Two of the methods I used for filtering genes were:

- 1. gene should be detected by both ara and tse and
- 2. both tse and ara should agree on functional class of the genes found by both of them

Here I will study all the cases in our filtered gene-set in which at least of one the functional classes (21 types excluding Z) is missing in a genome.

- 1) TtheileriEdinburgh missing D: both ara and tse were missing D
- 2) TTcruziTulacl2 missing Y: genes found by both ara and tse, however, they did not agree on the functional class. **Tse** reported Y, Ara reported N(as you see in the picture ARA is considering an intron in between the anticodon). In another case a gene was reported as Y but it was only found by ara.

```
> integrated tse aralintegrated tse araSsourceOrg=="TcruziTulacl2" & integrated tse araSarafunc=="Y
geneid sourceOrg tseidentity araidentity direction tsebegin tseend arabegin araend 1740 TcruziTulacl2_AQH001033301.1_66 TcruziTulacl2 notfound Tyr + -1 -1 477 602
                                                                                 602 notfound (gta) AQH001033301.1 notfound
                                                                                       arageneseg tsegeness
(((((tttttttt)))))))))))))
                                                            (((vvvv)))
                                                                                                   notfound
                                                                                                            notfound
araintronend tseacloc sourceSO tsenote aranote foundby arafunc tsefunc
1740 64 notfound contig notfound ara Y
integrated_tse_ara[integrated_tse_ara$sourceOrg=="TcruziTulacl2" & integrated_tse_ara$tsefunc==
genetd sourceOrg tseidentity araidentity direction tsebegin tseend arabegin araend tseac araac sourceseq tsescore arascore 10317 TcruziTulacl2_KB851309.1_82 TcruziTulacl2 Tyr Asn + 29977 30062 29977 30062 GTA (gtt) KB851309.1 72.3 108.4
                                                                             30062 GTA (gtt) KB851309.1
2232 GTA (gtt) KB851384.1
9623 TcruziTulacl2 KB851384.1 89 TcruziTulacl2
                                                                         2147
tsegeneseq
                                                                   (((((ttttttt)))))))))))))
                                                                                                               (((((ttttttt)))))))))))))
    tseintronbegin tseintronend araintronbegin araintronend tseacloc
                                                     sourceSO tsenote aranote foundby arafunc tsefunc
10317
                                 36
36
                                          49
           2184
                                              34-36 supercontig notfound
                                                                          both
```

- 3) TvivaxY486 missing Y: it was found only by tse.
- 4) PconfusumCUL13 missing D: it was only found by ara not tse
- 5) TcongolenseIL3000 missing S: both tse and ara found tRNA genes of type S. however they were exclusive to either ara or tse!

```
> integrated_tse_ara[integrated_tse_ara$sourceOrg=="TcongolenseIL3000"
                                                                          & integrated tse ara$arafunc==
     geneid sourceOrg
TcongolenseIL3000_T.congo_bin_contig_1162_4 TcongolenseIL3000
                                                             {\tt sourceOrg\ tseidentity\ araidentity\ direction\ tsebegin\ tseend\ arabegin\ araend}
                                                                                                                                                                           sourcesed tsescore
                                                                                                                                      19506 notfound (gct) T.congo_bin_contig_1162 notfound
1338 TcongolenseIL3000_T.congo_bin_contig_1205_9 TcongolenseIL3000
1240 TcongolenseIL3000_T.congo_bin_contig_1482_14 TcongolenseIL3000
1088 TcongolenseIL3000_T.congo_bin_contig_1482_16 TcongolenseIL3000
                                                                           notfound
                                                                                                                                1582
                                                                                                                                       1674 notfound (cga) T.congo_bin_contig_1205 notfound
                                                                                                                                       4003 notfound (cga) T.congo_bin_contig_1482 notfound 6007 notfound (cga) T.congo_bin_contig_1482 notfound
                                                                           notfound
                                                                                                                                3921
1729
              TcongolenseIL3000 T.congo.pschr.6 76 TcongolenseIL3000
                                                                           notfound
                                                                                                                             154282 154399 notfound (qct)
                                                                                                                                                                     T.congo.pschr.6 notfound
     arascore tsegenesed
                                                                                                                                                   arageneseg tsegeness
1432
        100.9
106.9
                notfound cytgtggcttgatgggaacttcgattgtgccatgcatgtgccttactttctctagtctggaatgtgtggcatgctgtggggggtaaattttatcagccgtctctgcaaattgagacccacgcgcc
                                                            \tt gccaagagaccatctggggttgcgagatcgtgggttcgagtcccacctcgtgccttgttgcgagatcgtgggttcgagtcccacctcgtgttt
1338
                notfound
                                                                                                                                                               notfound
                                                            tcgcgtctggtcacgggatcgtgggttcgagtccacctcgtgttttgttgcgagatcgtgggttcgagtcccacctcgcgtc
gccaagagaccatttggggttgcgagatcgtgggttcgagtccacctcgtgcccagttgcgagatcgtgggttcgagtcccacctcgtgttt
1240
        106.1
                 notfound
                                                                                                                                                               notfound
                                 gataaaactggagaaagattcctttttctgctgttgctgcgtctgttggcagcgtgataagccagtaaaagccgggtcaccaccttggctttttgcatgtgctaaccatgctttcatag
1729
        105.3
                                                                                                                             arageness tseintronbegin tseintronend araintronbegin
(((((tttttttt)))))))))))))
                                                                                                                                              notfound
                                                                                                                                                            notfound
                                                                                                                                                                           nointron
                                                                                                            notfound
                                                                                                                                                            notfound
                                                                                                             (((((ttttttt)))))..))).
                                                                                                                                              notfound
                                                                                                                                                            notfound
                                                                                                                                                                           nointron
                                                                                                            (((((ttttttt)))))))))))))
                                                                                                             (((((ttttttt))))))))))))))
                                                                                                                                              notfound
     araintronend tseacloc
                              sourceSO tsenote aranote foundby arafunc tsefunc
1432
               65 notfound
                                 contig notfound
         nointron notfound
1338
                                 contig notfound
                                                              ara
1240
         nointron notfound
                                contig notfound
                                                              ага
         nointron notfound
               ron notfound contig notfound
62 notfound chromosome notfound
1088
                                                              ara
                                                           ara s
ngolenseIL3000" & integrated_tse_ara$tsefunc=="S'
> integrated_tse_ara[integrated_tse_ara$source0;
geneid sourceOrg tseidentity araidentity direction tsebegin
5137 TcongolenseIL3000_T.congo_bin_contig_1482_13 TcongolenseIL3000 Ser notfound - 23214
                                                                                                              23214 23283
                                                                                                                                               GCT notfound T.congo_bin_contig_1482
                                                                                                                                                                                          22.7
4137 TcongolenseIL3000 T.congo bin contig 1637 23 TcongolenseIL3000
                                                                                       notfound
                                                                                                                125
                                                                                                                        191
                                                                                                                                          -1
                                                                                                                                               GCT notfound T.congo bin contig 1637
                                                                                                                                                                                          22.6
2136 TcongolenseIL3000_T.congo_bin_contig_2083_32 TcongolenseIL3000 arascore
                                                                                       notfound
                                                                                                                     5277
                                                                                                                                               GCT notfound T.congo_bin_contig_2083
                                                                                                                                                                                          25.4
                                                                                      neseq arageneseq
5137 notfound
                GaGAaGTGGCTCAGCGGTtaAGAGCACAGACTGCTCTTCCAaAGGtTCTGAGTTCAAATCCCAGCAACCA
                                                                                            notfound
                                                                                                         4137 notfound
                   GaTGGCTCAATGGGTaAGAGCACCCGACTGCTCTTCGGAAGGtTGGGAGTTCAAATCCCAGCAACCA
                                                                                             notfound
                                                                                                            >...>>>
2136 notfound GTGagaTGGCTCAGTGGTtaAGAGCGCCGACTGCTCTTCCAaAGGtCCTGAGTTCAAATCCCAGCAACCACA
                                                                                             notfound
                                                                                                                           ...<<<<....>>>......<
     tseintronbegin tseintronend araintronbegin araintronend tseacloc sourceSO tsenote
                                                                                            aranote foundby arafunc tsefunc
                                                                  33-35 notfound pseudo notfound
30-32 notfound pseudo notfound
5137
                                         notfound
                                                      notfound
2136
. I
                                         notfound
                                                      notfound
                                                                   33-35 notfound pseudo notfound
```

6) TcruziEsmeraldo missing Y: Ara and tse did not agree on the functional class. Tse reported Y, Ara reported N

7) TcruziSylvioX10-1-2012 missing Y: Ara and tse did not agree on the functional class. Tse reported Y, Ara

reported N

- 8) TcruzimarinkelleiB7 missing Y and W: Both Y and W were found only by ara and not tse
- 9) TcruziCLBrenerEsmeraldo-like missing Y: Ara and tse did not agree on the functional class. Tse reported Y, Ara reported N
- 9) TcruziCLBrenerNon-Esmeraldo-like missing Y: Ara and tse did not agree on the functional class. Tse reported Y,

 Ara reported N

Conclusions: Apparently the main reason for Y being missed is ARA and TSE not agreeing on the functional class of the tRNA genes found by both of them. This is the result of their intron displacement in anticodon loop which seems to be only one base.