

pCD33-CH Vector

Source	Constructed by Farrell MacKenzie
Company	Structural Genomics Consortium, Toronto
Description	pCD33-CH is an episomal mammalian expression vector derived from the pCEP4 vector (Invitrogen). It has a cytomegalovirus (CMV) early enhancer/promoter that drives expression of recombinant proteins with the addition of an N-terminal <i>Homo sapiens</i> CD33 signal peptide, which targets the protein for secretion, and a C-terminal 6xHis-tag.
Antibiotic resistance	Ampicillin (resistance in <i>E. coli</i>) Hygromycin B (resistance in mammalian cells)
Promoter	CMV
Cloning Methods	Insertion of a DNA sequence into the cloning/expression region is performed using Clontech's In-fusion enzyme-mediated directional recombination between complementary 15 nucleotide DNA sequences at the ends of the insert (PCR product) and vector linearized with KpnI & SphI. Insertion of a target sequence involves replacement of a SacB gene stuffer sequence, which provides for negative selection of the original plasmid on 5% sucrose.
N – terminal fusion sequence	MPLLLLLLPLLWAG
C – terminal fusion sequence	HHHHHH
5' primer tail for amplification of insert	5' CTTCTATGGGCAGGG --- 3'
3' primer tail for amplification of insert	5' ATGATGATGGTGATG --- 3'
5' sequencing primer pCEP4-fwd	5' AGCAGAGCTCGTTTAGTGAACCG 3'
3' sequencing primer EBV-rev	5' GTGGTTTGTCCAAACTCATC 3'

pCD33-CH sequence (12,191 bp):

GTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCC
ATATATGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAAC
GACCCCGGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTT
CCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGT
ATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTAT
GCCCAGTACATGACCTTACGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGC
TATTACCATGGTGTATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCAC
GGGATTTCCAAGTCTCACCCCACTTACGTCAATGGGAGTTTGTGGTGGCACCAAAATCAA
CGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTG
TACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTACTGTAACCGTCAGATCTCTAGAAGCTGG
ACCATGCCGCTGCTGCTACTGCTGCCCTTCTATGGGCGAGGGTACCCTGAAAGATCCATA
ACTTCGTATAGCATATTATACGAAGTTATGCGGCGCGACGTCCACATATACCTGCCGTT
CACTATTATTTAGTAAATGAGATATTATGATTTTTCTGAATTGTGATTAAGGCAACTTT
ATGCCCATGCAACAGAACTATAAAAAATACAGAGAATGAAAAGAAACAGATAGATTTTTAG
TTCTTTAGGCCGTAAGTCTGCAAATCCTTTTTATGATTTTCTATCAAACAAAAGAGGAAAATAG
ACCAGTTGCAATCCAAACGAGAGTCTAATAGAATGAGGTGAAAAGTAAATCGCGCGGGTTT
GTTACTGATAAAGCAGGCAAGACCTAAAATGTGTAAGGGCAAAGTGTATACTTTGGCGTCA
CCCCTTACATATTTTAGGTCTTTTTTTATTGTGCGTAACTAACTTGCCATCTTCAAACAGGAGG
GCTGGAAGAAGCAGACCGCTAACACAGTACATAAAAAAGGAGACATGAACGATGAACATCAA
AAAGTTTGCAAAACAAGCAACAGTATTAACCTTTACTACCGCACTGCTGGCAGGAGGCGCAA
CTCAAGCGTTTGCAAAAGAAACGAACCAAAAGCCATATAAGGAAACATACGGCATTTCAT
ATTACACGCCATGATATGCTGCAAATCCCTGAACAGCAAAAAAATGAAAAATATAAAGTTCCT
GAGTTCGATTCGTCCACAATTAATAATATCTCTTCTGAAAAGGCCTGGACGTTTGGGACAG
CTGGCCATTACAAAACACTGACGGCACTGTCGCAAATATCACGGCTACCACATCGTCTTTG
CATTAGCCGAGATCCTAAAATGCGGATGACACATCGATTTACATGTTCTATCAAAAAGTCG
GCGAAACTTCTATTGACAGCTGGAAAACGCTGGCCGCGTCTTAAAGACAGCGACAAATTC
GATGCAAATGATTCTATCCTAAAAGACCAAACACAAGAATGGTCAGGTTACAGCCACATTTACA
TCTGACGGAAAATCCGTTTATTCTACACTGATTTCTCCGGTAAACATTACGGCAAACAAACA
CTGACAACCTGCACAAGTAAACGTATCAGCATCAGACAGCTCTTTGAACATCAACGGTGTAGA
GGATTATAAATCAATCTTTGACGGTGACGGAAAACGTATCAAATGTACAGCAGTTCATCGA
TGAAGGCAACTACAGCTCAGGCGACAACCATACGCTGAGAGATCCTCACTACGTAGAAGAT
AAAGGCCACAATACTTAGTATTTGAAGCAAACACTGGAAGTGAAGATGGCTACCAAGGCGA
AGAATCTTTATTTAACAAAGCATACTATGGCAAAGCACATCATTCTTCCGTCAAGAAAGTCA
AAAACCTTCTGCAAAGCGATAAAAAACGCACGGCTGAGTTAGCAAACGGCGCTCTCGGTATGA
TTGAGCTAAACGATGATTACACACTGAAAAAAGTATGAAACCGCTGATTGCATCTAACACA
GTAACAGATGAAATTGAACGCGCGAAGCTCTTTAAAATGAACGGCAAATGGTACTCTGTTCC
TGACTCCCGCGGATCAAAAATGACGATTGACGGCACTTACGTCTAACGATTTTACATCTTGG
GTTATGTTTTCTAATTTCTTAACTGGCCCATACAAGCCGCTGAACAAAACCTGGCCTTGTGTTAA
AAATGGATCTTGATCCTAACGATGTAACCTTTACTTACTCACACTTCGCTGTACCTCAAGCGA
AAGGAAACAATGTCGTGATTACAAGCTATATGACAAACAGAGGATTCTACGCAGACAAACAA
TCAACGTTTGCGCCTAGCTTCTGCTGAACATCAAAGGCAAGAAAACATCTGTTGTCAAAGA
CAGCATCCTTGAACAAGGACAATTAACAGTTAACAAATAAAAACGCAAAAGAAAATGCCGATA
TCCTATTGGCATTGACGTCAGGTGGCACTTTTCGCATGCATCACCATCATCACTAAGATC
CAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAGTAAATGCAGTGAAAAAAAT
GCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACA
AGTTAACAAACAATTGCATTCATTTTATGTTTCAGGTTACGGGGGAGGTGTGGGAGGTTTT
TTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGCTGATTATGATCCGGCTGCCTCGCGG
TTTCGGTGTGACGGTAAAACCTCTGACACATGCAGCTCCCGGAGACGGTACAGCTTGT
CTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGCGTCAGCGGGTGTGGCGGGTGT
CGGGGCGCAGCCATGAGGTGACTCTAGAGGATCGATGCCCCGCCCGGACGAACTAAC
CTGACTACGACATCTCTGCCCTTCTTCGCGGGGAGTGCATGTAATCCCTTACAGTTGGTTG
GTACAACCTGCCAACTGGGCCCTGTTCCACATGTGACACGGGGGGGACCAAAACAAAAGG
GGTCTCTGACTGTAGTTGACATCCTTATAAATGGATGTGCACATTTGCCAACTGAGTGG

CTTTCATCCTGGAGCAGACTTTGCAGTCTGTGGACTGCAACACAACATTGCCTTTATGTGTAA
CTCTTGCTGAAGCTCTTACACCAATGCTGGGGGACATGTACCTCCCAGGGGCCAGGAAG
ACTACGGGAGGCTACACCAACGTCAATCAGAGGGGCCTGTGTAGCTACCGATAAGCGGACC
CTCAAGAGGGCATTAGCAATAGTGTATAAGGCCCCCTTGTTAACCCCTAAACGGGTAGCAT
ATGCTTCCCGGGTAGTAGTATATACTATCCAGACTAACCCTAATTCAATAGCATATGTTACCC
AACGGGAAGCATATGCTATCGAATTAGGGTTAGTAAAAGGGTCCTAAGGAACAGCGATATCT
CCCACCCCATGAGCTGTCACGGTTTTATTTACATGGGGTCAGGATTCCACGAGGGTAGTGAA
CCATTTTAGTCAACAAGGGCAGTGGCTGAAGATCAAGGAGCGGGCAGTGAACCTCCTGAAT
CTTCGCCTGCTTCTTCATTCTCCTTCGTTTAGCTAATAGAATAACTGCTGAGTTGTGAACAGT
AAGGTGTATGTGAGGTGCTCGAAAACAAGGTTTCAGGTGACGCCCCAGAATAAAATTTGGA
CGGGGGGTTTCAGTGGTGGCATTGTGCTATGACACCAATATAACCCTCACAAACCCCTTGGG
CAATAAATACTAGTGTAGGAATGAAACATTCTGAATATCTTTAACAATAGAAATCCATGGGGT
GGGACAAAGCCGTAAAGACTGGATGCCATCTCACACGAATTTATGGCTATGGGCAACACAT
AATCCTAGTGCAATATGATACTGGGTTATTAAGATGTGTCCAGGCAGGGACCAAGACAGG
TGAACCATGTTGTTACACTCTATTTGTAACAAGGGGAAAGAGAGTGGACGCCGACAGCAGC
GGACTCCACTGGTTGTCTCTAACACCCCCGAAAATTAACGGGGTCCACGCCAATGGGGC
CCATAAACAAAGACAAGTGGCCACTCTTTTTTTTGAATTGTGGAGTGGGGGCACGCGTCAG
CCCCACACGCCGCCCTGCGGTTTTGGACTGTAAAATAAGGGTGAATAACTTGGCTGATTG
TAACCCCGCTAACCCTGCGGTCAAACACTTGCCACAAAACCACTAATGGCACCCCGGG
GAATACCTGCATAAGTAGGTGGGCGGGCCAAGATAGGGGCGCGATTGCTGCGATCTGGAG
GACAAATTACACACACTTGCGCCTGAGCGCCAAGCACAGGGTTGTTGGTCTCATATTCAG
AGGTCGCTGAGAGCACGGTGGGCTAATGTTGCCATGGGTAGCATATACTACCCAAATATCT
GGATAGCATATGCTATCCTAATCTATATCTGGGTAGCATAGGCTATCCTAATCTATATCTGGG
TAGCATATGCTATCCTAATCTATATCTGGGTAGTATATGCTATCCTAATTTATATCTGGGTAGC
ATAGGCTATCCTAATCTATATCTGGGTAGCATATGCTATCCTAATCTATATCTGGGTAGTATAT
GCTATCCTAATCTGTATCCGGGTAGCATATGCTATCCTAATAGAGATTAGGGTAGTATATGCT
ATCCTAATTTATATCTGGGTAGCATATACTACCCAAATATCTGGATAGCATATGCTATCCTAAT
CTATATCTGGGTAGCATATGCTATCCTAATCTATATCTGGGTAGCATAGGCTATCCTAATCTA
TATCTGGGTAGCATATGCTATCCTAATCTATATCTGGGTAGTATATGCTATCCTAATTTATATC
TGGGTAGCATAGGCTATCCTAATCTATATCTGGGTAGCATATGCTATCCTAATCTATATCTGG
GTAGTATATGCTATCCTAATCTGTATCCGGGTAGCATATGCTATCCTCATGCATATACAGTCA
GCATATGATACCCAGTAGTAGAGTGGGAGTGTATCCTTTGCATATGCCGCCACCTCCCAAG
GGGGCGTGAATTTTCGCTGCTTGTCTTTTTCTGCTGGTTGCTCCCATTCTTAGGTGAATTTA
AGGAGGCCAGGCTAAAGCCGTGCGATGTCTGATTGCTCACCAGGTAATGTGCGTAATGTTT
TCCAACGCGAGAAGGTGTTGAGCGCGGAGCTGAGTGACGTGACAACATGGGTATGCCCAAT
TGCCCATGTTGGGAGGACGAAAATGGTGACAAGACAGATGGCCAGAAAACACCAACAGC
ACGCATGATGTCTACTGGGGATTTATCTTTAGTGCGGGGGAATACACGGCTTTTAATACGA
TTGAGGGCGTCTCCTAACAAGTTACATCACTCCTGCCCTTCCCTCACCCCTCATCTCCATCACC
TCCTTCTCCTCGTCTCCTCCGTATCACCCTCCGCGCAGCCCTTCCACCATAGGTGGAA
ACCAGGGAGGCCAAATCTACTCCATCGTCAAAGCTGCACACAGTCACCCTGATATTGCAGGTA
GGAGCGGGCTTTGTATAACAAGGTCTTAATCGCATCCTTCAAACCTCAGCAAATATATG
AGTTTGTAAAAAGACCATGAAATAACAGACAATGGACTCCCTTAGCGGGCCAGGTTGTGGGC
CGGGTCCAGGGGCCATTCCAAAGGGGAGACGACTCAATGGTGTAAAGACGACATTGTGGAAT
AGCAAGGGCAGTTCCTCGCCTTAGGTTGTAAAGGGAGGTCTTACTACCTCCATATACGAACA
CACGGGCGACCCAAGTTCCTTCGTGCGTAGTCTTTCTACGTGACTCCTAGCCAGGAGAGC
TCTTAAACCTTCTGCAATGTTCTCAAATTTGGGTTGGAACCTCCTTGACCACGATGCTTTCC
AAACCACCCTCCTTTTTTTCGCTGCTCCATCACCCTGACCCCGGGGTCCAGTGTGGG
CCTTCTCCTGGGTCTGCGGGGCCCTGCTCTATCGCTCCCGGGGGCACGTCAGGCTCA
CCATCTGGGCCACCTTCTGGTGGTATTCAAATAATCGGCTTCCCCTACAGGGTGGAAAAA
TGGCCTTCTACCTGGAGGGGGCCTGCGCGGTGGAGACCCGGATGATGACTGACTACT
GGGACTCCTGGGCCTCTTTTCTCCACGTCCACGACCTCTCCCCCTGGCTCTTTCACGACTTC
CCCCCTGGCTCTTTCACGTCTCTACCCCGGCGGCCTCCACTACCTCCTCGACCCCGGCC
TCCACTACCTCCTCGACCCCGGCCTCCACTGCCTCCTCGACCCCGGCCTCCACCTCCTGCT
CCTGCCCTCCTGCTCCTGCCCTCCTCCTGCTCCTGCCCTCCTGCCCTCCTGCTCCTG
CCCCTCCTGCCCTCCTGCTCCTGCCCTCCTGCCCTCCTGCTCCTGCCCTCCTGCCCTCCTGCT
TCTCCTGCTCCTGCCCTCCTGCCCTCCTGCTCCTGCCCTCCTGCCCTCCTGCT

CCTGCCCTCCTGCCCTCCTGCTCCTGCCCTCCTGCCCTCCTGCTCCTGCCCTCCTG
CTCCTGCCCTCCTGCTCCTGCCCTCCTGCTCCTGCCCTCCTGCCCTCCTGCCCTCCT
TCCTGCTCCTGCCCTCCTGCTCCTGCCCTCCTGCCCTCCTGCCCTCCTGCTCCTGCC
CCTCCTCCTGCTCCTGCCCTCCTGCCCTCCTGCCCTCCTCCTGCTCCTGCCCTCCTG
CCCCTCCTCCTGCTCCTGCCCTCCTCCTGCTCCTGCCCTCCTGCCCTCCTGCCCTCC
TCCTGCTCCTGCCCTCCTGCCCTCCTCCTGCTCCTGCCCTCCTCCTGCTCCTGCCCT
CCTGCCCTCCTGCCCTCCTCCTGCTCCTGCCCTCCTCCTGCTCCTGCCCTCCTGCC
CTCCTGCCCTCCTGCCCTCCTCCTGCTCCTGCCCTCCTCCTGCTCCTGCCCTCCTGC
TCCTGCCCTCCCGTCTCCTGCTCCTGCTCCTGTTCCACCGTGGGTCCCTTTGCAGCCAATG
CAACTTGGACGTTTTTGGGTCTCCGGACACCATCTCTATGTCTTGGCCCTGATCCTGAGCC
GCCGGGGTCTCCTGGTCTTCCGCCTCCTCGTCTCGTCTCTTCCCGTCTCGTCCATGG
TTATACCCCTCTTCTTTGAGGTCCACTGCCGCCGAGCCTTCTGGTCCAGATGTGTCTCC
CTTCTCTCCTAGGCCATTTCCAGGTCTGTACCTGGCCCTCGTCAGACATGATTCACACTA
AAAGAGATCAATAGACATCTTTATTAGACGACGCTCAGTGAATACAGGGAGTGCAGACTCCT
GCCCCCTCAAACAGCCCCCACCCTCATCCCCTTTCATGGTCTGTCAGACAGATCCAGG
TCTGAAAATTCCTCCTCCGAACCATCCTCGTCTCATACCAATTAATCGCAGCCCGGA
AAACTCCCGTGAACATCCTCAAGATTTGCGTCTGAGCCTCAAGCCAGGCCTCAAATTCCT
CGTCCCCCTTTTTGCTGGACGGTAGGGATTCTCGGGACCCCTCCTCTTCTCTTTC
AAGGTCACCAGACAGAGATGCTACTGGGGCAACGGAAGAAAAGCTGGGTGCGGCCTGTGA
GGATCAGCTTATCGATGATAAGCTGTCAAACATGAGAATTCTTGAAGACGAAAGGGCCTCGT
GATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCAC
TTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTAT
CCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAGGAAGAGTATGAGT
ATTCAACATTTCCGTGTGCGCCCTTATCCCTTTTTGCGGCATTTTGCCTTCTGTTTTGCTC
ACCCAGAAACGCTGGTGAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTA
CATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAAGTTTTTC
CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTGACGCCGGG
CAAGAGCAACTCGGTGCGCGCATACTATTCTCAGAATGACTTGGTTGAGTACTACCAGT
CACAGAAAAGCATTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCA
TGAGTGATAAACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAAC
CGTTTTTTGCACAACATGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGA
ATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTT
GCGCAAATTAATACTGGCGAACTACTTACTTAGCTTCCCGGCAACAATTAATAGACTGGA
TGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGTGGCTGGTTTAT
TGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCA
GATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATG
AACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGAC
CAAGTTTACTCATATACTTTAGATTGATTTAAACTTCATTTTTAATTTAAAGGATCTAGGT
GAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTAACGTGAGTTTTCTGTCGCGCTAATCTG
GTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGGATCCTTTTTTTCTGCGCGTAACTCTG
CTGCTTGCAAACAAAAAACCCAGCTACCAGCGGTGGTTTTGTTTGCCGGATCAAGAGCTAC
CAACTTTTTTCCGAAGGTAACCTGGCTTACGACAGAGCGCAGATACCAAATACTGTCTTCTA
GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACTCGCTCT
GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGAC
TCAAGACGATAGTTACCGGATAAGGCGCAGCGTGGGGCTGAACGGGGGGTTCGTGCACA
CAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAG
AAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCTG
GAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCTGT
CGGGTTTCCGACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGC
CTATGAAAAACGCCAGCAACGCGGCCTTTTTACGTTCTGTCCTTTTTGCTGGCCTTGAAG
CTGTCCCTGATGGTCTCATCTACCTGCCTGGACAGCATGGCCTGCAACGCGGGCATCCCG
ATGCCGCCGGAAGCGAGAAGAATCATAATGGGGAAGGCCATCCAGCCTCGCGTCGCGAAC
GCCAGCAAGACGTAGCCCAGCGCGTCCGCCCGAGATGCGCCGCGTCCGGCTGCTGGAG
ATGGCGGACGCGATGGATATGTTCTGCCAAGGGTTGGTTTGGCATTACAGTTCTCCGCA
AGAATTGATTGGCTCCAATCTTGGAGTGGTGAATCCGTTAGCGAGGTGCCGCCCTGCTTCA
TCCCCGTGGCCGTTGCTCGCGTTTGTGGCGGTGCCCGGAAGAAATATATTTGCATGT

CTTTAGTTCTATGATGACACAAACCCCGCCCAGCGTCTTGTTCATTGGCGAATTCGAACACGC
AGATGCAGTCGGGGCGGCGCGGTCCGAGGTCCACTTCGCATATTAAGGTGACGCGTGTGG
CCTCGAACACCGAGCGACCCTGCAGCGACCCGCTAACAGCGTCAACAGCGTGCCGCAGA
TCCCGGGGGCAATGAGATATGAAAAAGCCTGAACTCACCGCGACGTCTGTGAGAAGTTT
CTGATCGAAAAGTTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTC
GTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATATGTCTGCGGGTAAATAGCTGCGCCGA
TGGTTTCTACAAAGATCGTTATGTTTATCGGGCACTTTCATCGGCCGCGCTCCCGATTCCGG
AAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACA
GGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCAGCCGGTTCGC
GGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCAT
CGGACCGCAAGGAATCGGTCAATACTACATGGCGTGATTTTCATATGCGCGATTGCTGATC
CCCATGTGTATCACTGGCAAACGTGATGGACGACACCGTCAGTGCCTCCGTGCGCAGGC
TCTCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCGAAGTCCGGCACCTCGTGCACGC
GGATTTCCGGTCCAACAATGTCTGACGGACAATGGCCGCATAACAGCGGTTCATTGACTGG
AGCGAGGCGATGTTCCGGGATTCCCAATACGAGGTCGCCAACATCTTCTTGGAGGCCGT
GGTTGGCTTGTATGAGCAGCAGACCGCTACTTTCGAGCGGAGGCATCCGGAGCTTGCAG
GATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTT
GGTTGACGGCAATTTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGA
TCCGGAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACC
GATGGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGGATC
GGGAGATGGGGGAGGCTAACTGAAACACGGAAGGAGACAATACCGGAAGGAACCCGCGCT
ATGACGGCAATAAAAAGACAGAATAAAACGCACGGGTGTTGGGTCGTTTGTTCATAAACGCG
GGGTTCCGGTCCCAGGGCTGGCACTCTGTGATACCCACCGAGACCCCATTTGGGGCCAAT
ACGCCCGCTTTCTTCTTTTCCCCACCCACCCCAAGTTCGGGTGAAGGCCAGGGCT
CGCAGCCAACGTCGGGGCGGAGGCCCTGCCATAGCCACTGGCCCCGTGGGTTAGGGAC
GGGTCCCCCATGGGGAATGGTTTATGGTTCGTGGGGTTATTATTTGGGCGTTGCGTGG
GGTCAGGTCCACGACTGGACTGAGCAGACAGACCCATGGTTTTTGGATGGCCTGGGCATGG
ACCGCATGTACTGGCGCGACACGAACACCGGGCGTCTGTGGCTGCCAACACCCCGACC
CCAAAAACACCGCGCGGATTTCTGGCGTGCCAAGCTAGTCGACCAATTCTCATGTTTGAC
AGTTATCATCGCAGATCCGGGCAACGTTGTTGCCATTGCTGCAGGCGCAGAACTGGTAGG
TATGGAAGATCTATACATTGAATCAATATTGGCAATTAGCCATATTAGTCATTGGTTATATAGC
ATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCATAATATGTACATTTAT
ATTGGCTCATGTCCAATATGACCGCCAT