

OMTN, Volume 18

Supplemental Information

RNA_m5CPred: Prediction of RNA 5-Methylcytosine Sites Based on Three Different Kinds of Nucleotide Composition

Ting Fang, Zizheng Zhang, Rui Sun, Lin Zhu, Jingjing He, Bei Huang, Yi Xiong, and Xiaolei Zhu

Supporting Information S1. The benchmark dataset Met240 consists of a positive dataset and a negative dataset. The positive dataset contains 120 true m5C sites containing sequences with the m5C sites in the center, while the negative dataset contains 120 negative m5C sites containing sequences. Each of these segments is 41-bp long.

I. 120 true m5C sites containing sequences (S+)

>P_1
CGCCUCCCACGCGGGAGACCCGGGUCAAUCCCCGGCCAAU
>P_2
CCGGGUCAAUCCCCGGCCACUGCACGUGGUUGUUUUUCAC
>P_3
GGCCGUGGGUGUGUAGAGGCCUUGGUGGUGCAGUGGUAGAA
>P_4
GUGCAGUGGUAGAAUUCUCGCCUCCCACGUGGGAGACCCGG
>P_5
UGACUGCAGAUCAAGAGGUCCCCGGUCAAUCCGGGUGCC
>P_6
GCUGAAGGCAUCAAAGGUUCCGGGUUCGAGUCCCGGCGGA
>P_7
AGAUGUUGGUGGUAGUAGCACAUAUUCUAAAGAGAACUUCA
>P_8
AGCCAGGGCUGGGCCCCGGGCAGAGCCGCCGAGGUGCAGA
>P_9
CUGCGGUGAGUCUUGAAGCCCAGGGCUGGGCCCCGGGCAGAG
>P_10
UAUUUGAUCGGAUCGUGACCCCAGCCCCGCCGGGCCGACCC
>P_11
CGCUGUGGUCCCCGAGGUCCCGGAGCUGGCCCUUGCGGGGCC
>P_12
GAGACUCUAAUCUCAGGGUCGUGGGUUCGAGCCCCACGUU
>P_13
AGACGCCAGUAAUCAGUGCCCGGCUAGCUCAGUCGGUAGAG
>P_14
GGUUCGAAUCCGAGUCACGGCAGGUGGUUCUAAUUGCUGG
>P_15
UGCGUUGUGGCCGAGCAACCUCGGUUCGAAUCCGAGUCAC
>P_16
GCGCCGCCUGGUUAGUGGCUCGCCGUGAUCGUUAUAGUGGUU
>P_17
GCCUCACACGCGAAAGGUCCCCGGUUCGAAACUGGGCGGAA
>P_18
CAGGUUCCGCCUGUGGUUCCGUAGUGUAGUGGUUAUCACG
>P_19

AGCGUGCUGGGCCCAUAACCCAGAGGUCGAUGGAUCGAAAC
>P_20
GUCGAUGGAUCGAAACCAUCCUCUGCUAUCGGAGUUUUUUC
>P_21
GUCGCAGUCUCCCCUGGAGGCGUGGGUUCGAAUCCCACUCC
>P_22
AGUGCCCUUGCCGAGAGCGGCUCGUUGGUCUAGGGGUAUGA
>P_23
CUUCGGGAGCGCCCGGAUAGCUCAGUCGGUAGAGCAUCAGA
>P_24
UUAGUACUUGGAUGGGAGACCGCCUGGGAAUACCUGGUGCU
>P_25
AACACGCCAGAUCUCGGAAACUAAGCAGGGUCGGGCCUGGU
>P_26
UGGUCUUGUAAACCAGGGGUCGCGAGUUCGAUCCUCGCUGG
>P_27
GAAGCUAAGCAGGGUCGGGCCUGGUUAGUACUUGGACGGGA
>P_28
UGGCCGAGUGGUCUAAGGCGCCAGACUCAAGGUAAGCACCU
>P_29
CCAUACCACCCUGAACGCGCCCAAUCUCGUCUGAUCUCGGA
>P_30
AUCUCUGUCUACGGCCAUACCACCCUGAACGCGCCCAAUCU
>P_31
CUGCUUUACACGCAGAAGGUCCUGGGUUCAAGCCCCAGUGG
>P_32
AAUCCA AUGGGGUCUCCCCGCGCAGGUUCGAACCCUGCUCG
>P_33
CUGCUGCCGUGAUCGU AUAGCGGUUAGUAGUCUGCGUUGUG
>P_34
GUGGAACCAUGAGAUGUUACCUAGCGUUUUGUGAGCCAGGU
>P_35
AAGGCGUUGGACUUAAGAUCCAAUGGAUUCAUAUCCGCGUG
>P_36
AAAUAGCUCAGUUGGGAGAGCGUUAGACUGAAGAUCUAAAG
>P_37
GGUUUGGGUCCGAGAGGUCCCGGGUUCAAAUCCCGGACGAG
>P_38
GUGCCUCUCAUGUACAAGGCCUGAGUUUGACUCCAGCAC
>P_39
AAGGCGUUGGACUCGAAAUCCAAUGGGGUUUUCCCGCACAG
>P_40
CCCGACGGGGAGGUGUGUAGCUGCACUUUUUUGGCGACAGU
>P_41

CUGCGGAGCCACAGCCCAGCAGGACCUCGUGGCGCAACGG
>P_42
AACGGUAGCGCGUCUGACUCCAGAUCAGAAGGCUGCGUGUU
>P_43
UGC UUUGCAUGUAUGAGGCCCGGGUUCGAUCCCCGGCAUC
>P_44
CCCGACGGGGAGGCCAAGUACGUUUUUACCAUUCUUCGGUA
>P_45
CGUAACGGCUGCCGAAAUAGCUCAGUUGGGAGAGCGUUAGA
>P_46
AGGUGGCCCGGGUUCGACUCCCGGUAUGGGAACGCUUCCUU
>P_47
UGGUCUAGCGGUUAGGAUUCUGGUUUUCACCCAGGUGGCC
>P_48
UGGUUUUCACCCAGGCGGCCCGGGUUCGACUCCCGGUGUGG
>P_49
CCGCGGAGGUCAGACUGGGCAGGAGAUGCCGUGGACCCCG
>P_50
UAAGAGUAUCUGUAUCGCGGCUCGUUGGUCUAGGGGUAUGA
>P_51
AUUUGUGGACAUCUAGGUCGCUGGUUCGAUUCGGCUCG
>P_52
GUAGAGCGGAGGACUGUAGCCUGUAGAAACAUUUGUGGACA
>P_53
CAGUGUGCAUCCUUCGAUAGCUCAGCUGGUAGAGCGGAGGA
>P_54
UGACUACGGAUCAGAAGAUUCCAGGUUCGACUCCUGGCUGG
>P_55
GGGAUCGCGCCUGUGAAUAGCCACUGCACUCCAGCCUGGGC
>P_56
CUGGGCAACAUAGCGAGACCCCGUCUCUUUGAACAAUAAA
>P_57
GGUGCUAAUAACGCCAAGGUCGCGGGUUCGAUCCCCGUACG
>P_58
AACGGUACUUGUCAUCUCCACAUGGUCUAGCGGUUAGGA
>P_59
GGUUCGAAUCCGAGUCACGGCAUUGUGGGAACAAUGGCACG
>P_60
GUAUAGUGGUUAGUACUCUGCGUUGUGGCCGCAGCAACCUC
>P_61
GGACUCUGAAUCCAGCGAUCCGAGUUCAAAUCUCGGUGGAA
>P_62
CCGAGGCCGCCAGAGCCCUCCCCGGGAGACCCGAGGCCGC
>P_63

CAGUCGGUAGAGCAUGAGACCCUAAAUCUCAGGGUCGUGGG
>P_64
AACAUAUUGCAGCUGGGUAGCGUGGCCGAGCGGUCUAAGGC
>P_65
GGCCCCGGGUUCGAUCCCCGGCCAAUGCAGCAGCUGAAAGCU
>P_66
AUUCUCGCCUGCCAUGC GGGCGGCCGGGCUUCGAUCCUGG
>P_67
UGUCUAGUAAACAGGAGA UCCUGGGUUCGAAUCCCAGCGGU
>P_68
UGGUAUAGUGGUAAGCAUAGCUGCCUCCAAGCAGUUGACC
>P_69
GACUCCAGAUCAGAAGGUUGCGUGUUCAAGUCACGUCGGGG
>P_70
CUUAAGAAGCCAGCAGGUCCAUGGUGUAAUGGUUAGCACU
>P_71
UCCCAUGGUGUAAUGGUUAGCACUCUGGACUUUGAAUCCAG
>P_72
AAUCCCAGCGGUGCCUCAACCGAGCGUCCAAGCUCUUCCA
>P_73
AUUCUCAAAACUUUAAAUGGGCAAGAAGCCCAGCUAGCUCAG
>P_74
CAGUGCGAGCGGAGCAAUGCCGAGGUUGUGAGUUCGAUCCU
>P_75
GGUCACGACCCUCAGGCUUCAAUCUGAGGGUCCAGGGUUC
>P_76
AGGCUUCAAUCUGAGGGUCCAGGGUUCAUAUCCCUGUUCA
>P_77
CUCUUCGAAUGCACUUGCGGCCCGGGUCCUCCCAGGGCUA
>P_78
GUGGUCUAGUGGUUAGGAUUCAGUGCUCUCUCAUAAAUA
>P_79
AGCACUCUGGACUCUGAAUCCAGCGAUCUGAGUCACGGCAC
>P_80
AAGCGGUUACCUCCUCAUGCCGGACUUUCUAUCUGUCCAUC
>P_81
AGUUCUUAAAUGAAACAAGCAACCUGUCUGGGUUGUUCGA
>P_82
AGCAACCUGUCUGGGUUGUUCGAGACCCGCGGGCGCUCUCC
>P_83
AACCUCUGGGCUGGCUUUAGCUCAGCGGUACUUCGCGUGU
>P_84
CUUUAGCUCAGCGGUACUUCGCGUGUCAUCAAAACCACCUC
>P_85

AGUGGUUAUCACGUUCGCCUCACACGCGAAAGGUCCCCGGU
>P_86
GAUGUAGCUCAGUGGUAGAGCGCAUGCUUUGCAUGUAUGAG
>P_87
AGUUAGAAUACAGACAUGGCCGCCCGGCUAGCUCAGUCGGU
>P_88
AGGACUGUAGGCUCAUUAAGCAAGGUAUCCUUAGGUCGCUG
>P_89
AGCAAGCCCUCUUAGCGCAGCUGGCAGCGCUCAGUCUCAU
>P_90
AGUCUCAUAAUCUGAAGGUCCUGAGUUCAAGCCUCAGAGAG
>P_91
UCGCGGGUUCGAUCCCCGUACUGGCCAGACGUGACUUUUUA
>P_92
AGAAGAUUCUAGGUUCGACUCCUGGCUGGCUCGGGUGUUA
>P_93
AGGAAAUGCUAGUUCAGCUUCAGGUUCCAUGGUGUAAUGGU
>P_94
AAUCCAUGUGCUCUGCACGCGUGGGUUCGAAUCCCACCUU
>P_95
CAGACUCAAGUUGCUCUACUUCCCAGGUUUGGGGCUUCUGGUC
>P_96
GUGGUCUAAGGCGCCAGACUCAAGUUGCUCUACUUCCCAGGUU
>P_97
ACUGCUUCCUGUGUUCGGGUCUUCUGGUCUCCGUAUGGAGG
>P_98
AAAGGUCCCUGGAUCAAAACCAGGCGGAAACAAGUGGUUAC
>P_99
UCAAUUCUCACUCCCACCAGCAGGGUUGUUUCCGUAGUGUA
>P_100
GUUAGCUCAGUCGGCUAGAGCGUGGUGCUAAUAACGCCAAG
>P_101
AUCCUGGGUUCGAAUCCCAGCGGGGCCUUGAUUUCUGUACC
>P_102
UGGCUUAGCUGGUUAAAGCGCCUGUCUAGUAAACAGGAGAU
>P_103
GGAUGGCCGAGUGGUUAAGGCGUUGGACUUAAGAUCCAAUG
>P_104
AAAAGAAGAAGCUUUGUAACCGUUGGUUCCGUAGUGUAGU
>P_105
CCCAGCGGGGCCUUGGUUGGCAAGGUCAGUGUGCCUUUCUG
>P_106
AAGCCAACAGAACUCUCAUGCGGUCAGGGGGUGUAGCUCAG
>P_107

UGAGGCCCCGGGUUCGAUCCCCGGCACCUCCAAUUGGUGGU
>P_108
ACCUCCAAGUGAUGGUUUCUCCUCUGGCAGUUCUCAAGCGAC
>P_109
CGAUCCCCGACACCUCCAAGCGAUGGUUUUGCUCUGGUAGU
>P_110
CAAUCCCCGGCACCUCCAUCCUUUUGCUUUUAAUUUUUUU
>P_111
UGUCCCAGGUGUCAGGAUGGCCGAGUGGUCUAAGGCGCCAG
>P_112
UGAAGGUCCUGAGUUCGAACCUCAGAGGGGGCAAGGCGUCU
>P_113
GAACGCGCCCGAUCUGGUCUCAUCUCGGAAGCUAAGCAGGG
>P_114
GUCAGUUACAGAUCGAACUCCUUGUUCUACUCUUUCCCCC
>P_115
AUUACACUGUCCUCCUAUGACUGGGUGUAUGGCUCAGGGGU
>P_116
AGGGGACCCGACACACGUACACGUCCCUUCGAUAGCUCAG
>P_117
UUAGGUCGCUUGGUUCGAUUCCGGCUCGAAGGAGAGACACC
>P_118
UAUGUAUACAUGUGCCAUGCCCGAUCUCGUCUGAUCUCGGA
>P_119
AAAACUUAAAACUUUACAGUCAGAGGUUCAAUUCCUCUUCU
>P_120
AGAACUGCUAACUCAUGCCCCCAUGUCUAACAACAUGGCUU

II. 120 false **m5C** sites containing sequences (S_0^-)

>N_1
GGGAGUGGGAACAGGAUUUGCAAGACUCCUAGUACCUAAAU
>N_2
GAAAUGGCCUCAUUUGAUAAACUAGUAGGUUUUACACAGUGU
>N_3
GGGCAGCCUCCUUCUUGUCUCUGUUGUUGAGGAGUGGAAUG
>N_4
UCUUCAGGGUUUUCUAGGACCAAGGUGACAUUUCUACGAGG
>N_5
AAGAGAUAGGCUCUCACCUGCCUAAGGGCAGCAUCUUGUUU
>N_6
CGUGGAAAGGGAUUCAGAGACAGAGUCAAGAGAUGAUAGCA
>N_7
GGUGGUUCUAGUGAAGGCAGCAGCAAAGGGGGUAGAGGAGC

>N_8
GAUAAAUGCUGGGCCCAGGGCUGGGCAUAGCCUGCUGCACU
>N_9
CACUGGGUGUGCUGUAACAGCAAGGUUCAUCUGACUAUGUU
>N_10
GAAACAUUACCCCUCCCUUCCCGCACCAACGAUCUGCUUGU
>N_11
CUUUCUAGGGAAUAAGAUACUGUUUUUACACAGGCAAAGG
>N_12
AAACCUUCUGUGGGACAGGCCCUCCACCCAAGGCCCUUUA
>N_13
AAAUAAAAAAGAUCAAAGAGCUAAUAAGUGGUGGUCAGCAU
>N_14
CUGUCCUACUACUGCCCCCCCCCAUCGUAUCAUAGAAAAA
>N_15
AAUAAGAAAAGGUGUCAGAUCUGAUGGUGGAUAAAUAAGU
>N_16
AUCAGAAGCUGAAAUCACACCACAAGUAUCCACUAGCUCC
>N_17
UUAUUUAUUAAAAAGUAAAUCGUUACUGGUGCUUGAUUACA
>N_18
UUUGCAGAAGCAAGCUCAGUCUCCUGAGAGUUCACGCUGAG
>N_19
CCAUCAGACGAGAGAGCUUACUAUUUAGAGAGCCUUCUGCU
>N_20
CCUGCCAGGAUGUGCUACCACAGGCUGUGCUAACUGAACA
>N_21
GCAGACACUGCCAAAGGCAGCUGGAGCCCUACUAAGUGGAA
>N_22
CAUUGGUGCCUCAGACGUAACCAGUCUGUGGGAAGUAUGAA
>N_23
GCAACAUUCCUGCUGGAGGCUCCAGGAUCCUAAGGGACAC
>N_24
ACCUGAUCGAUAGGGAGGAACAACAUCAACCCUACCCCA
>N_25
GGUCCCGGAGAUUCCUUGUCAGCCACCCUACCCUACUUGA
>N_26
UGCCACAUGUGCACGCACACACACUCAUACCCGAACACAC
>N_27
UAUUGAAAUUCUAUCUCACCCUAGUCAGAAUGGCAAUCAUU
>N_28
GCAGAGAAGGGAGCCACGCACACUGUGGGUUGGAAUGUAAA
>N_29
UUUGCUCUGACACCACGAUCCACUGCAUCCCUCCUUAAC

>N_30
UCUAAGCCAAAUGAGCCCCACCCUCUUUAAGUAGAUUCUU
>N_31
CAAGGUGAUCUAACAGAAUUCAUAAGGGAGAAAGUCAGGGU
>N_32
AGGUAAUAGCUCUUUCUCUGCUUGUGACUUCCUGGCUC
>N_33
UGCCACAUGAACACGUAUGCCAAUGGACUCUGCUCGUGU
>N_34
UGCAGAGAAAUGGAUUAAACUAGCUUGGUGUGGAUCACUU
>N_35
UGUUUAUUCAUUUGGCUAUCCAUGUUCUGUACCAUCA
>N_36
UCUUUCUCAGUUUGUUGGAGCAUCUCAGCCAGACAAGACGC
>N_37
CUGUCCAAAGUGUUCUAUUCUUUAACACAUCAACGUUGUU
>N_38
AUUGAUCUUUUUAUGCCAUUCUAACUAAGUCAGAGCAGACU
>N_39
AUUCAGGAGGUCUCAGAACACAGGCACAAAUGAACAUGUUC
>N_40
UUUUUAAGGCUUUAUUGUCUCCUAAGACCUGCAACCAAAA
>N_41
GCUUGGGGAGGUGAUUGUGUCACAAGGGCAGAGUUGUAGAG
>N_42
CAUUUGGAGAAUAAAAGUUACAUUUACAAAGAGUCUUCAUG
>N_43
CUUUCUCACUUCUAUAGCUACAAGAAGUAAGCUUCUGUGAU
>N_44
UCACAGCUGUAACCCCCCCCCCAAGUACUAGGAUCAGUU
>N_45
GGUCCUAGUCUUAAAUGUGCUUUCUGCAUUUGUCACAGAA
>N_46
AUCACUAGCAUCCCCUUCUCUGGGGCAUCAAGCCUCCACA
>N_47
GGACCAAGUGCAUUCCUCCCAAUGAGGCCAGAAAAGGCAA
>N_48
AAACAGACUUUUUUUAUAGCCAAGUUUUUUUUUUCUUU
>N_49
CUUCUAUCUUCCUUUGCCUCAGAUAUUCAUCUUGGUUUGA
>N_50
UACAUUUUUUUUCUGACUAGCUGCAUGACCACUUUGCUUGU
>N_51
GAUUAUGGUUUACCACUCCUCAUCUCCUCCAGAUUCUCCC

>N_52
CAUCUCCCUCUGUCCAACUCCAUGCCUUAUAUGGUCUCUC
>N_53
CAGACACAUAUUUACACAUACACACACACACUUUAUUUU
>N_54
AUAAUCACAAAAACUAGAAACCAUGAUUAAGAGCAAACGU
>N_55
CCAGUAAGGUAAAAAAUACCCAAAGAAAGCAAUGUCAAUCA
>N_56
UGUAAGUGGCUGCACAACUUCACUCUUUUUCUGAUUUUAUC
>N_57
AAUGAGGUUUGACUUGAUUCCUAUUUUAGGUGCUUGUUUUG
>N_58
UUACAUGCAUUGCAGAUUUACUGGGUAGCUUGAAUGUAAGC
>N_59
CUUGAGCAAAAUUUUAUAGAUCUUGAAGUUGCAAAGUAGAC
>N_60
AUCGGUCCUCAUUCGCUAACUAUAGUUUUUUUCCUAAAAG
>N_61
AGACAAGAACUCCUGUUGAUCCAUAUAUGUGCCAAACCUCA
>N_62
GAGCUGUCAAGGGCACACAUCUGAGAUACAUGAAUGCUGAG
>N_63
UAAGACAACAACAGAGUCUUCUAGCGCAUCUUUGUGGUUU
>N_64
UGUCAGAAGAAUACCAUGCCCAUUUCUCAUUUGUCAUUUC
>N_65
CUCUUAUGACUACUAUUUCACUCUCUUCUAUGAGGUCAAUU
>N_66
CUUAUCAAAAGUCUAAAGUUACUUGGUAAUUGUAUUUCUUGG
>N_67
AUAAUUGGAAGAUUAUGAAACACUUUGUCAACCCACUAAA
>N_68
AUGACAAUGUCCUCCUAUUUCUGUGUUAUACCUGCUUGAAG
>N_69
GGCUUGUGAGUCUUUGGGGGCAUUUUCUAAUUAAGGGAGU
>N_70
CGAGGGAGACUUACCUCUUCCCAGGGCACUCAUAAGGUGUU
>N_71
GCCUUUACGUCAGUACCAGCCAUGUGGAAUGCUCUACAGGU
>N_72
ACCUCGAGGAAGGGGCACAACAACUGGAACUGUUUGCCAUA
>N_73
GGAACACAGGUCCAAGGCUGCUGUUUCAUUCACAAUUAUUC

>N_74
GGAAACACCCUUCUUCAGUACCCACCCUAGAAUGAACAAGA
>N_75
AUAGAGUUAUGCAAAAUUUUCUCUUGGAAUAUUCUUUUCUU
>N_76
UAGUGACAAUAGGCACAGAACAGUGGCAGCAGAGAAACAGU
>N_77
AUGUUAAGGACUGAAGAAUUCUGACAUCUAUGUGUUCUGU
>N_78
CUUGAAGUGGAAACUAAAAGCUUCUGUGGAAAAGGUGUUGG
>N_79
GAAUAUUUCACUGAAGCAGACACAGGAAAAAGAAUGUUCUG
>N_80
AAAAAACUUCUGGUGGUGUGCUGUGGCUUCUUGCUGCUUCU
>N_81
UAUAAAAGAACUCAACAGACAGUGGGAAGCAGAGCUUGGC
>N_82
CUCCAAAGAAUCCUUCUAAACAGGUCCAGUUCCACCAUAUC
>N_83
AGAGUAAAGCGGCAUAAAUCAGGCACUACAACCUCCAAC
>N_84
AAACAAUGAUAGCAUGUAUUCUCAAGGACACAGUGAUUUUU
>N_85
UUAAAUAAACAUGCUUGAGUCUGAGGUAAGAGAGGGCCUCC
>N_86
UUACUCUGAGAGUCCACUUUCAAAAAGCUUAGACAGCAAU
>N_87
GCACAGACUGUCUCUCCUGUCACAGGUAUGUAUGUUCAUGU
>N_88
GGAUCAUGAGUUACCUCUUUCUUCUCUCUCUGCAAGGCUA
>N_89
AGAGCUUCUCUCAAUUCUUCUUUUUUUUUUUUUUUUUUUU
>N_90
AUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU
>N_91
UUGCAUUUUCUCCCAUGUUACCUUUGAAAACACCUAGGU
>N_92
GUCUAAAGCCAUAACCAUCUACACAAACAUGUAGAAGUUUGA
>N_93
GAUCCAGGGGGCAUCCCAUCGUGCACACAGCCUCCCGAG
>N_94
CCAUGUCCUCUAUCCAUAAGCCUGAAAACACAGCCAGUAGG
>N_95
UUUGAAAUCCUUGUUCAUUCUUCUAUAAUCAGUCUGGCAC

>N_96
CUCAUGUACAUGGCGUAUUUCAGUUUUGUCACAAAAUUUGA
>N_97
CUACUAAUAAUGUAAGAGUCACGUAAUUGAGGAUGGUAUU
>N_98
UGUGAGGCCAGAGAAAAGGACCUUAGGCAUCAUUUCACAGG
>N_99
CAGCAAGCAGCCAUCUCUCUCUCUCUCUCUCUCUCUCUCUCUC
>N_100
CAAAAAAGAACAUAUAAAACCGAUUUACACUAUAUAUAUAU
>N_101
UCUCUUGGCAAUCCUCCUACCCUAGCUGCCAAGUAUUUGAC
>N_102
UUUAACUAAUAUACUGGAAACAUAAAAAUAUUUAUAUUACC
>N_103
GCACUUCUUUAAUCUCUUUUCGCCUUAUUGGUUGAUUAUAGG
>N_104
AUUUCUCACUGAAUCAAGCCCUACUGGUUGGUUACAAUGG
>N_105
CCCAAUGCUGACAUAACAUGCUUAUGCUGUCUCCUUGGAA
>N_106
UAGCACAGUAGUGUCAUCUGCAGCCAACUUGCUAUUAACCC
>N_107
UGCACCUGAACUUGUUGCUCUUUAACAGCAACAGUGCCC
>N_108
CAGCAGUGUCCCAGUGCAUUCAUUGCCCCAUGAGCAGCAUC
>N_109
UGGAUACAGAUGUCCUUUGCCUUACAGAAGCUUUGCAAUUU
>N_110
AAGAAAGAAAUUCACUUGGACUUGAGCUUUGUACAAGGACA
>N_111
GUAGGUGUGUGGGUUCAUUUCUGGGUCUCAAUUCUAUUUA
>N_112
GGAUAUCAGUCUGGCAGUUCCUCAGAAAAUUGGACACGAGG
>N_113
UUGCAGCCCCAUAGGAGGAGCAACAAUAUAAACCAACCAGU
>N_114
ACACAUGAAGGGACUCAUGUCUCCAGCUGCAUAUGUAGCAG
>N_115
AGGGUGGCCAUGUCAGUCAUCAUUGGGAGGAGAGGCCCUUG
>N_116
GUCCUGUGAAGGUUCUGUGCCUCAGUGUAGGAGAAUGCCAU
>N_117
ACAAAUGAAGAAUGAGCCAUCUUAUACACAGACUCAGUCUU

>N_118

AGCAGAUCAUUACUGAAU AACACUAGGAGAAGGCACUUGAA

>N_119

GAUUCUUUCCUUAAGAUUUGCUUUAAUUUUAAUUAUUUGUA

>N_120

ACUCAGGUCUCCUUUGAGAGCAGAACUCAGUCUUAUGUGCU