```
                       10        20        30        40        50        60        70        80        90
              ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|.
DehSerCAG     GATACGATGGCCGAGT-GGTT--AAGGCGGAGGATGCAGGTTCCTTTGGGC-----TCTGCCCG-CGCAGGTTCGAACCCTGCTCGTGTCG 82
DhhSerCAG     GATACGATGGCCGAGT-GGTT--AAGGCGAAGGATGCAGGTTCCTTTGGGC-----TCTGCCCG-CGCAGGTTCGAACCCTGCTCGTGTCG 82
CaSerCAG      GATACGATGGCCGAGT-GGTT--AAGGCGAAGGATGCAGGTTCCTTTGGGC-----ATTGCCCG-CGCAGGTTCGAACCCTGCTCGTGTCG 82
CadSerCAG     GATACGATGGCCGAGT-GGTT--AAGGCGAAGGATGCAGGTTCCTTTGGGC-----ATTGCCCG-CGCAGGTTCGAACCCTGCTCGTGTCG 82
CapSerCAG     GATGCGATGACCGAGT-GGTT--AAGGTGAAGGATGCAGGTTCCTTTGGGC-----ATTGCCCG-CGCAGGTTCGAACCCTGCTCGCGTCG 82
CaoSerCAG     GATGCGATGACCGAGT-GGTT--AAGGTGAAGGATGCAGGTTCCTTTGGGC-----ATTGCCCG-CGCAGGTTCGAACCCTGCTCGCGTCG 82
Ct_aSerCAGa   GATACGATGGCCGAGT-GGTT--AAGGCGGAGGATGCAGGTTCCTTTGGGC-----TCTGCCCG-CGCAGGTTCGAACCCTGCTCGTGTCG 82
ShsSerCAG     GATACGATGGCCGAGT-GGTT--AAGGCGAAGGATGCAGGTTCCTTTGGGC-----TCTGCCCG-CGCAGGTTCGAACCCTGCTCGTGTCG 82
CllSerCAG     GATGTGGTGGCCGAGT-GGTT--AAGGCGTGAGGTGCAGGTCCTCATGAGC-----TCTGCTCT-CGCAGGTTCGAATCCTGTCCACGTCG 82
MefSerCAG     GATGTGGTGGCCGAGT-GGTT--AAGGCGTGAGGTGCAGGTCCTCATGAGC-----TCTGCTCT-CGCAGGTTCGAATCCTGTCCACGTCG 82
ShpSerCAG     GATACGATGGCCGAGT-GGTT--AAGGCGGAGGATGCAGGTTCCTTTGGGC-----TCTGCCCG-CGCAGGTTCGAACCCTGCTCGTGTCG 82
MrgSerCAG     GATACGATGGCCGAGT-GGTT--AAGGCGGAGGATGCAGGTTCCTTTGGGC-----TCTGCCCG-CGCAGGTTCGAACCCTGCTCGTGTCG 82
CnmSerCAGa    GATACAATGGCCGAGT-GGTT--AAGGCGAAGGATGCAGGTTCCTTTGGGC-----ATTGCCCG-CGCAGGTTCGAACCCTGCTTGTGTCG 82
LoeSerCAG     GATGATGTGGCTGAGT-GGTTT-AAAGCGTCGGGTGCAGGTCCCGATGGGC-----ATTGCCCG-CGCAGGTTCGAACCCTGTCATCATCG 83
CatSerCAG     GATACGATGGCCGAGT-GGTT--AAGGCGAAGGATGCAGGTTCCTTTGGGC-----TCTGCCCG-CGCAGGTTCGAACCCTGCTCGTGTCG 82
MifSerCAG     GATACGATGGCCGAGT-GGTT--AAGGCGAAGGATGCAGGTTCCTTTGGGC-----TCTGCCCG-CGCAGGTTCGAACCCTGCTCGTGTCG 82
CameSerCAG    GATGCGATGACCGAGT-GGTT--AAGGTGAAGGATGCAGGTTCCTTTGGGC-----ATTGCCCG-CGCAGGTTCGAACCCTGCTCGCGTCG 82
StaSerCAG     GATACGATGGCCGAGT-GGTT--AAGGCGAGGGATGCAGGTTCCCTTGGGC-----TCTGCCCG-CGCAGGTTCGAACCCTGCTCGTGTCG 82
PtaSerUGAa    GGCACTATGGCCGAGC-GGTT--AAGGCGAAAGACTTGAAATCTTTTGGGA-----TTTTCCCG-CGCAGGTTCGAGTCCTGCTGGTGTCG 82
CaSerUGAa     GGCAACTTGTCCGAGC-GGTT--AAGGAGAAAGACTTGAAATCTTTTGGGC-----TTTGCCCG-CGCAGGTTCGAGTCCTGCAGTTGTCG 82
ErgSerUGAb    GGCAACATGGCCGAGT-GGTT--AAGGCGACAGACTTGAAATCTGTTGGGC-----TCTGCCCG-CGCTGGTTCAAATCCTGCTGTTGTCG 82
SpSerUGAa     GTCACTATGTCCGAGT-GGTT--AAGGAGTTAGACTTGAAATCTAATGGGC-----TTTGCCCG-CGCAGGTTCAAATCCTGCTGGTGACG 82
LoeSerUGAa    GGCAACTTGTCCGAGT-GGTT--AAGGAGAAAGACTTGAAATCTTTTGGGC-----TCTGCCCG-CGTAGGTTCGAGTCCTGCAGTTGTCG 82
PiuSerUGAb    GTCAGTTTGTCCGAGT-GGTT--AAGGAGAAAGACTTGAAATCTTTTGGCC-----TCTGGCCG-CACAGGTTCGAATCCTGTAGCTGACG 82
Ct_aSerUGAa   GGCGGTATGTCCGAGT-GGTT--AAGGAGAAAGACTTGAAATCTTTTGGGC-----TCTGCCCG-CGCAGGTTCGAGTCCTGCTGCCGTCG 82
PtaSerAGAa    GGCAACGTGGCCGAGC-GGTT--AAGGCGAAAGATTAGAAATCTTTTGGGA-----TTTTCCCG-CGCAGGTTCGAGTCCTGCCGTTGTCG 82
CaSerAGAa     GGCAACTTGTCCGAGT-GGTT--AAGGAGAAAGATTAGAAATCTTTTGGGC-----TTTGCCCG-CGCAGGTTCGAGTCCTGCAGTTGTCG 82
SpSerAGAa     GACACTATGCCTGAGT-GGTT--AAAGGGACAGACTAGAAATCTGTTGCGG-----TCTCCGCG-CGCAGGTTCAAATCCTGCTGGTGTCG 82
LwSerAGAb     GGCAATTTGGCCGAGT-GGTT--AAGGCGTAAGATTAGAAATCTTTTGGGC-----TCTGCCCG-CGCAGGTTCGAATCCTGCAGTTGTCG 82
ErgSerAGAa    GGCAACATGGCCGAGC-GGTT--AAGGCGAAAGATTAGAAATCTTTTGGGC-----TATGCCCG-CGCAGGTTCGAGTCCTGCTGTTGTCG 82
YlSerAGAa     GGCATCCTGGCCGAGC-GGTT--AAGGCGAGAGATTAGAAATCTCTTGGGG-----TTTCCCCG-CGTAGGTTCGAATCCTGCGGATGTCG 82
WaSerAGAa     GGCAACTTGGCCGAGT-GGTT--AAGGCGAAAGATTAGAAATCTTTTGGG------TTTACCCG-CGCAGGTTCGAATCCTGCAGTTGTCG 81
PtaSerCGA     GACACTATGGCCGAGT-GGTT--AAGGCGAAAGACTCGAAATCTTTTGGGA-----TCTTCCCG-CGCAGGTTCAAATCCTGCTGGTGTCG 82
ErgSerCGAc    GACACTATGGCCGAGT-GGTT--AAGGCGGGAGACTCGAAATCTCTTGGGC-----TCTGCCCG-CGCAGGTTCAAATCCTGCTGGTGTCG 82
SpSerCGAa     GTCACTATGTCCGAGT-GGTT--AAGGAGTTAGACTCGAAATCTAATGGGC-----TCTGCCCG-CGCAGGTTCAAATCCTGCTGGTGACG 82
Kop_bSerCGAa  GGCACTATGGCCGAGT-GGTT--AAGGCGAAAGACTCGAAATCTTTTGGGC-----TCTGCCCG-CGCAAGTTCAAATCCTGCTGGTGTCG 82
PiuSerCGAa    GGCAATTTGTCCGAGT-GGTT--AAGGAGGAAGACTCGAAATCTTCTGGGC-----TTTGCCCG-CGCAGGTTCGAATCCTGCAGTTGTCG 82
YlSerCGAa     GGCAACTTGCCGGAGT-GGTT--AACGGGACAGATTCGAAATCTGTTGGGC-----TTTGCCCG-CGCAGGTTCGAATCCTGCAGTTGTCG 82
CaSerGCUa     GTCATGATGTCCGAGT-GGTT--AAGGAGTTGCCCTGCTAAGGCAATGGGC-----TCTGCCTT-CGCAGGTTCGAATCCTGTTCATGACG 82
PtaSerGCUa    GTCCCAGTGGCCGAGT-GGTT--AAGGCGACGCCCTGCTAAGGCGTTGTGC-----TCTGCATG-CGCAGGTTCGAATCCTGTCTGTGACG 82
CllSerGCUa    GTTACAGTGGCCGAGT-GGTT--AAGGCGACACCCTGCTAAGGTGTTGGGC-----ATTGCCTG-CGCAGGTTCGAATCCTGTCTGTGACG 82
ErgSerGCUa    GTCAAAGTGGCCGAGT-GGTT--AAGGCGATGCCCTGCTAAGGCATTGGGT-----TTTACCTG-CGCAGGTTCGAATCCTGTCTTTGACG 82
LwSerGCUb     GTCCCAGTGGCCGAGT-GGTT--AAGGCGATGCCCTGCTAAGGCATTGGGT-----TTTACCTG-CGCAGGTTCGAATCCTGTCTGTGACG 82
YlSerGCUa     GTCCCGGTGGCCGAGT-GGTT--AAGGCGTTGCCCTGCTAAGGCAGTGGGT-----TTTGCCCG-CGCAGGTTCGAATCCTGTCCGTGACG 82
Ct_aSerGCUa   GTCCCAGTGTCCGAGT-GGTT--AAGGAGTTGCCCTGCTAAGGCAATGGGC-----TTTGCCTT-CGCAGGTTCGAATCCTGTCTGTGACG 82
KcLeuCAG      GGACGTGTGGCCGAGT-GGTCT-AAGGCGCTCGCTTCAGGTGCGGGTCTC------GCAAGAGG-CGTGGGTTCGAACCCCACCGCGTTCA 82
YlLeuCAGa     GGCATTCTGGCCGAGT-GGTCT-AAGGCGCCAGGTTCAGGTCCTGGTCTC------TCCGGAGG-CGAGAGTTCGAATCTCTCGGATGTCA 82
YlLeuCAGc     GGGAGTTTGGCCGAGT-GGTCT-AAGGCGCTAGCTTCAGGTGCTAGTCTC------GAAAGAGG-CGTGAGTTCGAACCTCACAGCTCTCA 82
YlLeuCAGd     GGGAGTTTGGCCGAGT-GGTCT-AAGGCGCTAGCTTCAGGTTCTAGTGTC------GAAAGAGG-CGTGAGTTCGAACCTCCCAGCTCTCT 82
YlLeuCAGe     GGCAACCTGGCCGAGT-GGTCT-AAGGCGCCAGGTTCAGGTCCTGGTCTC------TTCGGAGG-CGAGAGTTCGAATCTCTCGGTTGTCA 82
LatLeuCAGa    GGGAGTTTGGCCGAGT-GGTT--AAGGCGTCAGATCCAGGTTCTGATATC------TAAGGATG-CAAGGGTTCGAACCCCTTAGCTCTCA 81
LwLeuCAGa     GAGAGTTTGGCCGAGT-GGTT--AAGGCGTCAGATCCAGGTTCTGATATC------TACGGATG-CGCGGGTTCGAACCCCGTAGCTCTCA 81
CyjLeuCAG     GGGAGAGTGGCCGAGT-GGTT--AAGGCGCCCGCTTCAGGTGCTGGTAACG-----TAAAGTTG-CGCGAGTTCGAACCTCGTCTCTTCCA 82
Kop_bLeuCAGa  GGATGTGTGGCCGAGT-GGTCT-AAGGCGCTAGCTTCAGGTGCTAGTCTC------TTCGGAGG-CGTGAGTTCGAACCTCACTGCATTCA 82
Kop_bLeuCAGb  GGACGTGTGGCCGAGA-GGTCT-AAGGCGCTAGCTTCAGGTGCTAGTCTC------TTCGGAGG-CGTGAGTTCGAACCTCACCGCGTTCA 82
OgpLeuCAG     GTGGGTTTGGCCGAGT-GGTT--AAGGCGTCACGTTCAGGTCGTGATTTCG-----TTAAGAAG-CAAGAGTTCGAACCTCTTAGCCCACA 82
PiuLeuCAGa    GTGAGTTTGGCCGAGT-GGTCT-AAGGCGTTGCGTTCAGGTCGCAATATC------TTCGGATG-CAAGAGTTCGAACCTCTTAGCTCACA 82
PiuLeuCAGb    GGGTGTGTGGCCGAGA-GGTT--AAGGCGTCAGGATCAGAACCTGATTTTC-----TTAGGATTTCCTGGGTTCGATCCCCAGCACACTCA 83
WaLeuCAGa     GGGAGTGTGGCCGAGT-GGTT--AAGGCGCCCGCTTCAGGTGCAGGTCTCTT----TTAAGAGG-CGCGAGTTCGAACCTCGTCACTTCCA 83
WicLeuCAG     GGGAGTGTGGCCGAGT-GGTT--AAGGCGCCCGCTTCAGGTGCAGGTCTCTA----TTAAGAGG-CGCGAGTTCGAACCTCGTCACTTCCA 83
KaLeuCAG      GGGTGTGTGGCCGAGT-GGTT--AAGGCGCCTGCTTCAGGTGCAGGTCTGTTAAAGTTAACAGG-CGCGAGTTCGAACCTCGCCGCATCCA 87
KlwLeuCAG     GGGTGTGTGGCCGAGT-GGTT--ATGGCGCCTGCTTCAGGTGCAGGTCTGTTAATGTTAACAGG-CGCGAGTTCGAACCTCGCCGCATCCA 87
KmmLeuCAG     GGGTGTGTGGCCGAGT-GGTT--AAGGCGCCTGCTTCAGGTGCAGGTCTGTTAATGTTAACAGG-CGCGAGTTCGAACCTCGCCGCATCCA 87
KlLeuCAG      GGGTGTGTGGCCGAGT-GGTT--AAGGCGCCTGCTTCAGGTGCAGGTCTGTTAATGTTAACAGG-CGCGAGTTCGAACCTCGCCGCATCCA 87
LakLeuCAG     GGGTGTGTGGCCGAGT-GGTT--AAGGCGCCTGCTTCAGGTGCAGGTCTGTTT-CGTTAACAGG-CGCGAGTTCGAACCTCGCCGCATCCA 86
ErgLeuCAG     GGGTGTGTGGCCGAGT-GGTT--AAGGCGCCTGCGTCAGGTGCAGGTCTGATT-AGTTATCAGG-CGCGAGTTCGAATCTCGCCACATCCA 86
ErcLeuCAG     GGGTGCGTGGCCGAGT-GGTT--AAGGCGCCTGCGTCAGGTGCAGGTCTGTTT-CGTTAACAGG-CGCGAGTTCGAGTCTCGCCACATCCA 86
DebLeuCAG     GGGAGCGTGGCCGAGA-GGTTT-AAGGCGTCAGGATCAGAACCTGATTTTCTCC--ATGAGATTCCCTGGGTTCGAACCCCAGCGCTCCCA 87
ShoLeuCAG     GGCGAAGTGGCCGAGT-GGTCT-ATGGCGCTAGCTTCAGGTGCTAGTCTACG----TTTGTGGG-CGTGGGTTCGAACCCCACCTTCGTCA 84
ShjLeuCAGa    GGCGAAGTGGCCGAGT-GGTCT-ATGGCGCTAGCTTCAGGTGCTAGTTCACA----TATGTGAG-CGTGGGTTCGAACCCCACCTTCGTCA 84
ShcLeuCAG     GGCGAAGTGGCCGAGT-GGTCT-ATGGCGCTAGCTTCAGGTGCTAGTCTACG----TTTGTGGG-CGTGGGTTCGAACCCCACCTTCGTCA 84
SpLeuCAG      GGCGAAGTGGCCGAGT-GGTCT-ATGGCGCTAGCTTCAGGTGCTAGTCTACG----TATGTGGG-CGTGGGTTCGAACCCCACCTTCGTCA 84
CaLeuAAGa     GACTGGATGGCCGAGC-GGTC--AAGGCGCCAGGGTAAGGTCCTGGTCTACG----AAAGTGGG-CGCGAGTTCGAATCTCGTTTCAGTCA 83
PtaLeuUAGa    GAGGGTTTGGCCGAGT-GGTT--AAGGCGCCATGTTTAGGCCGTGGTATCA-----TTGCGATG-CGCGGGTTCGAATCCCGCATCCCTCA 82
PtaLeuAAGa    GGCAATCTGGCCGAGT-GGTCT-AAGGCGCCAGGTTAAGGTCCTGGTCTC------TTCGGAGG-CGTGGGTTCGAACCCCACGATTGTCA 82
PiuLeuAAGa    GGCAGCTTGGCCGAGT-GGTCT-AAGGCGTCAGGTTAAGGTCCTGGTCTC------TTCGGAGG-CGTGGGTTCGAACCCCACAGCTGTCA 82
KaaLeuAAGa    GGTACTCTGGCCGAGT-GGTCT-AAGGCGTCAGGTTAAGGTCCTGGTCTC------TACGGAGG-CGCGGGTTCGAACCCCGCGGGTATCA 82
YlLeuAAGa     GGCAACCTGGCCGAGC-GGTCT-AAGGCGCCAGGTTAAGGTCCTGGTCTC------TTCGGAGG-CGAGAGTTCGAATCTCTCGGTTGTCA 82
LoeLeuAAGa    GGATGGATGGCCGAGCTGGTTT-AAGGCGTCAGGGTAAGGTCCTGATCTC------TTCGGAGG-CGCGAGTTCGAATCTCGTTCCATTCA 83
MrgLeuAAGa    GGTACTCTGGCCGAGTTGGTCT-AAGGCGCCAGGGTAAGGTCCTGGTCTC------TTCGGAGG-CGCGAGTTCGAATCTCGCGGGTATCA 83
KnLeuAAGa     GGGAGTTTGGCCGAGT-GGTTT-AAGGCGTCAGATTAAGGCTCTGATATC------TTCGGATG-CGCGGGTTCGAATCCCACAGCTCTCA 82
CglLeuGAG     GGTACTCTGGCCGAGT-GGTCT-AAGGCGCCAGGTCGAGGTCCTGGTCTC------TTCGGAGG-CGCGGGTTCGAACCCCGCGGGTATCA 82
KlLeuGAG      GGCACTATGGCCGAGT-GGTTT-AAGGCGCCAGGTCGAGGTCCTGGTATC------TCCGGATG-CGCGAGTTCGAATCTCGCTGGTGTCA 82
ErgLeuGAG     GGTACTCTGGCCGAGC-GGTT--AAGGCGCCAGGTCGAGGTCCTGGTCTC------TTCGGAGG-CGCGAGTTCGAATCTCGCGGGTATCA 81
YlLeuUAGa     GGGACTTTGGCCGAGC-GGTCT-AAGGCGCCAGGTTTAGGCCCTGGTATC------GCAAGATG-CGCGAGTTCGAATCTCGCAGGTCTCA 82
LatLeuUAGc    GGGAGTTTGGCCGAGT-GGTT--AAGGCGTCAGATTTAGGTTCTGATATC------TGCGGATG-CGCGGGTTCGAATCCCGTAGCTCTCA 81
PiuLeuUAGa    GTGGGTTTGGCCGAGT-GGTCT-AAGGCGTCACGTTTAGGCCGTGATATC------TTCGGATG-CAAGAGTTCGAACCTCTTAGCCCACA 82
KaaLeuUAGa    GGGAGTTTGGCCGAGT-GGTTT-AAGGCGTCAGATTTAGGTTCTGATATC------TTCGGATG-CGTGGGTTCGAATCCCATAGCTCTCA 82
PiuLeuUAGb    GGCAGTTTGGCCGAGT-GGTCT-AAGGCGTCAGGTTTAGGTCCTGATCTT------TTCGAAGG-CGTGGGTTCGAACCCCACAGCTGTCA 82
ErgLeuUAGa    GAGAGTTTGGCCGAGT-GGTT--AAGGCGTCAGATTTAGGCTCTGATATC------TACGGATG-CGCGGGTTCAAATCCCGTAGCTCTCA 81
```

```
                       10        20        30        40        50        60        70        80        90
              ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|.
CaLeuCAAa     GGCACTATGGCCGAGTTGGTCT-AAGGCGGTAGACTCAAGTTCTACTATC------GTAAGATG-CAGGAGTTCGAATCTCCTTGGTGTCA 83
PtaLeuCAAa    GGATGTGTGGCCGAGA-GGTCT-AAGGCGTCTGACTCAAGATCAGATCTC------GAAAGAGG-CGCGAGTTCGAACCTCGTCGCATCCA 82
PtaLeuCAAb    GGATGAGTGGCCGAGG-GGTCT-AAGGCGTCTGACTCAAGTTCAGATCTC------GCAAGAGG-CGCGAGTTCGAATCTCGTCTCATTCA 82
DehLeuCAAa    GGCTCTGTGGCCGAGT-GGTT--AAGGCGGTAGACTCAAGTTCTACTATCG-----TAAAGATG-CACGAGTTCGAACCTCGTCGGAGTCA 82
MrgLeuCAAa    GGCTCTGTGGCCGAGT-GGTTT-AAGGCGGCAGACTCAAGTTCTGTTATCG-----TAAAGATG-CACGCGTTCGAATCGCGTCGGAGTCA 83
ShoLeuCAAa    GGCCGGTTGGCCGAGC-GGTCT-ATGGCGTAGCACTCAAGTTGCTATATC------GAAAGATG-CGAGGGTTCGAATCCCTCACCGGTCA 82
SpLeuCAAa     GGCCAGTTGGCCGAGC-GGTCT-ATGGCGCAGCACTCAAGTTGCTGTATC------GAAAGATG-CGAGGGTTCGAATCCCTCACTGGTCA 82
CapLeuCAAa    GGAAAGGTGGCCGAGTTGGTCT-AAGGCGCTAGACTCAAGTTCTAGTATC------GTAAGATG-CGCGAGTTCGAATCTCGTCCTTTTCA 83
WicLeuCAAa    GGATCTTTGGCCGAGT-GGTTT-AAGGCGTTCGACTCAAGATCGAATATC------GCAAGATG-CAGGAGTTCGAATCTCCTAGGATCCA 82
KaLeuCAAa     GGTTGTGTGGCCGAGC-GGTCT-AAGGCGCCTGATTCAAGCTCAGGTATC------GTAAGATG-CAAGAGTTCGAATCTCTTCGCAACCA 82
CaLeuUAAa     GGAGGGTTGACCGAGTTGGTCT-AAGGTGACAGACTTAAGATCTGTTGGTCA----GTTGACCG-CGCGAGTTCGAATCTCGTACCCTTCA 85
PtaLeuUAAa    GGAGGGTTGGCCGAGT-GGTCT-AAGGCGGTAGACTTAAGATCTACTGGACA----GACGTCCG-CGCGAGTTCGAACCTCGCATCCTTCA 84
Ct_aLeuUAAa   GGAGGGTTGGCCGAGTTGGTCT-AAGGCGGCAGACTTAAGATCTGCTGGTGT----GTTCACCG-CGCGAGTTCGAATCTCGCATCCTTCA 85
LatLeuUAAa    GGTGGGTTGGCCGAGTTGGTCT-AAGGCGGCAGACTTAAGATCTGTTGGACG----GTTGTCCG-CGCGAGTTCGAACCTCGCATCCACCA 85
LoeLeuUAAc    GGAGAGTTGACCGAGTTGGTT--AAGGTGACAGACTTAAGATCTGTTGGTAT----GTCTACCG-CGCGAGTTCGAATCTCGTACTCTTCA 84
KlLeuUAAa     GGAGGGTTGGCCGAGT-GGTCT-AAGGCGGCAGACTTAAGATCTGTTGGACA----ATTGTCCG-CGCGAGTTCGAACCTCGCATCCTTCA 84
ErgLeuUAAa    GGAGGGTTGGCCGAGT-GGTCT-AAGGCGGCAGACTTAAGATCTGTTGGACG----ATTGTCCG-CGCGGGTTCGAAACCCGCACCCTTCA 84
KaaLeuUAAa    GGAGGGTTGGCCGAGT-GGTCT-AAGGCGGCAGACTTAAGATCTGTTGGAC-----GGTGTCCG-CGCGAGTTCGAACCTCGCATCCTTCA 83
SpLeuUAAa     GCGGCTATGCCCGAGT-GGTCT-AAGGGGGCAGATTTAAGCCCTGCTGTT------GTAAAACG-CGAGAGTTCGAACCTCTCTGGCCGCA 82
SjLeuUAAa     GCAGCGATGCCCGAGT-GGTCT-AAGGGGGCAGATTTAAGCTCTGCTGTT------GATAAACG-CGCGAGTTCGAACCTCGCTCGCTGCA 82
CaValAACa     GGTTTCGTGGTCTAGTTGGTT--ATGGCATATGCTTAACACGCATAAC--------------GT-CCCCAGTTCGATCCTGGGCGGAATCA 74
PtaValAACa    GGTTTTGTGGTCTAGTTGGTT--ATGGCATCTGCTTAACACGCAGAAC--------------GT-CCCCAGTTCGATCCTGGGCAAAATCA 74
PiuValAACa    GGTCGGATGGTCTAGTTGGTT--ATGGCATATGCTTAACACGCATAAC--------------GT-CCCCAGTTCGATCCTGGGTTCGATCA 74
DehValAACc    GGTTTTGTGGTCTAGTTGGTT--ATGGCATCTGCTTAACACGCAGAAC--------------GT-CCCCAGTTCGATCCTGGGCAAAATCA 74
SpValAACa     GGTCGTGTGGTTTAGATGGTT--ATAATTTCTGCTTAACACGCAGACG--------------GT-CCCAAGTTCGAGTCTTGGCACGATCA 74
CaValCAC      GGTTTAGTGGTGTAGC-GGT---ATCACGTTTCCTTCACACGGAAAAG--------------GT-CTCGAGTTCGATCCTCGGCTAGATCA 72
WicValCAC     GGTTTAGTAGTGTAGT-GGTT--ATCACGTCTGTTTCACCCACAGAAT--------------GT-CGCGAGTTCGAATCTCGCCTGAACCT 73
PiuValCAC     AGTCGTTTGGTCTAGT-GGTT--ACGACGTCCCCTTCACACGGGGAAG--------------GT-CGCGAGTTCGAACCTCGCATCGACTA 73
SpValCAC      AGTCCCGTGTTCTAGT-GGT---ATGATTTCTCCTTCACACGGAGAAG--------------GT-CCCGGGTTCGATCCCCGGCGGGACTA 72
CaValUAC      GGTTGAGTAGTGTAGC-GGCCT-ATCACGTTTCGTTTACACCGAAAAG--------------GT-CCCGAGTTCGATCCTCGGCTCGATCA 74
PtaValUACa    GGTCTTGTAGTGTAGT-GGTTC-ATCACGTTTCGTTTACACCGAGAAG--------------AT-CGCGAGTTCGAACCTCGCCTGGGCCA 74
PiuValUACa    AGTTGTTTGGTCTAGT-GGTT--ACGACGTCCCCTTTACACGGGGAAG--------------GT-CGCGAGTTCGAACCTCGCATCAACTA 73
DehValUACa    GGTTTAGTGGTGTAGC-GGCT--ATCACGTTTCGTTTACACCGAAAAG--------------GC-CCCGAGTTCGATCCTCGGCTAGATCA 73
ErgValUACa    GGTCCAATGGTCTAGT-GGTTC-ACGACGTCGCCTTTACACGGCGAAG--------------AT-CCCGAGTTCGAACCTCGGTTGGATCA 74
SpValUACa     GCATCTGTAGTCTAGT-GGTT--ATGATTTCTGCTTTACACGCAGACG--------------GT-CCCAGGTTCGATCCCTGGTAGATGCA 73
YlValUACa     GGTTTCTTGGTCTAGTTGGTC--ATGGCATCCGCTTTACACGCGGAAC--------------GT-CGGCAGTTCGATCCTGTCAGAAATCA 74
PtaAlaCAG     GGGCGTGTGGCGCAGT-GGTT--AGCGCGTTTTCTTCAGGTGAAAAA---------------GA-CCCCAGTTCGAACCCGGGCACGTCCA 72
ErgAlaCGCb    GGGCGTGTGGCGCAGTTGGT---AGCGCGCTCCCTTCGCATGGGAGAG--------------GT-CATCGGTTCGATTCCGGTCTCGTCCA 73
KnAlaCGCa     GGGTGTGTGGCGCAGTTGGT---AGCGCGCTCCCTTCGCATGGGAGAG--------------GC-CTCTGGTTCGATTCCTGACTCATCCA 73
LatAlaCGCa    GGGCACGTGGCGCAGATGGT---AGCGCGCTTCCTTCGCAAGGAAGAG--------------GT-CATCGGTTCGAGTCCGGTCTTGTCCA 73
YlAlaCGCa     GGGCTTGTAGTGTAGT-GGCT--ATCGCGTCCGTTTCGCATACGGAAG--------------GT-CTCCGGTTCGACTCCGGACTGGTCCA 73
KcAlaCGCa     GGGCGTGTGGCGCAGT-GGT---AGCGCAATTGCTTCGCATGCAAGAG--------------GT-CGCCGGTTCGACTCCGGCCTCGTCCA 72
DhhAlaCGCa    GGGCGTGTGGCGTAGTTGGT---AGCGCGTTCGCTTCGCACGCGAAAG--------------GT-CTCTGGTTCGACTCCTGACTCGTCCA 73
CaAlaAGCa     GGGCGTGTGGCGTAGTTGGT---AGCGCGTTCCCTTAGCATGGGAAAG--------------GT-CATGAGTTCGACTCTTATCTCGTCCA 73
PtaAlaAGCa    GGGCGTGTGGCGTAGTTGGT---AGCGCGTTCGCTTAGCATGTGAAAG--------------GT-CTCGGGTTCAATTCCTGACTCGTCCA 73
YlAlaAGCa     GGGCGTCTGGTGTAGTTGGTT--ATCACATGCGCTTAGCATGCGCGAG--------------GT-CCCCGGTTCGATTCCGGTGTCGTCCA 74
SpAlaAGCa     GGGCATGTGGTGTAGATGGTT--ATCACGCTTCCTTAGCATGGAAGAG--------------GT-CCCAGATTCGAGTTCTGGCTTGTCCA 74
CllAlaAGCa    GGGCGTGTGGCGTAGTTGGT---AGCGCGTTCCCTTAGCATGGGAAAG--------------GT-CTCCGGTTCGACTCCGGACTCGTCCA 73
DebAlaAGCa    GGGCGTGTGGCGTAGTTGGT---AGCGCGTTCGCTTAGCATGCGAAAG--------------GT-CCTGGGTTCGACTCCCAGCTCGTCCA 73
ZrAlaAGCa     GGGCATGTGGCGTAGTTGGT---AGCGCGCCCCCTTAGCATGGGGGAG--------------GT-CTTCGGTTCGAATCCGGACTTGTCCA 73
LoeAlaAGCa    GGGCGTGTAGCTCAGTTGGC---AGAGCGTTTCCTTAGCATGGAAAAG--------------GT-CTTGAGTTCGATTCTTAACTCGTCCA 73
CaAlaUGCa     GGGCGTGTGGCGTAGTTGGT---AGCGCGTTTGCCTTGCAAGCAAAAG--------------GT-CATGAGTTCGACTCTCATCTCGTCCA 73
ErgAlaUGCc    GGGCATATGGCGCAGTTGGT---AGCGCGCTTCCCTTGCAAGGAAGAG--------------GT-CATCGGTTCGATTCCGGTTATGTCCA 73
KaAlaUGCb     GGGCGTGTGGCGCAGTTGGT---AGCGCGCTTCCCTTGCAAGGAAGAG--------------GT-CTTCGGTTCGATTCCGGACTCGTCCA 73
DehAlaUGCa    GGGCGTGTGGCGTAGTTGGT---AGCGCGTCTGCCTTGCAAGCAGAAG--------------GT-CACTGGTTCGATTCCTGTCTCGTCCA 73
LoeAlaUGCa    GGGCGUGUAGCUCAGUUGGU---AGAGCGCUUGCCUUGCAAGCAAGAG--------------GU-CUAGAGUUCGAUUCUCUACUCGUCCA 73
LwAlaUGCa     GGGCTAGTAGCACAGT-GGT---AGCGCGCTTCCCTTGCAAGGAAGAG--------------GT-CACCGGTTCGATCCCGGTCTAGTCCA 72
YlAlaUGCa     GGGCGTGTAGCACAGT-GGT---AGCGCGTGGGCTTTGCATGCTTGAG--------------GT-CACCGGTTCGAACCCGGTCTCGTCCA 72
PtaAlaUGCa    GGGCGTGTGGCGTAGTTGGT---AGCGCGTCTGCCTTGCAAGCAGAAG--------------GT-CATTGGTTCGATTCCGATCTCGTCCA 73
PiuAlaUGCa    GGGCGTGTGGCGTAGTTGGT---AGCGCGTTCGCCTTGCAAGCGAAAG--------------GT-CATCGGTTCGACTCCGGTCTCGTCCA 73
CaPheGAA      GCAGCTTTAGCTCATCTGGG---AGAGCGTCAGACTGAAGATCTGGAG--------------GT-ACTGTGTTCGAGCCACAGAAGCTGCA 73
SpPheGAAa     GTCGCAATGGTGTAGTTGGG---AGCATGACAGACTGAAGATCTGTTG--------------GT-CATCGGTTCGATCCCGGTTTGTGACA 73
SjPheGAAb     GTCGCAATGGTGTAGTTGG----AGCATGACAGACTGAAGATCTGTTG--------------GT-CACCGGTTCGATCCCGGTTTGTGACA 72
KlPheGAAa     GCTGCTTTAGCTCAGTTGGG---AGAGCGTATGACTGAAGTTCATAAG--------------GT-CCTGTGTTCGATCCACAGAAGCAGCA 73
MefPheGAAa    GCGGCTTTAGCTCAGTTGGG---AGAGCGTCAGACTGAAGATCTGAAG--------------GT-CCTGTGTTCGATCCACAGAAATCGCA 73
CaMetCAU      AGCATTGTGACGCAGT-GGA---AGCGTGCGAGGCTCATAACCTCGAA--------------GT-CCCTGGATCGAAACCAGGCAATGCTA 72
PtaMetCAUf    AGTATAATAGTGTAAT-GGT---AACACATTAGACTCATAATCTAAA---------------GA-TTCATGTTCAAATCATGATTATGCAA 71
WicMetCAUg    AGCACTATAATGTAAT-GGAT--AACATATTAGACTCATGTTCTATT---------------AA-TATAGGTTCAAGTCCTATTAGTGCAT 72
CapMetCAUe    AGCAATATAATGTAATTGGT---AACATATAAGATTCATGCTTTTAT---------------TA-TGATAGTTCGAGTCTATCTATTGCAA 72
CaArgACGa     CTGGCTGTGGCCTAAT-GGT---AAGGCACTTGACTACGGATCAAGGG--------------AT-TGCAGGTTCGAGTCCTGTCAGCCAGG 72
PiuArgACGa    CTCCACATAGCGTAAT-GGTG--ATCGCGTCTGACTACGGATCAGAAG--------------AT-TCTAGGTTCGAGTCCTAGTGTGGAGG 73
SpArgACGa     GGTCTCGTGGCCCAAT-GGTT--AAGGCGCTTGACTACGGATCAAGAG--------------AT-TCCAGGTTCGACTCCTGGCGGGATCG 73
PtaArgACGa    TTCCTCATGGCCCAAT-GGTC--ACGGCGTCTGGCTACGAACCAGAAG--------------AT-TCCAGGTTCGACTCCTGGTGGGGAAG 73
CaIleAAUa     GGTCCCTTGGCCCAGTTGGTT--AAGGCGTGGTGCTAATAACGCCAAG--------------AT-CAGCAGTTCGATCCTGCTAGGGACCA 74
PtaIleAAUa    GGTCCCTTGGCCCAGTTGGTT--AAGGCGCCGTGCTAATAACGCGGAG--------------AT-CAGCAGTTCGATCCTGCTAGGGACCA 74
YlIleAAUa     GGTCCCTTGGCCCAGTTGGTT--AAGGCGTCGTGCTAATAACGCGAAG--------------AT-CAGCAGTTCGATCCTGCTAGGGACCA 74
CaThrAGUa     GCTCTTATGGCCAAGTTGGT---AAGGCACCTCACTAGTAATGAGGAG--------------AT-CGTAGGTTCGAATCCTGCTGAGAGCA 73
CaThrCGU      GCCTACTTGGCCAAGT-GGT---AAGGCACTCCACTCGTAATGGAGCG--------------AT-CGTCGGTTCAATCCCGACAGTGGGCA 72
CaThrUGUa     GCCTTTTTAGCTTAGT-GGT---AGAGCGTTGCACTTGTAATGCAAAG--------------GT-CGCTAGTTCAATTCTGGCAAGAGGCA 72
ErgThrUGUa    GCCCTGTTAGCTTAGT-GGT---AGAGCGTCGCACTTGTAATGCGAAG--------------GT-CACTAGTTCAATTCTGGTACAGGGCA 72
PiuThrUGUa    GCCTCTTTAGCTTAGT-GGT---AGAGCGTCGCACTTGTAATGCGGAG--------------GT-CGCTAGTTCAATTCTGGCAAGTGGCA 72
PtaThrUGUa    GCCGTTATGGCCAAGT-GGT---AAGGCACCGCACTTGTAATGCGGCG--------------AT-CCGCGGTTCAATCCCGCGTGACGGCA 72
```

Fig. S1: Sequence alignment of the representative tRNAs used in the phylogenetic tree reconstructions. All introns have been removed. tRNAs are named by species abbreviation (Table S1), tRNA type, and anticodon. Gene duplicates are denoted by lowercase letters. Gene lengths are printed at the 3’ ends.

Fig. S2: Phylogenetic tree computed on the alignment of 172 tRNA sequences (Fig. S1) using the Maximum-Likelihood approach as implemented in FastTree v. 2.1.7 with 1000 replicates. Support values are shown as relative numbers. Branches are colored according to Fig. 1B. The scale bar denotes substitutions per site.

Fig. S3: Phylogenetic tree computed on the alignment of 172 tRNA sequences (Fig. S1) using the Neighbour-Joining approach as implemented in ClustalW v. 2.1 with 1000 replicates. Support values are shown as absolute numbers. Branches are colored according to Fig. 1B. The scale bar denotes substitutions per site.

Fig. S4: Phylogenetic tree computed on the alignment of 172 tRNA sequences (Fig. S1) using the Bayesian approach as implemented in Phase v. 2 with 1,000,000 iterations. Branches are colored according to Fig. 1B. The scale bar denotes substitutions per site.

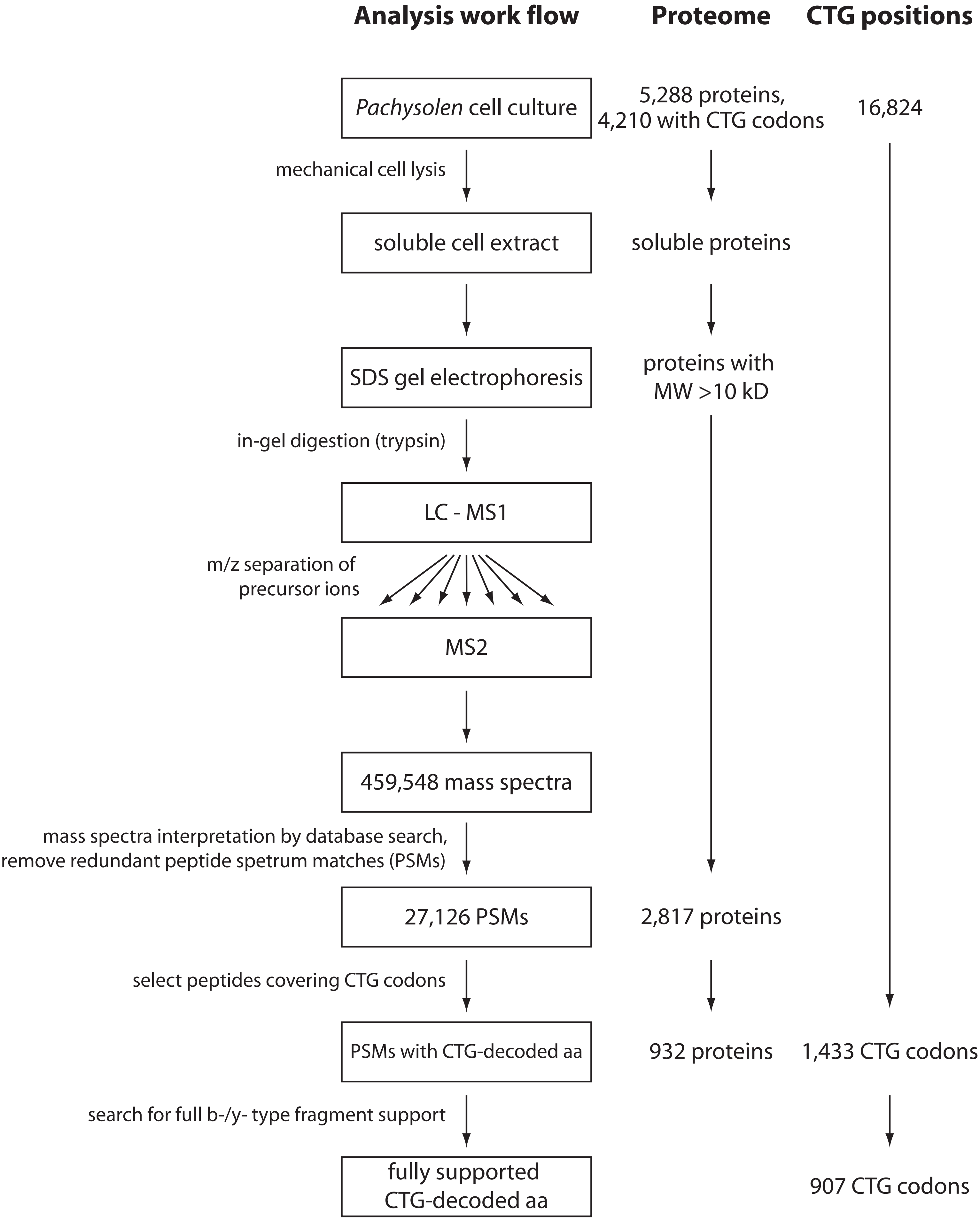


Fig. S5: Summary of the proteome analysis workflow. For the database search we generated a *Pachysolen* gene prediction dataset, in which all codons were iteratively translated into all amino acids (except isoleucine and leucine, which have identical molecular weights). Therefore, the database consists of 1,216 (64 x 19) times 5,288 proteins. The peptides with differently translated codons distinguish in their total MW, accordingly have different precursur ion masses and result in different spectra.

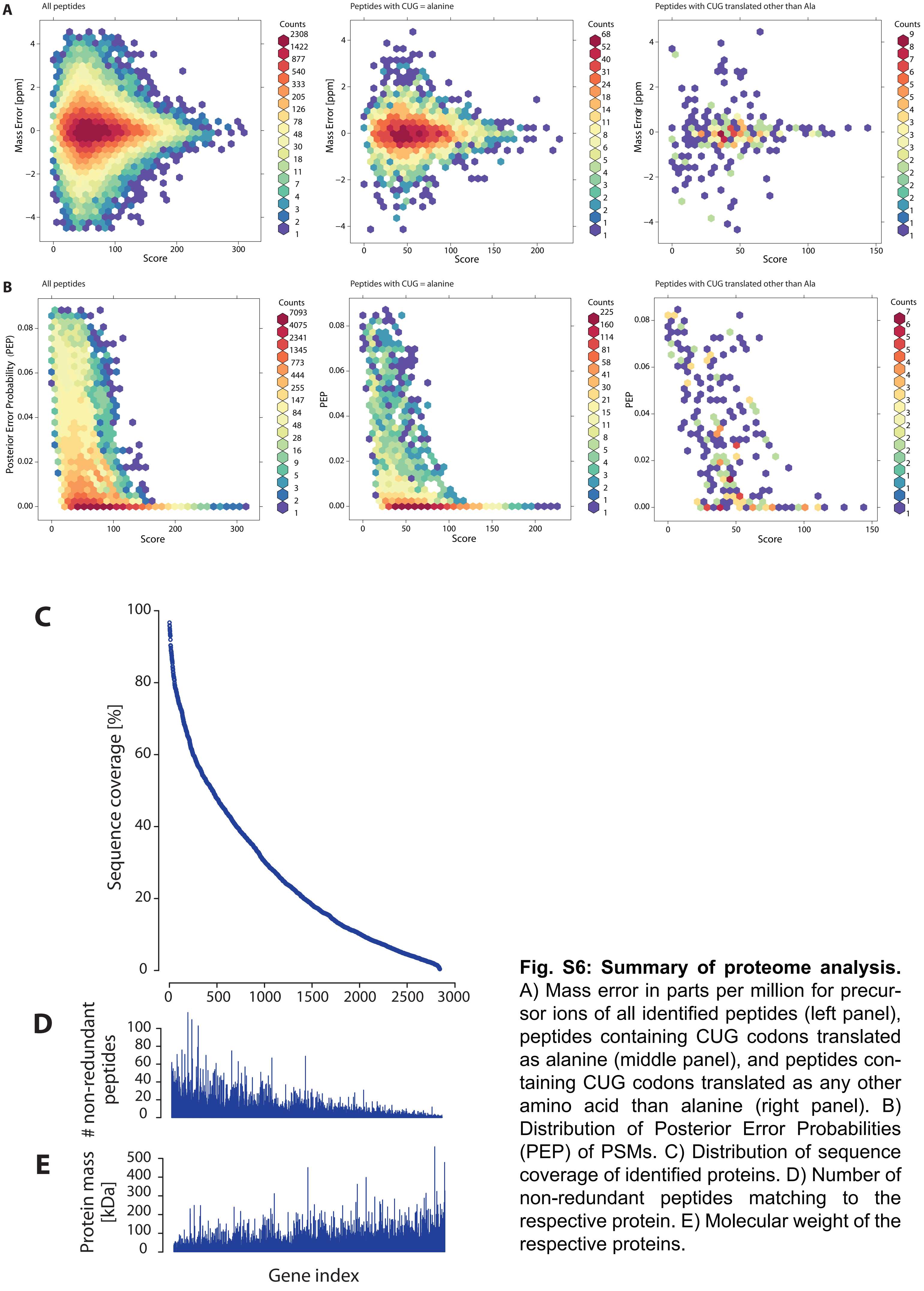


Fig. S6: Summary of proteome analysis. A) Mass error in parts per million for precursor ions of all identified peptides (left panel), peptides containing CUG codons translated as alanine (middle panel), and peptides containing CUG codons translated as any other amino acid than alanine (right panel). B) Distribution of Posterior Error Probabilities (PEP) of PSMs. C) Distribution of sequence coverage of identified proteins. D) Number of non-redundant peptides matching to the respective protein. E) Molecular weight of the respective proteins.

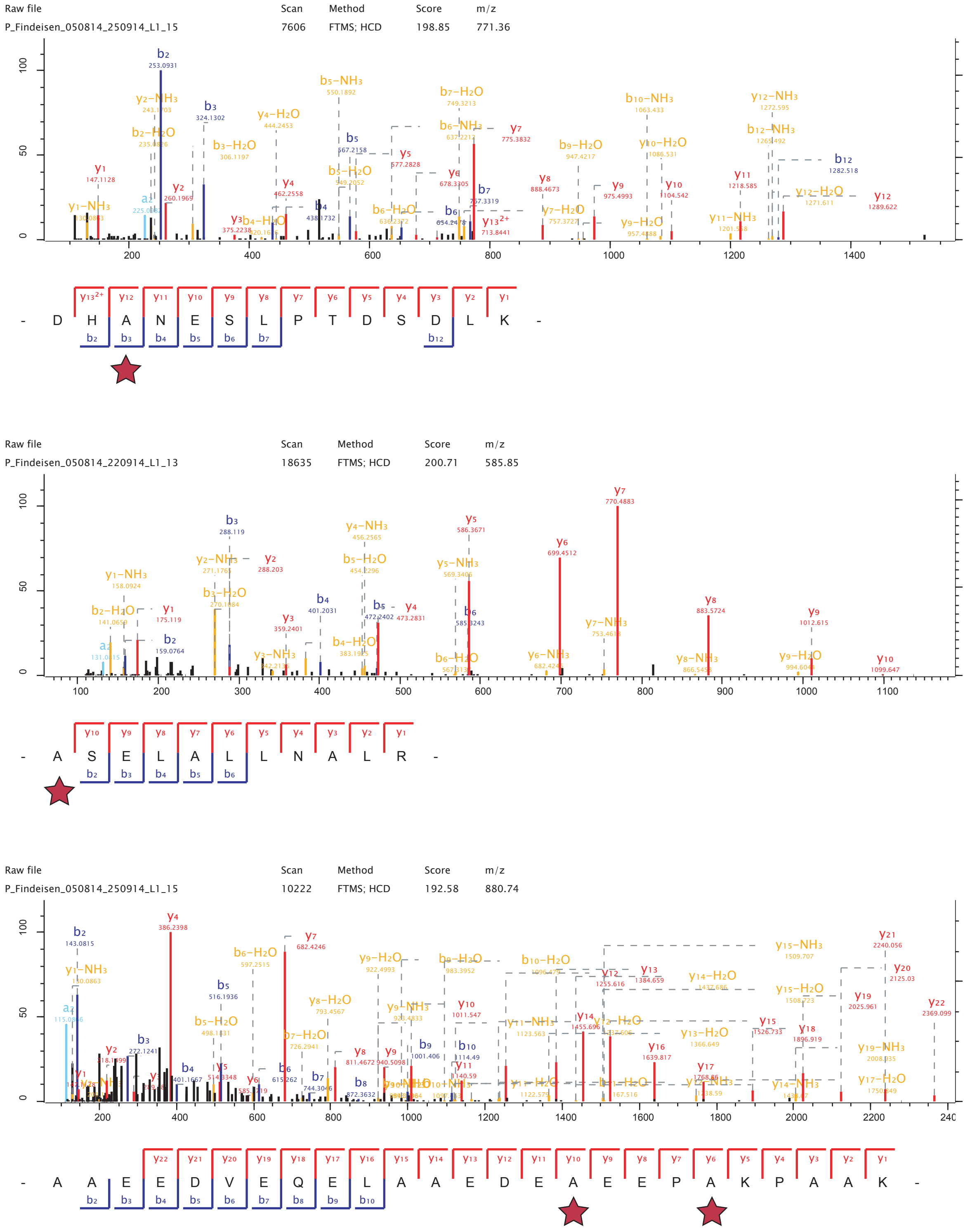
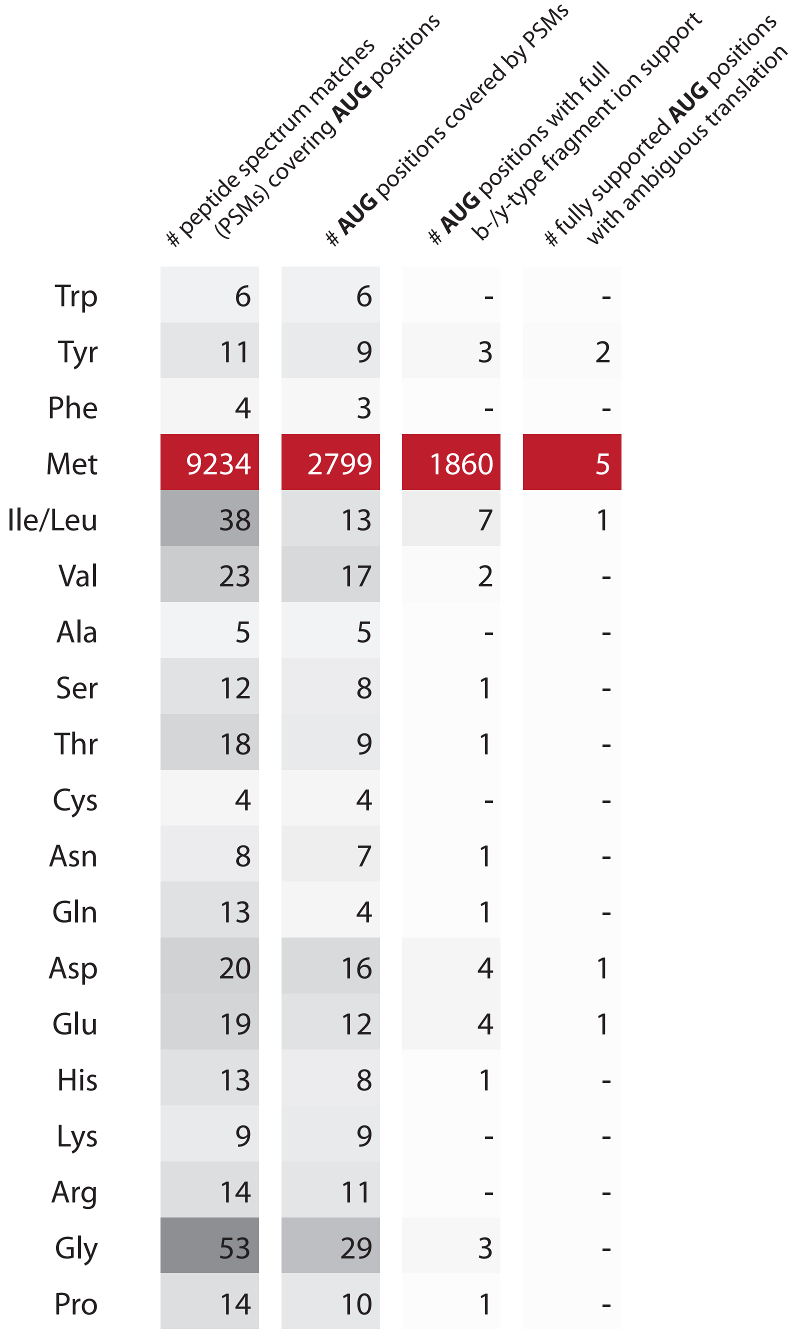


Fig. S7: Representative LC-MS/MS spectra featuring CUG codons translated as alanine (marked with stars).



	# peptide spectrum matches (PSMs) covering **AUG** positions	# **AUG** positions covered by PSMs	# **AUG** positions with full b-/y-type fragment ion support	# fully supported **AUG** positions with ambiguous translation
Trp	6	6	-	-
Tyr	11	9	3	2
Phe	4	3	-	-
Met	9234	2799	1860	5
Ile/Leu	38	13	7	1
Val	23	17	2	-
Ala	5	5	-	-
Ser	12	8	1	-
Thr	18	9	1	-
Cys	4	4	-	-
Asn	8	7	1	-
Gln	13	4	1	-
Asp	20	16	4	1
Glu	19	12	4	1
His	13	8	1	-
Lys	9	9	-	-
Arg	14	11	-	-
Gly	53	29	3	-
Pro	14	10	1	-

Fig. S8: **Decoding of the unambiguously decoded AUG codon.** Number and distribution of observed AUG translations in all peptide spectrum matches from MS/MS analysis.

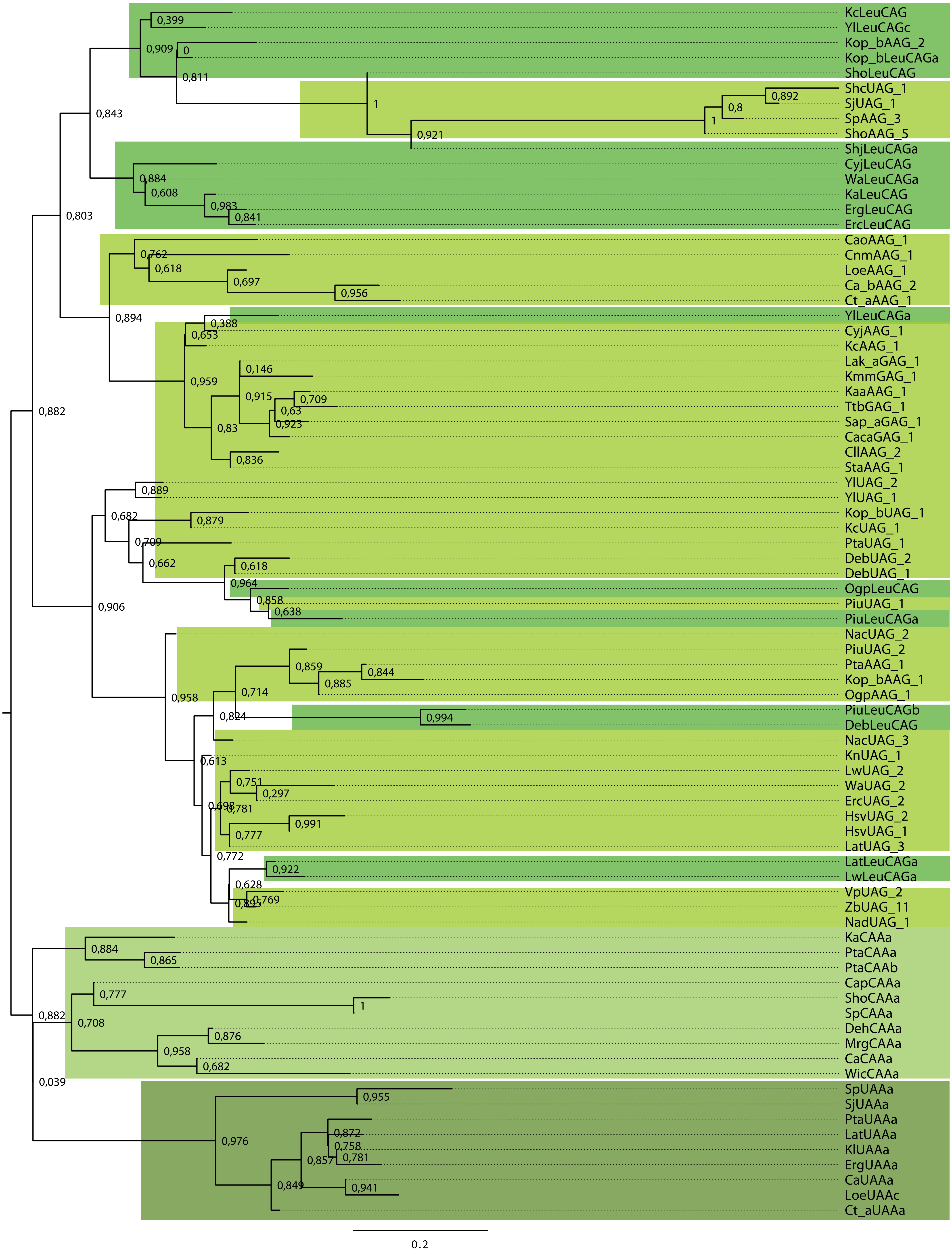


Fig. S9: Phylogenetic tree computed on the alignment of 80 representative Leu-tRNA sequences using the Maximum-Likelihood approach as implemented in FastTree v. 2.1.7 with 1000 replicates. Support values are shown as relative numbers. Branches are colored according to Fig. 1B. The scale bar denotes substitutions per site.

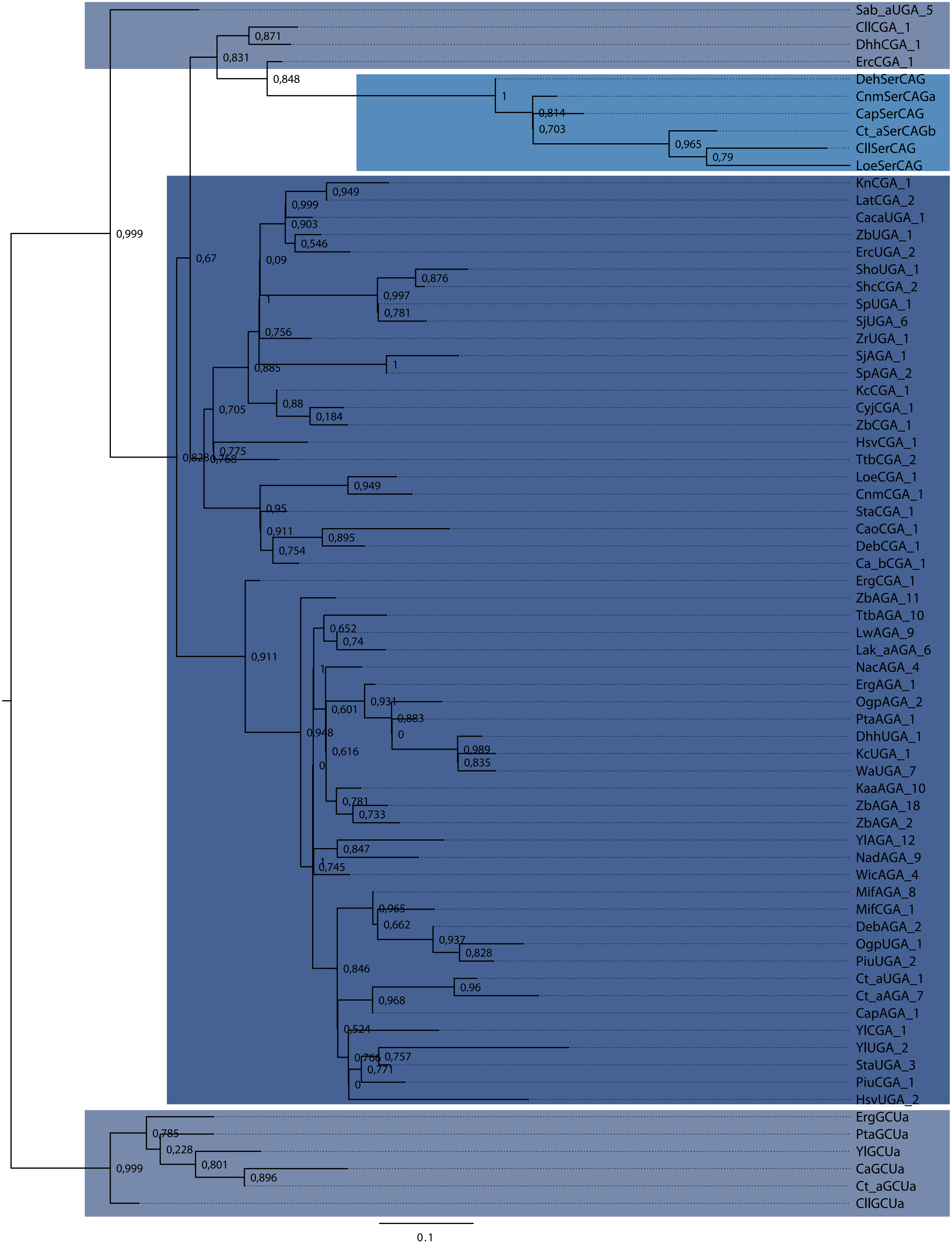


Fig. S10: Phylogenetic tree computed on the alignment of 70 representative Ser-tRNA sequences using the Maximum-Likelihood approach as implemented in FastTree v. 2.1.7 with 1000 replicates. Support values are shown as relative numbers. Branches are colored according to Fig. 1B. The scale bar denotes substitutions per site.

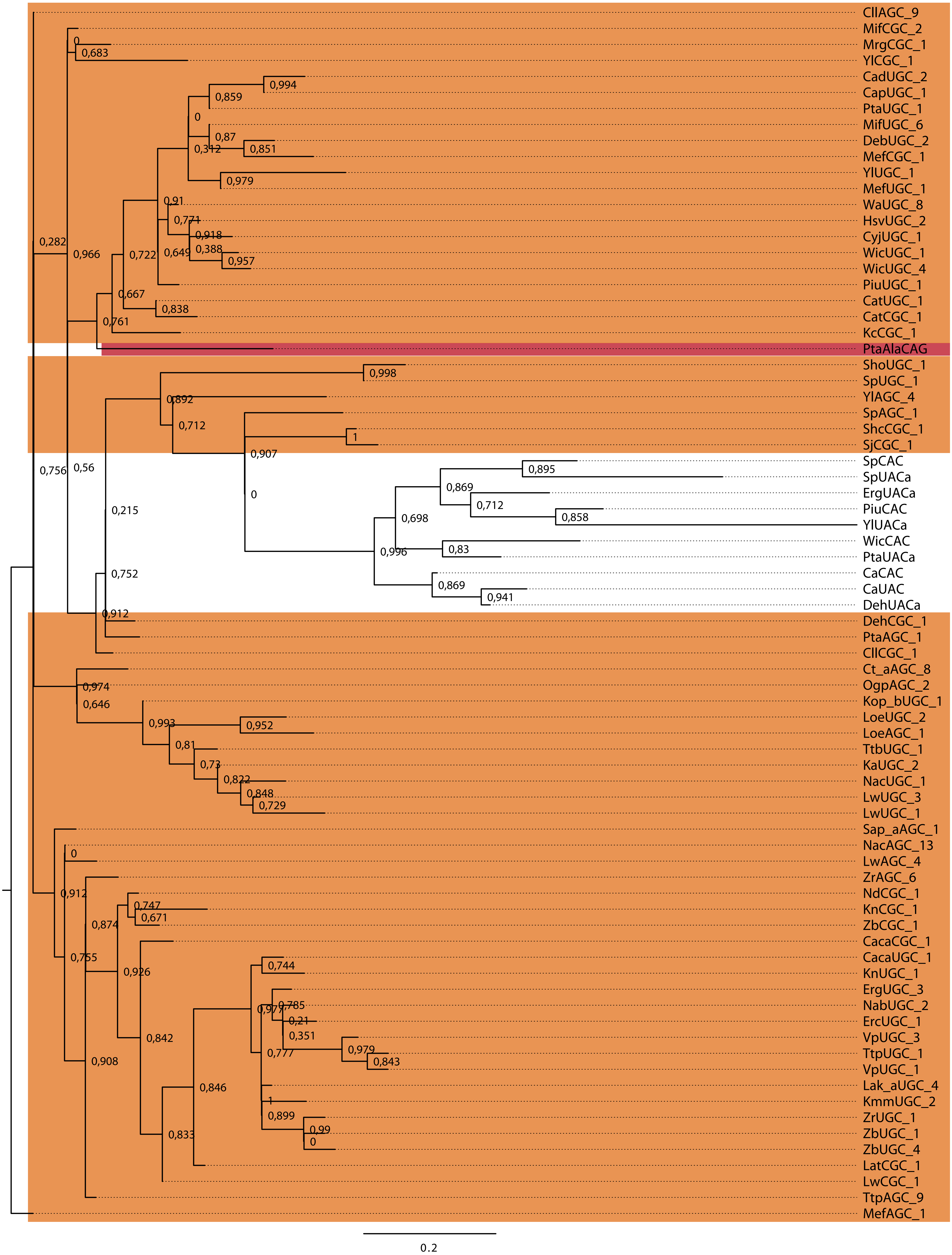
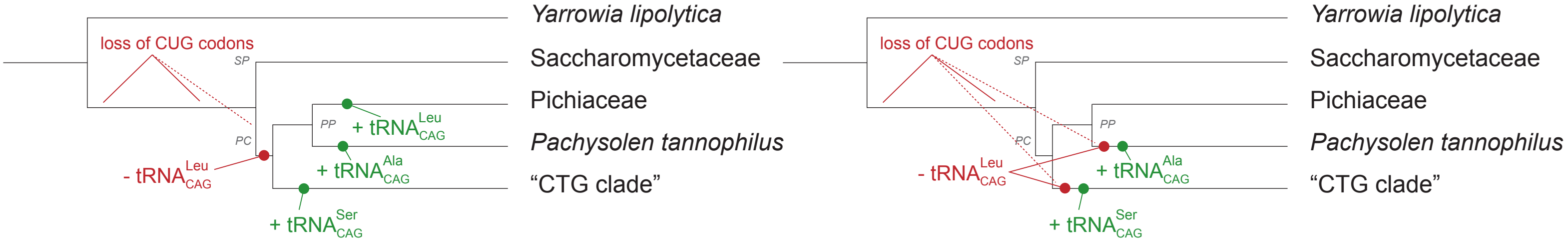
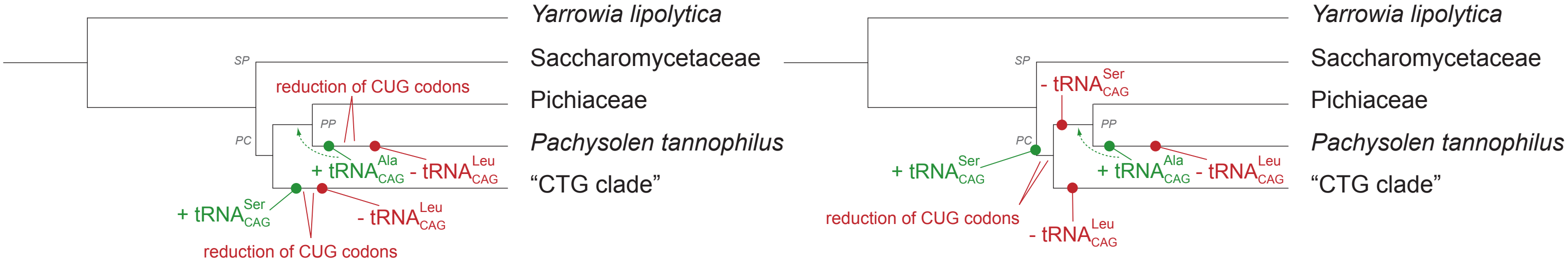


Fig. S11: Phylogenetic tree computed on the alignment of 76 representative Ala-tRNA sequences, with Val-tRNAs as outgroup, using the Maximum-Likelihood approach as implemented in FastTree v. 2.1.7 with 1000 replicates. Support values are shown as relative numbers. Branches are colored according to Fig. 1B. The scale bar denotes substitutions per site.

Codon capture



Ambiguous intermediate



tRNA loss driven codon reassignment

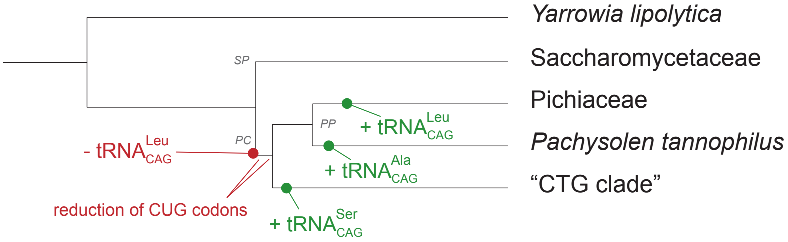


Fig. S12: **The mechanism of CUG codon reassignment.** The scheme contrasts the presence of tRNA_{CAG}s and evolution of CUG codons according to the *tRNA loss driven codon reassignment* hypothesis with assumptions based on the *codon capture* and *ambiguous intermediate* theories. The schemes present alternative but less likely scenarios compared to the most probable sequence of events as shown for each hypothesis in Fig. 4. Mixed models seem extremely unlikely and are not shown.

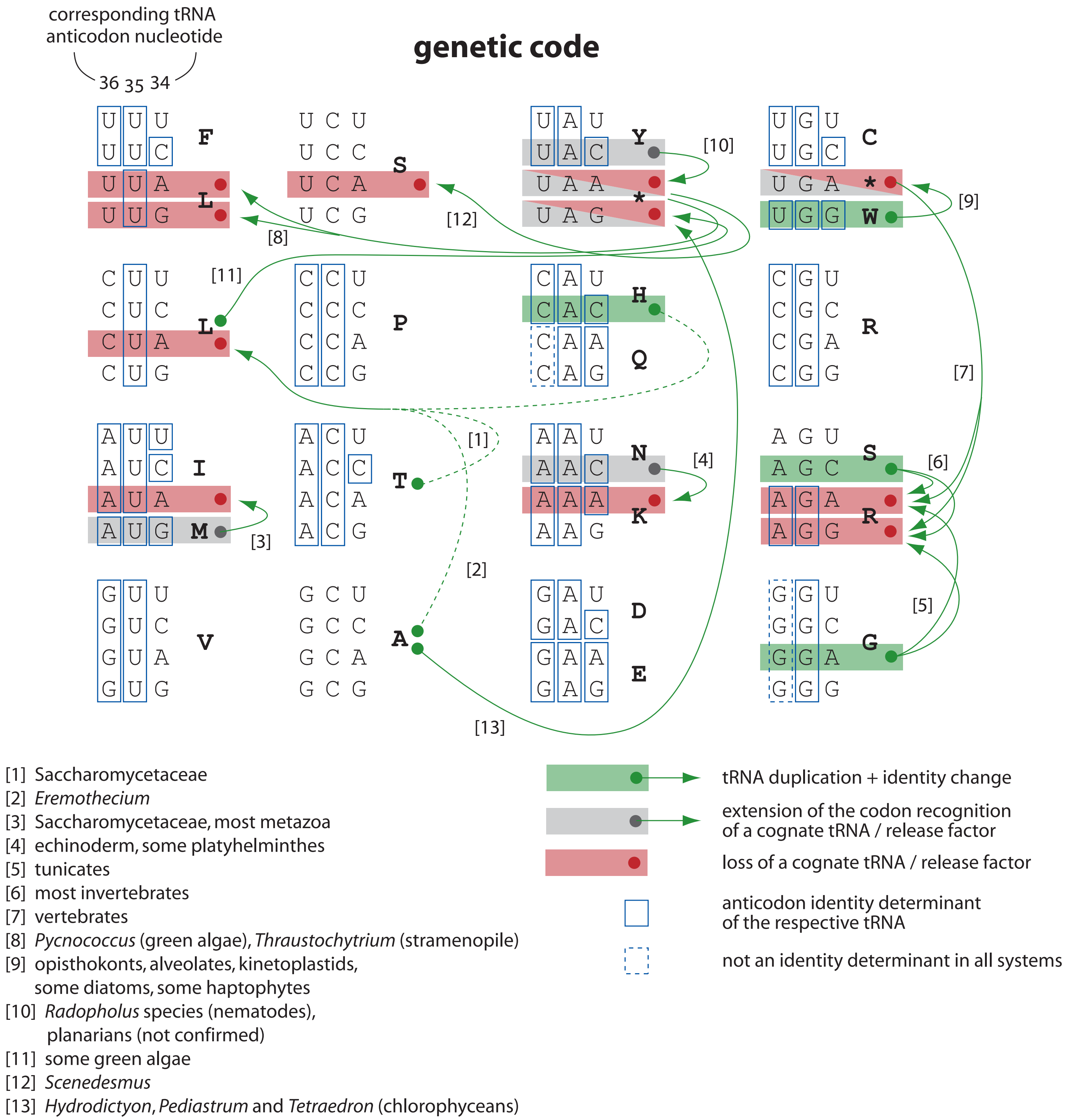


Fig. S13: Codon reassignment in mitochondria. In mitochondria, codon reassignments have happened rather frequently. However, mitochondrial genomes are extremely small and changes in decoding only affect a few genes. In contrast to most nuclear genomes, where four-codon families require at least two tRNAs, most mitochondria require only one tRNA with a U in the wobble position for decoding all four codons. Therefore, reassignments mainly affect codons in mixed codon boxes such as stop codons, AUA isoleucine, AGR arginine, and AAA lysine. These reassignments might have happened according to the *tRNA loss driven codon reassignment* hypothesis because all mentioned codons could be translated, although rather inefficiently, by another tRNA of the codon family box. Usually, the ambiguous decoding is not competitive and happens only occasionally. However, it preserves the organism's viability after loss of one of the respective tRNAs. For example, the TGA stop codon in mitochondrial genomes can be translated by the tryptophan tRNA, a decoding which has also been shown to happen at low levels under selective pressure in *Bacillus subtilis*. Subsequently to tRNA loss, the inefficient capturing tRNAs are often optimized for decoding both the cognate and the captured codon by mutations (e.g. the tryptophan tRNAs in those mitochondria, where they also decode TGA codons, have a C to U mutation in the wobble position of the anticodon creating a UCU that can pair with both UGA and UGG) and/or by nucleotide modifications such as 5'-taurinomethylation of the wobble uridine in mitochondrial methionine tRNA(CAU) and lysidine and agmatidine modification of the wobble C in isoleucine tRNA(AUA) in bacteria and archaea, respectively. Instead of adjusting the tRNA for higher decoding fidelity, the aaRS might become adapted by compensatory mutations to slightly different amino acid identity patterns. This could happen by enabling two different tRNAs be charged by the same aaRS, or by gene duplication and subsequent adaptation of the two aaRSs to the respective tRNAs. An example for the latter case is the decoding of the CUN codons in yeast mitochondria by threonine.

Table S1: Summary of the leucine, serine, and alanine tRNA data.

Species	Abbreviation	GenBank accession numbers	Reference	TTA	TTG	CTT	CTC	CTA	CTG	TCT	TCC	TCA	TCG	AGT	AGC	GCT	GCC	GCA	GCG
Candida albicans	Ca_b	CH672346, CH672349-CH672350, CH672354, CM000309-CM000313, GG670278, GG670282	Butler G, Rasmussen MD, Lin MF et. al. (2009) Evolution of pathogenicity and sexual reproduction in eight Candida genomes. Nature 459, 657-662.	5	6	2	0	0	1	4	0	3	1	0	2	6	0	2	0
Candida bracarensis	Cnb	CAPU01000001-CAPU01000251	Gabaldon T, Martin T, Marcet-Houben M et. al. (2013) Comparative genomics of emerging pathogens in the Candida glabrata clade. BMC Genomics 14, 623.	4	9	0	1	4	0	9	0	3	1	0	3	11	0	6	0
Candida castelii	Caca	CAPW01000001-CAPW01000101	Gabaldon T, Martin T, Marcet-Houben M et. al. (2013) Comparative genomics of emerging pathogens in the Candida glabrata clade. BMC Genomics 14, 623.	2	5	0	1	2	0	5	0	2	1	0	2	5	0	3	1
Candida dubliniensis	Cad	FM992688-FM992695	Jackson AP, Gamble JA, Yeomans T et. al. (2009) Comparative genomics of the fungal pathogens Candida dubliniensis and Candida albicans. Genome Res 19, 2231-2244.	5	5	2	0	0	1	4	0	3	1	0	1	6	0	2	0
Candida glabrata	Cgl	CR380947-CR380959	Dujon B, Sherman D, Fischer G et. al. (2004) Genome evolution in yeasts. Nature 430, 35-44.	3	8	0	1	4	0	9	0	2	1	0	3	10	0	5	0
Candida maltosa	Cnm	AOGT01000001-AOGT01002947	http://www.ncbi.nlm.nih.gov/nuccore/AOGT00000000.1	1	0	1	0	0	1	0	0	2	1	0	0	0	0	1	0
Candida metapsilosis	Came	CBZN020000001-CBZN020000096	http://www.ncbi.nlm.nih.gov/nuccore/CBZN000000000.2	3	3	1	0	0	1	2	0	3	1	0	1	3	0	3	0
Candida nivariensis	Cdn	CAPV01000001-CAPV01000123	Gabaldon T, Martin T, Marcet-Houben M et. al. (2013) Comparative genomics of emerging pathogens in the Candida glabrata clade. BMC Genomics 14, 623.	4	10	0	1	4	0	10	0	3	1	0	3	12	0	6	0
Candida orthopsilosis	Cao	HE681719-HE681726	Riccombeni A, Vidanes G, Proux-Wera E et. al. (2012) Sequence and analysis of the genome of the pathogenic yeast Candida orthopsilosis. PLoS One 7, e35750.	3	1	1	0	0	1	2	0	2	1	0	1	3	0	2	0
Candida parapsilosis	Cap	CABE01000001-CABE01000024	Butler G, Rasmussen MD, Lin MF et. al. (2009) Evolution of pathogenicity and sexual reproduction in eight Candida genomes. Nature 459, 657-662.	3	3	1	0	0	1	2	0	3	1	0	1	3	0	3	0
Candida tenuis	Cat	GL996507-GL996531	Wohlbach DJ, Kuo A, Sato TK et. al. (2011) Comparative genomics of xylose-fermenting fungi for enhanced biofuel production. Proc Natl Acad Sci U S A 108, 13212-13217.	2	5	1	0	0	1	4	0	2	1	0	3	4	0	2	1
Candida tropicalis	Ct_a	GG692395-GG692418	Butler G, Rasmussen MD, Lin MF et. al. (2009) Evolution of pathogenicity and sexual reproduction in eight Candida genomes. Nature 459, 657-662.	6	6	2	0	0	1	7	0	4	1	0	3	9	0	5	0
Clavispora lusitaniae	Cll	CH408076-CH408083, GG670277	Butler G, Rasmussen MD, Lin MF et. al. (2009) Evolution of pathogenicity and sexual reproduction in eight Candida genomes. Nature 459, 657-662.	2	11	3	0	0	1	7	0	1	2	0	3	9	0	3	1
Cyberlindnera jadinii	Cyj	DG000065-DG000077	Tomita Y, Ikeo K, Tamakawa H et. al. (2012) Genome and Transcriptome Analysis of the Food-Yeast Candida utilis. PLoS One 7, e37226.	1	7	3	0	1	1	6	0	2	1	0	2	6	0	2	0

Species	Abbreviation	GenBank accession numbers	Reference	TTA	TTG	CTT	CTC	CTA	CTG	TCT	TCC	TCA	TCG	AGT	AGC	GCT	GCC	GCA	GCG
Debaryomyces hansenii	Deh	CR382133-CR382139	Dujon B, Sherman D, Fischer G et. al. (2004) Genome evolution in yeasts. Nature 430, 35-44.	9	4	2	0	0	1	6	0	3	1	0	3	7	0	4	1
Debaryomyces hansenii var hansenii	Dhh	AHBE01000001-AHBE01000541	Kumar S, Randhawa A, Ganesan K et. al. (2012) Draft Genome Sequence of Salt-Tolerant Yeast Debaryomyces hansenii var. hansenii MTCC 234. Eukaryot Cell 11, 961-962.	1	0	2	0	0	1	0	0	2	1	0	2	0	0	0	1
Dekkera bruxellensis	Deb	KB454452-KB454483	Piskur J, Ling Z, Marcet-Houben M et. al. (2012) The genome of wine yeast Dekkera bruxellensis provides a tool to explore its food-related properties. Int J Food Microbiol 157, 202-209.	1	2	2	0	2	1	2	0	0	1	0	2	3	0	3	0
Eremothecium cymbalariae	Erc	CP002497-CP002504	Wendland J, Walther A (2011) Genome evolution in the eremothecium clade of the Saccharomyces complex revealed by comparative genomics. G3 (Bethesda) 1, 539-548.	2	8	0	1	2	1	5	0	2	1	0	3	5	0	3	1
Eremothecium gossypii	Erg	AE016814-AE016820	Dietrich FS, Voegeli S, Brachat S et. al. (2004) The Ashbya gossypii genome as a tool for mapping the ancient Saccharomyces cerevisiae genome. Science 304, 304-7.	2	7	0	1	6	1	7	0	2	3	0	3	7	0	4	2
Hanseniaspora vineae	Hsv	JFAV02000001-JFAV02000305	Giorello FM, Berna L, Greif G et. al. (2014) Genome Sequence of the Native Apiculate Wine Yeast Hanseniaspora vineae T02/19AF. Genome Announc 2, e00530-14.	5	9	0	0	3	0	7	0	3	1	0	3	7	0	5	0
Kazachastania naganishii	Kn	HE978314-HE978326	Gordon JL, Armisen D, Proux-Wera E et. al. (2011) Evolutionary erosion of yeast sex chromosomes by mating-type switching accidents. Proc Natl Acad Sci U S A 108, 20024-20029.	3	5	1	0	3	0	7	0	1	1	0	3	7	0	1	1
Kazachstania africana	Kaa	HE650821-HE650832	Gordon JL, Armisen D, Proux-Wera E et. al. (2011) Evolutionary erosion of yeast sex chromosomes by mating-type switching accidents. Proc Natl Acad Sci U S A 108, 20024-20029.	14	2	2	0	2	0	11	0	4	1	0	4	12	0	5	0
Kluyveromyces aestuarii	Ka	AEAS01000001-AEAS01000336	Baker CR, Tuch BB, Johnson AD (2011) Extensive DNA-binding specificity divergence of a conserved transcription regulator. Proc Natl Acad Sci U S A 108, 7493-7498.	3	7	0	1	2	1	6	0	3	1	0	2	7	0	5	0
Kluyveromyces lactis	Kl	CR382121-CR382126	Dujon B, Sherman D, Fischer G et. al. (2004) Genome evolution in yeasts. Nature 430, 35-44.	3	7	0	1	2	1	6	0	2	1	0	2	7	0	3	0
Kluyveromyces marxianus var marxianus	Kmm	JH924896-JH924903	Jeong H, Lee DH, Kim SH et. al. (2012) Genome Sequence of the Thermotolerant Yeast Kluyveromyces marxianus var. marxianus KCTC 17555. Eukaryot Cell 11, 1584-1585	3	8	0	1	2	1	7	0	2	1	0	2	7	0	4	0
Kluyveromyces wickerhamii	Klw	AEAV01000001-AEAV01000510	Baker CR, Tuch BB, Johnson AD (2011) Extensive DNA-binding specificity divergence of a conserved transcription regulator. Proc Natl Acad Sci U S A 108, 7493-7498.	4	6	0	1	2	1	6	0	3	1	0	1	8	0	4	0
Komagataella pastoris	Kop_b	FN392319-FN392325	De Schutter K, Lin YC, Tiels P et. al. (2009) Genome sequence of the recombinant protein production host Pichia pastoris. Nat Biotechnol 27, 561-566.	1	3	2	0	1	2	4	0	2	1	0	2	5	0	2	0
Kuraishia capsulata	Kc	HG793125-HG793131	Morales L, Noel B, Porcel B et. al. (2013) Complete DNA sequence of Kuraishia capsulata illustrates novel genomic features among budding yeasts (Saccharomycotina). Genome Biol Evol 5, 2524-2539.	1	3	2	0	1	1	2	0	2	1	0	1	3	0	2	1
Lachancea kluyveri	Lak_a	CM000687-CM000694	Cliften, P., Sudarsanam et. al. (2003) Finding functional features in Saccharomyces genomes by phylogenetic footprinting. Science 301, 71-76.	3	11	0	1	3	1	11	0	2	1	0	4	13	0	4	0

Species	Abbreviation	GenBank accession numbers	Reference	TTA	TTG	CTT	CTC	CTA	CTG	TCT	TCC	TCA	TCG	AGT	AGC	GCT	GCC	GCA	GCG
Lachancea thermotolerans	Lat	CU928165-CU928171, CU928180	Souciet JL, Dujon B, Gaillardin C et. al. (2009) Comparative genomics of protoploid Saccharomycetaceae. Genome Res 19, 1696-1709.	2	8	0	2	3	3	9	0	2	2	0	3	10	0	3	2
Lachancea waltii	Lw	AADM01000001-AADM01000713	Manolis Kellis, Bruce W. Birren, Eric S. Lander (2004) Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae. Nature 428, 617-624.	2	0	0	2	4	2	9	0	3	2	0	5	11	0	4	1
Lodderomyces elongisporus	Loe	CH981524-CH981534, DS236866-DS236881	Butler G, Rasmussen MD, Lin MF et. al. (2009) Evolution of pathogenicity and sexual reproduction in eight Candida genomes. Nature 459, 657-662.	3	3	2	0	0	1	2	0	3	1	0	2	4	0	2	0
Metschnikowia fructicola	Mef	ANFW01000001-ANFW01008430	Hershkovitz V, Sela N, Taha-Salaime L et. al. Right (2013) De-novo assembly and characterization of the transcriptome of Metschnikowia fructicola reveals differences in gene expression following interaction with Penicillium digitatum and grapefruit peel. BMC Genomics 14, 168.	3	18	4	0	0	4	15	0	4	6	0	4	15	0	3	3
Meyerozyma guilliermondii	Mrg	CH408155-CH408162, DS236883	Butler G, Rasmussen MD, Lin MF et. al. (2009) Evolution of pathogenicity and sexual reproduction in eight Candida genomes. Nature 459, 657-662.	2	5	2	0	0	1	4	0	3	2	0	3	5	0	3	1
Millerozyma farinosa	Mif	FO082046-FO082059	Louis VL, Despons L, Friedrich A et. al. (2012) Pichia sorbitophila, an Interspecies Yeast Hybrid, Reveals Early Steps of Genome Resolution After Polyploidization. G3 (Bethesda) 2, 299-311.	4	14	4	0	0	2	10	0	4	2	0	6	12	0	8	2
Nakaseomyces bacillisporus	Nab	CAPX01000001-CAPX01000186	Gabaldon T, Martin T, Marcet-Houben M et. al. (2013) Comparative genomics of emerging pathogens in the Candida glabrata clade. BMC Genomics 14, 623.	7	4	0	1	1	0	6	0	3	1	0	3	7	0	5	0
Nakaseomyces delphensis	Nd	CAPT01000001-CAPT01000179	Gabaldon T, Martin T, Marcet-Houben M et. al. (2013) Comparative genomics of emerging pathogens in the Candida glabrata clade. BMC Genomics 14, 623.	4	7	0	1	4	0	9	0	3	1	0	3	4	0	6	1
Naumovozyma castellii	Nac	AACF01000001-AACF01000570	Cliften, P., Sudarsanam et. al. (2003) Finding functional features in Saccharomyces genomes by phylogenetic footprinting. Science 301, 71-76.	7	9	2	0	4	0	13	0	4	1	0	4	13	0	5	0
Naumovozyma dairenensis	Nad	HE580267-HE580278	Gordon JL, Armisen D, Proux-Wera E et. al. (2011) Evolutionary erosion of yeast sex chromosomes by mating-type switching accidents. Proc Natl Acad Sci U S A 108, 20024-20029.	9	8	1	0	4	0	11	0	3	1	0	5	10	0	5	0
Ogataea parapolymorpha	Ogp	AEOI01000001-AEOI01000013	Ravin NV, Eldarov MA, Kadnikov VV et. al. (2013) Genome sequence and analysis of methylotrophic yeast Hansenula polymorpha DL1. BMC Genomics 14, 837.	1	3	1	0	1	1	2	0	1	1	0	2	3	0	2	0
Pachysolen tannophilus	Pta	CAHV01000001-CAHV01000267	Liu X, Kaas RS, Jensen PR et. al. (2012) Draft genome sequence of the yeast Pachysolen tannophilus CBS 4044/NRRL Y-2460. Eukaryot Cell 11, 827.	5	2	2	0	1	1	3	0	4	1	0	3	3	0	4	0
Pichia kudriavzevii	Piu	ALNQ01000001-ALNQ01000621	Chan GF, Gan HM, Ling HL et. al. (2012) Genome Sequence of Pichia kudriavzevii M12, a Potential Producer of Bioethanol and Phytase. Eukaryot Cell 11, 1300-1301.	2	3	3	0	2	2	5	0	2	1	0	2	0	0	11	0
Saccharomyces arboricola	Saa	CM001563-CM001578	Liti G, Nguyen Ba AN, Blythe M et. al. (2013) High quality de novo sequencing and assembly of the Saccharomyces arboricolus genome. BMC Genomics 14, 69.	8	9	0	1	3	0	10	0	4	1	0	4	11	0	5	0

Species	Abbreviation	GenBank accession numbers	Reference	TTA	TTG	CTT	CTC	CTA	CTG	TCT	TCC	TCA	TCG	AGT	AGC	GCT	GCC	GCA	GCG
Saccharomyces bayanus	Sab_a	AACG02000001-AACG02000586	Cliften, P., Sudarsanam et. al. (2003) Finding functional features in Saccharomyces genomes by phylogenetic footprinting. Science 301, 71-76.	8	11	0	1	3	0	15	0	5	1	0	4	12	0	5	0
Saccharomyces cerevisiae	Sc_c	NC_001133-NC_001148	Otero JM, Vongsangnak W, Asadollahi MA et. al. (2010) Whole genome sequencing of Saccharomyces cerevisiae: from genotype to phenotype for improved metabolic engineering applications. BMC Genomics 11, 723.	7	10	0	1	3	0	11	0	3	1	0	2	11	0	5	0
Saccharomyces kudriavzevii	Sak	JH796010-JH798063	Cliften, P., Sudarsanam et. al. (2003) Finding functional features in Saccharomyces genomes by phylogenetic footprinting. Science 301, 71-76.	7	10	0	1	3	0	9	0	4	1	0	4	11	0	5	0
Saccharomyces mikatae	Smi	AACH01000001-AACH01002808	Kellis, M., Patterson et. al. (2003) Sequencing and comparison of yeast species to identify genes and regulatory elements. Nature 423, 241-254.	6	8	0	1	3	0	10	0	2	1	0	3	9	0	5	0
Saccharomyces paradoxus	Sap_a	AABY01000001-AABY01000832	Kellis, M., Patterson et. al. (2003) Sequencing and comparison of yeast species to identify genes and regulatory elements. Nature 423, 241-254.	7	0	0	1	3	0	11	0	3	1	0	4	11	0	5	0
Scheffersomyces stipitis	Shs	CM000437, CP000496-CP000502	Jeffries TW, Grigoriev IV, Grimwood J et. al. Right (2007) Genome sequence of the lignocellulose-bioconverting and xylose-fermenting yeast Pichia stipitis. Nat Biotechnol 25, 319-326.	3	8	3	0	0	1	6	0	2	2	0	2	7	0	3	0
Spathaspora arborariae	Sta	KI547134-KI547163	Lobo FP, Goncalves DL, Alves SL Jr et. al. (2014) Draft Genome Sequence of the D-Xylose-Fermenting Yeast Spathaspora arborariae UFMG-HM19.1AT. Genome Announc 2, e01163-13.	8	3	2	0	0	1	6	0	4	1	0	2	6	0	4	0
Spathaspora passalidarum	Shp	GL996499-GL996506	Wohlbach DJ, Kuo A, Sato TK et. al. (2011) Comparative genomics of xylose-fermenting fungi for enhanced biofuel production. Proc Natl Acad Sci U S A 108, 13212-13217.	6	5	2	0	0	1	6	0	3	1	0	2	7	0	3	0
Tetrapisispora blattae	Ttb	HE806316-HE806325	Gordon JL, Armisen D, Proux-Wera E et. al. (2011) Evolutionary erosion of yeast sex chromosomes by mating-type switching accidents. Proc Natl Acad Sci U S A 108, 20024-20029.	12	9	0	1	3	0	16	0	4	2	0	6	19	0	5	0
Tetrapisispora phaffii	Ttp	HE612856-HE612871	Gordon JL, Armisen D, Proux-Wera E et. al. (2011) Evolutionary erosion of yeast sex chromosomes by mating-type switching accidents. Proc Natl Acad Sci U S A 108, 20024-20029.	9	2	0	1	2	0	8	0	2	1	0	3	9	0	3	0
Torulaspora delbrueckii	Tod	HE616742-HE616749	Gordon JL, Armisen D, Proux-Wera E et. al. (2011) Evolutionary erosion of yeast sex chromosomes by mating-type switching accidents. Proc Natl Acad Sci U S A 108, 20024-20029.	3	8	0	1	4	0	7	0	2	1	0	3	8	0	4	0
Vanderwaltozyma polyspora	Vp	DS480378-DS480658	Scannell DR, Frank AC, Conant GC et. al. (2007) Independent sorting-out of thousands of duplicated gene pairs in two yeast species descended from a whole-genome duplication. Proc Natl Acad Sci U S A 104, 8397-8402.	6	8	0	1	2	0	10	0	3	1	0	4	10	0	4	0
Wickerhamomyces anomalus	Wa	JH119215-JH119570	Schneider J, Rupp O, Trost E et. al. (2012) Genome sequence of Wickerhamomyces anomalus DSM 6766 reveals genetic basis of biotechnologically important antimicrobial activities. FEMS Yeast Res 12, 382-386.	12	6	2	0	2	2	3	0	7	1	0	4	10	0	8	0
Wickerhamomyces ciferrii	Wic	CAIF01000001-CAIF01000364	Schneider J, Andrea H, Blom J et. al. (2012) Draft genome sequence of Wickerhamomyces ciferrii NRRL Y-1031 F-60-10.	7	3	1	0	1	1	4	0	5	1	0	2	6	0	4	0

Species	**Abbreviation**	**GenBank accession numbers**	**Reference**	**TTA**	**TTG**	**CTT**	**CTC**	**CTA**	**CTG**	**TCT**	**TCC**	**TCA**	**TCG**	**AGT**	**AGC**	**GCT**	**GCC**	**GCA**	**GCG**
			Eukaryot Cell 11, 1582-1583.																
Yarrowia lipolytica	Yl	CR382127-CR382132	Dujon B, Sherman D, Fischer G et. al. (2004) Genome evolution in yeasts. Nature 430, 35-44.	1	3	21	0	2	13	21	0	2	4	0	6	30	0	4	2
Zygosaccharomyces bailii	Zb	CBTC010000001-CBTC010000154	Galeote V, Bigey F, Devillers H et. al. (2013) Genome Sequence of the Food Spoilage Yeast Zygosaccharomyces bailii CLIB 213T. Genome Announc 1.	5	19	0	3	12	0	21	0	4	5	0	9	28	0	9	2
Zygosaccharomyces rouxii	Zr	CU928173-CU928176, CU928178-CU928179, CU928181	Souciet JL, Dujon B, Gaillardin C et. al. (2009) Comparative genomics of protoploid Saccharomycetaceae. Genome Res 19, 1696-1709.	4	10	0	1	5	0	12	0	3	2	0	5	12	0	9	0
Schizosaccharomyces cryophilus	Shc	KE546988-KE547017	Rhind N, Chen Z, Yassour M et. al. (2011) Comparative functional genomics of the fission yeasts. Science 332, 930-936.	3	4	6	0	1	1	11	0	2	2	0	5	11	0	3	1
Schizosaccharomyces japonicus	Sj	KE651166-KE651197	Rhind N, Chen Z, Yassour M et. al. (2011) Comparative functional genomics of the fission yeasts. Science 332, 930-936.	26	3	0	0	2	8	1	0	32	8	0	8	0	0	12	3
Schizosaccharomyces octosporus	Sho	KE503206-KE503210	Rhind N, Chen Z, Yassour M et. al. (2011) Comparative functional genomics of the fission yeasts. Science 332, 930-936.	3	5	8	0	1	1	11	0	2	2	0	5	15	0	3	1
Schizosaccharomyces pombe	Sp	NC_003421, NC_003423-NC_003424	Rhind N, Chen Z, Yassour M et. al. (2011) Comparative functional genomics of the fission yeasts. Science 332, 930-936.	2	4	5	0	1	1	7	0	2	1	0	3	9	0	2	1