

## pEMBL Sequence

AATTCCCATCATCAATAATATACCTTATTTTTGGATTGAAGCCAATATGATAATGAGGGGGTGGAGTTTGTGACGTGGC  
GCGGGGCGTGGGAACGGGGCGGGTACGTAGTAGTCTCTAGAGTCTGTATTAGAGGTCACGTGAGTGTGTTTTCGCA  
CATTTTTGCGACACCATGTGGTACGCTGGGTATTTAAGCCGAGTGAGCACGCAGGGTCTCCATTTTGAAGCGGGAG  
GTTTGAACGCGCAGCCGCCATGCCGGGGTTTTACGAGATTGTGATTAAGGTCCCAGCGACCTTGACGGGCATCTG  
CCCGGCATTTCTGACAGCTTTGTGAACTGGGTGGCCGAGAAGGAATGGGAGTTGCCGCCAGATTCTGACATGGATC  
TGAATCTGATTGAGCAGGCACCCCTGACCGTGGCCGAGAAGCTGCAGCGCGACTTTCTGACGGAATGGCGCCGTGT  
GAGTAAGGCCCGGAGGCCCTTTTTCTTTGTGCAATTTGAGAAGGGAGAGAGCTACTTCCACATGCACGTGCTCGTGG  
AAACCACCGGGGTAAATCCATGGTTTTGGGACGTTTCTGAGTCAGATTCGCGAAAACTGATTGAGAGAATTTACC  
GCGGGATCGAGCCGACTTTGCCAACTGGTTCGCGGTACAAAAGACCAGAAATGGCGCCGGAGGCCGGAACAAGG  
TGGTGGATGAGTGCTACATCCCAATTACTTGTCCCAAAACCCAGCCTGAGCTCCAGTGGGCGTGGACTAATATG  
GAACAGTATTTAAGCGCCTGTTTGAATCTCACGGAGCGTAAACGGTTGGTGGCGCAGCATCTGACGCACGTGTGCGA  
GACGCAGGAGCAGAACAAGAGAATCAGAATCCCAATTCTGATGCGCCGGTATCAGATCAAAAACCTTACGCCAGGT  
ACATGGAGCTGGTCGGGTGGCTCGTGGACAAGGGGATTACCTCGGAGAAGCAGTGGATCCAGGAGGACCAGGCCT  
CATACATCTCCTTCAATGCGGCCTCCAACCTCGCGGTCCAAATCAAGGCTGCCTTGGACAATGCGGGAAGATTATG  
AGCCTGACTAAAACCGCCCCGACTACCTGGTGGGCCAGCAGCCCGTGGAGGACATTTCCAGCAATCGGATTTATA  
AAATTTTGAACATAACGGGTACGATCCCAATATGCGGCTTCCGTCTTTCTGGGATGGGCCACGAAAAAGTTCGGC  
AAGAGGAACACCATCTGGCTGTTTGGGCCTGCAACTACGGGAAGACCAACATCGCGGAGGCCATAGCCCACACTG  
TGCCCTTCTACGGGTGCGTAACTGGACCAATGAGAATTTCCCTTCAACGACTGTGTGACAAGATGGTGTATCTGG  
TGGGAGGAGGGGAAGATGACCGCCAAGGTGCTGGAGTCGGCCAAAGCCATTCTCGGAGGAAGCAAGGTGCGCGT  
GACCAGAAATGCAAGTCTCGGCCAGATAGACCCGACTCCCGTATCGTACCTCCAACACCAACATGTGCGCCG  
TGATTGACGGAACTCAACGACCTTGAACACCAGCAGCCGTTGCAAGACCGGATGTTCAATTTGAACTCACCCG  
CGTCTGGATCATGACTTTGGGAAGGTCACCAAGCAGGAAGTCAAAGACTTTTTCCGGTGGCAAGGATCAGT  
GGTTGAGGTGAGCATGAATCTACGTCAAAAAGGGTGGAGCCAAAGAAAAGACCCGCCAGTGGACGATATA  
AGTGAGCCCAAACGGGTGCGGAGTCAGTTGCGCAGCCATCGACGTGACGCGGAAAGCTTTCGATCAACTACGCAG  
ACAGGTACCAAACAATGTTCTCGTCACGTGGGCATGAATCTGATGCTGTTTCCCTGCAGACAATGCGAGAGAATG  
AATCAGAATTCAAATATCTGCTTCACTCACGGACAGAAAGACTGTTTAGAGTGCTTTCCCGTGTGAGAATCTCAACCC  
GTTTCTGTGCTCAAAAAGGCGTATCAGAACTGTGCTACATTCATCATATCATGGGAAAGGTGCCAGACGCTTGCCT  
GCCTGCGATCTGGTCAATGTGGATTTGGATGACTGCATCTTTGAACAATAAATGATTTAAATCAGGTATGGCTGCCGA  
TGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAAGGAATAAGACAGTGGTGGAAAGCTCAAACCTGGCCAC  
CACCACAAAGCCCGCAGAGCGGCATAAGGACGACAGCAGGGTCTTGTGCTTCCCTGGGTACAAGTACCTCGGACC  
CTTCAACGGACTCGACAAGGGAGAGCCGGTCAACGAGGCAGACGCCGCGGCCCTCGAGCACGTACAAAGCCTACG  
ACCGGCAGCTCGACAGCGGAGACAACCCGTACCTCAAGTACAACCACGCCGACGCGGAGTTTCAGGAGCGCCTTAA  
AGAAGATACGTCTTTTGGGGCAACCTCGGACGAGCAGTCTTCCAGGCGAAAAGAGGGTCTTGAACCTCTGGGC  
CTGGTTGAGGAACCTGTTAAGACGGCTCCGGGAAAAAAGAGGCCGGTAGAGCACTCTCCTGTGGAGCCAGACTCCT  
CCTCGGGAACCGGAAAGCGGGCCAGCAGCCTGCAAGAAAAGATTGAATTTTGGTTCAGACTGGAGACGCAGACTC  
AGTACCTGACCCCGAGCCTCTCGGACAGCCACCAGCAGCCCTCTGGTCTGGGAATAATACGATGGCTACAGGC  
AGTGGCGCACCAATGGCAGACAATAACGAGGGCGCCGACGGAGTGGGTAATTCCTCCGGAATTTGGCATTGCGATT  
CCACATGGATGGGCGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTGCCACCTACAACAACCACCTCTA  
CAAACAATTTCCAGCCAATCAGGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGGGGGTATTTTGA  
CTTCAACAGATTCCACTGCCACTTTTACCACGTGACTGGCAAAGACTCATCAACAACAACCTGGGGATTCCGACCCAA  
GAGACTCAACTTCAAGCTCTTTAACATTCAAGTCAAAGAGGTCACGCAGAATGACGGTACGACGACGATTGCCAATAA  
CCTTACCAGCACGGTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCTCCTCGGCTCGGCGCATCAAGGAT  
GCCTCCCGCCGTTCCAGCAGACGTCTTATGGTGGCACAGTATGGATACCTCACCTGAACAACGGGAGTCAGGC  
AGTAGGACGCTCTTCAATTTACTGCCTGGAGTACTTTCTTCTCAGATGCTGCGTACCGGAAACAACCTTACCTCAG  
CTACACTTTTGGAGACGTTCCCTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCAT  
CGACCAGTACCTGTATTACTTGAGCAGAACAACACTCCAAGTGAACCACCACGACGTCAGTCAAGGCTTTCAGTTTTCTCA  
GGCCGGAGCGAGTGACATTCGGGACCAGTCTAGGAAGTGGCTTCTGACCCTGTTACCAGCAGCAGGATATCA  
AAGACATCTGCGGATAACAACAACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTC  
TCTGGTGAATCCGGCCATGGCAAGCCACAAGGACGATGAAGAAAAGTTTTTCTCAGAGCGGGTCTCATCTTTG  
GGAAGCAAGGCTCAGAGAAAACAATGTGAACATTGAAAAGTTCATGATTACAGACGAAGAGGAAATCGGAACAACC  
AATCCCGTGGCTACGGAGCAGTATGGTCTGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGA  
TGTAACACACAAGGCGTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTACGGGGCCCATCTGGGCA  
AAGATTCCACACACGGACGGACATTTTACCCCTCTCCCTCATGGGTGGATTTCGGACTTAAACACCCCTCCTCCACA  
GATTCTCATCAAGAACACCCCGTACCTGCGAATCCTTCGACCACCTTTCAGTGCAGGCAAGTTTGTCTTCTCATCAC  
ACAGTACTCCACGGGACACGGTACGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAAACAGCAAACGCTGGAATCC  
CGAAATTCAGTACACTTCCAACATAACAAGTCTGTTAATCGTGGACTTACCGTGGATACTAATGGCGTGTATTTCAGA  
GCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTGTAATTGCTTGTAAATCAATAAACCCTTAAATTCGTTTCA  
GTTGAACTTTGGTCTCTGCGTATTTCTTTCTTATCTAGTTTCCATGCTCTAGACTACTACGTCACCCGCCCGTTCCCA

CGCCCCGCGCCACGTACAAACTCCACCCCTCATTATCATATTGGCTTCAATCCAAAATAAGGTATATTATTGATGA  
TGCATCGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATG  
GAATCCAGACGATTGAGCGTCAAATGTAGGTATTTCCATGAGCGTTTTTCTGTTGCAATGGCTGGCGGTAATATT  
GTTCTGGATATTACCAGCAAGGCCGATAGTTTGAGTTCTTCTACTCAGGCAAGTGATGTTACTACTAATCAAAGAAGTA  
TTGCGACAACGGTTAATTTGCGTGATGGACAGACTCTTTTACTCGGTGGCCTCACTGATTATAAAAAACTTCTCAGG  
ATTCTGGCGTACCGTTCCTGTCTAAAATCCCTTTAATCGGCCTCCTGTTTAGCTCCCGCTCTGATTCTAACGAGGAAA  
GCACGTTATACGTGCTCGTCAAAGCAACCATAGTACGCGCCCTGTAGCGGCGCATTAAAGCGCGGGCGGGTGTGGTGG  
TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTTCTTCCCTTCTTTCTCGCC  
ACGTTTCGCCGGCTTTCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCT  
CGACCCCAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTTCGCCCTTTGA  
CGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCAAAATGGAACAACACTCAACCCTATCTCGGTCTATTCTTT  
TGATTTATAAGGGATTTTGGCGATTTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTA  
ACAAAATATTAACGTTTACAATTTAAATATTTGCTTATAAATCTTCTGTTTTTGGGGCTTTTCTGATTATCAACCGGG  
GTACATATGATTGACATGCTAGTTTTACGATTACCGTTCATCGATTCTCTTGTGTTGCTCCAGACTCTCAGGCAATGACC  
TGATAGCCTTTGTAGAGACCTCTCAAAAATAGCTACCCTCTCCGGCATGAATTTATCAGCTAGAACGGTTGAATATCAT  
ATTGATGGTGATTTGACTGTCTCCGGCCTTTCTCACCCGTTTGAATCTTTACCTACACATTACTAGGCATTGCAATTTA  
AAATATATGAGGGTTCTAAAAATTTTTATCCTTGC GTTGAATAAAGGCTTCTCCCGCAAAAGTATTACAGGGTCATAA  
TGTTTTTGGTACAACCGATTTAGCTTTATGCTCTGAGGCTTTATTGCTTAATTTTGTAAATTTTGCCTTGCCTGTATG  
ATTTATTGGATGTTGGAATTCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTTACACCCGCATATGGTGA  
CTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCGACACCCGCCAACACCCCGCTGACGCGCCCTG  
ACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTT  
CACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAA  
TGGTTTTCTTAGACGTCAGGTGGCACTTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTC  
AAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAA  
CATTTCCGTGTCGCCCTTATCCCTTTTTTTCGGGCATTTTGCCTTCTGTTTTTGTCTCACCCAGAAACGCTGGTGAAG  
TAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAG  
AGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTA  
TTGACGCCGGGCAAGAGCAACTCGGTGCGCCGATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCGATCACA  
GAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGCGGC  
CAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCT  
GCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCAGATGCCTGTAGCAAT  
GGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGA  
GGCGGATAAAGTTGCAAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAATCTGGAGCCG  
GTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACAC  
GACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGT  
AACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACCTTCATTTTTAAATTTAAAAGGATCTAGGTGAAG  
ATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACTGAGTTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGA  
TCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGC AAACAAAAAAACCACCGCTACCAGCGG  
TGGTTTTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATA  
CTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAA  
TCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACC GGTTGGACTCAAGACGATAGTTACC GGAT  
AAGGCGCAGCGTCCGGCTGAACGGGGGTTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCGAACTG  
AGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCG  
GCAGGGTCCGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCCGGTT  
TCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAAC  
GCGGCCTTTTTACGGTTCCTGGCCTTTTTGCTGGCCTTTTTGCTCACATGTTCTTTCTGCGTTATCCCCTGATTCTGTG  
GATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGACCGGAACGACCGAGCGCAGCGAGTCAGTGA  
GCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCGCGCGTTGGCCGAT TCATTAATGCAG