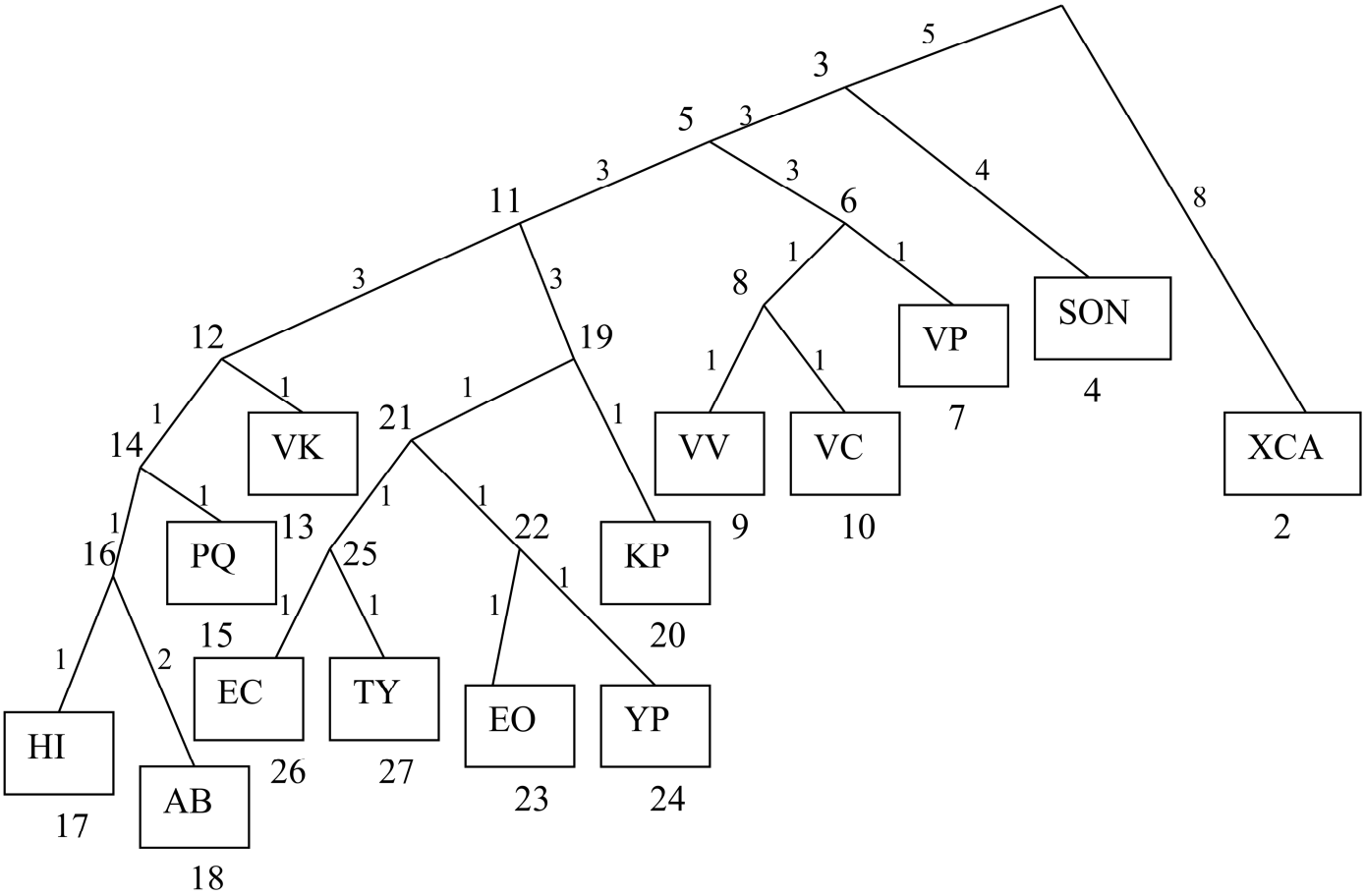
Пример 1. Классическая аттенюаторная регуляция биосинтеза треонина у гамма-протеобактерий.

Дерево видов:

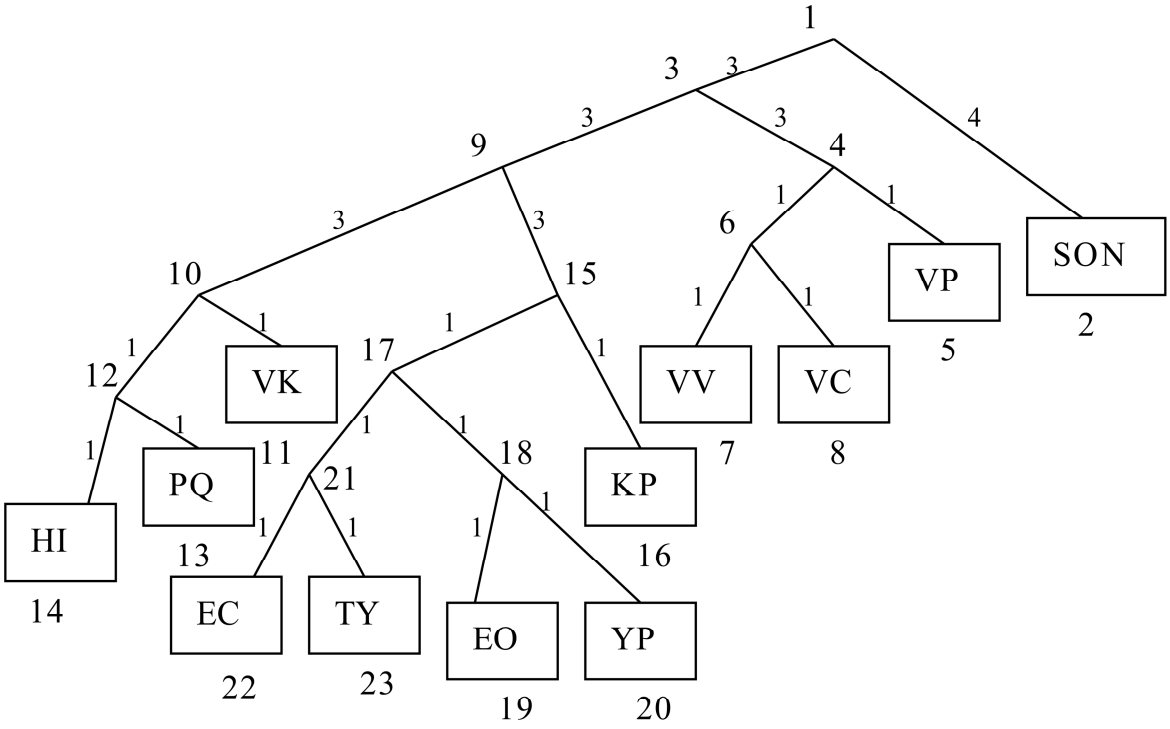


Полученное выравнивание:

```
1:    UGUCGGGGCGGGCUGUCGUAUUCGCCUUAAAGAAGAAAACGACGGCAAAA-GCCCGCACUUCCGACAAAGGA-GUGCGGGC--UUCUUUGUC
2:    GCCCGGUGCGGGCCGUCGUCUUCGCGUAACUUCCGAAAACAACGGC------CCCGCAC--CCGAAUCAGGAUGCGCGGGG--UUCUUUCUC
XCA:  GCCCGGTGCGGTCCGTCGTCTTCGCGTAACTTCCGAAAACAACGGC------CCCGCAC--CCGGATCAGGATGCG-GGGG---TCTCCCTC
3:    UGUGGGGGCGGGCUGCU--AUACACCCUAAAGAAUAUAACGACG--AAAAGGCCCGUACUUCCAACAAAGAA-GUACGGGC-UUUUUUUGUU
4:    AGUGGGGGCGGGCUG----AUACACCCUAAAGAAUAUAACGACG-----AG-CCCG--CUUCCCACAAAGAA---GCGGGC--UUUUUUGUU
SON:  AGTGGGGGCGGGCTG----ATACACCCTAAAGAATTTAACGACG-----AG-CCCG--CTTCCCACAAAGAA---GCGGGC--TTTTTTGTT
5:    UGUUGGGGCAGGCUGCUGAGCGCACCCAAAAG---A-AAUUCAGAAAAAAGGCCUGUACC-CCAACAAGA---GUACAGGCCUUUUUUU-UA
6:    UGUUGGGGCAGGCUGCUGAGCG------AAAG---A-AAUUCACAAAAAAGGCCUGUAUC-CCAACAAGA----UACAGGCCUUUUUUU--A
7:    UGUUGGGGCAGGCUGCUGAGCG------AAAG---A-AAUUCACAAAAAAGGCCUGUAUC-C-AACAAGA----UACAGGCCUUUUUUU--A
VP:   TGTTGGGGCAGGCTGCTGAGCG------AAAG---A-AATTCACAAAAAAGGCCTGTATC-C-AACAAGA----TACAGGCCTTTTTTT--A
8:    UGUUGGGGCAGGCUGCUGAGCG------AAAG---A-AAUUCACAAAAAAGGCCUGUAUC-CCAACAAGA----UACAGGCCUUUUUUU--A
9:    UGUUGGGGCAGGCUGCUGAGCG------AAAGAACA-AAUUUCAAAAAAAGGCCUGUAUC-C-AACAAGA----UACAGGCCUUUUUUU--A
VV:   TGTTGGGGCAGGCTGCTGAGCG------AAAGAACA-AATTTCAAAAAAAGGCCTGTATC-C-AACAAGA----TACAGGCCTTTTTTT--A
10:   UGUUGGGGCAGGCUGCUGAGCG------AAA----A-AAUUCACAAAAAAGGCCUGUAUC-CCAACAAGA----UACAGGCCUUUUUUU--A
VC:   TGTTGGGGCAGGCTGCTGAGCG------CAA----A-ATTTCACAAAAAAGGCCTGTATC-CCAACC-GA----TACAGGCCTTTTTTT--A
11:   UGUUGGUGCGGGCUGCUGUGCGUACCAAAAAGA-CA-AAUUCACAAAAAAG-CCCGUACC-U-AACCUGA--AGUACGGGCUUUUUUUU-UA
12:   AUAGUGUGCGGGUU--AGUGCGUAACAAAAAGAUCGAAUUCCAC--AAAA--CCCGUAC--UGAAUCAAA--AGUGCGGG--UUUUUUUAUG
13:   AUAGUGUGCGGGUU--AGUGCGUAACAAAAAGAUCGAAUUCCAC--AAAA--CCCGUAC--UGAAUAAAA--AGUGCGGG--UUUUUUUAUG
VK:   ATAGTGTGCGGGTT--AGTGCGTAACAAAAAGATCGAATTCCAC--AAAA--CCCGTAC--TGAATAAAA--AGTGCGGG--TTTTTTTATG
14:   -UAGAGUGCGGGUUU-AGUGCGCUACAAAAAGAUCAGAAUAAAC--AAAA--CCCGUAC--UACA-AAAA--A-UGCGGG--UUUUUUUGUA
15:   -CAUAGUGCGGGUUUUAAUGCGCUGAAAUAAUGAAAGAAUAAACCGAAAA--CCCG--C--UACA---------AGCGGG--UUUUUUUGUA
PQ:   -CATAGTGCGGGTTTTAATTGGCTGAAATAATGAAAGAATAAACCGAAAA--CCCG--C--TACA---------AGCGGG--TTTTTTTGTA
16:   A-AUAGUGCGGGUU--AGUGCGCC-AAAAAAGAACAAAAUACA--GAAAA--CCCG--C-AUUCA-AAGA--AUAGCGGG--UUUUUUUAUA
17:   A-AUGGUGCGGGUU--AGUGCAGC-AAA----AACAAGAUACA--GAAAA--CCCG--CGAUUCAACUGA--AUAGCGGG--UUUUUUUAUA
HI:   A-ATGGTGCGGGTT--AGTGCAGC-AAA----AACAAGATACA--GAAAA--CCCG--CGATTCAACTGA--ATAGCGGG--TTTTTTTATA
18:   A-AUGGUGCGGGUU--AGUGCGUUGAAA----AACAGAAUACA--GAAAA--CCCG--C-AUUUACCCGA--GUAGCGGG--UUUUUUUAUA
AB:   A-ATGGGGCGGGCT--AGTGCGTTGAAG----AATAGAATTCAT-G--AA--CCCG--C-ATTT-CCCGA--G-AGCGGG--TTTTTTTTATG
19:   UAACGGUGCGGGCU--GACGCGUACAGGAAACA-CAGAA-------AAAAG-CCCG--C-----ACCUGAACAGUGCGGGCUUUUUUUUUGA
20:   UAACGGUGCGGGCU--GACGCGUACAGGAAACA-CAGAA-------AAAAG-CCCG--C-----ACCUGAACAGUGCGG--UUUUUUUUUGA
KP:   TAACGGTGCGGGCT--GACGCGTACAGGAAACA-CAGAA-------AAAAG-CCCG--C-----ACCTGAACAGTGCGG--TTTTTTTTTGA
21:   UAACGGUGCGGGCU--GACGCGUACAGGAAACA-CAGAA-------AAAAG-CCCG--C-----ACCUGAACAGUGCGGGCUUUUUUUUUGA
22:   UAACGGUGCGGGCU--GACGCAUACAGGAAACA-CAGAA-------AAAAG-CCCG--C-----ACCUGAACAGUGCGGGCUUUUUUUUU--
23:   UAACGGUGCGGGCU--GACGCAUACAGGAAACACCAGAA-------AAAAG-CCCG--C-----ACCGA-ACAGUGCGGGCUUUUUUUUU--
EO:   TAACGGTGCGGGCT--GACGCATACAA-AGATTCCAGAA-------AA-AG-CCCG--C-----ACCGA-ACAGTGCGGGCTTTTTTTTT--
24:   UAACGGGGCGGGCU--GACGCGUACAGGAAACAACAGAA-------AAAAG-CCCG--C-----ACCUAGACAGUGCGGGCUUUUUUUUU--
YP:   TTACGGGGCGGGCT--GACGCGTACAGGAAACAATAGAA-------AAAAG-CCCG--C-----ACCTAGACAGTGCGGGCTTTTTTTTT--
25:   UAACGGUGCGGGCU--GACGCGUACAGGAAACA-CAGAA-------AAAAG-CCCG--C-----ACCUGAACAGUGCGGGCUUUUUUUUU--
26:   UAACGGUGCGGGCU--GACGCGUACAGGAAACA-CAGAA-------AAAAG-CCCG--C-----ACCUGAACAGUGCGGGCUUUUUUUUU--
EC:   TAACGGTGCGGGCT--GACGCGTACAGGAAACA-CAGAA-------AAAAG-CCCG--C-----ACCTGA-CAGTGCGGGCTTTTTTTTT--
27:   UAACGGUGCGGGCU--GACGCGUACAGGAAACA-CAGAA-------AAAAG-CCCG--C-----ACCUGAACAGUGCGGGCUUUUUUUUU--
TY:   TAACGGTGCGGGCT--GACGCGTACAGGAAACA-CAGAA-------AAAAG-CCCG--C-----ACCTGAACAGTGCGGGCTTTTTTTTC--
```

Пример 2. Классическая аттенюаторная регуляция биосинтеза лейцина у гамма-протеобактерий.

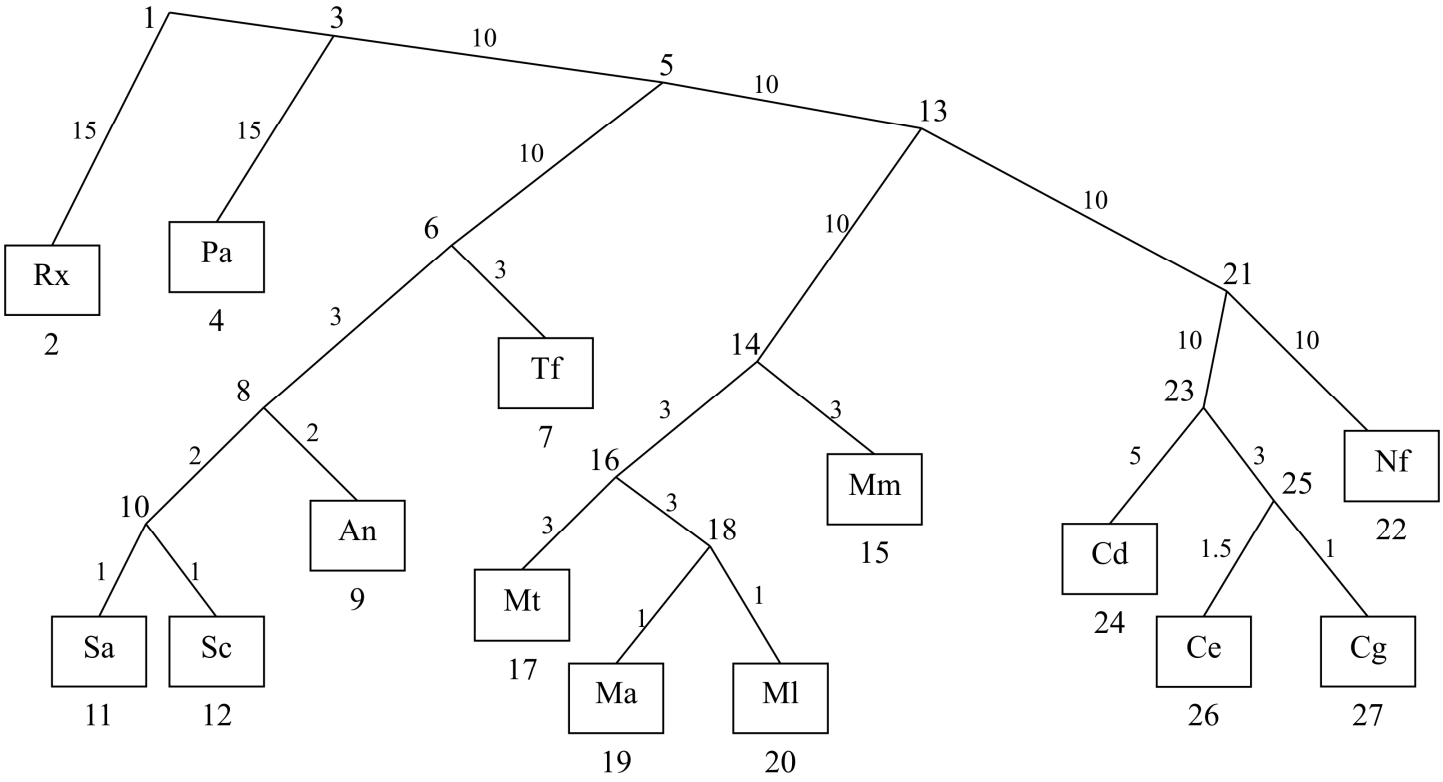
Дерево видов:



Полученное выравнивание:

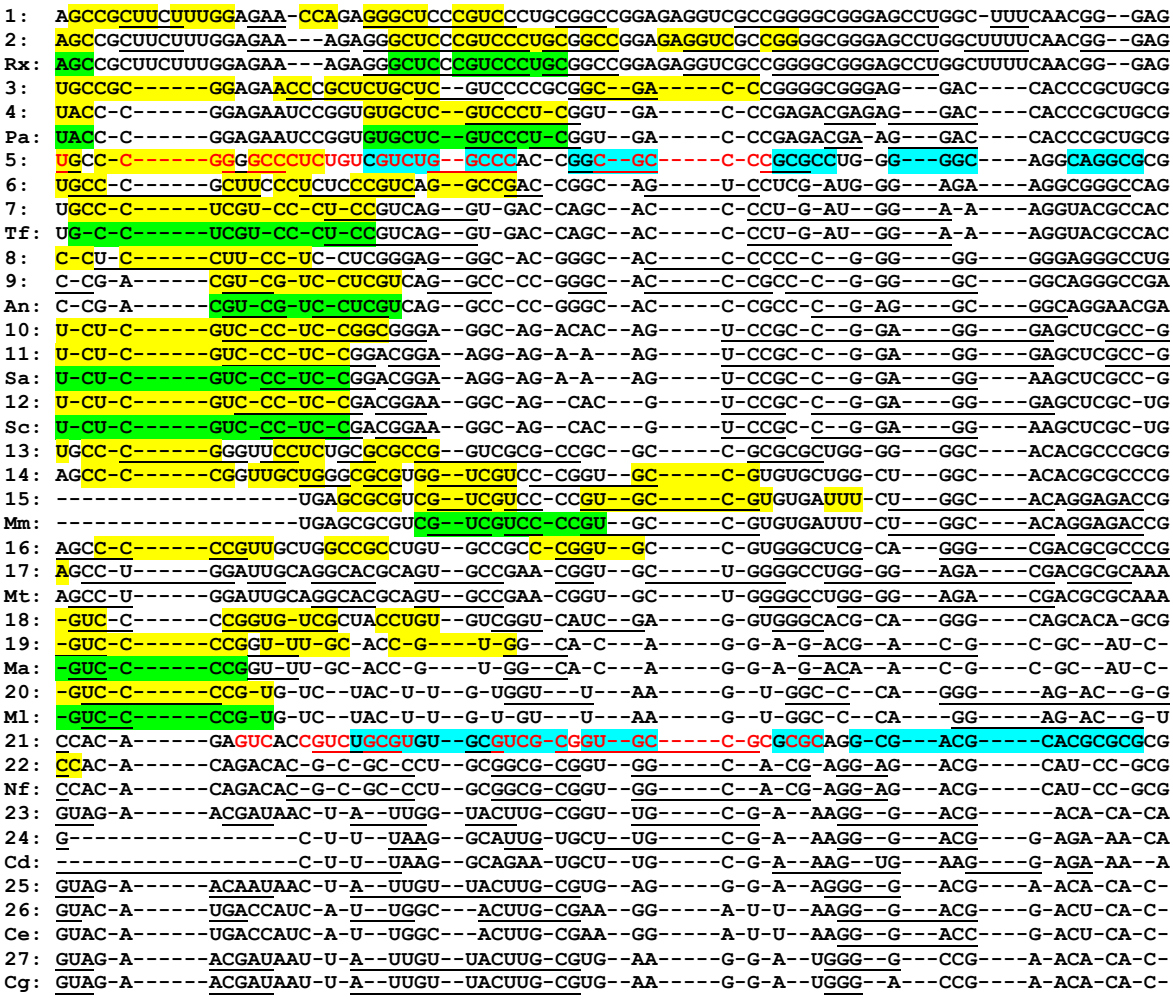
```
1:    AAAACGCGCA----------GCGCGGGUGGGCUGAGGGUUGCAAACUCGCCAGAUAGGCAAAACACUCAUAAACCCCGCACGUAAUGUU--GCGGGGUUUUUUUGUA
2:    AAAACGCGGA----------------GGUCUUAGUGUUGCUCGCUCGAUAGAUAGGCAAAACACUCAUAAACCCCGCAC-UAAUGUU--GCGGGGU-UUUUUGUA
SON:  AAAACGCGGA----------------GGTCTTAGTGTTGCTCGCTCGATAGATAGGCAAAACACTCATAAACCCCGCAC-TAATGTT--GCGGGGT-TTTTTTGTA
3:    UAUUCGCGCA--G------CGCGCGGGUGGGCUGUGGAAGAAAAAUUCGCCAGACCGAAUUUACACUAAAAAACCC-GCAUGUAAUAUUGCGCGGG-UUUUUUUAUA
4:    UAU----------------CGCGCGGGUAGGCUGUGGAAGAAAAAUA-ACCACACCAA-UUUCAUAUAGAA-ACCC-GCAUG-AAUAAU--GCGGG-UUUUUUUAUA
5:    U-U----------------CGCGCGGGUAGGCUGUGGAAGAAAAAUA-ACCACACCAA-UUUC-U-UAGAA-ACCC-GCAUG-AA-AAU--GCGGG-UUUUUUUAUA
VP:   T-T----------------CGCGCGGGTAGGCTGTGGAAGAAAAATA-ACCACACCAA-TTTC-T-TAGAA-ACCC-GCATG-AA-AAT--GCGGG-TTTTTTTATA
6:    UAU----------------CGCGCGGGUAGGCUGUGGAAGAAAAAUA-ACCACACCAAA-UU-AUAUAAAA-ACCC-GCACAU-AGAAU--GCGGG-UUUUUUUAUA
7:    -AU----------------CGCGCGGGUAGGCUGUGGAAGAAAAAUAAACCACACAGAA-UA-ACA-ACUA-GCCC-GCACAU-CG-AU--GCGGG-CUUUUUUAUA
VV:   -AT----------------CGCGCGGGTAGGCTGTGGAAGAAAAATAAACCACACAGAA-TA-ACA-ACTA-GCCC-GCACAT-CG-AT--GCGGG-CTTTTTTATA
8:    -AU----------------CACGCGGGUAGGCUGUGGACAAAAACACCAC------ACA-AG-AUA-AAAA-ACCC-GCAGCU--G-AU--GCGGG-UUUUUUUAUA
VC:   -AT----------------CACGCGGGTAGGCTGTGGACAAAAACACCAC------ACA-AG-ATA-AAAA-ACCC-GCAGCT--G-AT--GCGGG-TTTTTTTATA
9:    UAUUUG-GCGCGGUUAGGGUGUUGGGUUAGGUUGUGGAAUAAAAAAUCGAAUUAUUCCAUUUUGAUUAGAAAACCC-GCGCGUUAUUGUAUGCGGG-UUUUUUUAUA
10:   UCAUUGUGCG--GCUA-GGU--UG--UGGA-UA----AAAAAAAA-UAAAA-U-AUCC-CACAAAUUAG-AAACCC-GCACCUUAAAAGAUGCGGG-UUUUUUUUUA
11:   UCAUUGUGCG--GCUA-GGU--UG--UGGA-UA----AAAAACAG-UAAAA-U-AUCC-CACAAAUUAG---ACCC-GCAC----AUGUAAGCGGG-UCUUUUUUUA
VK:   TCATTGTGCG--GCTA-GGT--TG--TGGA-TA----AAAAACAG-TAAAA-T-ATCC-CACAAATTAG---ACCC-GCAC----ATGTAAGCGGG-TCTTTTTTTA
12:   UCUUUGUGCG--GCUAAGGUUGUG--UGGAAUA----AAAAAAAAGUAAA--UAAUCCACACAACUUUG-AAACCC-GCACCUUAAAAGUUGCGGG-UUUUUUUUUA
13:   AUUUUGUGCG--GAUAAGGUAUUGAUUGGAA-A----AAGUAAAUG-GAC--UAUCC-ACAUUACUU-G---ACCC-GCACCUUAAAA--UGCGGG-UUUUUUUUAU
PQ:   ATTTTGTGCGAGGATAAAG-ATTGATTGGAA-A----AAGTAAATG-GAC--TATCC-ACATT-CTT-G----CCC-GCACCTTAAAA--TGCGGG-CTTTTTTTAT
14:   UUUGUGUGCG--GCUAAG-UUGUGGAUAAAA-A----AACAAUCAG-AUG--UAAAU-ACAC-AAUUUG-AAACCC-GCACUUUAUAAGUUGCGGG-UUUUUUUUCU
HI:   TTTGTGTGCG--GCTAAG-TTGTGGATAAAA-A----AACAGTCAG-ATG--TAAAT-ACCC-AATTT-TAAACCC-GCACTTTATAAGTTGCGGG-TTTTTTATCT
15:   CAUUUGUGCGCGGUUAAGCUGUUGGGCGGCGUUCAGCGUUAAGUCAUCGCUUCAUUCCAGCAGGACUAUAAAACCC-GCGCCU---UGUGCGCGGG-UUUUUUUAUG
16:   CAUUUG-GCGCGGU-AGGCUGUUGGGCGACGUUCAGCGUUAAGUCAUC------UUCCAGCAAGACUAUAAAACCC-GCGCCU---UG-GCGCGGG-UUUUUUUAUG
KP:   CATTT--GCGCGGT-AGGCTGTTGGGCGACGTTCAGCGTTAAGTCATC------TTCCAGCAAGACTATAAAACCC-GCGCCT---TG-GCGCGGG-TTTTTTTATG
17:   CAUCUGUGCGCGGGUAAGAUGUUGUGCGAUGCUUUCAGUUCAGUAAUUGAAUCAUCGCAGAUGAAUCAAAAAACCC-GCGCCU---UGUGCGCGGG-UUUUUUUAUG
18:   UCUCU-UG-CGCGGU-AGACCGAGUGUGCGGCAUUCCAAUCAGUAAGU-CAGCAUCGCAAGUCAAACAAAAAACCC-GCGCCG---U-UGCGCGGG-UUUUUUUAUG
19:   UCU-U-UG-CGCGGU-AGAC-GAGUGAGCGGCAU-CCA-GCAUUAAG-CCAGCA-CGC-AGUCAAACAAAAAACCC-GCGCCA---U-UGCGCGGG-UUUUUUUAUG
EC:   TCT-T-TG-CGCGGT-AGAC-GAGTGAGCGGCAT-CCA-GCATTAAG-CCAGCA-CGC-AGTCAAACAAAAAACCC-GCGCCA---T-TGCGCGGG-TTTTTTTATG
20:   AUUGU-UG-CGCGGU-AGACCGGGUGGGCGGCAUU-CAA-CAUUAAGU-CAGC-UCG-AAGUCAAACAAAA--CCC-GCGCCG-----UGUGCGGG-UUUUUUUAUG
TY:   ATTG--TG-CGCGGT-AGACC-GGTGGGCGGCATT-CAA-CATTAAGT-CAGC-TCG-AAGTCAAACAAAA--CCC-GCGCCG-----TGTGCGGG-TTTTTTTATG
21:   UCUCU-UG-CGCGGU-AUGGUUUGUGUGGGGCACAGAAAUCAGUAUAUGAUUCUUGCCACAAGAAUCAAAAGACCC-GCGCAA----AUGCGCGGG-UUUUUUUAUG
22:   U-UCU-UG-CGCGGU-A-GGUU--UGUGG-GCAGACUUCAGAAC-UAAGUUUCUCGCCACACGAUACAAAA-ACCC-GCGCUG----AUGCGCGGG-UUUUUUUUUA
EO:   T-TCT-TG-CGCGGT-A-GGTT--TGTGG-GCAGACTTCAGAAC-TAAGTTTCTCGCCACACGATACAAAA-ACCC-GCGCTG----ATGCGCGGG-TTTTTTTTTA
23:   -CUCU-UG-CGCGGU-AUG-----UGUGGUGGACGGAAAUCAG-AACUGAUUCAGCCAUCAAGAUUAACAAG-CCC-GCGCAA----AUGCGCGGG-UUUUUUUUGU
YP:   -CTCT-TG-CGCGGT-ATG-----T-TGGTGGACGGAAATCAG-AACTGATTCAGCCATCAAGATTAACAAG-CCC-GCGCAA----ATGCGCGGG-TTTTTTTTGT
```

Пример 3. Т-боксовая регуляция гена *ileS* у актинобактерий.
Дерево видов:



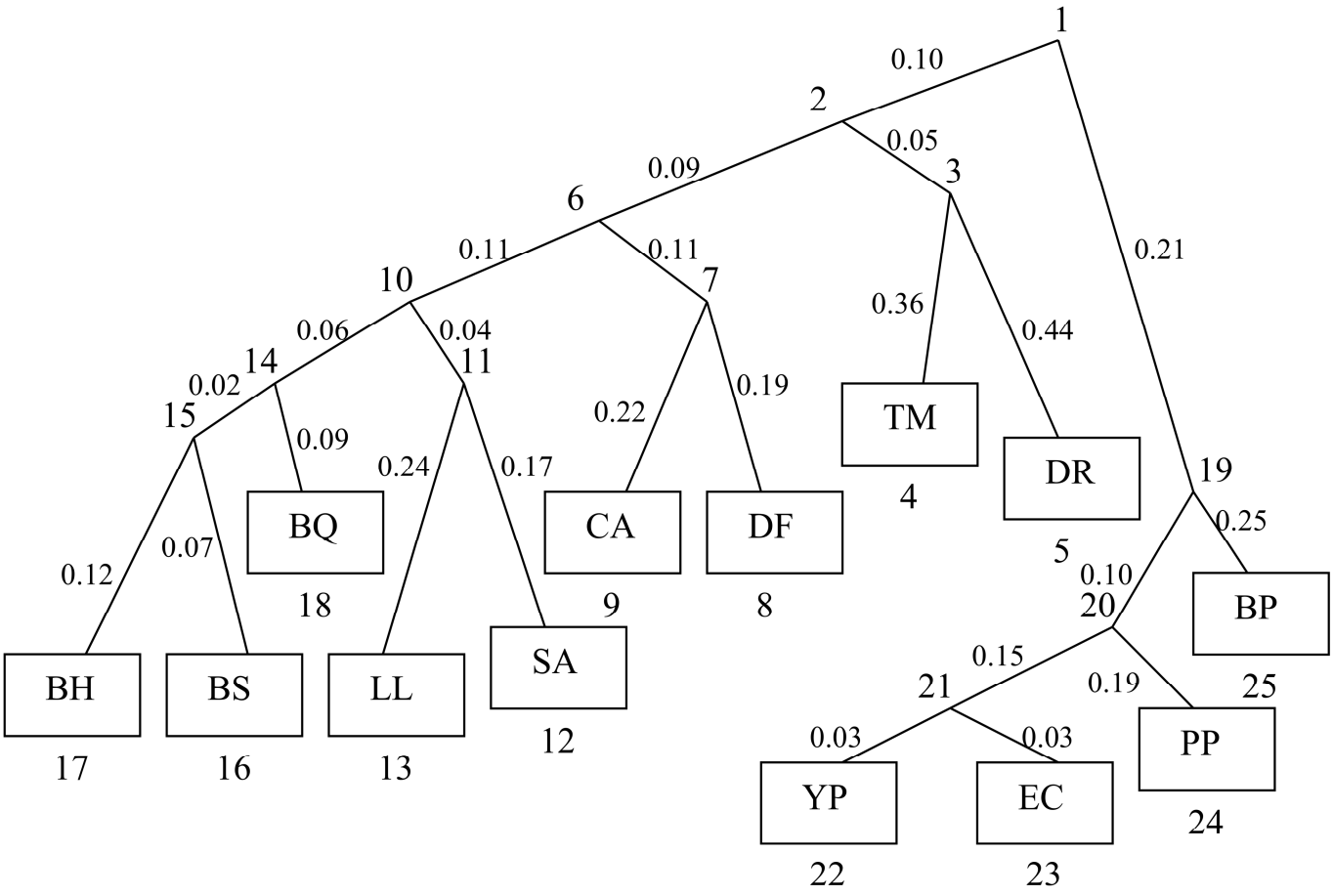
Полученное выравнивание:

```
1:  CGGUGCCG-CGAGGCCUCGU----GGCCAAGCAGGGUGGUACCGCG-----------------------------UGGUACCGCGGG
2:  GGG-GCCG-CGAGGCCUCG-----GG-CAAGCAGGG-----------------------------------------UGGUACCGCGAG
Rx: GGG-GCCG-CGAGGCCUCG-----GG-CAAGCAGGG-----------------------------------------UGGUACCGCGAG
3:  CGGUGCCGACGAGG--UCGUGCAAGG---AGCAGGGUGGUACCGCG------GCGC-------------------UGGUACCGCGGG
4:  CGACGUCGUUGACG--UCGUGCAAGG---AG--GG-----------------------------------------UGGUACCGCGGG
Pa: CGACGUCGUUGACG--UCGUGCAAGG---AG--GG-----------------------------------------UGGUACCGCGGG
5:  CGGGGGCGACCGGG--CCGCGCGGGGCAAGCAGGGUGGUACCGCG------GCGCUGCACCGGCCCGCCGCCAGCUGGUGCCGCGUG
6:  AGAGAGCGAGCGGC--CCGCG-GCGGCCAAGGAGGGUGGUACCGCG------GGGCUGCACCAGCCGGGC-CCAGCCC-UGUCGCGUG
7:  AG-GA-CGA-CGG---CCGC--GCGGCCAAGGAGGGUGGUACCGCG------GGG--GC-------------------------GUC
Tf: AG-GA-CGA-CGG---CCGC--GCGGCCAAGGAGGGUGGUACCGCG------GG---GC-------------------------GUC
8:  A-AU-GAG-GCG-C--CCG-G-GGGGCCAAGGAGGGUGGUACCGCG------GGGCGGCACCAGCCGGGCACCAGCCC-GGUCGCGGG
9:  G-AU-GGG-GCG-C--GCA-G-UACGCCAAGCGAGGUGGUACCGCG------GUGCGGCACCAGCCGGGCACCAGCCCCGGUCGGGAG
An: G-AU-GGG-GCG-C--GCA-G-UACGGCAAGCGAGGUGGUACCGCG------GUGCGGCACCAGCCGGGCACCAGCCCCGGUCGGGAG
10: A-AC-GAG-GCC-C--CCG-G-GGGGCCAAGGAGGGUGGUACCGCG------GGAGCGCGCCGC-CACGGCGUACG----G-CUCGGC
11: A-CA-CAG-GGC-G--CCG-G-GGAGCCAAGGAGGGUGGUACCGCG------GGAGCGCGCCGCACACGGCGUACGGAAAGACUCGGC
Sa: A-CA-CAG-GGC-G--CCG-G-GGAGCCAAGGAGGGUGGUACCGCG------GGAGCGCGCCGCACACGGCGUACGGAAAGACUCGGC
12: C-AC-GAC-GCA-C--CGG-C-CGGGCCAAGGAGGGUGGUACCGCG------GGAGCACGCCG-----GGCG---GG-------CGGC
Sc: C-AC-GAC-GCA-C--CGG-C-CGGGCCAAGGAGGGUGGUACCGCG------GGAGCA-------------------------CGGC
13: CGGCGGCGUCCGGG--GCGCGCGGGGGCAAGCGGGGUGGUACCGCG----C-GCGCUCCGGGCGCGCACCGACGUCGGGUCCCGCGUG
14: CGGCCGCGACUAUC--GCGGGU-GCGGCAAGCGGGGUGGUACCGCG------GCGCUC-----GCGCACCGGCGCGGCGUCGUCCCCG
15: CGGCCGC-ACU------CAGGU-GCGGCAAGCGGGGUGGUACCGCG------GCGCUC-----GCGCAC--------------------
Mm: CGGCCGC-ACU------CAGGU-GCGGCAAGCGGGGUGGUACCGCG------GCGCUC-----GCGCAC--------------------
16: CGGCCGCGACUAAC--GCCGGU-GCGGCAAGCGGGGUGGUACCGCG------GCGCUC-----GCGCACCGGCGCGGCGUCGUCCCCG
17: CGGCCGCG-C-AUC--GGCGG--GCGGCAAGCGGGGUGGUACCGCG------GCGCUC-----GCGCACCGGCGUGGCGUCGUCCCCG
Mt: CGGCCGCG-C-AUC--GGCG----UGGCAAGCGGGGUGGUACCGCG------GCGUUC-----GCGCACCGGCGUGGCGUCGUCCCCG
18: GGGCCGCG-CGAAU--GCGCGU-GCGGCAAGCGGGGUGGUACCGCG------GCGCUC-----GCGCACCC-AGC-GCGUCGUC----
19: UGGCCACG-CGAAA--GCGCG--GC---AAGCGGGGUGGUACCGCG------GCGCUC-----GCGCAGCC-AGC-GCGUCGUC----
Ma: UGGCCACG-CGAAA--GCGCG--GC---AAGCGGGGUGGUACCGCG------GCGCUC-----GCGCAGCC-AGC-GCGUCGUC----
20: GCCGUGCG-----U--UCGCGU-GCGGCAAGCGGGGUGGUACCGCG------GCGCUC-----GCGCACC-UAGC-GCGUCGUC----
Ml: GCCGUGCG-----U--UCGCGU-GCGGCAAGCGGGGUGGUACCGCG------GCGCUC-----GCGCACC-UAGC-GCGUCGUC----
21: CGGGCGCGUCCGGA--GCGCUC-GCGACAAGCGGGGUGGUACCGCGGUC-CGGCGCUCCUGGCGC----CGAGGUCGUCCCCGCUAGG
22: CGGU-GCGUCC-GA---CGC-C-G-GACAAACGGGGUGGUACCGCGGUUUCGGCGCUCCGGGCGC----CGAGGUCGUCCCCG-U-GC
Nf: CGGU-GCGUCC-GA---CGC-C-G-GACAAACGGGGUGGUACCGCGGUUUCGGCGCACCGGGCGC----CGAGGUCGUCCCCG-U-GC
23: AGGGC-UAGG-GAA--G-A-UA-GC-UCAAGCGGGGUGGUACCGCGCUC-C-GUU-UUUUAGGGC--------GU---CCCCGC-A-G
24: AUGCC-UCGG-GUA--G-A-AU-GC-UCAAGCGGGGUGGUACCGCGCUC-C-GAA-U----GGGC--------GU---CCCCGC-A--
Cd: AUGCC-UCUG-GUG--G-A-AU-GC-UCAAGCGGGGUGGUACCGCGC--G-GAAA-------CGC--------GU---CCCCGC-A--
25: UGUGC-UAGG-GAA--G-U-UA-GC-UCAAGCGGGGUGGUACCGCG-UC-C-GUU-UUUUAGGGC--------GC---CCCCGC-A-G
26: UGUUG-GUGG-GCC--G-C-AG-GU-UCAAGCGGGGUGGUACCGCG-UC-C-GGA-UCAAGGGGC--------GU---CCCCGC-A-A
Ce: UGUUG-GUGG-GCC--G-C-AG-GU-UCAAGCAGGGUGGUACCGCG-UC-C-GGA-UCAAGGGGC--------GU---CCCCGC-A-A
27: GGAGC-UAGU-UAA--U-U-UA-GC-UCAAGCGGGGUGGUACCGCG-UC-C-GUU-UUUUAGGGC--------GC---CCCCGC-A-G
Cg: GGAGC-UAGU-UAA--U-U-UA-GC-UCAAGCUGGGUGGUACCGCG-UC-C-GUU-UUUUAGGGC--------GC---CCCCGC-A-G
```



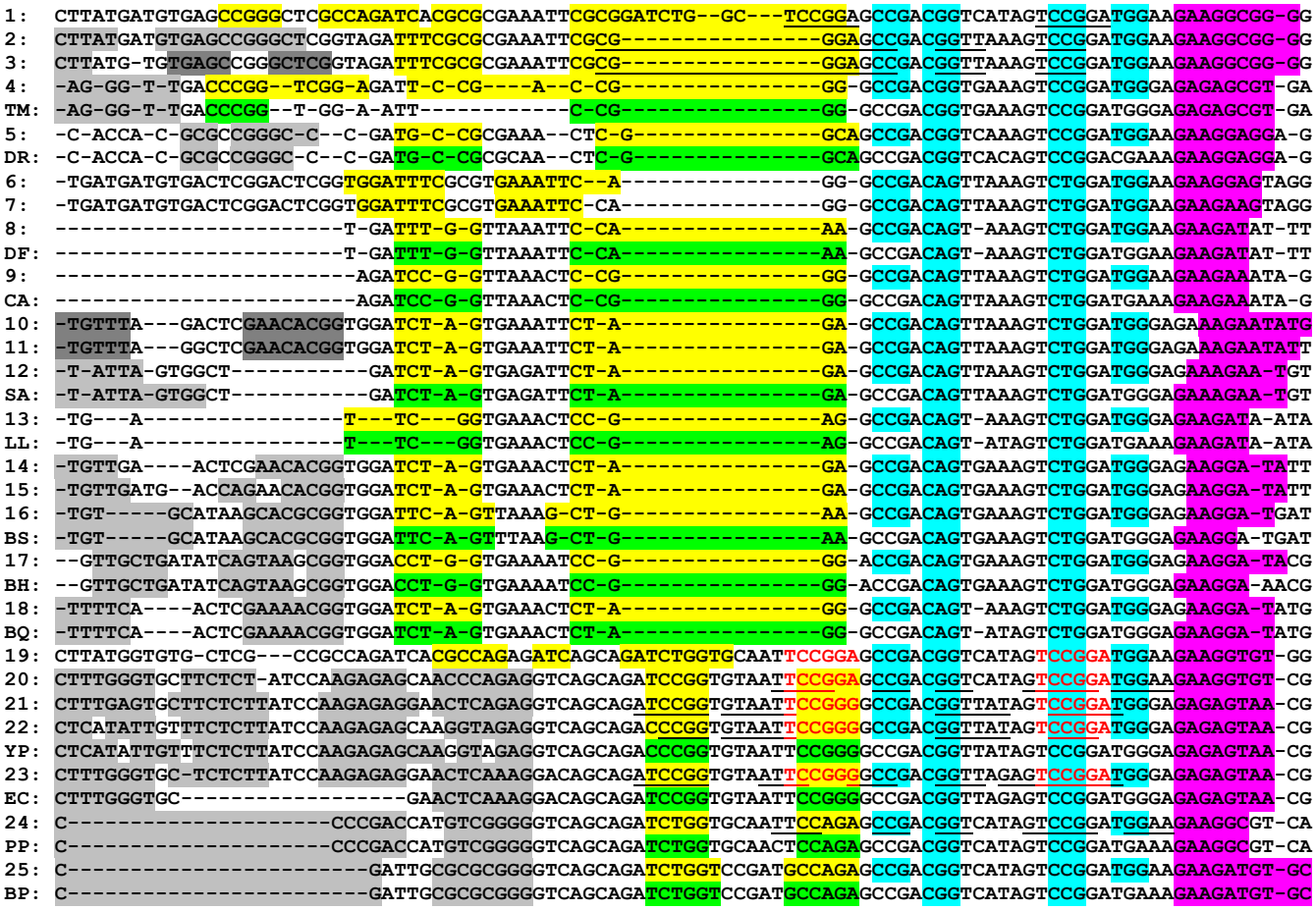
Пример 4. RFN- регуляция экспрессии генов биосинтеза и транспорта рибофлавина у эубактерий (ген *ribB* у BP, EC, PP, YP и ген *ribD* у остальных видов).

Дерево видов:



Полученное выравнивание:

```
1:  AGATCTGTCTTCAGGGCGGG----G-GGTGAAATTCCC-CACCGGCGGTAAATCGGAATGTCGCC-GCTAGCCCAGCGG-GCCCTCTGTAGCCCGCCAG--
2:  AGATCTGTCTTCAGGGCGGG-C--G-GGTGAAATTCCCGCACCGGCGGTAAAT---AATCGCGCC-GCAAGCCCAGCGA-GCCCCGTGTAGCCCGCCAG--
3:  AGATCTGTCTTCAGGGCGGG-C--G-GGTGAAATTCCCGCACCGGCGGTGAAT---AATCGCGCC-GCAAGCCCTGCGA-GCCCCGTGTATCCGG------
4:  A-AACGCTC-TC---G-GGG-C--G-GGTGGAATTCCCGCACCGGCGGTGA----------------AAGCCC-GCGA-GCCC----TCTCCGG------
TM: A-AACGCTC-TC---G-GGG-CA-G-GGTGGAATTCCCG-ACCGGCGGTGA----------------AAGCCC-GCGA-GCCC----TCTCAGG------
5:  -GACCTCT-TTCGGGGCGGG----G-GATGAAATTCCC-CACCGGCGGTAAGT---TCTCCCG-AA-CAAGCCC-GCGA-GGCCCGCGCAAACCG------
DR: -GACCTCT-TTCGGGGCGGG----GCGA---AATTCCC-CACCGGCGGTAAGT---TCTCCCG-AA-CAAGCCC-GCGA-AGCCCGCGCAAACCG------
6:  -TATCT-TCTTC-G-G-GGGTC--G-GGTGAAATTCCCG-ACCGGCGGTAA-----AATCGCCC--GCGAGCC-AAGGAT-A-CCCTGTGGTCCG------
7:  -TATCT-TCTTC-G---GGGTCATG-GGTGAAATTCCCA-ACCGGCGGTAA-----AATAGCCC--GCGAGCC-AAGG-TAA-CCCT-TGGTCCG------
8:  C-TTAA-TCTTC-G---GGGT-ATG-GGTGAAATTCCCA-ATCGGCGGT--------ATAGCCC--GCGAGCC-AAGG-TAAAACCT-TGGT---------
DF: C-TTAA-TCTTC-G---GGGT-A-G-GGTGAAATTCCCA-ATCGGCGGT--------ATAGCCC--GCGAGCC-AAGG-TAAAA-CT-TGGT---------
9:  -GATGT-TCTTCAG---GGG-A-TG-GGTGAAATTCCCA-ATCGGCGGT-------AA-AGCCC--GCAAGCC-A--G-------TT-TGG-C--------
CA: -GATGT-TCTTCA----GGG-A-TG-GGTGAAATTCCCA-ATCGGCGGT-------AA-AGCCC--GCAAGCC-A----------TT-TGG-C--------
10: -TATAT-TCTTCCG---GGG-CA-G-GGTGAAATTCCCG-ACCGGCGGTAATT---AATCGCCT--GCGACCT-AAGG-T---C---GTGACCCG------
11: -TATAT-TCTTCCG---GGG-CA-G-GGTGAAATTCCCT-ACCGGCGGTAATT---AATCGCCT--GCGACCT-AAGG-T---C---GTGATTCG------
12: -TA-AT-TCTTTCG---GGG-CA-G-GGTGAAATTCCCA-ACCGGCAGTAAAT---AA-AGCCT--GCGA-CT-T-GG-TAATAT--GT--TTCA------
SA: -TA-AT-TCTTTCG---GGG-CA-G-GGTGAAATTCCCA-ACCGGCAGTAAAT---AA-AGCCT--GCGA-CC-T-GC-TAATAT--GT--TTCA------
13: ATA-AA-TCTTCAG---GGG-CA-G-GGTGAAATTCCCT-ACCGGCGGT--------ATAGCCC--GCGA-GC-T-GC-T----T--G-GA-GCA------
LL: ATA-AA-TCTTCA----GGG-CA-G-GGTGTAATTCCCT-ACCGGCGGT--------ATAGCCC--GCGA-GC-T-GC-T----T--G-CA-GCA------
14: AT-TAT-CCTTC-G---GGGTCA-G-GGTGAAATTCCCG-ACCGGCGGTAATG---AAGCGCAT--TCG-CCT-TA-G-T---C---GTGACCCG------
15: AT-TAT-CCTTC-G---GGGTCA-G-GGTGAAATTCCCG-ACCGGCGGTAATG---AAGCGCAT--TCG-CCT-TA-G-T-CCC---GTGACCCG------
16: TT-GAT-TCTTC-G---GGG-CA-G-GGTGGAAATCCCG-ACCGGCGGTAGTA---AAGCGCAT--TTG-CTT-TA-G-AGCCC---GTGACCCG------
BS: TT-GTA-TCTTC-G---GGG-CA-G-GGTGGAAATCCCG-ACCGGCGGTAGTA---AAGCACAT--TTG-CTT-TA-G-AGCCC---GTGACCCG------
17: TT-TAT-CCTTC-G---GGG-CATG-GGTGGAAATCCCG-ACCGGCGGTGATG---AAGCGCAT--GCT-TCT-TA-G-T---CC--GTGACCCG------
BH: TT-TAT-CCTTC-G---GGG-C-TG-GGTGGAAATCCCG-ACCGGCGGTGATG---AAGCGAAT--GCT-TCT-TA-G-T---CC--GTGACCCG------
18: AT-CAT-CCTTC-G---GGGTC--G-GGTGAAATTCCCA-ACCGGCGGTGATG---AAGTGCAT--ACT-TCT-TA-G-T---CC--GTGACCCG------
BQ: AG-CAT-CCTTC-G---GGGTC--G-GGTGAAATTCCCA-ACCGGCGGTGATG---AAGTGCAA--ACT-TCT-AA-G-T---CC--GTGACCCG------
19: --GTGTGTCTTCAGGGCGGG----G-GGTGAAATTCCC-CACCGGCGGTAAATCGGAAGGTGGCC-GCTAGCCCGCGGGCTGCCTCGGTAGCCGCCAGCG
20: --GTGTGTCTTCAGGGCGG-------GGTGAAATTCCC-CACCGGCGGTAAATCGAATGATGACCATCTAGCCGGTGGCTGCCTCGAAAGCCCGCGAGCG
21: --GCTTATTCTCAGGGCGG-------GGTGAAAGTCCC-CACCGGCGGTAAATTGTATTGCGACGATATAGTC-CGTCGCTCGGTCGAAAGCCCGCGAGCG
22: --GCTTATTCTCAGGGCGG-------GGTGAAAGTCCC-CACCGGCGGTAAATTGTATTGCGACGATATAGTCACGTCGGTGGTGCAAAAGCCCGCGAGCG
YP: --GCTTATTCTCAGGGCGG-------GGTGAAAGTCCC-CACCGGCGGTAAATTGTATTGCGACGATATAGTCACGTCGGTGGTGCAAAAGCCCGCGAGCG
23: --GCTTATTCTCAGGGCGG-------GGCGAAATTCCC-CACCGGCGGTAAAT----------------------CAACTCAGTTGAAAGCCCGCGAGCG
EC: --GCTTATTCTCAGGGCGG-------GGCGAAATTCCC-CACCGGCGGTAAAT----------------------CAACTCAGTTGAAAGCCCGCGAGCG
24: --GTCGGTCTTCAGGGCGG-------GGTGTAAGTCCC-CACCGGCGGTAAATCGAAAGATG--------------------------AGCCCGCGAGCG
PP: --GTCGGTCTTCAGGGCGG-------GGTGTAAGTCCC-CACCGGCGGTAAATCGAAAGATG--------------------------AGCCCGCGAGCG
25: --GTGCGTCTTCAGGGCGG-------GGCGAAATTCCC-CACCGGCGGTAGGCCGGCATGTTGCCGGCGAGCCCGCGAG-------------CGCCCGCG
BP: --GTGCGTCTTCAGGGCGG-------GGCGAAATTCCC-CACCGGCGGTAGGCCGGCATGTTGCCGGCGAGCCCGCGAG-------------CGCCCGCG
```



ДОПОЛНИТЕЛЬНЫЕ ПРИМЕРЫ

Пример 1. Рассмотрим дерево видов с шестью листьями, показанное на рисунке 1, у которого три ребра, смежные с вершиной 2, имеют длины равные 10 и остальные ребра имеют длины равные 1. Вершины нумерованы, как показано на рисунке 1. Листья этого дерева помечены обозначениями видов. Использованы сокращения: Ma – *Mycobacterium avium*, Ml – *Mycobacterium leprae*, Mm – *Mycobacterium marinum*, Mt – *Mycobacterium tuberculosis,* Mb – *Mycobacterium bovis.* Нуклеотидные последовательности в листьях представляют собой гомологичные участки Т-боксовой регуляции.

Рисунок 1. Дерево видов микобактерий.

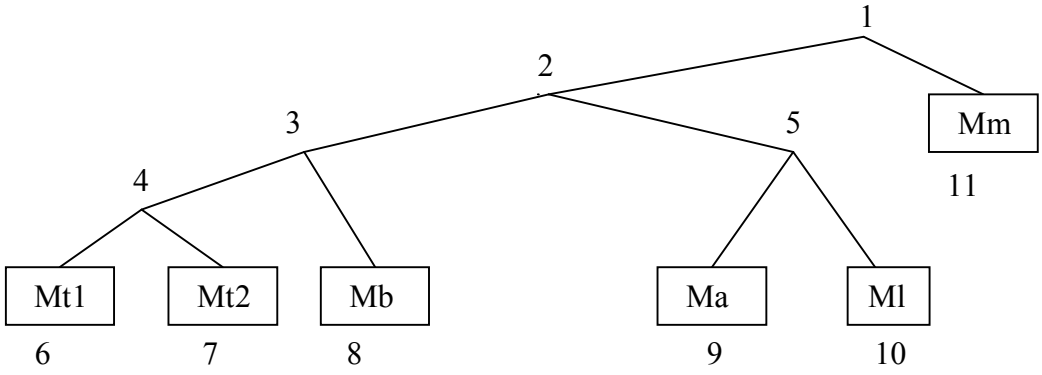


Рисунок 2. Результат работы алгоритма на примере 1. Первый столбец указывает номер вершины, за которой прописными буквами указана последовательность из найденной алгоритмом минимальной конфигурации. Малыми буквами указаны современные последовательности там, где они отличаются от выданных алгоритмом.

```
1:  CGGCCG---CAC-U------CAGGUG-CGGCAAGCGGGGUGGUACCGCGGCGCU-CGCGCA-CUG--AGCGCGU--CGUCG
2:  CGGCCG---CAC-U------CAGGUG-CGGCAAGCGGGGUGGUACCGCGGCGCU-CGCGCA-CUG--AGCGCGU-GCGUCG
3:  CGGCCG---CGC-UCG---GC--GU--CGGCAAGCGGGGUGGUACCGCGGCGCUUCGCGCA-C--C-AGCGCGUGGCGUCG
4:  CGGCCG---CGCAUCG---GC--GU---GGCAAGCGGGGUGGUACCGCGGCG-UUCGCGCA-C--C--G-GCGUGGCGUCG
5:  CGGCCG---CGC-UCG---GC--GU-GCGGCAAGCGGGGUGGUACCGCGGCGCU-CGCGCA-C--C-AGCGCGU--CGUCG
6:  CGGCCG---CGCAUCG---GC--GU---GGCAAGCGGGGUGGUACCGCGGCG-UUCGCGCA-C--C--G-GCGUGGCGUCG
7:  CGGCCG---CGCAUCG---GC--GU---GGCAAGCGGGGUGGUACCGCGGCG-UUCGCGCA-C--C--G-GCGUGGCGUCG
8:  CGGCCG---CGCAUCG---GC--GU---GGCAAGCGGGGUGGUACCGCGGCG-UUCGCGCA-C--C--G-GCGUGGCGUCG
9:  --GCCGUG-CGU-UCG----C--GU-GCGGCAAGCGGGGUGGUACCGCGGCGCU-CGCGCA-C--CUAGCGCGU--CGUCG
    uggcc---acg---cgaaagc--g---cggcaagcggggugguaccgcggcgcu-cgcgcagc--c-agcgcgu--cgucg
10: UGGCC---ACG---CGAAAGC--G---CGGCAAGCGGGGUGGUACCGCGGCGCU-CGCGCAGC--C-AGCGCGU--CGUCG
    --gccgug-cgu-ucg----c--gu-gcggcaagcggggugguaccgcggcgcu-cgcgca-c--cuagcgcgu--cgucg
11: CGGCCG---CGC-UCG----C--GU--CGGCAAGCGGGGUGGUACCGCGGCGCU-CGCGCA-C--C-AGCGCGU--CGUCG
    cggccg---cac-u------caggug-cggcaagcggggugguaccgcggcgcu-cgcgca-cug--agcgcgu--cgucg

1:    UCCCCG--------------UG-------------CCGU----GUGAUUUCUGG--CAC--AGGAGAC--CG------
2:    UCCCCG-G-C----------UG-------------CCGU----GUGAUUUCUGG--CAC--AGGAGAC--CG------
3:    UCCCCG-GCCUG--U-----UGCAGGCAC--A---CCG-----GUG----CUGGGGC-C--AGGAGACGACGCGCAAA
4:    UCCCCGAGCCUGGAU-----UGCAGGCACGCAGUGCCG-AACGGUG----CUGGGGC-CUGGGGAGACGACGCGCAAA
5:    UCCCCG----UG--U-----UGC-G--AC--A---CCG-----GUG----CUGG--CAC--AGGAGAC-ACGCGCAU-
6:    UCCCCGAGCCUGGAU-----UGCAGGCACGCAGUGCCG-AACGGUG----CUGGGGC-CUGGGGAGACGACGCGCAAA
7:    UCCCCGAGCCUGGAU-----UGCAGGCACGCAGUGCCG-AACGGUG----CUGGGGC-CUGGGGAGACGACGCGCAAA
8:    UCCCCGAGCCUGGAU-----UGCAGGCACGCAGUGCCG-AACGGUG----CUGGGGC-CUGGGGAGACGACGCGCAAA
9:    UCCCCG----UG--UCUACUUGU-GUUA---A---CCG-------------UGG--CCC--AGGAGAC---G----U-
leaf: uccccg-----g--u----uug------c--a---ccg-------------ugg--cac--aggagacaacgcgcauc
10:   UCCCCG-----G--U----UUG------C--A---CCG-------------UGG--CAC--AGGAGACAACGCGCAUC
leaf: uccccg----ug--ucuacuugu-guua---a-----g-------------ugg--ccc--aggagac---g----u-
11:   UCCCCG-----G--U-----UG---------A---CCG-----GUG----CUGG--CAC--AGGAGAC--CG------
leaf: uccccg--------------ug-------------ccgu----gugauuucugg--cac--aggagac--cg------
```

Итак, в примере 1 минимальная конфигурация дает набор последовательностей близких к исходным данным в листьях и весьма консервативных по вторичной структуре. В предковой последовательности (в вершине 1) имеется одна шпилька (вторичная структура), которая иногда с несколькими близкими вариантами проходит вдоль всего дерева эволюции и переходит в заранее выделенную подчеркиванием на рисунке 2 одну из шпилек в современных данных. Вторая, альтернативная шпилька родилась в вершине 3. Поэтому алгоритм указывает на древнее происхождение первой шпильки и относительную новизну второй шпильки, которая, возникнув в вершине 3, сохраняется до листьев. Алгоритм предлагает концевые последовательности в листьях 9, 10 и 11, которые отличаются от современных последовательностей по первичной структуре только на краях и не затрагивают шпилек.

Пример 2. Рассмотрим дерево видов, рисунок 3, содержащее 9 листьев, которые помечены обозначениями видов. Вершины дерева нумерованы, как показано на этом рисунке. Листьям приписаны современные последовательности. Эти последовательности - гомологичные участки с терминатором и участком остатков урацилов у классической аттенюаторной регуляции перед опероном *ilv* из видов актиномицетов (таблица 1).

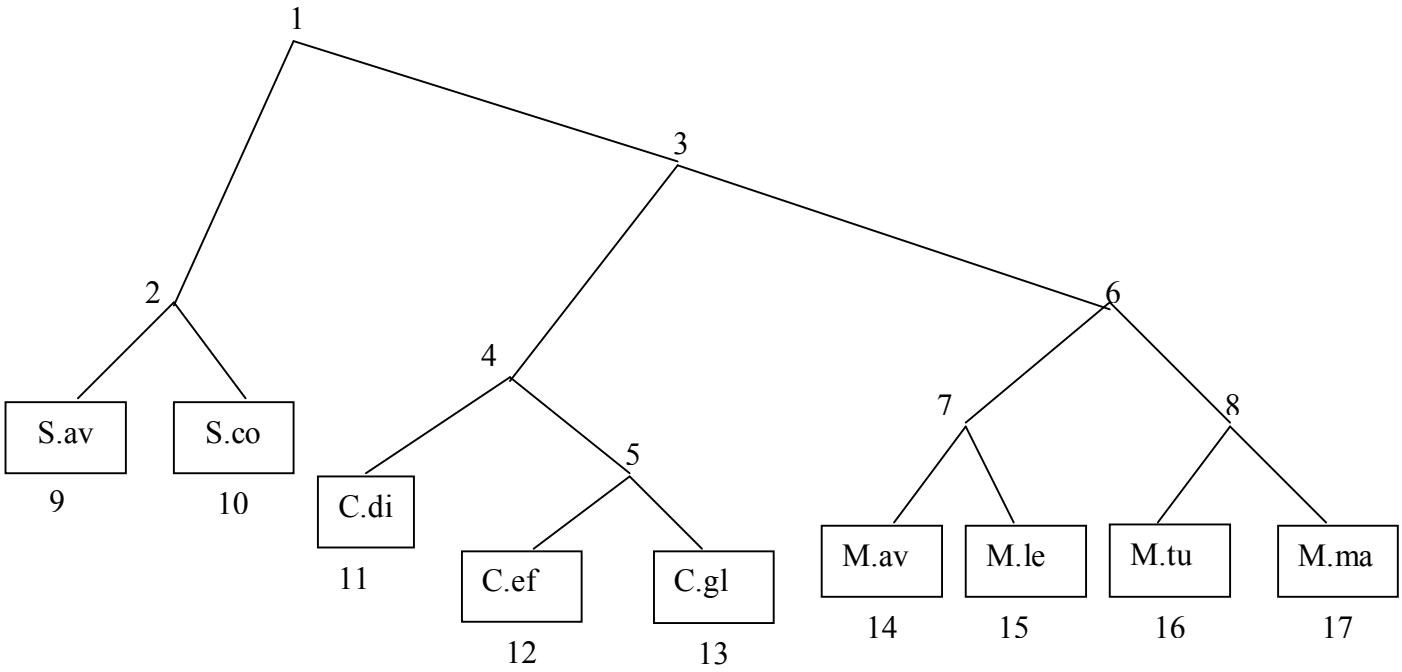


Рисунок 3. Дерево видов к примеру 2. Длины ребер ведущих в вершины 12, 13, 14 и 15 равны 1; в вершины 9, 10, 11, 5, 8, 16, 17 равны 2; в вершину 7 равно 2.5; в вершины 4, 6 равны 4; в вершину 2 равно 8; в вершину 3 равно 7.

Таблица 1. Исходные данные к примеру 2.

Сокращение	Вид	Участок с терминатором
C.di	*C. diphtheriae*	aaaagcGCCCUCGaCAGCAccacacaUGCUGAGCGGGGGCuuuccuuau
C.ef	*C. efficiens*	caagcGCCCUCGACAGUACccaccacaGUGCUGuuUCGAGGGCuuuguugu
C.gl	*C.glutamicum*	caagcGCCCUCGaCAACACUcaccacAGUGUUGgaaCGAGGGCuuucuuguu
M.av	*M. avium*	caacgcgcAACCCUCGugCAGCacaaGCUGuCGGGGGUUuuuuguu
M.tu	*M.tuberculosis*	caacgcgACCCUCGugCAGCagcugaGCUGgCGAGGGUuuuuucuu
M.bo	*M. bovis*	caacgcgACCCUCGugCAGCagcugaGCUGgCGAGGGUuuuuucuu
M.le	*M. leprae*	caacgcgcAACCCUCGugCAGCUagucAGCUGuCGAGGGUUuuuuguu
M.ma	*M. marinum*	caacgcgcAACCCUCGUgCAGCagcugaGCUGACG*GGGGUUuuuuguu
S.av	*S. avermitilis*	cggcgcgcuCCCCUCGcuUGCCucacGGCACGAGGGGuuuuuuguu
S.co	*S. coelicolor*	cgacgcgcuCCCCUCGcuUGCCuuacGGCACGAGGGGuuuuuuguu

Рисунок 4. Результат работы алгоритма на примере 2. В современных последовательностях терминатор показан зеленым цветом.

```
1(1):   -CGACGCGCUCCCCUCG--C-UUG--C-CUCA-CGGCA-U--C-G-G-------AGAGGG-UUUUU--UGUU--
2(1):   -CGACGCGCUCCCCUCG--C-UUG--C-CUCA-CGGCA----C-G-G------CAGAGGG-UUUUU--UGUU--
9(1):   -CGACGCGCUCCCCUCG--C-UUG--C-CUCA-CGGCA----C-G-G------CAGAGGG-UUUUU--UGUU--
S.av:   -CGGCGCGCUCCCCUCG--C-UUG--C-CUCA-CGGCA----C-G---------AG-GGG-UUUUU--UGUU--
10(1):  -CGACGCGCUCCCCUCG--C-UUG--C-CUCA-CGGCA----C-G-G------CAGAGGG-UUUUU--UGUU--
S.co:   -CGACGCGCUCCCCUCG--C-UUG--C-CUUA-CGGCA----C-G---------AG-GGG-UUUUU--UGUU--
3(2):   -CAACGCGC--CCCUCG-ACA--G--CACUCAGCCACAGUG-CUGAG------C-GAGGG-UUUUU--UGUU--
4(2):   -CAA-GCG---CCCUCG-ACA--G--CACUCAGCCACAGUG-CUGAG-----UC-GAGGG-CUUUC--UGUU-U
11(1):  -CAA-GCG---CCCUCG-ACA--G--CAC-CA-CCACAGUG-CUGAG------C-GAGGG-CUUUCUCUGUU-U
C.di:   AAAA-GCG---CCCUCG-ACA--G--CAC-CA-C-ACA-UG-CUGAG------C-G-GGGGCUUUC-CU--UAU
5(2):   -CAA-GCG---CCCUCG-ACA--G-ACACUCA-CCACAGUGUCUGAGG----UC-GAGGG-CUUUCUCUGUU-U
12(2):  -CAA-GCG---CCCUCG-ACA--G-ACAC-CA-CCACAGUG-CU--G-----UC-GAGGG-CUUUC--UGUU-U
C.ef:   -CAA-GCG---CCCUCG-ACA--GUAC-C-CA-CCACAGUG-CU--G---UUUC-GAGGG-CUU----UGUUGU
13(2):  -CAA-GCG---CCCUCG-ACA--G-ACACUCA-CCACAGUGUCU--G------C-GAGGG-CUUUCUCUGUU-U
C.gl:   -CAA-GCG---CCCUCG-ACA----ACACUCA-CCACAGUGU-U--GGAA---C-GAGGG-CUUUCU-UGUU--
6(1):   -CAACGCGCAACCCUCGUGCA--G----CU-A-CCACAG---CUG-------UC-GAGGG-UUUUU--UGUU--
7(1):   -CAACGCGCAACCCUCGUGCA--G----CU-AGCCACAG---CUG-------UC-GAGGG-UUUUU--UGUU--
14(1):  -CAACGCGCAACCCUCGUGCA--G----CU-AGUCACAG---CUG-------UC-GAGGG-UUUUU--UGUU--
M.av:   -CAACGCGCAACCCUCGUGCA--G----C--A--CA-AG---CUG-------UC-G-GGGGUUUUU--UGUU--
15(1):  -CAACGCGCAACCCUCGUGCA--G----CU-AGUCACAG---CUG-------UC-GAGGG-UUUUU--UGUU--
M.le:   -CAACGCGCAACCCUCGUGCA--G----CU-AGUCA--G---CUG-------UC-GAGGG-UUUUU--UGUU--
8(2):   -CAACGCGC-ACCCUCGUGCA--G----CU-AGCUGCAGUG-CUG-----CG-C-GAGGG-UUUUU--UGUU--

16(3):  -CAACGCGCAACCCUCGUGCA--G----C--AGCUGCAG---CUG------G-C-GAGGG-UUUUU--UCUU--
M.tu:   -CAACGCG--ACCCUCGUGCA--G----C--AGCUG-AG---CUG------G-C-GAGGG-UUUUU--UCUU--

17(3):  -CAACGCGCAACCCUCGUGCA--G----C--AGCUGCAG---CUG-----CG-C-GAGGG-UUUUU--UGUU--
M.ma:   -CAACGCGCAACCCUCGUGCA--G----C--AGCUG-AG---CUG----ACG---G-GGG-UUUUU--UGUU--
```

Для всех листьев имеет место небольшое отличие от современных последовательностей. В отличие от предыдущего примера, здесь отличия в листьях затрагивают в основном среднюю часть последовательности. Это связано с возможностью улучшить качество шпилек.