggacct

 1621 gcatgcccct gacgggctgg cggtagactg gatccaccgc aacatctact ggacagattc

 1681 agtcccaggc agcgtatctg tggctgacac caagggcgta aagaggagga cactgttcca

 1741 agaggcaggg tccagaccca gagccatcgt agtggaccct gtgcatggct tcatgtactg

 1801 gacagattgg ggaacacccg ccaagatcaa gaaagggggt ttgaatggtg tggacatcca

 1861 ctcactggtg accgaaaaca tccagtggcc aaatggcatc acactagatc tttccagtgg

 1921 ccgtctctat tgggttgatt ccaaactcca ctctatctcc agcatcgatg tcaatggggg

 1981 caatcggaaa accattttgg aggatgagaa ccggctggcc caccccttct ccttggccat

 2041 ctatgaggac aaagtgtatt ggacagatgt cataaacgaa gccattttca gtgccaatcg

 2101 actcacgggt tcagatgtga atttggtggc tgaaaacctc ttgtccccgg aggacattgt

 2161 cctgttccac aaggtcacac agcctagagg ggtgaactgg tgtgagacaa cagccctcct

 2221 ccccaatggt ggttgccagt acctgtgcct gcccgcccca cagatcggtc cccactcgcc

 2281 caaattcacc tgcgcctgcc ctgatggcat gctgctggcc aaggacatgc ggagctgcct

 2341 cacagaagtc gacactgtac tgaccaccca ggggacatcc gccgtccggc ctgtggtcac

 2401 cgcatcagct accaggccac cgaagcacag tgaggatctc tcagctccca gtactcctag

 2461 gcagcctgtg gacaccccag ggctcagcac agtggcgtca gtgacagtgt cccaccaagt

 2521 ccagggtgac atggctggca gagggaatga ggagcagcca catggtatga ggttcctgtc

 2581 catcttcttc cctattgcac tggttgccct ccttgtcctt ggggccgtcc tgctgtggag

 2641 gaactggcgg ctgaagaaca tcaacagcat aaactttgac aacccagtct accagaagac

 2701 cacagaggac gagctccaca tttgccgaag ccaggatggc tatacctacc cctcaagaca

 2761 gatggtcagc ctggaggacg atgtggcatg agcagccggg agagccgtct ctttccggga

 2821 tccattgcca agcttaggca gaaaagacac tctctccaga cctccccatc cagcactggt

 2881 cctgccacct ccctgggctc tgtgttgctc aaagcaagat aagagcaaag ctgggctggg

 2941 ggccaagctc agctgcctgt ctgccccagg ttctgtttta tatatttatt gtctggggac

 3001 agaaaaggct actggctgtg cttgaaattc gaattctgcc tgggaatttt tttttttctc

 3061 ccatcttcat ttccttggaa gccaaatagg ctgtcccaga agcttctgaa ttctactctt

 3121 ttcaacacta acacggagga ggaagagacc cacccaggca gagagaagcc tgctggacat

 3181 gagacagcag gtactactga ccgggtcctg ggtaccatag ccagtcccta acctcaggac

 3241 ttcctgcatt gaccttcatc ccagagcctt cctgcctgac tctgccgccc ccatcagaag

 3301 ccaggaaagt gactcgtgac attcggtact gatcttgctg gatcttgagg ccggcagctg

 3361 ccagtgatgg atacccgctc atgcctttgc caaaacgtca cccgttcctg gagggacaag

 3421 cccaagtcgc cattctccct taatatttat caagtgcctg agacaactgg ttgccttgct

 3481 caggagtcct ggcctgctca gtgtcctgct gctcaggggt ggccggcagg acccctgtcc

 3541 tgtgcccagc ccccgagtct ccgagtgagg cttgcacttc tctctgcgca cttttcccgt

 3601 tctgtgctgg gcacccacgt ggtccacatt tgtactccta ggttgcactg accctggtgc

 3661 attggatggg acctcaggcc ggaagtggcc ccgaatcatt gaccacagaa gataattgcc

 3721 aggctccaag catccattgc cttcttcagt actgggggga actcagggcc tctgtctggt

 3781 gtttagttag tctgtttgtt tgtttgtttg tttgtttgtt tggtggtttc tggctctgca

 3841 cagcagagtt gtttagctcc tctggtaacg tcatcctctc tcacagagac atccagcctg

 3901 ctgaagggga gagatggggc gaggggaggg atacacgcat ggtttctcag gtgtggtggc

 3961 acttggcttt aatcctatgc tcaagaggaa gaggtgggca catgtctgag gtccaagcct

 4021 acatagtgag accttatata tctaaaaacc aaaagaaatc ccagactatt tggtgggact

 4081 tggaagaata cttagaaatt gccacagatt tgtcacatgg gtaacctaag aagacagtgg

 4141 tggccttact tggattcctg ggtagatcca cctgtccact actcaagaca caagggcaca

 4201 ctgcccagcc cactgggaca cttgtaaaga gctctgtaat gtgccgatgg cacctggctc

 4261 ggttttcatt ctgtatattc aaggatatca cacatatgta ttaaatctat ttatttttgc

 4321 aaaccctgat tgctgcacct ctctgcaatt tctccaggtt aggctggggt gatactctgg

 4381 gatgctctgt atacaaagat gatttgaatg ggctgatctc aggcctgtcc tgtgcgttgc

 4441 tttgagtggg tgggaactgc tttgaaaccc ttgttcagat gtttttatag gctgaaaata

 4501 tcatactgtg atggattaaa ttcttttttg gaaaaaaaaa aaaaaaaaa

*Sequence 3:* Human Superoxide Dismutase 1 (SOD1)

*Source:* <https://www.ncbi.nlm.nih.gov/nuccore/NM_000454.4>

 1 gtttggggcc agagtgggcg aggcgcggag gtctggccta taaagtagtc gcggagacgg

 61 ggtgctggtt tgcgtcgtag tctcctgcag cgtctggggt ttccgttgca gtcctcggaa

 121 ccaggacctc ggcgtggcct agcgagttat ggcgacgaag gccgtgtgcg tgctgaaggg

 181 cgacggccca gtgcagggca tcatcaattt cgagcagaag gaaagtaatg gaccagtgaa

 241 ggtgtgggga agcattaaag gactgactga aggcctgcat ggattccatg ttcatgagtt

 301 tggagataat acagcaggct gtaccagtgc aggtcctcac tttaatcctc tatccagaaa

 361 acacggtggg ccaaaggatg aagagaggca tgttggagac ttgggcaatg tgactgctga

 421 caaagatggt gtggccgatg tgtctattga agattctgtg atctcactct caggagacca

 481 ttgcatcatt ggccgcacac tggtggtcca tgaaaaagca gatgacttgg gcaaaggtgg

 541 aaatgaagaa agtacaaaga caggaaacgc tggaagtcgt ttggcttgtg gtgtaattgg

 601 gatcgcccaa taaacattcc cttggatgta gtctgaggcc ccttaactca tctgttatcc

 661 tgctagctgt agaaatgtat cctgataaac attaaacact gtaatcttaa aagtgtaatt

 721 gtgtgacttt ttcagagttg ctttaaagta cctgtagtga gaaactgatt tatgatcact

 781 tggaagattt gtatagtttt ataaaactca gttaaaatgt ctgtttcaat gacctgtatt

 841 ttgccagact taaatcacag atgggtatta aacttgtcag aatttctttg tcattcaagc

 901 ctgtgaataa aaaccctgta tggcacttat tatgaggcta ttaaaagaat ccaaattcaa

 961 actaaaaaaa aaaaaaaaaa a

*Sequence 4:* Human Iodothyronine Deiodinase 1

*Source:* <https://www.ncbi.nlm.nih.gov/nuccore/NM_000792.6>

 1 agagcttact ctggctttgc cgagatgggg ctgccccagc cagggctgtg gctgaagagg

 61 ctctgggtgc tcttggaggt ggctgtgcat gtggtcgtgg gtaaagtgct tctgatattg

 121 tttccagaca gagtcaagcg gaacatcctg gccatgggcg agaagacggg tatgaccagg

 181 aacccccatt tcagccacga caactggata ccaacctttt tcagcaccca gtatttctgg

 241 ttcgtcttga aggtccgttg gcagcgacta gaggacacga ctgagctagg gggtctggcc

 301 ccaaactgcc cggtggtccg cctctcagga cagaggtgca acatttggga gtttatgcaa

 361 ggtaataggc cactggtgct gaattttgga agttgtacct gaccttcatt tatgttcaaa

 421 tttgaccagt tcaagaggct tattgaagac tttagttcca tagcagattt tcttgtcatt

 481 tacattgaag aagcacatgc atcagatggc tgggctttta agaacaacat ggacatcaga

 541 aatcaccaga accttcagga tcgcctgcag gcagcccatc tactgctggc caggagcccc

 601 cagtgccctg tggtggtgga caccatgcag aaccagagca gccagctcta cgcagcactg

 661 cctgagaggc tctacataat ccaggagggc aggatcctct acaagggtaa atctggccct

 721 tggaactaca acccagagga agttcgtgct gttctggaaa agctccacag ttaatctgga

 781 cagatacctc aattctaggt gaccaacggg agggcttctc aaggcttagc tctccctgag

 841 acccagctgg cttttaccct tgacctgtgt ccctagctga atcactagct cagatttttc

 901 tgatctaagc aaacaactcc cagctgagga atgcaggcca cagcacccaa tcaagacaaa

 961 ttgttattat cagaaaatga agcaacactt gagctgttca ggccagttcc ctgttgaaga

 1021 aacagttccc tgttgaagaa agtagagcct gacactgctc ccactttgga gaccacattc

 1081 cctgcacacg gtctttgaga gagcagttgc actctacagg cacacttctg aggtacggta

 1141 tctctctcca gccactctga taccaagtaa ttcaagctgg cattccttct attagggaaa

 1201 ttcattttac ccaatttgca tttatggaat tgatcattta agacactaaa ttagttttta

 1261 gaaccaatta tgggaagaat tccagttgtt aggaagagat gaggagttgg aagaggaggg

 1321 attagaaaca ggaggaggca gtcatcctct ccttgccaaa agatttaaac ctgtccacat

 1381 tggtggtgat gatgggtgag tttccatggt aacacatccc taattttacc agggaagagg

 1441 agagtactca ctttaccatc tttgaatata tttcatagaa atctagctct ctgtaccctg

 1501 aaatcttcca ctagcctcac ttttcaacag agtcatctag aagggagggt tggcttccca

 1561 aaagcataac cttgaccaaa ccaaacaata ggcaccagca atgctgtcat tcagttatgc

 1621 agaagctcat ttgtgaaatt ctgtttctct gatttcttcg caagtctctt aatggtcatt

 1681 tgtgttagat tacatcaaac tgatggatag ccattggtat tcatctattt taactctgtg

 1741 tctttacata tttgtttatg atggccacag cctaaagtac acacggctgt gacttgattc

 1801 aaaagaaaat gttataagat gcagtaaact aataacagaa ttattaaaat atatcaggct

 1861 aaaaaaaaa

*Sequence 5:* Human Selenoprotein N1 (SEPN1, Reference Sequence)

*Source:* <https://www.ncbi.nlm.nih.gov/nuccore/NM_206926.2>

 1 gcttcccggg ccgccggcag ccgccgccag ccgcagccat gggccgggcc cggccgggcc

 61 aacgcgggcc gcccagcccc ggccccgccg cgcagcctcc cgcgccaccg cgccgccgcg

 121 cccgttccct ggcgctgctc ggagccctgc tggccgccgc cgctgccgcc gccgtccggg

 181 tctgcgcccg ccacgccgag gcccaggcgg ccgcgcggca ggaactggcg ctgaagaccc

 241 tggggacaga tggccttttt ctcttttcct ccttggacac tgacggggat atgtacatca

 301 gccctgagga gttcaaaccc attgctgaga agctaacagg gtcaactccc gcggccagct

 361 gcgaggagga ggagttgccc cctgacccta gcgaggagac gctcaccata gaagcccgat

 421 tccagcctct gctcccggag accatgacca agagcaaaga tggcttccta ggggtctccc

 481 gcctcgccct gtccggcctc cgaaactgga cagccgccgc ctcaccaagt gcagtgtttg

 541 ccacccgcca cttccagccc ttccttcccc cgccaggcca ggagctgggt gagccctggt

 601 ggatcatccc cagtgagctg agcatgttca ctggctacct gtccaacaac cgcttctatc

 661 caccgccgcc caagggcaag gaggtcatca tccaccggct cctgagcatg ttccaccctc

 721 ggccctttgt gaagacccgc tttgcccctc agggagctgt ggcctgcctg actgccatca

 781 gcgacttcta ctacactgtg atgttccgga tccatgccga gttccagctc agtgagccgc

 841 ccgacttccc cttttggttc tcccctgctc agttcaccgg ccacatcatc ctctccaaag

 901 acgccaccca cgtccgcgac ttccggctct tcgtgcccaa ccacaggtct ctgaatgtgg

 961 acatggagtg gctttacggg gccagtgaaa gcagcaacat ggaggtggac atcggctaca

 1021 taccccagat ggagctggag gccacgggcc cctctgtgcc ctccgtgatc ctggatgagg

 1081 atggcagcat gatcgacagc cacctgcctt caggggagcc cctgcagttt gtgtttgagg

 1141 agatcaagtg gcagcaggag ctgagctggg aggaggctgc ccggcgcctg gaggtggcca

 1201 tgtacccctt caagaaggtc tcctacttgc cgttcactga ggccttcgac cgagccaagg

 1261 ctgagaacaa gctggtgcac tcaatcctgc tgtggggggc cctggatgac cagtcctgct

 1321 gaggttcagg gcggactctc cgggagactg tcctggaaag ttcgcccatc ctcaccctgc

 1381 tcaacgagag cttcatcagc acctggtccc tggtgaagga gctggaggaa ctgcagaaca

 1441 accaggagaa ctcgtcccac cagaagctgg ctggcctgca cctggagaag tacagcttcc

 1501 ccgtggagat gatgatctgc ctgcccaatg gcaccgtggt ccatcacatc aatgccaact

 1561 acttcttgga catcacctcc gtgaagcccg aggaaatcga gagcaatctc ttcagcttct

 1621 catccacctt tgaagacccg tccacggcca cctacatgca gttcctgaag gagggactcc

 1681 ggcgtggcct gcccctcctc cagccctaga gtgcctggac gggatctgat gcacaggccc

 1741 ccacgcctca gagccagagt ggtcctcagc ccatttcaga ctgcagatgc cgcccactcc

 1801 caccccactc ctaggctgcc ttggagggta caagatccac tgagggtggc caccacagcc

 1861 ttggctccat ggtggcgggt agacaaggga tgcctgggct gactgggcag aggaacctct

 1921 agctctgact gtcactcggc tctccctacc catttggctc tggaagctgc ttggcccccc

 1981 cagatcaggg cctgggtgaa ctccctggac ctttcctagc cagccgcaca gtctaggccc

 2041 ttgtggggtg aagaatggag ggaggagcag gctaggaaga cggggccacc accctctcct

 2101 tgctttcagc ccttcccaca ggaaacatca agaagcccca gccaggaggg gccaggctgc

 2161 caaggcggct cccctgttta tctagagcct tcgttcctgg ccataccccg gactgccctc

 2221 ctgtgcctga tgtccccagc tggggtcagt ctcaacagga gccagtcttc tggagcctct

 2281 gggcagaacc ctccatcaga gtggaaatca gacgggaccc cctgcagctt ccctgaccac

 2341 gccactgacc agctatctgg ggaagtttac tgtgaagggg tttctgcctt tagcaatggg

 2401 gttcactaag ggggttcccg aggcccaggg ccaaggcact cccaccgcct accttagcac

 2461 agggtctctg caggactgcg ggagccagcg ctcctgccgc ccctcttgcc cctcagacct

 2521 tgcatccaca gaagcacaac ccagccaaac accacagcct tctccagagc cggcactgtc

 2581 ccggcaacca ggggtgcccc aggctagctc ttctacctct ggggcaccac ggactcccct

 2641 tggccactct tgggactttg gtccacgtcc tgagccactg accacggcca gtctctcttt

 2701 ttatatgtgc agaaaagtgt ttttacacaa actttctcat ggtttgtagg tattttttta

 2761 taaccccagt gctgaggaga aaggaggggc agtggcttcc ccggcagcag ccccatgatg

 2821 gctgaatccg aaatcctcga tgggtccagc ttgatgtctt tgcagctgca cctatgggaa

 2881 gaagtagtcc tctcttcctt ctcctcttca gctttttaaa aacagtcctc agaggatcca

 2941 tgatccccag cactgtccca tcctccacaa aggcccacag gcatgcctgt actctctttc

 3001 attaaggtct tgaagtcagg ctgccccctc cccagccccc agttctctcc ccaccccctc

 3061 accccacccg gggctcactc agcctggcag aggaagaagg aaggcagaca tctccgcagc

 3121 cactcctggg ccttttatgt gccgagttac cccacttgcc ttgggcgtgt ccactgagcc

 3181 ttccccagcc agtcttgttc tcaattttgt tttgttttgt tttgagacgg agtcttgctc

 3241 tgtcacccag gctggagtgc tatggctcga tcttggctca ctgcaacctc cacctcccag

 3301 gttcaagcaa ttctcttgcc tcagcctccc gagtagctgg gattacaggt gcatgccacc

 3361 atggctggct aatttttgta tttttagtag agatggggtt tcaccatatt ggtcaggctg

 3421 atctggaact tctgacctca ggtgatccac ctgcctcagc ctcccaaagt gctgggatta

 3481 caggcgtgag caatcgtgcc cagccttgtt cttaattttg tatcatccag tcatcgctaa

 3541 tattacacgc accttctcac ttaatcctca cgacaagcct gtgaggcaga tgctcattgt

 3601 tcccatcttg atgaaacttg agtctcaggg aagtgaagtg acttgcccag ggtcactcag

 3661 gtagagttga gattcaaacc cacatgtggc tccaaagtct gcatctggat ttgggggtgt

 3721 tttttggcat ggcaccctca cctctctccc tgcctgtttt ccccaaagtg gaaaggaagg

 3781 cctttcaaac cagagtgtct cactcccctc tgacctccag accagatggg gcatgagcca

 3841 gccagctcag ccaggctccc tgtgtcctgg gaggaagtgt ccccatcccc catgcccctt

 3901 atggggaggg agggcgtctg atgctctctc tctgcctccc cccccatcct gtcaggcaca

 3961 ggtgacgggg gcagcccatg cgagcccttc tcctgctgct ctgggagggc cagttccaca

 4021 ttgagccagc ctggtcccat ggaaaatgat ggcctgggct ttctgaggcc ttatctgatg

 4081 cctctgcagt tcatgtcccc caccaggcct cgaggctcag ggtgggagag ggccccgggc

 4141 tgccctgtca ctcctctaac acttccctcc cctgtcccca acatgccctg taataaaatt

 4201 agagaagact aa

*Sequence 6:* Human Selenoprotein N1 (Mutated SEPN1 Sequence from Maya, Missing Nucleotides 2825–2828 (aatc).

*No link to National Center for Biotechnology Information (NCBI) is available. This sequence is nearly identical to Sequence 5, but is missing nucleotides 2825-2828 and is replaced with spaces and highlighted in yellow to indicate this deletion.*

 1 gcttcccggg ccgccggcag ccgccgccag ccgcagccat gggccgggcc cggccgggcc

 61 aacgcgggcc gcccagcccc ggccccgccg cgcagcctcc cgcgccaccg cgccgccgcg

 121 cccgttccct ggcgctgctc ggagccctgc tggccgccgc cgctgccgcc gccgtccggg

 181 tctgcgcccg ccacgccgag gcccaggcgg ccgcgcggca ggaactggcg ctgaagaccc

 241 tggggacaga tggccttttt ctcttttcct ccttggacac tgacggggat atgtacatca

 301 gccctgagga gttcaaaccc attgctgaga agctaacagg gtcaactccc gcggccagct

 361 gcgaggagga ggagttgccc cctgacccta gcgaggagac gctcaccata gaagcccgat

 421 tccagcctct gctcccggag accatgacca agagcaaaga tggcttccta ggggtctccc

 481 gcctcgccct gtccggcctc cgaaactgga cagccgccgc ctcaccaagt gcagtgtttg

 541 ccacccgcca cttccagccc ttccttcccc cgccaggcca ggagctgggt gagccctggt

 601 ggatcatccc cagtgagctg agcatgttca ctggctacct gtccaacaac cgcttctatc

 661 caccgccgcc caagggcaag gaggtcatca tccaccggct cctgagcatg ttccaccctc

 721 ggccctttgt gaagacccgc tttgcccctc agggagctgt ggcctgcctg actgccatca

 781 gcgacttcta ctacactgtg atgttccgga tccatgccga gttccagctc agtgagccgc

 841 ccgacttccc cttttggttc tcccctgctc agttcaccgg ccacatcatc ctctccaaag

 901 acgccaccca cgtccgcgac ttccggctct tcgtgcccaa ccacaggtct ctgaatgtgg

 961 acatggagtg gctttacggg gccagtgaaa gcagcaacat ggaggtggac atcggctaca

 1021 taccccagat ggagctggag gccacgggcc cctctgtgcc ctccgtgatc ctggatgagg

 1081 atggcagcat gatcgacagc cacctgcctt caggggagcc cctgcagttt gtgtttgagg

 1141 agatcaagtg gcagcaggag ctgagctggg aggaggctgc ccggcgcctg gaggtggcca

 1201 tgtacccctt caagaaggtc tcctacttgc cgttcactga ggccttcgac cgagccaagg

 1261 ctgagaacaa gctggtgcac tcaatcctgc tgtggggggc cctggatgac cagtcctgct

 1321 gaggttcagg gcggactctc cgggagactg tcctggaaag ttcgcccatc ctcaccctgc

 1381 tcaacgagag cttcatcagc acctggtccc tggtgaagga gctggaggaa ctgcagaaca

 1441 accaggagaa ctcgtcccac cagaagctgg ctggcctgca cctggagaag tacagcttcc

 1501 ccgtggagat gatgatctgc ctgcccaatg gcaccgtggt ccatcacatc aatgccaact

 1561 acttcttgga catcacctcc gtgaagcccg aggaaatcga gagcaatctc ttcagcttct

 1621 catccacctt tgaagacccg tccacggcca cctacatgca gttcctgaag gagggactcc

 1681 ggcgtggcct gcccctcctc cagccctaga gtgcctggac gggatctgat gcacaggccc

 1741 ccacgcctca gagccagagt ggtcctcagc ccatttcaga ctgcagatgc cgcccactcc

 1801 caccccactc ctaggctgcc ttggagggta caagatccac tgagggtggc caccacagcc

 1861 ttggctccat ggtggcgggt agacaaggga tgcctgggct gactgggcag aggaacctct

 1921 agctctgact gtcactcggc tctccctacc catttggctc tggaagctgc ttggcccccc

 1981 cagatcaggg cctgggtgaa ctccctggac ctttcctagc cagccgcaca gtctaggccc

 2041 ttgtggggtg aagaatggag ggaggagcag gctaggaaga cggggccacc accctctcct

 2101 tgctttcagc ccttcccaca ggaaacatca agaagcccca gccaggaggg gccaggctgc

 2161 caaggcggct cccctgttta tctagagcct tcgttcctgg ccataccccg gactgccctc

 2221 ctgtgcctga tgtccccagc tggggtcagt ctcaacagga gccagtcttc tggagcctct

 2281 gggcagaacc ctccatcaga gtggaaatca gacgggaccc cctgcagctt ccctgaccac

 2341 gccactgacc agctatctgg ggaagtttac tgtgaagggg tttctgcctt tagcaatggg

 2401 gttcactaag ggggttcccg aggcccaggg ccaaggcact cccaccgcct accttagcac

 2461 agggtctctg caggactgcg ggagccagcg ctcctgccgc ccctcttgcc cctcagacct

 2521 tgcatccaca gaagcacaac ccagccaaac accacagcct tctccagagc cggcactgtc

 2581 ccggcaacca ggggtgcccc aggctagctc ttctacctct ggggcaccac ggactcccct

 2641 tggccactct tgggactttg gtccacgtcc tgagccactg accacggcca gtctctcttt

 2701 ttatatgtgc agaaaagtgt ttttacacaa actttctcat ggtttgtagg tattttttta

 2761 taaccccagt gctgaggaga aaggaggggc agtggcttcc ccggcagcag ccccatgatg

 2821 gctg cg aaatcctcga tgggtccagc ttgatgtctt tgcagctgca cctatgggaa

 2881 gaagtagtcc tctcttcctt ctcctcttca gctttttaaa aacagtcctc agaggatcca

 2941 tgatccccag cactgtccca tcctccacaa aggcccacag gcatgcctgt actctctttc

 3001 attaaggtct tgaagtcagg ctgccccctc cccagccccc agttctctcc ccaccccctc

 3061 accccacccg gggctcactc agcctggcag aggaagaagg aaggcagaca tctccgcagc

 3121 cactcctggg ccttttatgt gccgagttac cccacttgcc ttgggcgtgt ccactgagcc

 3181 ttccccagcc agtcttgttc tcaattttgt tttgttttgt tttgagacgg agtcttgctc

 3241 tgtcacccag gctggagtgc tatggctcga tcttggctca ctgcaacctc cacctcccag

 3301 gttcaagcaa ttctcttgcc tcagcctccc gagtagctgg gattacaggt gcatgccacc

 3361 atggctggct aatttttgta tttttagtag agatggggtt tcaccatatt ggtcaggctg

 3421 atctggaact tctgacctca ggtgatccac ctgcctcagc ctcccaaagt gctgggatta

 3481 caggcgtgag caatcgtgcc cagccttgtt cttaattttg tatcatccag tcatcgctaa

 3541 tattacacgc accttctcac ttaatcctca cgacaagcct gtgaggcaga tgctcattgt

 3601 tcccatcttg atgaaacttg agtctcaggg aagtgaagtg acttgcccag ggtcactcag

 3661 gtagagttga gattcaaacc cacatgtggc tccaaagtct gcatctggat ttgggggtgt

 3721 tttttggcat ggcaccctca cctctctccc tgcctgtttt ccccaaagtg gaaaggaagg

 3781 cctttcaaac cagagtgtct cactcccctc tgacctccag accagatggg gcatgagcca

 3841 gccagctcag ccaggctccc tgtgtcctgg gaggaagtgt ccccatcccc catgcccctt

 3901 atggggaggg agggcgtctg atgctctctc tctgcctccc cccccatcct gtcaggcaca

 3961 ggtgacgggg gcagcccatg cgagcccttc tcctgctgct ctgggagggc cagttccaca

 4021 ttgagccagc ctggtcccat ggaaaatgat ggcctgggct ttctgaggcc ttatctgatg

 4081 cctctgcagt tcatgtcccc caccaggcct cgaggctcag ggtgggagag ggccccgggc

 4141 tgccctgtca ctcctctaac acttccctcc cctgtcccca acatgccctg taataaaatt

 4201 agagaagact aa