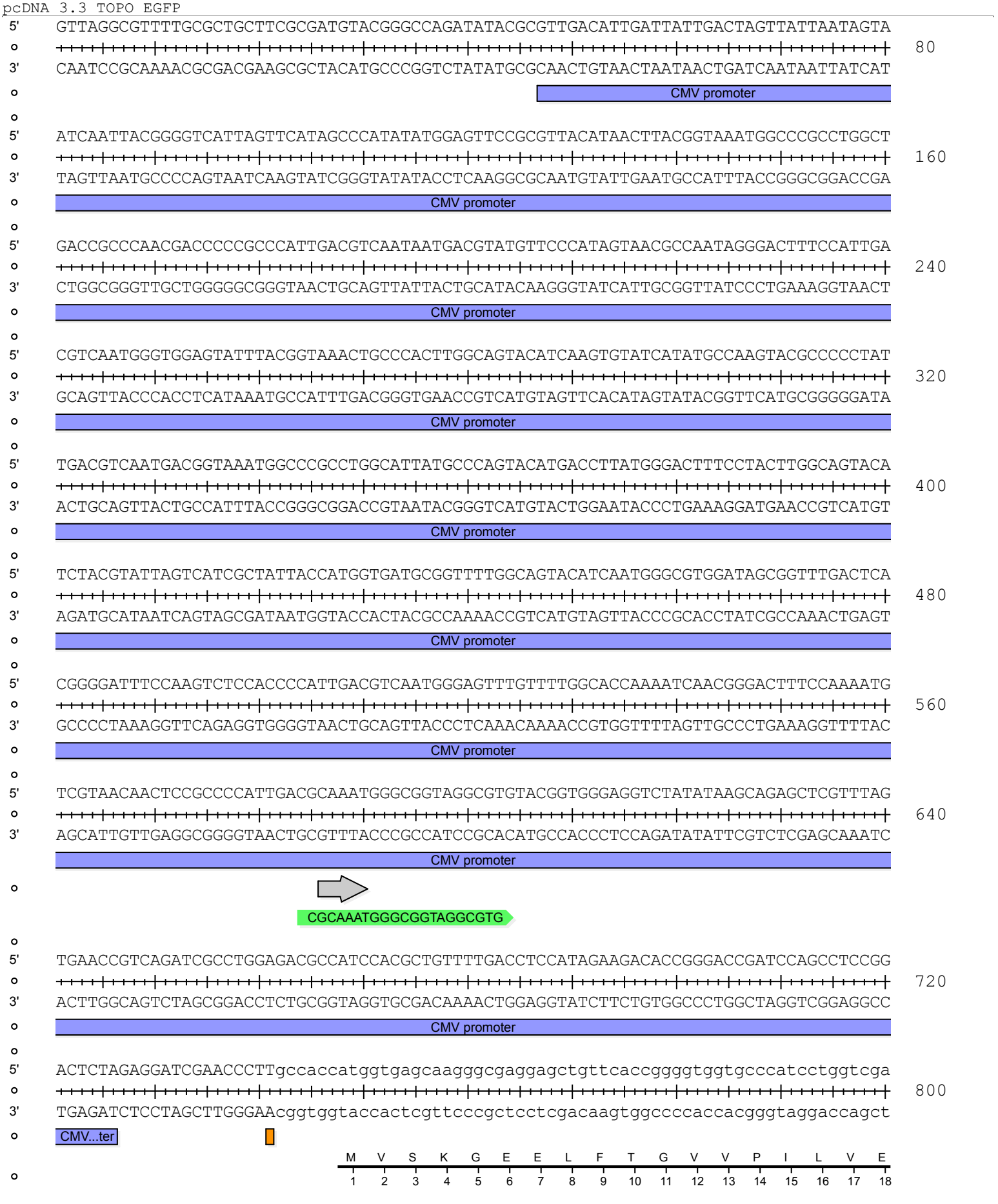
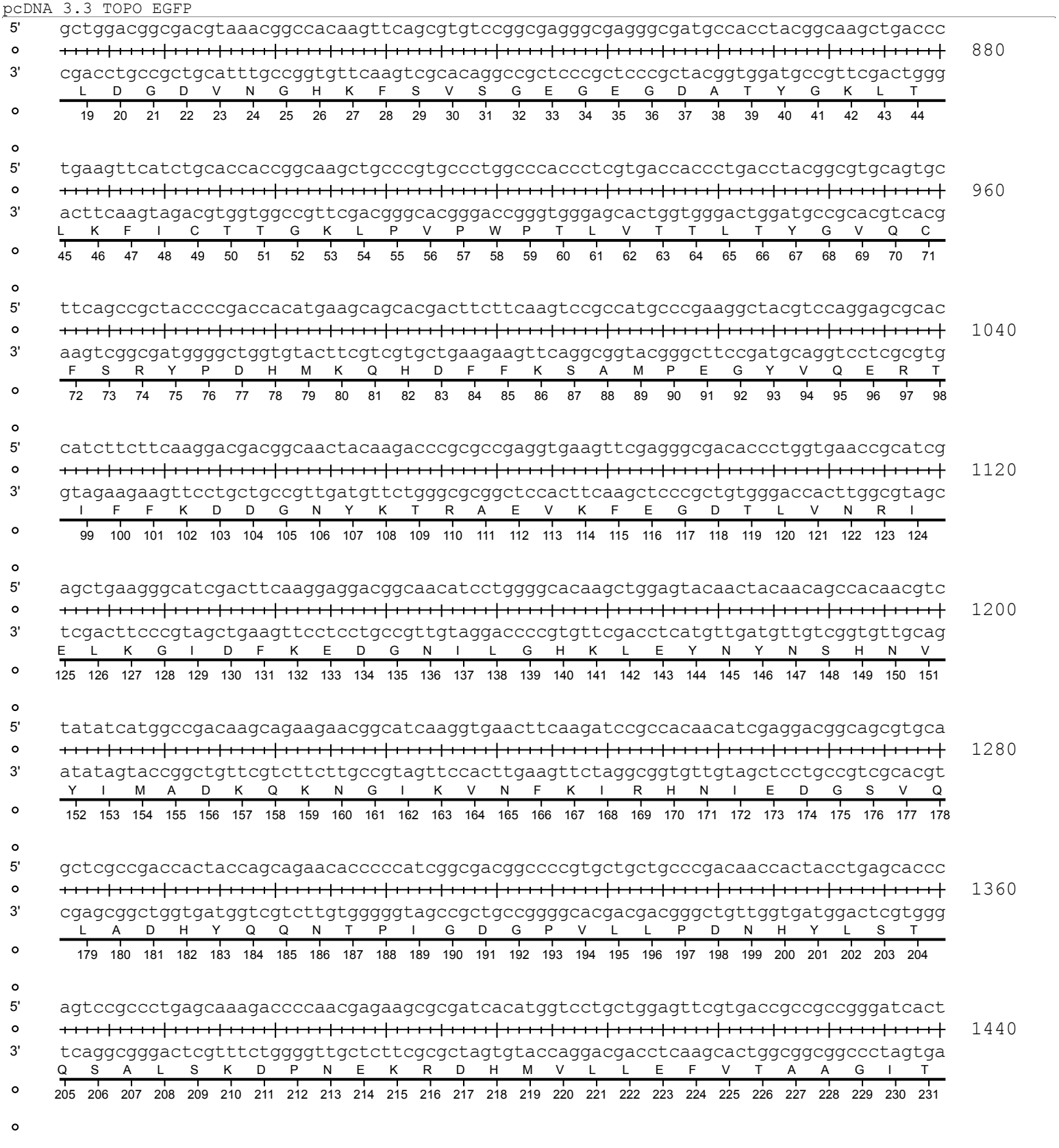

5' GTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTA
3' CAATCCGCAAAACGCGACGAAGCGCTACATGCCCGGTCTATATGCGCAACTGTAACTAATAACTGATCAATAATTATCAT
80
CMV promoter
5' ATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCT
3' TAGTTAATGCCCCAGTAATCAAGTATCGGGTATATACCTCAAGGCGCAATGTATTGAATGCCATTTACCGGGCGGACCGA
160
CMV promoter
5' GACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGA
3' CTGGCGGGTTGCTGGGGGCGGGTAACTGCAGTTATTACTGCATACAAGGGTATCATTGCGGTTATCCCTGAAAGGTAACT
240
CMV promoter
5' CGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTAT
3' GCAGTTACCCACCTCATAAATGCCATTTGACGGGTGAACCGTCATGTAGTTCACATAGTATACGGTTCATGCGGGGGATA
320
CMV promoter
5' TGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACA
3' ACTGCAGTTACTGCCATTTACCGGGCGGACCGTAATACGGGTCATGTACTGGAATACCCTGAAAGGATGAACCGTCATGT
400
CMV promoter
5' TCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCA
3' AGATGCATAATCAGTAGCGATAATGGTACCACTACGCCAAAACCGTCATGTAGTTACCCGCACCTATCGCCAAACTGAGT
480
CMV promoter
5' CGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATG
3' GCCCCTAAAGGTTCAGAGGTGGGGTAACTGCAGTTACCCTCAAACAAAACCGTGGTTTTAGTTGCCCTGAAAGGTTTTAC
560
CMV promoter
5' TCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAG
3' AGCATTGTTGAGGCGGGGTAACTGCGTTTACCCGCCATCCGCACATGCCACCCTCCAGATATATTCGTCTCGAGCAAATC
640
CMV promoter
CGCAAATGGGCGGTAGGCGTG
5' TGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGG
3' ACTTGGCAGTCTAGCGGACCTCTGCGGTAGGTGCGACAAAACTGGAGGTATCTTCTGTGGCCCTGGCTAGGTCGGAGGCC
720
CMV promoter
5' ACTCTAGAGGATCGAACCCTTgccaccatggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcga
3' TGAGATCTCCTAGCTTGGGAAcggtggtaccactcgttcccgctcctcgacaagtggccccaccacgggtaggaccagct
800
CMV...ter
M V S K G E E L F T G V V P I L V E
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18


5' gctggacggcgacgtaaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccacctacggcaagctgaccc 880
3' cgacctgccgctgcatttgccggtgttcaagtcgcacaggccgctcccgctcccgctacggtggatgccgttcgactggg
L D G D V N G H K F S V S G E G E G D A T Y G K L T
19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44
5' tgaagttcatctgcaccaccggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcgtgcagtgc 960
3' acttcaagtagacgtggtggccgttcgacgggcacgggaccgggtgggagcactggtgggactggatgccgcacgtcacg
L K F I C T T G K L P V P W P T L V T T L T Y G V Q C
45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71
5' ttcagccgctaccccgaccacatgaagcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcac 1040
3' aagtcggcgatggggctggtgtacttcgtcgtgctgaagaagttcaggcggtacgggcttccgatgcaggtcctcgcgtg
F S R Y P D H M K Q H D F F K S A M P E G Y V Q E R T
72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98
5' catcttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctggtgaaccgcatcg 1120
3' gtagaagaagttcctgctgccgttgatgttctgggcgcggctccacttcaagctcccgctgtgggaccacttggcgtagc
I F F K D D G N Y K T R A E V K F E G D T L V N R I
99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124
5' agctgaagggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccacaacgtc 1200
3' tcgacttcccgtagctgaagttcctcctgccgttgtaggaccccgtgttcgacctcatgttgatgttgtcggtgttgcag
E L K G I D F K E D G N I L G H K L E Y N Y N S H N V
125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151
5' tatatcatggccgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgca 1280
3' atatagtaccggctgttcgtcttcttgccgtagttccacttgaagttctaggcggtgttgtagctcctgccgtcgcacgt
Y I M A D K Q K N G I K V N F K I R H N I E D G S V Q
152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178
5' gctcgccgaccactaccagcagaacacccccatcggcgacggccccgtgctgctgcccgacaaccactacctgagcaccc 1360
3' cgagcggctggtgatggtcgtcttgtgggggtagccgctgccggggcacgacgacgggctgttggtgatggactcgtggg
L A D H Y Q Q N T P I G D G P V L L P D N H Y L S T
179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204
5' agtccgccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgccgccgggatcact 1440
3' tcaggcgggactcgtttctggggttgctcttcgcgctagtgtaccaggacgacctcaagcactggcggcggccctagtga
Q S A L S K D P N E K R D H M V L L E F V T A A G I T
205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231

5' ctcggcatggacgagctgtacaagtaaAAGGGTTCGATCCCTACCGGTTAGTAATGAGTTTAAACGGGGGAGGCTAACTG
3' gagccgtacctgctcgacatgttcattTTCCCAAGCTAGGGATGGCCAATCATTACTCAAATTTGCCCCCTCCGATTGAC
1520
TK poly A signal
CGAT...TTC
L G M D E L Y K .
232 233 234 235 236 237 238 239 240
5' AAACACGGAAGGAGACAATACCGGAAGGAACCCGCGCTATGACGGCAATAAAAAGACAGAATAAAACGCACGGGTGTTGG
3' TTTGTGCCTTCCTCTGTTATGGCCTTCCTTGGGCGCGATACTGCCGTTATTTTTCTGTCTTATTTTGCGTGCCCACAACC
1600
TK poly A signal
CGATT...CCTTC
5' GTCGTTTGTTCATAAACGCGGGGTTCGGTCCCAGGGCTGGCACTCTGTCGATACCCCACCGAGACCCCATTGGGGCCAAT
3' CAGCAAACAAGTATTTGCGCCCCAAGCCAGGGTCCCGACCGTGAGACAGCTATGGGGTGGCTCTGGGGTAACCCCGGTTA
1680
TK poly A signal
5' ACGCCCGCGTTTCTTCCTTTTCCCCACCCCACCCCCCAAGTTCGGGTGAAGGCCCAGGGCTCGCAGCCAACGTCGGGGCG
3' TGCGGGCGCAAAGAAGGAAAAGGGGTGGGGTGGGGGGTTCAAGCCCACTTCCGGGTCCCGAGCGTCGGTTGCAGCCCCGC
1760
TK poly A signal
5' GCAGGCCCTGCCATAGCAGATCTGCGCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGC
3' CGTCCGGGACGGTATCGTCTAGACGCGTCGACCCCGAGATCCCCCATAGGGGTGCGCGGGACATCGCCGCGTAATTCGCG
1840
TK poly A signal
5' GGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTT
3' CCGCCCACACCACCAATGCGCGTCGCACTGGCGATGTGAACGGTCGCGGGATCGCGGGCGAGGAAAGCGAAAGAAGGGAA
1920
5' CCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGCATCCCTTTAGGGTTCCGATTTAGTGCTTTA
3' GGAAAGAGCGGTGCAAGCGGCCGAAAGGGGCAGTTCGAGATTTAGCCCCGTAGGGAAATCCCAAGGCTAAATCACGAAAT
2000
5' CGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCC
3' GCCGTGGAGCTGGGGTTTTTTGAACTAATCCCACTACCAAGTGCATCACCCGGTAGCGGGACTATCTGCCAAAAAGCGGG
2080
5' TTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATT
3' AAACTGCAACCTCAGGTGCAAGAAATTATCACCTGAGAACAAGGTTTGACCTTGTTGTGAGTTGGGATAGAGCCAGATAA
2160
5' CTTTTGATTTATAAGGGATTTTGGGGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAAT
3' GAAAACTAAATATTCCCTAAAACCCCTAAAGCCGGATAACCAATTTTTTACTCGACTAAATTGTTTTTAAATTGCGCTTA
2240
5' TAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCAT
3' ATTAAGACACCTTACACACAGTCAATCCCACACCTTTCAGGGGTCCGAGGGGTCGTCCGTCTTCATACGTTTCGTACGTA
2320
5' CTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTA
3' GAGTTAATCAGTCGTTGGTCCACACCTTTCAGGGGTCCGAGGGGTCGTCCGTCTTCATACGTTTCGTACGTAGAGTTAAT
2400

```
5'  GTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATG
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2480
3'  CAGTCGTTGGTATCAGGGCGGGGATTGAGGCGGGTAGGGCGGGGATTGAGGCGGGTCAAGGCGGGTAAGAGGCGGGGTAC
o

5'  GCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTT
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2560
3'  CGACTGATTAAAAAAAATAAATACGTCTCCGGCTCCGGCGGAGACGGAGACTCGATAAGGTCTTCATCACTCCTCCGAAA
o

5'  TTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGGAGCTTGTATATCCATTTTCGGATCTGATCAAGAGACAGGATGAGG
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2640
3'  AAACCTCCGGATCCGAAAACGTTTTTCGAGGGCCCTCGAACATATAGGTAAAAGCCTAGACTAGTTCTCTGTCCTACTCC
o

5'  ATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACT
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2720
3'  TAGCAAAGCGTACTAACTTGTTCTACCTAACGTGCGTCCAAGAGGCCGGCGAACCCACCTCTCCGATAAGCCGATACTGA
o

5'  GGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAG
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2800
3'  CCCGTGTTGTCTGTTAGCCGACGAGACTACGGCGGCACAAGGCCGACAGTCGCGTCCCCGCGGGCCAAGAAAAACAGTTC
o

5'  ACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTG
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2880
3'  TGGCTGGACAGGCCACGGGACTTACTTGACGTCCTGCTCCGTCGCGCCGATAGCACCGACCGGTGCTGCCCGCAAGGAAC
o

5'  CGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGT
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  2960
3'  GCGTCGACACGAGCTGCAACAGTGACTTCGCCCTTCCCTGACCGACGATAACCCGCTTCACGGCCCCGTCCTAGAGGACA
o

5'  CATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACC
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3040
3'  GTAGAGTGGAACGAGGACGGCTCTTTCATAGGTAGTACCGACTACGTTACGCCGCCGACGTATGCGAACTAGGCCGATGG
o

5'  TGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGA
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3120
3'  ACGGGTAAGCTGGTGGTTCGCTTTGTAGCGTAGCTCGCTCGTGCATGAGCCTACCTTCGGCCAGAACAGCTAGTCCTACT
o

5'  TCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATC
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3200
3'  AGACCTGCTTCTCGTAGTCCCCGAGCGCGGTCGGCTTGACAAGCGGTCCGAGTTCCGCGCGTACGGGCTGCCGCTCCTAG
o

5'  TCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGC
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3280
3'  AGCAGCACTGGGTACCGCTACGGACGAACGGCTTATAGTACCACCTTTTACCGGCGAAAAGACCTAAGTAGCTGACACCG
o

5'  CGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGC
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3360
3'  GCCGACCCACACCGCCTGGCGATAGTCCTGTATCGCAACCGATGGGCACTATAACGACTTCTCGAACCGCCGCTTACCCG
o

5'  TGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCT
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3440
3'  ACTGGCGAAGGAGCACGAAATGCCATAGCGGCGAGGGCTAAGCGTCGCGTAGCGGAAGATAGCGGAAGAACTGCTCAAGA
o

5'  TCTGAGCGGGACTCTGGGGTTCGCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGC
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3520
3'  AGACTCGCCCTGAGACCCCAAGCGCTTTACTGGCTGGTTCGCTGCGGGTTGGACGGTAGTGCTCTAAAGCTAAGGTGGCG
o

5'  CGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGC
o   ++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  3600
3'  GCGGAAGATACTTTCCAACCCGAAGCCTTAGCAAAAGGCCCTGCGGCCGACCTACTAGGAGGTCGCGCCCCTAGAGTACG
o
```

```
5'    TGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACA
o     +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  3680
3'    ACCTCAAGAAGCGGGTGGGGTTGAACAAATAACGTCGAATATTACCAATGTTTATTTCGTTATCGTAGTGTTTAAAGTGT
o
5'    AATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTC
o     +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  3760
3'    TTATTTCGTAAAAAAAGTGACGTAAGATCAACACCAAACAGGTTTGAGTAGTTACATAGAATAGTACAGACATATGGCAG
o
5'    GACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACA
o     +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  3840
3'    CTGGAGATCGATCTCGAACCGCATTAGTACCAGTATCGACAAAGGACACACTTTAACAATAGGCGAGTGTTAAGGTGTGT
o
5'    ACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCA
o     +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  3920
3'    TGTATGCTCGGCCTTCGTATTTCACATTTCGGACCCCACGGATTACTCACTCGATTGAGTGTAATTAACGCAACGCGAGT
o
5'    CTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCG
o     +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  4000
3'    GACGGGCGAAAGGTCAGCCCTTTGGACAGCACGGTCGACGTAATTACTTAGCCGGTTGCGCGCCCCTCTCCGCCAAACGC
o
5'    TATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACT
o     +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  4080
3'    ATAACCCGCGAGAAGGCGAAGGAGCGAGTGACTGAGCGACGCGAGCCAGCAAGCCGACGCCGCTCGCCATAGTCGAGTGA
o
5'    CAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCC
o     +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  4160
3'    GTTTCCGCCATTATGCCAATAGGTGTCTTAGTCCCCTATTGCGTCCTTTCTTGTACACTCGTTTTCCGGTCGTTTTCCGG
o
5'    AGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTC
o     +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  4240
3'    TCCTTGGCATTTTTCCGGCGCAACGACCGCAAAAAGGTATCCGAGGCGGGGGGACTGCTCGTAGTGTTTTTAGCTGCGAG
o
5'    AAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTG
o     +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  4320
3'    TTCAGTCTCCACCGCTTTGGGCTGTCCTGATATTTCTATGGTCCGCAAAGGGGGACCTTCGAGGGAGCACGCGAGAGGAC
o
5'    TTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGT
o     +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  4400
3'    AAGGCTGGGACGGCGAATGGCCTATGGACAGGCGGAAAGAGGGAAGCCCTTCGCACCGCGAAAGAGTTACGAGTGCGACA
o
5'    AGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGC
o     +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  4480
3'    TCCATAGAGTCAAGCCACATCCAGCAAGCGAGGTTCGACCCGACACACGTGCTTGGGGGGCAAGTCGGGCTGGCGACGCG
o
5'    CTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA
o     +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  4560
3'    GAATAGGCCATTGATAGCAGAACTCAGGTTGGGCCATTCTGTGCTGAATAGCGGTGACCGTCGTCGGTGACCATTGTCCT
o
5'    TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTA
o     +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  4640
3'    AATCGTCTCGCTCCATACATCCGCCACGATGTCTCAAGAACTTCACCACCGGATTGATGCCGATGTGATCTTCCTGTCAT
o
5'    TTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGC
o     +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  4720
3'    AAACCATAGACGCGAGACGACTTCGGTCAATGGAAGCCTTTTTCTCAACCATCGAGAACTAGGCCGTTTGTTTGGTGGCG
o
5'    TGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTT
o     +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  4800
3'    ACCATCGCCACCAAAAAAACAAACGTTCGTCGTCTAATGCGCGTCTTTTTTTCCTAGAGTTCTTCTAGGAAACTAGAAAA
o
```

```
5'  CTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACC
o   +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  4880
3'  GATGCCCCAGACTGCGAGTCACCTTGCTTTTGAGTGCAATTCCCTAAAACCAGTACTCTAATAGTTTTTCCTAGAAGTGG
o
5'  TAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATG
o   +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  4960
3'  ATCTAGGAAAATTTAATTTTTACTTCAAAATTTAGTTAGATTTCATATATACTCATTTGAACCAGACTGTCAATGGTTAC
o
5'  CTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAA
o   +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  5040
3'  GAATTAGTCACTCCGTGGATAGAGTCGCTAGACAGATAAAGCAAGTAGGTATCAACGGACTGAGGGGCAGCACATCTATT
o
5'  CTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTA
o   +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  5120
3'  GATGCTATGCCCTCCCGAATGGTAGACCGGGGTCACGACGTTACTATGGCGCTCTGGGTGCGAGTGGCCGAGGTCTAAAT
o
5'  TCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAA
o   +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  5200
3'  AGTCGTTATTTGGTCGGTCGGCCTTCCCGGCTCGCGTCTTCACCAGGACGTTGAAATAGGCGGAGGTAGGTCAGATAATT
o
5'  TTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGG
o   +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  5280
3'  AACAACGGCCCTTCGATCTCATTCATCAAGCGGTCAATTATCAAACGCGTTGCAACAACGGTAACGATGTCCGTAGCACC
o
5'  TGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTG
o   +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  5360
3'  ACAGTGCGAGCAGCAAACCATACCGAAGTAAGTCGAGGCCAAGGGTTGCTAGTTCCGCTCAATGTACTAGGGGGTACAAC
o
5'  TGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTAT
o   +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  5440
3'  ACGTTTTTTCGCCAATCGAGGAAGCCAGGAGGCTAGCAACAGTCTTCATTCAACCGGCGTCACAATAGTGAGTACCAATA
o
5'  GGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAT
o   +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  5520
3'  CCGTCGTGACGTATTAAGAGAATGACAGTACGGTAGGCATTCTACGAAAAGACACTGACCACTCATGAGTTGGTTCAGTA
o
5'  TCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACT
o   +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  5600
3'  AGACTCTTATCACATACGCCGCTGGCTCAACGAGAACGGGCCGCAGTTATGCCCTATTATGGCGCGGTGTATCGTCTTGA
o
5'  TTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGAT
o   +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  5680
3'  AATTTTCACGAGTAGTAACCTTTTGCAAGAAGCCCCGCTTTTGAGAGTTCCTAGAATGGCGACAACTCTAGGTCAAGCTA
o
5'  GTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGC
o   +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  5760
3'  CATTGGGTGAGCACGTGGGTTGACTAGAAGTCGTAGAAAATGAAAGTGGTCGCAAAGACCCACTCGTTTTTGTCCTTCCG
o
5'  AAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGC
o   +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  5840
3'  TTTTACGGCGTTTTTTCCCTTATTCCCGCTGTGCCTTTACAACTTATGAGTATGAGAAGGAAAAAGTTATAATAACTTCG
o
5'  ATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCAC
o   +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  5920
3'  TAAATAGTCCCAATAACAGAGTACTCGCCTATGTATAAACTTACATAAATCTTTTTATTTGTTTATCCCCAAGGCGCGTG
o
5'  ATTTCCCCGAAAAGTGCCACCTGACGTCGACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTG
o   +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++  6000
3'  TAAAGGGGCTTTTCACGGTGGACTGCAGCTGCCTAGCCCTCTAGAGGGCTAGGGGATACCAGCTGAGAGTCATGTTAGAC
o
```

```
5'   CTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTT
o    +++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++++++++|   6080
3'   GAGACTACGGCGTATCAATTCGGTCATAGACGAGGGACGAACACACAACCTCCAGCGACTCATCACGCGCTCGTTTTAAA
o
5'   AAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGG
o    +++++++++|+++++++++|+++++++++|+++++++++|+++++++++|+++                              6133
3'   TTCGATGTTGTTCCGTTCCGAACTGGCTGTTAACGTACTTCTTAGACGAATCC
o
```