

The dataset from *E.coli* composed of 34 true D modification site sequences (positive dataset) and 127 false D modification site sequences (negative dataset)

Positive dataset_ *E.coli*

```
>P_E.coli_1
UAUAGCUCAGCUGGGAGAGCGCU
>P_E.coli_2
CGUAGCUCAGCUGGAUAGAGUAC
>P_E.coli_3
CGUAGCUCAGCUGGAUAGAGCGC
>P_E.coli_4
CUGUAGUUCAGUCGGUAGAACGG
>P_E.coli_5
CGGUAGUUCAGUCGGUUAGAAUA
>P_E.coli_6
GUUCAGUCGGUUAGAAUACCGC
>P_E.coli_7
UAACAAAGCGGUUAUGUAGCGGA
>P_E.coli_8
UCGCCAAGCGGUAAGGCACCGGU
>P_E.coli_9
UAGUUCAAUGGUAGAACGAGAGC
>P_E.coli_10
GAAUAGCUCAGUUGGUAGAGCAC
>P_E.coli_11
CUAUAGCUCAGUUGGUAGAGCCC
>P_E.coli_12
UGUAGCUCAGGUGGUAGAGCGC
>P_E.coli_13
UAGCUCAGUGGUUAGAGCAGGCG
>P_E.coli_14
AGGUGGCGGAAUUGGUAGACGCG
>P_E.coli_15
AGGUGGUGGAAUUGGUaGACACG
>P_E.coli_16
GGAUGGUGGAAUCGGUaGACACA
>P_E.coli_17
GGCGAAAUCGGUaGACGCAGUUG
>P_E.coli_18
CGUUAGCUCAGUUGGUAGAGCAG
>P_E.coli_19
ACGUAGCUCAGUUGGUUAGAGCA
>P_E.coli_20
GCUCAGUUGGUUAGAGCACAUCA
>P_E.coli_21
GGAUAGCUCAGUCGGUAGAGCAG
>P_E.coli_22
GCGCAGCCUGGUAGCGCACUUCG
>P_E.coli_23
GUCGUCUCCGGUGAGGCGGCUGG
>P_E.coli_24
GGCCGAGCGGUUgAAGGCACCGG
>P_E.coli_25
GCCGGAGCGGCUgAACGGACCGG
>P_E.coli_26
GGCCGAGAGGCUgAAGGCGCUCC
>P_E.coli_27
GUCCGAGUGGCUgAAGGAGCACG
```

>P_E.coli_28
AUAUAGCUCAGUUGGUAGAGCGC
>P_E.coli_29
GCGUAGUUCAAUUGGUAGAGCAC
>P_E.coli_30
GGGUGAUUAGCUCAGCUGGGAGA
>P_E.coli_31
UCAUAGCUCAGUUGGUUAGAGCA
>P_E.coli_32
GCUCAGUUGGUUAGAGCACCACC
>P_E.coli_33
GAGUGGCGAAAUUGGUaGACGCA
>P_E.coli_34
AGCUCAGUAGGUAGAGCAACUGA

Negative dataset_ *E.coli*

>N_E.coli_1
UGGGAGAGCGCUUGCAUGGCAUG
>N_E.coli_2
GAGCGCUUGCAUGGCAUGCAAGA
>N_E.coli_3
CUUGCAUGGCAUGCAAGAGGUCA
>N_E.coli_4
CAUGCAAGAGGUCAGCGGUUCGA
>N_E.coli_5
GAGGUCAGCGGUUCGAUCCCGCU
>N_E.coli_6
GCUGGAUAGAGUACUCGGCUACG
>N_E.coli_7
GAGUACUCGGCUACGAACCGAGC
>N_E.coli_8
GAACCGAGCGGUCGGAGGUUCGA
>N_E.coli_9
GCGGUCGGAGGUUCGAAUCCUCC
>N_E.coli_10
GGAUAGAGCGCUGCCCUCCGGAG
>N_E.coli_11
GAGCGCUGCCCUCCGGAGGCAGA
>N_E.coli_12
GGAGGCAGAGGUCUCAGGUUCGA
>N_E.coli_13
GAGGUCUCAGGUUCGAAUCCUGU
>N_E.coli_14
GAACGGCGGACUGUUAUCCGUA
>N_E.coli_15
CGGACUGUUAUCCGUAUGUCAC
>N_E.coli_16
GUUAAUCCGUAUGUCACUGGUUC
>N_E.coli_17
CCGUAUGUCACUGGUUCGAGUCC
>N_E.coli_18
GUCGGUUAGAAUACCGCCUGUC
>N_E.coli_19
GAAUACCGCCUGUCACGCAGGG
>N_E.coli_20
CACGCAGGGGGUCGCGGGUUCGA
>N_E.coli_21
GGGGUCGCGGGUUCGAGUCCCGU
>N_E.coli_22

UAUGUAGCGGAUUGCAAUCCGU
>N_E.coli_23
CGGAUUGCAAUCCGUCUaGUCC
>N_E.coli_24
GCAAUCCGUCUaGUCCGGUUCG
>N_E.coli_25
GUCUaGUCCGGUUCGACUCCGGA
>N_E.coli_26
UAAGGCACCGUUUUUGAUACCG
>N_E.coli_27
CCGGUUUUUGAUACCGGCAUUC
>N_E.coli_28
UGAUACCGGCAUUCUUGGUUCG
>N_E.coli_29
CCGGCAUUCUUGGUUCGAAUCC
>N_E.coli_30
AGAACGAGAGCUUCCCAAGCUCU
>N_E.coli_31
GCUUCCCAAGCUCUaCGAGGG
>N_E.coli_32
CUAUaCGAGGGUUCGAUUCUU
>N_E.coli_33
AGAGCACGACCUUGCCAAGGUCG
>N_E.coli_34
CCUUGCCAAGGUCGGGUCGCGA
>N_E.coli_35
CAAGGUCGGGUCGCGAGUUCGA
>N_E.coli_36
GGGUCGCGAGUUCGAGUCUCGU
>N_E.coli_37
UGGUAGAGCCCUGGAUUGUGAUU
>N_E.coli_38
GAGCCUUGGAUUGUGAUUCCAGU
>N_E.coli_39
CUGGAUUGUGAUUCCAGUUGUCG
>N_E.coli_40
UGUGAUUCCAGUUGUCGUGGGUU
>N_E.coli_41
UCCAGUUGUCGUGGGUUCGAAUC
>N_E.coli_42
GAGCGACCCCUGAUAAAGGGUGA
>N_E.coli_43
CCUGAUAAAGGGUGAGGUCGGUGG
>N_E.coli_44
UAAGGGUGAGGUCGGUGGUUCA
>N_E.coli_45
GAGGUCGGUGGUUCAAGUCCACU
>N_E.coli_46
GAGCAGGCGACUCAUAAUCGCUU
>N_E.coli_47
GCGACUCAUAAUCGCUUGGUCGC
>N_E.coli_48
UCAUAAUCGCUUGGUCGCUUGUU
>N_E.coli_49
CGCUUGGUCGCUUGGUUCAAGUCC
>N_E.coli_50
GGUaGACGCGCUAGCUUCAGGUG
>N_E.coli_51
ACGCGCUAGCUUCAGGUGUUAGU
>N_E.coli_52
CUAGCUUCAGGUGUUAGUGUCCU
>N_E.coli_53
UCAGGUGUUAGUGUCCUACGGA

>N_E.coli_54
UGUUAGUGUCCUACGGACGUGG
>N_E.coli_55
CCUUACGGACGUGGGGGUUCAAG
>N_E.coli_56
GGACGUGGGGGUUCAAGUCCCC
>N_E.coli_57
GGUaGACACGCUACCUUGAGGUG
>N_E.coli_58
ACACGCUACCUUGAGGUGGUAGU
>N_E.coli_59
CUACCUUGAGGUGGUAGUGCCCA
>N_E.coli_60
UGAGGUGGUAGUGCCCAAUAGGG
>N_E.coli_61
GUAGUGCCCAAUAGGGCUUACGG
>N_E.coli_62
CCCAAUAGGGCUUACGGGUUCA
>N_E.coli_63
GGGCUUACGGGUUCAAGUCCCGU
>N_E.coli_64
GACACAAGGGAUUUAAAAUCCCU
>N_E.coli_65
GGGAUUUAAAAUCCCU CGGCGUU
>N_E.coli_66
AUCCCU CGGCGUUCGCGCUgUGC
>N_E.coli_67
GGCGUUCGCGCUgUGCGGGUUCA
>N_E.coli_68
CGCUgUGCGGGUUCAAGUCCCGC
>N_E.coli_69
GGUaGACGCAGUUGAUUCAAAAU
>N_E.coli_70
ACGCAGUUGAUUCAAAAUCAACC
>N_E.coli_71
UUGAUUCAAAAUCAACCGUAGAA
>N_E.coli_72
AAAAUCAACCGUAGAAUACGUG
>N_E.coli_73
AACCGUAGAAAUACGUGCCGGUU
>N_E.coli_74
AUACGUGCCGGUUCGAGUCCGGC
>N_E.coli_75
UGGUAGAGCAGUUGACUUUUAAU
>N_E.coli_76
GAGCAGUUGACUUUUAAUCAAUU
>N_E.coli_77
UUGACUUUUAAUCAAUUGGUCGC
>N_E.coli_78
GUUAGAGCACAUACUCAUAAUG
>N_E.coli_79
CACAUACUCAUAAUGAUGGGGU
>N_E.coli_80
ACUCAUAAUGAUGGGGUCACAGG
>N_E.coli_81
UAAUGAUGGGGUCACAGGUUCGA
>N_E.coli_82
GGGGUCACAGGUUCGAAUCCCGU
>N_E.coli_83
AGAGCAGGGGAUUGAAAAUCCCC
>N_E.coli_84
GGGAUUGAAAAUCCCGUGUCCU
>N_E.coli_85

GAAAAUCCCCGUGUCCUUGGUUC
>N_E.coli_86
UCCCCGUGUCCUUGGUUCGAUUC
>N_E.coli_87
UGGUAGCGCACUUCGUUCGGGAC
>N_E.coli_88
GCGCACUUCGUUCGGGACGAAGG
>N_E.coli_89
GGACGAAGGGGUCGGAGGUUCGA
>N_E.coli_90
GGGGUCGGAGGUUCGAAUCCUCU
>N_E.coli_91
gAAGGCACCGGUCUUGAAAACCG
>N_E.coli_92
ACCCGAAAGGGUUCAGAGUUCG
>N_E.coli_93
GGGUUCCAGAGUUCGAAUCUCUG
>N_E.coli_94
UCAAAUCCAGUUGGGGCCGCCAG
>N_E.coli_95
GCCGCCAGCGGUCCCAGGGCAGGU
>N_E.coli_96
UCCCGGGCAGGUUCGACUCCUGU
>N_E.coli_97
GAAAACCGGAGUAGGGGCAACUC
>N_E.coli_98
GUAGGGGCAACUCUACCGGGGU
>N_E.coli_99
UCUACCGGGGUUCAAAUCCCC
>N_E.coli_100
AGGCGCUCCCCUGCUAAGGGAGU
>N_E.coli_101
UGCUAAGGGAGUAUGCGGUCAA
>N_E.coli_102
GGAGUAUGCGGUCAAAaGCUGCA
>N_E.coli_103
CGGUCAAAaGCUGCAUCCGGGGU
>N_E.coli_104
UGCAUCCGGGGUUCGAAUCCCCG
>N_E.coli_105
AGGAGCACCCUUGGUAAGGGUG
>N_E.coli_106
CGCCUGGAAAGUGUGUAUACGGC
>N_E.coli_107
GAAAGUGUGUAUACGGCAACGUA
>N_E.coli_108
AUACGGCAACGUAUCGGGGGUUC
>N_E.coli_109
CGUAUCGGGGGUUCGAAUCCCC
>N_E.coli_110
AGAGCGCACCCUUGGUAAGGGUG
>N_E.coli_111
CUUGGUAAGGGUGAGGUCGGCAG
>N_E.coli_112
UAAGGGUGAGGUCGGCAGUUCGA
>N_E.coli_113
GAGGUCGGCAGUUCGAAUCUGCC
>N_E.coli_114
UAGAGCACCGGUCUCCAAAACCG
>N_E.coli_115
CCAAAACCGGGUGUUGGGAGUUC
>N_E.coli_116
GGUGUUGGGAGUUCGAGUCUCUC

>N_E.coli_117
GGGAGAGCACCUCUUACAAGG
>N_E.coli_118
AGCACCUCUUACAAGGAGGGG
>N_E.coli_119
CAAGGAGGGGGUCGGCGGUUCGA
>N_E.coli_120
GGGGUCGGCGGUUCGAUCCCGUC
>N_E.coli_121
CCAGAUUUAGGUUCUGGCGCCGC
>N_E.coli_122
GCGCCGCAAGGUGUGCGAGUUCA
>N_E.coli_123
AGGUGUGCGAGUUCAAGUCUCGC
>N_E.coli_124
AGCAACUGACUUGUAAUCAGUAG
>N_E.coli_125
CUGACUUGUAAUCAGUAGGUCAC
>N_E.coli_126
UAAUCAGUAGGUCACCAGUUCGA
>N_E.coli_127
UAGGUCACCAGUUCGAUUCGGU