Electronic Supplementary Material (ESI) for Molecular BioSystems.
This journal is © The Royal Society of Chemistry 2016

Supporting Information S1. The benchmark dataset consists of a positive dataset and a negative dataset. The former contains 120 true m^5C site containing sequences with the m^5C site in the center, while the latter contains 120 false m^5C site containing sequences. Each of these segments is 41-bp long.

I. 120 true m^5C site containing sequences

>P_1
CGCCUCCCACGCGGGAGACCCGGGUUCAAUUCCCGGCCAAU
>P_2
CCGGGUUCAAUUCCCGGCCACUGCACGUGGUUGUUUUUCAC
>P_3
GGCCGUGGGUGUGUAGAGGCCUUGGUGGUGCAGUGGUAGAA
>P_4
GUGCAGUGGUAGAAUUCUCGCCUCCCACGUGGGAGACCCGG
>P_5
UGACUGCAGAUCAAGAGGUCCCCGGUUCAAAUCCGGGUGCC
>P_6
GCUGAAGGCAUUCAAAGGUUCCGGGUUCGAGUCCCGGCGGA
>P_7
AGAUGUUGGUGGUAGUAGCACAUAUUCUAAAGAGAACUUCA
>P_8
AGCCCAGGGCUGGGCCCGGGCAGAGCCGCCGCAGGUGCAGA
>P_9
CUGCGGUGAGUCUUGAAGCCCAGGGCUGGGCCCGGGCAGAG
>P_10
UAUUUGAUCGGAUCGUGACCCCAGCCCCGCCGGGCCGACCC
>P_11
CGCUGUGGUCCCCGAGGUCCCGGAGCUGGCCCUGCGGGGCC
>P_12
GAGACUCUUAAUCUCAGGGUCGUGGGUUCGAGCCCCACGUU
>P_13
AGACGCCAGUAAUCAGUGCCCGGCUAGCUCAGUCGGUAGAG
>P_14
GGUUCGAAUCCGAGUCACGGCAGGUGGUUCUAACUUGCUGG
>P_15
UGCGUUGUGGCCGCAGCAACCUCGGUUCGAAUCCGAGUCAC
>P_16
GCGCCGCCUGGUUAGUGGCUCGCCGUGAUCGUAUAGUGGUU
>P_17
GCCUCACACGCGAAAGGUCCCCGGUUCGAAACUGGGCGGAA
>P_18
CAGGUUCCGCCUGUGGUUUCCGUAGUGUAGUGGUUAUCACG
>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
AAUCCAAUGGGGUCUCCCCGCGCAGGUUCGAACCCUGCUCG
>P_33
CUGCUGCCGUGAUCGUAUAGCGGUUAGUAGUCUGCGUUGUG
>P_34
GUGGAACCAUGAGAUGUUACCUAGCGUUUUGUGAGCCAGGU
>P_35
AAGGCGUUGGACUUAAGAUCCAAUGGAUUCAUAUCCGCGUG
>P_36
AAAUAGCUCAGUUGGGAGAGCGUUAGACUGAAGAUCUAAAG
>P_37
GGUUUGGGUCCGAGAGGUCCCGGGUUCAAAUCCCGGACGAG
>P_38
GUGCCUCUCAUGUACAAGGCCCUGAGUUUGACUCCCAGCAC
>P_39
AAGGCGUUGGACUCGAAAUCCAAUGGGGUUUCCCCGCACAG
>P_40
CCCGACGGGGAGGUGUGUAGCUGCACUUUUUUGGCGACAGU
>P_41
CUGCGCGAGCCACAGCCCAGCAGGACCUCGUGGCGCAACGG
>P_42

AACGGUAGCGCGUCUGACUCCAGAUCAGAAGGCUGCGUGUU
>P_43
UGCUUUGCAUGUAUGAGGCCCCGGGUUCGAUCCCCGGCAUC
>P_44
CCCGACGGGGAGGCCAAGUACGUUUUUACCAUUCUUCCGUA
>P_45
CGUAACGGCUGCCGAAAUAGCUCAGUUGGGAGAGCGUUAGA
>P_46
AGGUGGCCCGGGUUCGACUCCCGGUAUGGGAACGCUUCCUU
>P_47
UGGUCUAGCGGUUAGGAUUCCUGGUUUUCACCCAGGUGGCC
>P_48
UGGUUUUCACCCAGGCGGCCCGGGUUCGACUCCCGGUGUGG
>P_49
CCGCGCGAGGUCAGACUGGGCAGGAGAUGCCGUGGACCCCG
>P_50
UAAGAGUAUCUGUAUCGCGGCUCGUUGGUCUAGGGGUAUGA
>P_51
AUUUGUGGACAUCCUUAGGUCGCUGGUUCGAUUCCGGCUCG
>P_52
GUAGAGCGGAGGACUGUAGCCUGUAGAAACAUUUGUGGACA
>P_53
CAGUGUGCAUCCUUCGAUAGCUCAGCUGGUAGAGCGGAGGA
>P_54
UGACUACGGAUCAGAAGAUUCCAGGUUCGACUCCUGGCUGG
>P_55
GGGAUCGCGCCUGUGAAUAGCCACUGCACUCCAGCCUGGGC
>P_56
CUGGGCAACAUAGCGAGACCCCGUCUCUUUUGAACAAUAAA
>P_57
GGUGCUAAUAACGCCAAGGUCGCGGGUUCGAUCCCCGUACG
>P_58
AACGGUUACUUGUCAUCUCCCACAUGGUCUAGCGGUUAGGA
>P_59
GGUUCGAAUCCGAGUCACGGCAUUGUGGGAACAAUGGCACG
>P_60
GUAUAGUGGUUAGUACUCUGCGUUGUGGCCGCAGCAACCUC
>P_61
GGACUCUGAAUCCAGCGAUCCGAGUUCAAAUCUCGGUGGAA
>P_62
CCGAGGCCGCCCAGAGCCCUCCCCGGGAGACCCGAGGCCGC
>P_63
CAGUCGGUAGAGCAUGAGACCCUUAAUCUCAGGGUCGUGGG
>P_64

AACAUAUUGCAGCUGGGUAGCGUGGCCGAGCGGUCUAAGGC
>P_65
GGCCCGGGUUCGAUUCCCGGCCAAUGCAGCAGCUGAAAGCU
>P_66
AUUCUCGCCUGCCAUGCGGGCGGCCGGGCUUCGAUUCCUGG
>P_67
UGUCUAGUAAACAGGAGAUCCUGGGUUCGAAUCCCAGCGGU
>P_68
UGGUAUAGUGGUAAGCAUAGCUGCCUUCCAAGCAGUUGACC
>P_69
GACUCCAGAUCAGAAGGUUGCGUGUUCAAGUCACGUCGGGG
>P_70
CUUAAGAAGCCAGCAGGUCCCAUGGUGUAAUGGUUAGCACU
>P_71
UCCCAUGGUGUAAUGGUUAGCACUCUGGACUUUGAAUCCAG
>P_72
AAUCCCAGCGGUGCCUCAACCGAGCGUCCAAGCUCUUUCCA
>P_73
AUUCUCAAACUUUAAAUGGGCAAGAAGCCCAGCUAGCUCAG
>P_74
CAGUGCGAGCGGAGCAAUGCCGAGGUUGUGAGUUCGAUCCU
>P_75
GGUCACGACCCUCAGGCUUUCAAUCUGAGGGUCCAGGGUUC
>P_76
AGGCUUUCAAUCUGAGGGUCCAGGGUUCAUAUCCCUGUUCA
>P_77
CUCUUCGAAUGCACUUGCGGCCCGGGUUCCUCCCAGGGCUA
>P_78
GUGGUCUAGUGGUUAGGAUUCAGUGCUCUCUCAUAAAAUAA
>P_79
AGCACUCUGGACUCUGAAUCCAGCGAUCUGAGUCACGGCAC
>P_80
AAGCGGUUACCUCCUCAUGCCGGACUUUCUAUCUGUCCAUC
>P_81
AGUUCUUUAAUUGAAACAAGCAACCUGUCUGGGUUGUUCGA
>P_82
AGCAACCUGUCUGGGUUGUUCGAGACCCGCGGGCGCUCUCC
>P_83
AACCUCUGGGCUGGCUUUAGCUCAGCGGUUACUUCGCGUGU
>P_84
CUUUAGCUCAGCGGUUACUUCGCGUGUCAUCAAACCACCUC
>P_85
AGUGGUUAUCACGUUCGCCUCACACGCGAAAGGUCCCCGGU
>P_86

GAUGUAGCUCAGUGGUAGAGCGCAUGCUUUGCAUGUAUGAG
>P_87
AGUUAGAAUACAGACAUGGCCGCCCGGCUAGCUCAGUCGGU
>P_88
AGGACUGUAGGCUCAUUAAGCAAGGUAUCCUUAGGUCGCUG
>P_89
AGCAAGCCCUCUUAGCGCAGCUGGCAGCGCGUCAGUCUCAU
>P_90
AGUCUCAUAAUCUGAAGGUCCUGAGUUCAAGCCUCAGAGAG
>P_91
UCGCGGGUUCGAUCCCCGUACUGGCCAGACGUGACUUUUUA
>P_92
AGAAGAUUCUAGGUUCGACUCCUGGCUGGCUCGGGUGUUAA
>P_93
AGGAAAUGCUAGUUCAGCUUCAGGUUCCAUGGUGUAAUGGU
>P_94
AAUCCAUUGUGCUCUGCACGCGUGGGUUCGAAUCCCACCUU
>P_95
CAGACUCAAGUUGCUACUUCCCAGGUUUGGGGCUUCUGGUC
>P_96
GUGGUCUAAGGCGCCAGACUCAAGUUGCUACUUCCCAGGUU
>P_97
ACUGCUUCCUGUGUUCGGGUCUUCUGGUCUCCGUAUGGAGG
>P_98
AAAGGUCCCUGGAUCAAAACCAGGCGGAAACAAGUGGUUAC
>P_99
UCAAUUCUCACUCCCACCAGCAGGGUUGUUUCCGUAGUGUA
>P_100
GUUAGCUCAGUCGGCUAGAGCGUGGUGCUAAUAACGCCAAG
>P_101
AUCCUGGGUUCGAAUCCCAGCGGGGCCUUGAUUUCUGUACC
>P_102
UGGCUUAGCUGGUUAAAGCGCCUGUCUAGUAAACAGGAGAU
>P_103
GGAUGGCCGAGUGGUUAAGGCGUUGGACUUAAGAUCCAAUG
>P_104
AAAAGAAGAAGCUUUGUAACCGUUGGUUUCCGUAGUGUAGU
>P_105
CCCAGCGGGGCCUUGGUUGGCAAGGUCAGUGUGCCUUUCUG
>P_106
AAGCCAACAGAACUCUCAUGCGGUCAGGGGGUGUAGCUCAG
>P_107
UGAGGCCCCGGGUUCGAUCCCCGGCACCUCCAAAUGGUGGU
>P_108

ACCUCCAAGUGAUGGUUUCCCUCUGGCAGUUCUCAAGCGAC
>P_109
CGAUCCCCGACACCUCCAAGCGAUGGUUUUGCUCUGGUAGU
>P_110
CAAUCCCCGGCACCUCCAUUCCUUUUGCUUUUAAUUUUUUU
>P_111
UGUCCCAGGUGUCAGGAUGGCCGAGUGGUCUAAGGCGCCAG
>P_112
UGAAGGUCCUGAGUUCGAACCUCAGAGGGGGCAAGGCGUCU
>P_113
GAACGCGCCCGAUCUGGUCUCAUCUCGGAAGCUAAGCAGGG
>P_114
GUCAGUUACAGAUCGAACUCCUUGUUCUACUCUUUCCCCCC
>P_115
AUUACACUGUCCUCCUAUGACUGGGUGUAUGGCUCAGGGGU
>P_116
AGGGGACGCCGACACACGUACACGUCCCUUCGAUAGCUCAG
>P_117
UUAGGUCGCUGGUUCGAUUCCGGCUCGAAGGAGAGACACCC
>P_118
UAUGUAUACAUGUGCCAUGCCCGAUCUCGUCUGAUCUCGGA
>P_119
AAAACUUAAAACUUUACAGUCAGAGGUUCAAUUCCUCUUCU
>P_120
AGAACUGCUAACUCAUGCCCCCAUGUCUAACAACAUGGCUU

II. 120 false m^5C site containing sequences

>N_1
GGGAGUGGGAACAGGAUUUGCAAGACUCCUAGUACCUAAAU
>N_2
GAAAUGGCCUCAUUUGAUAACUAGUAGGUUUUACACAGUGU
>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
CUUUCCUAGGGAAUAAGAUACUGUUUUUACACAGGCAAAGG
>N_12
AAACCUUCUGUGGGACAGGCCCUCCACCCAAGGCCCUUUAA
>N_13
AAAUAAAAAAGAUCAAAGAGCUAAUAAGUGGUGGUCAGCAU
>N_14
CUGUCCUACUACUGCCCCCCCCCCAUCGUAUCAUAGAAAAA
>N_15
AAUAAGAAAAGGUGUCAGAUCUGAUGGUGGAUAAAUAAAGU
>N_16
AUCAGAAGCUGAAAUCACACCACAAGUAUCCCACUAGCUCC
>N_17
UUAUUUAUUAAAAAGUAAAUCGUUACUGGUGCUUGAUUACA
>N_18
UUUGCAGAAGCAAGCUCAGUCUCCUGAGAGUUCACGCUGAG
>N_19
CCAUCAGACGAGAGAGCUUACUAUUUAGAGAGCCUUCUGCU
>N_20
CCUGCCAGGAUGUGCUACCACAGGCUGUGCUCAACUGAACA
>N_21
GCAGACACUGCCAAAGGCAGCUGGAGCCCUACUAAGUGGAA
>N_22
CAUUGGUGCCUCAGACGUAACCAGUCUGUGGGAAGUAUGAA
>N_23
GCAACAUUCCCUGCUGGAGGCUCCAGGAUCCUAAGGGACAC
>N_24
ACCUGAUCGAUAGGGAGGAACAACACAUCAACCCUACCCCA
>N_25
GGUCCCGGAGAUUUCCUUGUCAGCCACCCUACCCUACUUGA
>N_26
UGCCCACAUGUGCACGCACACACACUCAUACCCGAACACAC
>N_27
UAUUGAAAUUCUAUCUCACCCUAGUCAGAAUGGCAAUCAUU
>N_28
GCAGAGAAGGGAGCCACGCACACUGUGGGUUGGAAUGUAAA
>N_29
UUUGCUCCUGACACCACGAUCCACUGCAUCCCUCCUUAAAC
>N_30
UCUAAGCCAAAUGAGCCCCACCCUCUUUUAAGUAGAUUCUU
>N_31

CAAGGUGAUCUAACAGAAUUCAUAAGGGAGAAAGUCAGGGU
>N_32
AGGUAAUAGCUCCCCUUUCUCUGCUUGUGACUUCCUGGCUC
>N_33
UGCCACAUGAACACGUAUGCCAAUGGACUCUGCUCCCGUGU
>N_34
UGCAGAGAAAAUGGAUUAAACUAGCUUGGUGUGGAUCACUU
>N_35
UGUUUAUUCAUUUGGCUAUUCCAUUGUUCCUGUACCAUCAA
>N_36
UCUUUCUCAGUUUGUUGGAGCAUCUCAGCCAGACAAGACGC
>N_37
CUGUUCCAAAGUGUUCUAUUCUUUAACACAUCAACGUUGUU
>N_38
AUUGAUCUUUUUAUGCCAUUCUAACUAAGUCAGAGCAGACU
>N_39
AUUCAGGAGGUCUCAGAACACAGGCACAAAUGAACAUGUUC
>N_40
UUUUUAAGGCUUUAUUGUCUCCUAAGACCUGCAACCAAAAA
>N_41
GCUUGGGGAGGUGAUUGUGUCACAAGGGCAGAGUUGUAGAG
>N_42
CAUUUGGAGAAUAAAAGUUACAUUUACAAAGAGUCUUCAUG
>N_43
CUUUCUCACUUCUAUAGCUACAAGAAGUAAGCUUCUGUGAU
>N_44
UCACAGCUGUAACCCCCCCCCCCCAAGUACUAGGAUCAGUU
>N_45
GGUCCUAGUCUUAAAUUGUGCUUUCUGCAUUUGUCACAGAA
>N_46
AUCACUAGCAUCCCCCUUCUCUGGGGCAUCAAGCCUCCACA
>N_47
GGACCAAGUGCAUUCCCUCCCAAUGAGGCCAGAAAAGGCAA
>N_48
AAACAGACUUUUUUUUAUAGCCAAGUUAUUUUUUUUUCUUU
>N_49
CUUCUAUCUUUCCUUUGCCUCAGAUAUUCAUCUUGGUUUGA
>N_50
UACAUUUAUUUUCUGACUAGCUGCAUGACCACUUUGCUUGU
>N_51
GAUUAUGGUUUACCACUCCUCAUCUCCUUCCAGAUCUUCCC
>N_52
CAUCUCCCUCCUGUCCAACUCCAUGCCUUAUAUGGUCUCUC
>N_53

CAGACACAUAUUUACACAUACACACACACACACUUUAUUUU
>N_54
AUAAUCACAAAAACUAGAAACCAUGAUAUAAGAGCAAACGU
>N_55
CCAGUAAGGUAAAAAAUACCCAAAGAAAGCAAUGUCAAUCA
>N_56
UGUAAGUGGCUGCACAACUUCACUCUUUUUCUGAUUUUAUC
>N_57
AAUGAGGUUUGACUUGAUUCCUAUUUUAGGUGCUUGUUUUG
>N_58
UUACAUGCAUUGCAGAUUUACUGGGUAGCUUGAAUGUAAGC
>N_59
CUUGAGCAAAAUUUAUAGAUCCUUGAAGUUGCAAAGUAGAC
>N_60
AUCGGUCCUCCAUUCGCUAACUAUAGUUUUUUUCCUAAAAG
>N_61
AGACAAGAACUCCUGUUGAUCCAUAUAUGUGCCAAACCUCA
>N_62
GAGCUGUCAAGGGCACACAUCUGAGAUACAUGAAUGCUGAG
>N_63
UAAGACAACAACAGAGUCUUCUUAGCGCAUCUUUGUGGUUU
>N_64
UGUCAGAAGAAUUACCAUGCCCAUUUCUCAUUUGUCAUUUC
>N_65
CUCUUAUGACUACUAUUUCACUCUCUUCUAUGAGGUCAAUU
>N_66
CUUAUCAAAGUCUAAAGUUACUUGGUAAUUGUAUUUCUUGG
>N_67
AUAAUUGGAAGAUUAUGAAACACUUUGUCAACCCACUUAAA
>N_68
AUGACAAUGUCCUCCUAUUUCUGUGUUAUACCUGCUUGAAG
>N_69
GGCUUGUGAGUCUUUGGGGGCAUUUUCUAAUUAAUGGGAGU
>N_70
CGAGGGAGACUUACCUCUUCCCAGGGCACUCAUAAGGUGUU
>N_71
GCCUUUACGUCAGUACCAGCCAUGUGGAAUGCUCUACAGGU
>N_72
ACCUCGAGGAAGGGGCACAACAACUGGAACUGUUUGCCAUA
>N_73
GGAACACAGGUCCAAGGCUGCUGUUUCAUUCACAAUUAUUC
>N_74
GGAAACACCCUUCUUCAGUACCCACCCUAGAAUGAACAAGA
>N_75

AUAGAGUUAUGCAAAAUUUUCUCUUGGAAUAUUCUUUUCUU
>N_76
UAGUGACAAUAGGCACAGAACAGUGGCAGCAGAGAAACAGU
>N_77
AUGUUAAGGACUGAAGAAUUCUGACAUUCUAUGUGUUCUGU
>N_78
CUUGAAGUGGAAACUUAAAGCUUCUGUGGAAAAGGUGUUGG
>N_79
GAAUAUUUCACUGAAGCAGACACAGGAAAAAGAAUGUUCUG
>N_80
AAAAAACUUCUGGUGGUGUGCUGUGGCUUCUUGCUGCUUCU
>N_81
UAUAAAAAGAACUCAACAGACAGUGGGAAGCAGAGCUUGGC
>N_82
CUCCAAAGAAUCCUUCUAAACAGGUCCAGUUCCACCAUAUC
>N_83
AGAGUAAAGCGGCAUUAAAUCAGGCACUUACAACCUCCAAC
>N_84
AAACAAUGAUAGCAUGUAUUCUCAAGGACACAGUGAUUUUU
>N_85
UUAAAUAAACAUGCUUGAGUCUGAGGUAAGAGAGGGCCUCC
>N_86
UUACUCUGAGAGUCCACUUUCAAAAAGCUUAGACAGCAAAU
>N_87
GCACAGACUGUCUCUCCUGUCACAGGUAUGUAUGUUCAUGU
>N_88
GGAUCAUGAGUUACCUCUUUCUUUCUCUCUCUGCAAGGCUA
>N_89
AGAGCUUCUCUCAAUUCUUCCCCCUCCCCCAUAGCUAUUUC
>N_90
AUAUUUUUCUAAUACUUCAUCCAACUCUCAAAGGAGAUGGC
>N_91
UUGCAUUUUCUCCCAUGUUACCUUUGAAAACACCUGUAGGU
>N_92
GUCUAAAGCCAUACCAUCUACACAAACAUGUAGAAGUUUGA
>N_93
GAUCCAGGGGGCAUCCCCAUCGUGCACACAGCCUCCCCGAG
>N_94
CCAUGUCCUCUAUUCCAUAGCCUGAAAACACAGCCAGUAGG
>N_95
UUUGAAAUUCCUUGUUCAUUCUUCUAUAAUCAGUCUGGCAC
>N_96
CUCAUGUACAUGGCGUAUUUCAGUUUUGUCACAAAAUUUGA
>N_97

CUACUUAAUAAUGUAAGAGUCACGUAAUUGAGGAUGGUAUU
>N_98
UGUGAGGCCAGAGAAAAGGACCUUAGGCAUCAUUUCACAGG
>N_99
CAGCAAGCAGCCAUCUCUCUCUCUCUCUCUCUCUCUCUCUC
>N_100
CAAAAAAGAACAAUAUAAACCGAUUUACACUAUAUAUAUAU
>N_101
UCUCUUGGCAAUCCUCCUACCCUAGCUGCCAAGUAUUUGAC
>N_102
UUUAACUAAUAUACUGGAAACAUAAAAAUAUUUAUAUUACC
>N_103
GCACUUCUUUAAUCUCUUUUCGCCUUAUUGGUUGAUAUAGG
>N_104
AUUUCUCACUGAAUCAAGCCCUUACUGGUUGGUUACAAUGG
>N_105
CCCAAUGCUGACAUAACAUGCUUAUGCUGUCUUCCUUGGAA
>N_106
UAGCACAGUAGUGUCAUCUGCAGCCAACUUGCUAUUAACCC
>N_107
UGCACCUGAACUUGUUGCUCCUUUUAACAGCAACAGUGCCC
>N_108
CAGCAGUGUCCCAGUGCAUUCAUUGCCCCAUGAGCAGCAUC
>N_109
UGGAUACAGAUGUCCUUUGCCUUACAGAAGCUUUGCAAUUU
>N_110
AAGAAAGAAAUUCACUUGGACUUGAGCUUUGUACAAGGACA
>N_111
GUAGGUGUGUGGGUUCAUUUCUGGGUCUUCAAUUCUAUUUA
>N_112
GGAUAUCAGUCUGGCAGUUCCUCAGAAAAUUGGACACGAGG
>N_113
UUGCAGCCCCAUAGGAGGAGCAACAAUAUAAACCAACCAGU
>N_114
ACACAUGAAGGGACUCAUGUCUCCAGCUGCAUAUGUAGCAG
>N_115
AGGGUGGCCAUGUCAGUCAUCAAUGGGAGGAGAGGCCCUUG
>N_116
GUCCUGUGAAGGUUCUGUGCCUCAGUGUAGGAGAAUGCCAU
>N_117
ACAAAUGAAGAAUGAGCCAUCUUAUACACAGACUCAGUCUU
>N_118
AGCAGAUCAUUACUGAAUAACACUAGGAGAAGGCACUUGAA
>N_119

GAUUCUUUCCUUAAGAUUUGCUUUAAUUUUAAUUAUUUGUA
>N_120
ACUCAGGUCUCCUUUGAGAGCAGAACUCAGUCUUAUGUGCU