A

```
mip10      TGCCGATTTGGGGAAGAATTTTAAAAATCAAGGCATAGATGTTAATCCGGAAGCAATGGC     60
mip18      TGCCGATTTGGGAAAGAATTTTAAAAATCAAGGCATTGATGTTAGTCCGGAAGCATTGGC     60
           ************ ********************** ******* ********** ****

mip10      TAAAGGCATGCAAGACGCTATGAGTGGCGCTCAATTGGCTTTAACCGAACAGCAAATGAA     120
mip18      TAAAGGCATGCAAGACGCTATGAGTGGCGCTCAATTGGCTTTAACCGAACAGCAAATGAA     120
           ************************************************************

mip10      AGACGTTCTTAACAAGTTTCAGAAAGATTTGATGGCAAAGCGTACTGCTGAATTCAATAA     180
mip18      AGACGTTCTTAACAAATTTCAAAAAGATTTGATGGCAAAACGTACAGCTGAATTCAATAA     180
           *************** ***** ***************** ***** **************

mip10      GAAAGCGGATGAAAATAAAGTAAAAGGGGAAGCCTTTTTAACTGAAAACAAAAACAAGCC     240
mip18      GAAAGCAGATGAAAATAAAGTAAAAGGGGAAGCCTTTTTAACTGAAAACAAAAACAAGCC     240
           ****** *****************************************************

mip10      AGGCGTTGTTGTATTGCCAAGTGGTTTGCAATACAAAGTAATCAATTCTGGAAATGGTGT     300
mip18      AGGCGTTGTTGTATTACCAAGTGGTTTGCAATATAAAGTAATCAATGCTGGAAATGGTGT     300
           *************** ***************** ************ *************

mip10      TAAACCCGGAAAATCGGATACAGTCACTGTCGAATATACTGGTCGTCTGATTGATGGTAC     360
mip18      TAAACCTGGTAAATCGGATACAGTGACTGTCGAATACACTGGTCGTCTGATTGATGGTAC     360
           ****** ** ************** *********** ***********************

mip10      CGTTTTTGACAGTACCGAAAAAACTGGTAAGCCAGCAACGTT       402
mip18      CGTTTTTGACAGTACCGAAAAAACTGGTAAGCCAGCTACTTT       402
           ********************************** ** **
```

B

```
neuA2      GATACTTCGGATGTGATTCATACTGTTATTGACCTCCTGTTTAAGTTTCAGCAAATGGAG   60
neuA201    GATTCATCAGATGTAGTATACGCCGTTATTGACTTGTTGAATAATTATAAGAAAATCAAC   60
           *** * ** *****  *  *  * ********* *  **  *** * * ** ****  *

neuA2      GTTGTTTTTGACAGTGTATTGCTGTTACAACCAACTTCTCCATTTAGGAAGCCAGAAACC   120
neuA201    GTGTTTTTTGACAGCGTTTTACTTCTGCAACCTACCTCCCCTTTTAGAAAACCTGAGACA   120
           **  ********** ** ** **  * ***** ** ** ** ***** ** ** ** **

neuA2      ATAAGACATGCTGTTGAAATACATCAAGCAACGGGAAAAAGTGTCGTTTCAGTTAGTCCC   180
neuA201    ATTAGAAAGGCCGTTCTAATGCATAGAGACATTGGTAATAGTGTTGTATCTATCAATAAA   180
           ** *** * ** ***  *** ***  **  *  ** ** ***** ** **  * * *

neuA2      ATCTCTTTAAAGCCTTCTTGGTGCAGAAGCATTGATAGCCAAGGCAATTTGGTTAAGCCA   240
neuA201    GTTAGCTTTAAACCTTCTTGGTATAGGACAGTAGATGATCTAGGTAACTTATGTTCCCCT   240
            *    ** ** **********  ** *   * ***   * *** ** **   *   **

neuA2      GAACTTTTTCAGGATCTTGAAATTTATTGCAACGAGAATCCAATTTATAAATTAAATGGG   300
neuA201    AATATTTTTAGAAATTCTGATGATTCTGAAGAAGGTGAACCCATTTATAAACTAAATGGT   300
            *  *****    **  ***   ** *    * *   * ** ********* *******

neuA2      TCAATTTATATTGCAACAGCTAAACAAATTATTGAAAATAAGAGTTTTTATAGC         354
neuA201    GCAATTTATATAGCTACTACCGAACAGATCATATCTAATAAGTCATTTTATAGT         354
            ********** ** **  *  **** ** **    ******   ********
```

Figure S1

Comparison of the mismatched alleles for the ST1335 and ST1395 strains. Sequences of *mip* (A) and *neuA* (B) alleles were aligned using Clustal Omega version 1.2.4. Identical nucleotides are marked below the alignment by asterisks.

A

```
mompS_15      ACGACAGTGATCACTGGTTTGATTTTGCCAACTGGCATAACTACAACAACAAGTGGGATG      60
mompS_7       ACGACAGTGATCACTGGTTTGATTTTGCCAACTGGCATAACTACAACAACAAGTGGGATG      60
              ************************************************************

mompS_15      CTGTTAATGCTGAACTGGGTCAATTCGTAGATTTCAGTGCTAACAAGAAAATGCGTTTCC     120
mompS_7       CTATTAATGCTGAACTGGGTCAATTCGTAGATTTCAGTGCTAACAAGAAAATGCGTTTCC     120
              * ***********************************************************

mompS_15      ATGGCGGTGTTCAATATGCTCGCATTGAAGCGGATGTGAACCGTTATTTCAATAACTTTG     180
mompS_7       ATGGCGGTGTTCAATATGCTCGCATTGAAGCGGATGTGAACCGTTATTTCAATAACTTTG     180
              ************************************************************

mompS_15      TCTTTAACGGATTCAACTCCAAGTTCAATGGCTTTGGTCCTCGCACTGGTTTAGACATGA     240
mompS_7       TCTTTAACGGATTCAACTCCAAGTTCAATGGCTTTGGTCCTCGCACTGGTTTAGACATGA     240
              ************************************************************

mompS_15      ACTATGTATTTGGCAATGGCTTTGGTATTTATGCTAAAGGAGCTGCCGCTATTCTGGTTG     300
mompS_7       ACTATGTATTTGGCAATGGCTTTGGTATTTATGCTAAAGGAGCTGCCGCTATTCTGGTTG     300
              ************************************************************

mompS_15      GTACCAGTGATTTCTACGATGGAGTTGGTTTCATCAGCGGTTCTAAAAATGC  352
mompS_7       GTACCAGTGATTTCTACGATGGAGTTGGTTTCATCAGCGGTTCTAAAAATGC  352
              ************************************************
```

B

```
pilE_14       ATGCTCAATAATCGACATTTTATGAAGAATAAGCGCATGAGACAATCTGCGTTTACCCTG      60
pilE_4        ATGCTCAGTCATGTACATTTTATGAAGAATAGCCGTATGAAACAATCCGCGTTTACCCTG      60
pilE_6        ATGCTCGGTCATGTACATTTTATGAAGAATAGCCGTATGAAACAATCCGCGTTTACCCTG      60
              ****** * **  *****************  ** **** ****** ************

pilE_14       GTTGAAGTTCTGATCAGCATGGTCATTATGGGAATTCTGGTTTCAATTGCCTATCCATCC     120
pilE_4        GTTGAAGTTCTGATCAGCATGGTCATTATGGGCATTCTGGTTTCAATTGCCTATCCATCC     120
pilE_6        GTTGAAGTTCTGATCAGCATGGTCATTATGGGAATTCTGGTTTCAATTGCCTATCCATCC     120
              ********************************************************************

pilE_14       TATTTACAATATATCCAAAAGTCCCGTCGTGCCGATGCCCATGCTACATTGACACAAGAT     180
pilE_4        TATTTACAATATATCCAAAAATCCCGTCGTGCCGATGCTCACGCCACATTGACACAAGAT     180
pilE_6        TATTTACAATATATCCAAAAATCCCGTCGTGCCGATGCTCACGCCACATTGACACAAGAT     180
              ******************** **************** ** ** ***************

pilE_14       CAAATCATTTTAGAACGCTGTTATTCACAGAATTTTTCCTATGCTGCGGCGTGTGGCGCC     240
pilE_4        CAAATTATTTTAGAACGCTGTTATTCACAGAATTTTTCTTATGCTGCGGCGTGTGGCGCC     240
pilE_6        CAAATTATTTTAGAACGCTGTTATTCACAGAATTTTTCTTATGCTGCGGCGTGTGGCGCC     240
              ***** ******************************** *********************

pilE_14       TTACCGGCATTTCCGCAAACAACACCAAATGGGTACTATACGATCAATATTTCAAACCTT     300
pilE_4        TTACCAGCATTTCCTCAAACAAGCCCGAACGGGTACTATACTATCAATATTTCAAACCTG     300
pilE_6        TTACCAGCATTTCCTCAAACAAGTCCGAACGGGTACTATACTATCAATATTTCAAACCTG     300
              ***** ******** ********** ** *********** *****************

pilE_14       ACAGCTACAACGTATACCTTAACGGCAACCCCT      333
pilE_4        ACAGCCACAACGTATAGCTTAACTGCAACCCCT      333
pilE_6        ACAGCCACAACGTATAGTTTAACTGCAACCCCT      333
              ***** ***********  ***** *********
```

Figure S2

Comparison of the predominant *mompS* 15 (A) and *pilE* 14 (B) alleles of subspecies *fraseri* with the corresponding alleles that occurred less frequently. The alignment was performed by Clustal Omega version 1.2.4. Identical nucleotides are marked below the alignment by asterisks. The mismatched nucleotides are marked by black rectangles.

A

```
flaA21      TCAAACCAAACAATCAACTTCAGTATCGGTAGCATAAAAGCTTCTTCTATCGGAGGTATT    60
flaA34      TCAAATCAAACAATCAACTTCAGTATCGGTAGCATAAAAGCTTCTTCTATCGGAGGTATT    60
            ****** ************************************************************

flaA21      GCCACAGCTACAGGAACAGAAGTTGCTGGTGCTGCAGCAACAGATATTACTATTGCAATT    120
flaA34      GCCACAGCTACAGGAACAGAAGTTGCTGGTGCTGCAGCAACAGATATTACTATTGCAATT    120
            ************************************************************

flaA21      GGAGGAGGTGCAGCCACCAGTATTAACTCTTCTGCCAATTTTACCGGGGCGCTAAACGGA    180
flaA34      GGAGGAGGTGCAGCCACCAGTATTAACTCTTCTGCCAATTTTACCGGGGCGCTAAACGGA    180
            ************************************************************

flaA21      CA 182
flaA34      CA 182
            **
```

B

```
asd28       CGCTGTTGGTATTAGTCGTATTAATGTGGCGACTTACCAGTCTGTTTCCGGGACTGGCAA     60
asd56       CGCTGTTGGTATTAGTCGTATCAATGTAGCGACTTATCAGTCTGTTTCCGGGACTGGCAA     60
            ********************* ***** ******** ************************

asd28       AAAAGCCATCAGCGAGCTGGTTGCTCAGGTAGGCGATCTTTTAAATGGGAGGCCAGCTAA    120
asd56       AAAAGCCATCAGCGAGCTGGTTGCTCAGGTAGGCGATCTTTTAAATGGGAGGCCAGCTAA    120
            ************************************************************

asd28       TGTTCAAGTTTATCCTCAGCAAATTGCTTTTAATGCGCTTCCTCATATTGATCAGTTCGA    180
asd56       TGTTCAAGTTTATCCTCAGCAAATTGCTTTTAATGCGCTTCCTCATATTGATCAGTTCGA    180
            ************************************************************

asd28       AGACAATGGTTATACCCGAGAAGAGATGAAGATGGTCTGGGAAACTCGCAAGATTATGGA    240
asd56       AGACAATGGTTATACCCGAGAAGAGATGAAGATGGTCTGGGAAACTCGCAAGATTATGGA    240
            ************************************************************

asd28       AGATGACAGCATTATGGTTAACCCTACAGCCGTCAGAGTTCCTGTTATTTATGGGCATTC    300
asd56       AGATGACAGCATTATGGTTAACCCTACAGCCGTCAGAGTTCCTGTTATTTATGGGCATTC    300
            ************************************************************

asd28       TGAAGCGGTTCATCTGGAGTTAAAAAAGCCTTTGACGGCTGACGATGCTCGCGCGCTTTT    360
asd56       TGAAGCGGTTCATCTGGAGTTAAAAAAGCCTTTGACGGCTGACGATGCTCGCGCGCTTTT    360
            ************************************************************

asd28       GGCAAAGGCACCCGGCGTTACTGTAGTGGATAATCTTTCTAAAGCAAGTTATCCCACTGC    420
asd56       GGCAAAGGCACCCGGCGTTACTGTAGTGGATAATCTTTCTAAAGCAAGTTATCCCACTGC    420
            ************************************************************

asd28       AATTAAGGATGCAGTCGGACATGATGATGTTTTTGTAGGGCGTATAAGACAGG            473
asd56       AATTAAGGATGCAGTCGGACATGATGATGTTTTTGTAGGGCGTATAAGACAGG            473
            ***************************************************
```

Figure S3

Comparison of the *flaA* (A) and *asd* (B) alleles found in Lp strains predicted to belong to subspecies *raphaeli*. The alignment was performed by Clustal Omega version 1.2.4. The mismatched nucleotides are marked by black rectangles. C. Clonal complex D formed by STs that had either *pilE* 27 or *proA* 29 and may, potentially, belong to subspecies *raphaelli*. ST819 is the predicted founder. ST799, ST701 and ST2116 are subgroup founders.

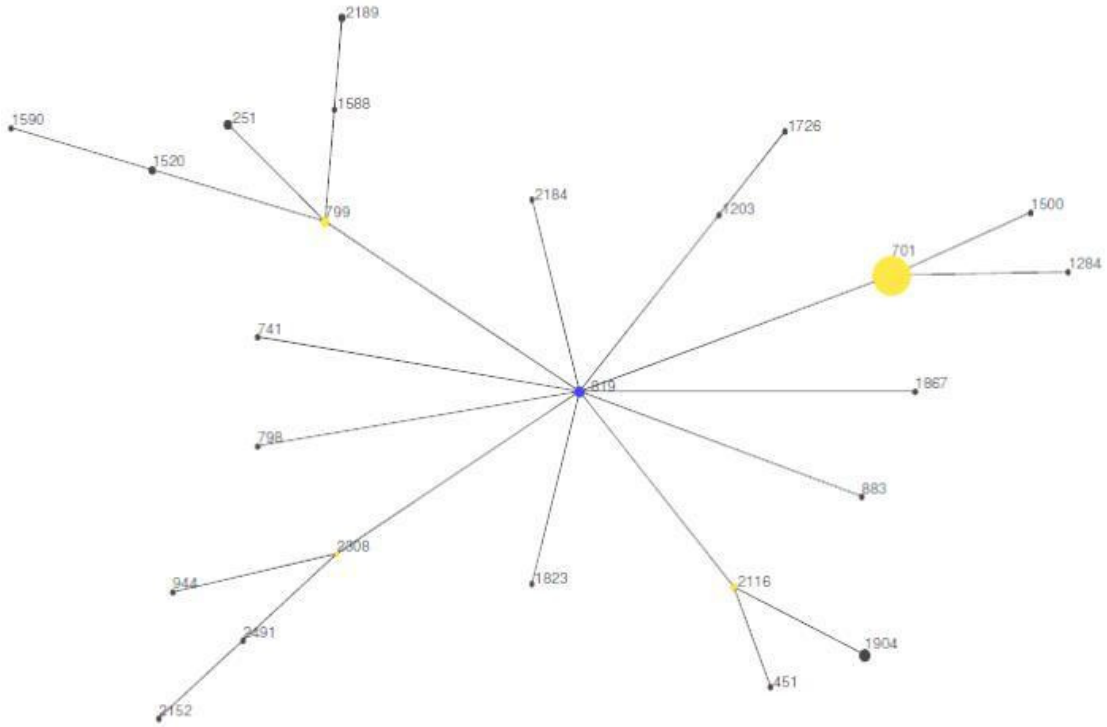

Figure S4

Clonal complex D formed by STs that may potentially belong to subspecies *raphaeli*. The input data on STs was downloaded from the ESGLI database on December 14, 2017. ST819 is the predicted primary founder and is identified as blue circle. ST701, ST799 and ST2116 are subgroup founders (STs that have at least two descendant single locus variants) and are denoted by yellow circles. The area of each circle represents the prevalence of the ST in the input data.

A

```
flaA_11     TCAAACCAAACAATCAACTTCAGTATCGGCAGCATAAAAGCTTCTTCTATTGGAGGTATT   60
flaA_30     TCAAACCAAACAATCAACTTCAGTATCGGCAGCATAAAAGCTTCTTCTATTGGAGGTATT   60
            ************************************************************

flaA_11     GCCACAGCTACAGGAACAGAAGTTGCTGGTGCAGCAGCAACAGATATTACTATTGCAATT   120
flaA_30     GCCACAGCTACAGGAACAGAAGTTGCTGGTGCTGCAGCAACAGATATTACTATTGCAATT   120
            ******************************** **************************

flaA_11     GGAGGAGGTGCAGCCACCAGTATTAACTCTTCTGCCAATTTTACCGGGGCGCTAAACGGG   180
flaA_30     GGAGGAGGTGCAGCCACCAGTATTAACTCTTCTGCCAATTTTACCGGGGCGCTAAACGGG   180
            ************************************************************

flaA_11     CA182
flaA_30     CA182
            **
```

B

```
asd_16      CGCTGTTGGTATTAGTCGTATCAATGTAGCGACTTATCAGTCTGTTTCCGGGACTGGCAA    60
asd_44      CGCTGTTGGTATTAGTCGTATCAATGTAGCGACTTATCAGTCTGTTTCCGGGACTGGCAA    60
            ************************************************************

asd_16      AAAAGCCATCAGCGAGCTGGTTGCTCAGGTAGGTGATCTTTTAAATGGAAGACCAGCTAA    120
asd_44      AAAAGCCATCAGCGAGCTGGTTGCTCAGGTAGGTGATCTTTTAAATGGAAGACCAGCTAA    120
            ************************************************************

asd_16      TGTTCAAGTTTATCCTCAGCAAATTGCTTTTAATGCACTTCCGCATATTGATCAATTTGA    180
asd_44      TGTTCAAGTTTATCCTCAGCAAATTGCTTTTAATGCGCTTCCTCATATTGATCAGTTCGA    180
            ********************************** ***** *********** ** **

asd_16      AGACAATGGCTATACCCGAGAAGAGATGAAGATGGTCTGGGAAACTCGCAAGATTATGGA    240
asd_44      AGACAATGGTTATACCCGAGAAGAGATGAAGATGGTCTGGGAAACCCGCAAAATTATGGA    240
            ********* ********************************** ***** ********

asd_16      AGATGACAGCATTATGGTTAACCCTACAGCCGTCAGGGTTCCTGTTATTTATGGGCATTC    300
asd_44      AGATGACAGCATTATGGTTAACCCTACAGCCGTCAGGGTTCCTGTTATTTATGGGCATTC    300
            ************************************************************

asd_16      TGAAGCGGTTCATCTGGAATTAAAAAAACCTTTGACGGCTGACGATGCTCGTGCGCTTTT    360
asd_44      TGAAGCGGTTCATCTGGAATTAAAAAAACCTTTGACGGCTGACGATGCTCGTGCGCTTTT    360
            ************************************************************

asd_16      GGCAAAAGCACCTGGAGTTACTGTAGTGGATAATACTGCTAAAGCAAGTTATCCCACTGC    420
asd_44      GGCAAAAGCACCTGGAGTTACTGTAGTGGATAATACTGCTAAAGCAAGTTATCCCACTGC    420
            ************************************************************

asd_16      AATTAAGGATGCAGTCGGACATGATGATGTTTTTGTAGGGCGTATAAGACAGG           473
asd_44      AATTAAGGATGCAGTCGGACATGATGATGTTTTTGTAGGGCGTATAAGACAGG           473
            **************************************************
```

Figure S5

Alignment of the *flaA* (A) and *asd* (B) alleles that are predominant in subspecies *fraseri* with those from D-7708. The alignment was performed using Clustal Omega version 1.2.4. Identical nucleotides are marked below the alignment by asterisks. The mismatched nucleotides are marked by black rectangles**.**

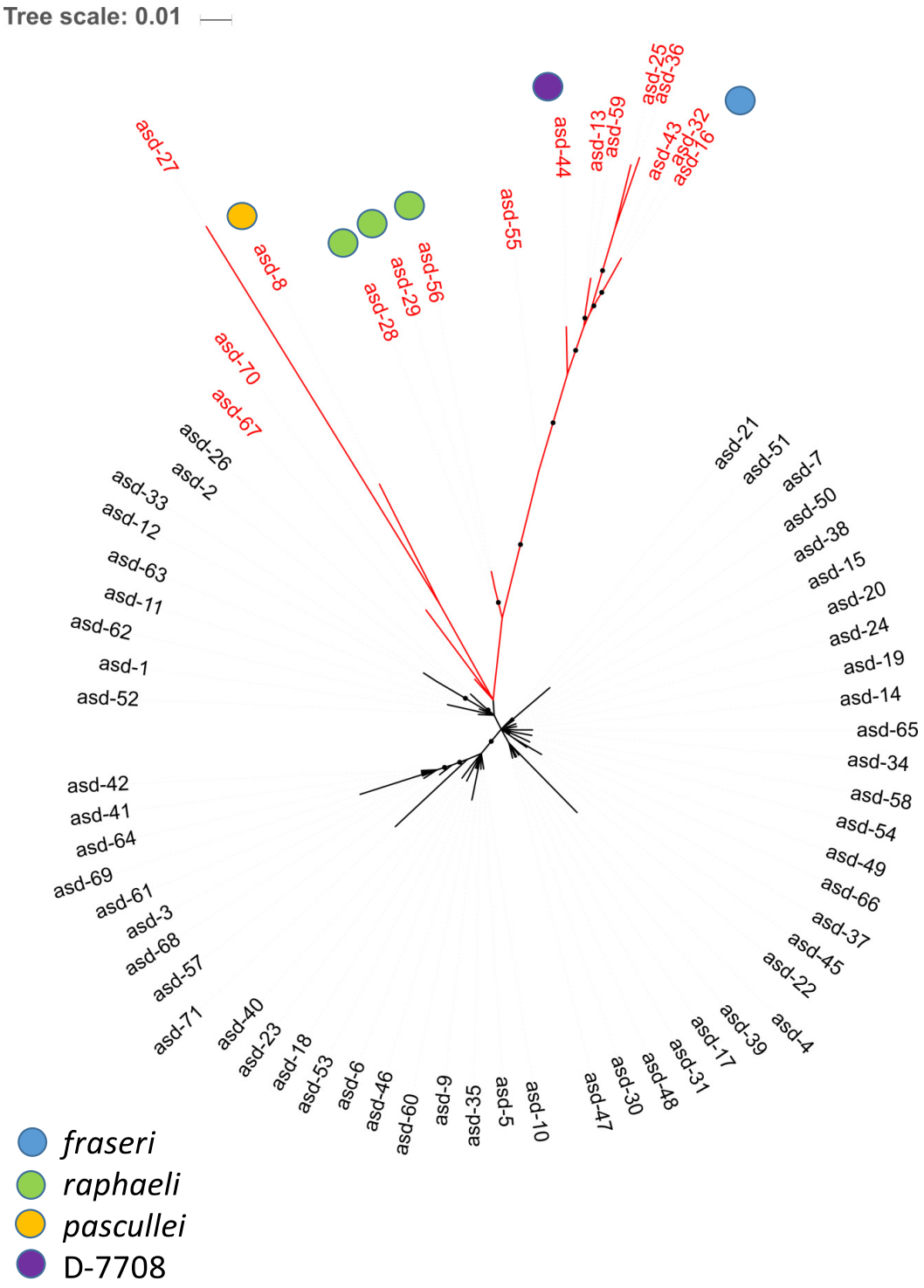

Figure S6

A: *asd*

Phylogenetic trees of seven gene fragment sequences used in the SBT scheme. The sequences for each SBT locus were extracted from the ESGLI database on October 13, 2016. The alleles were aligned using Clustal Omega version 1.2.4. Individual locus alignments were analyzed in RAxML version 8.9.2 (1) using a general time reversible (GTR) model with a Γ distributed rate heterogeneity parameter. 100 bootstraps were conducted using the rapid bootstrapping option and the bipartition recovery data was assigned to the best tree to provide branch support values.

1. **Stamatakis A.** 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics **30:**1312-1313.

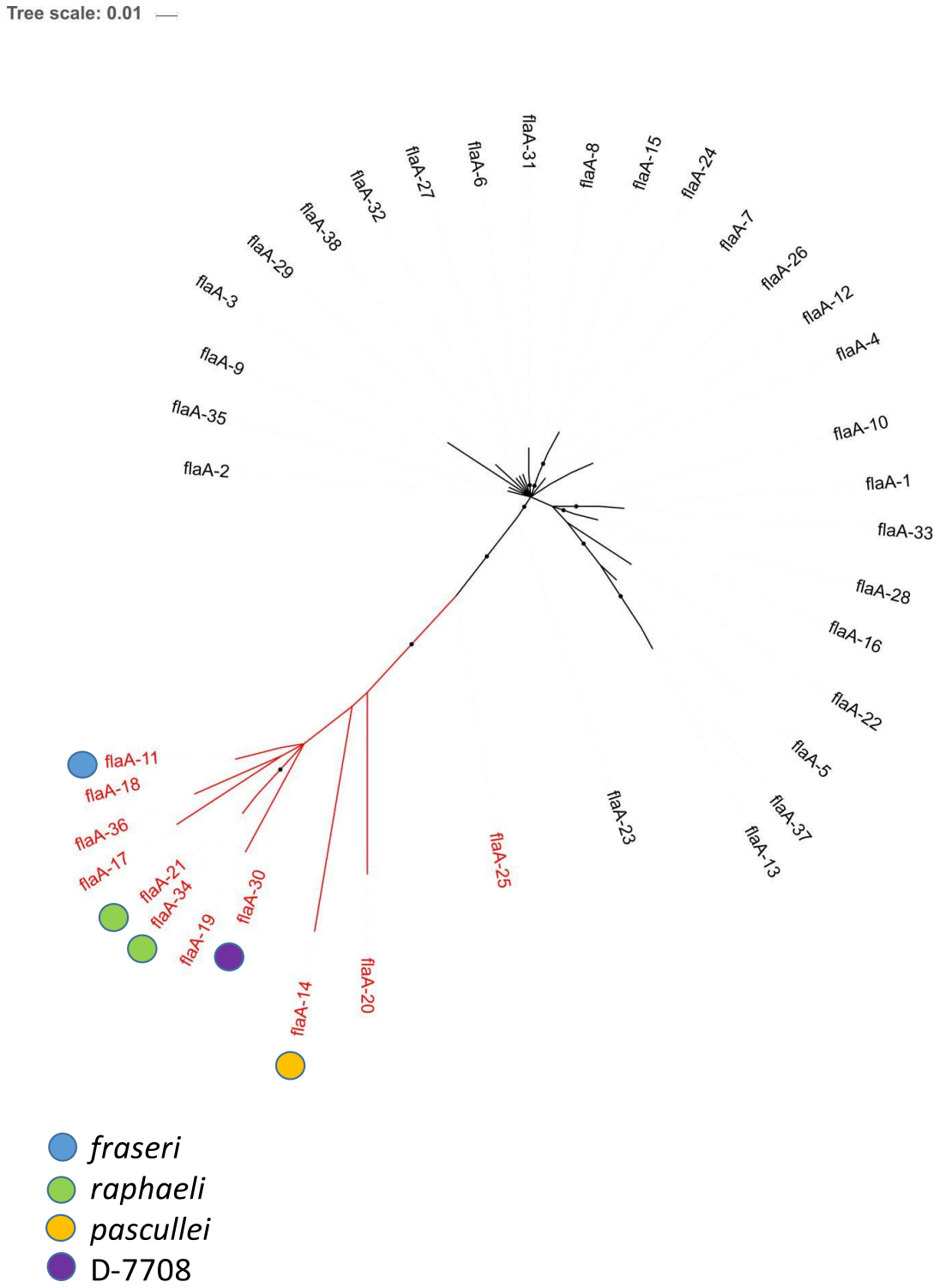

Figure S6
B: *flaA*

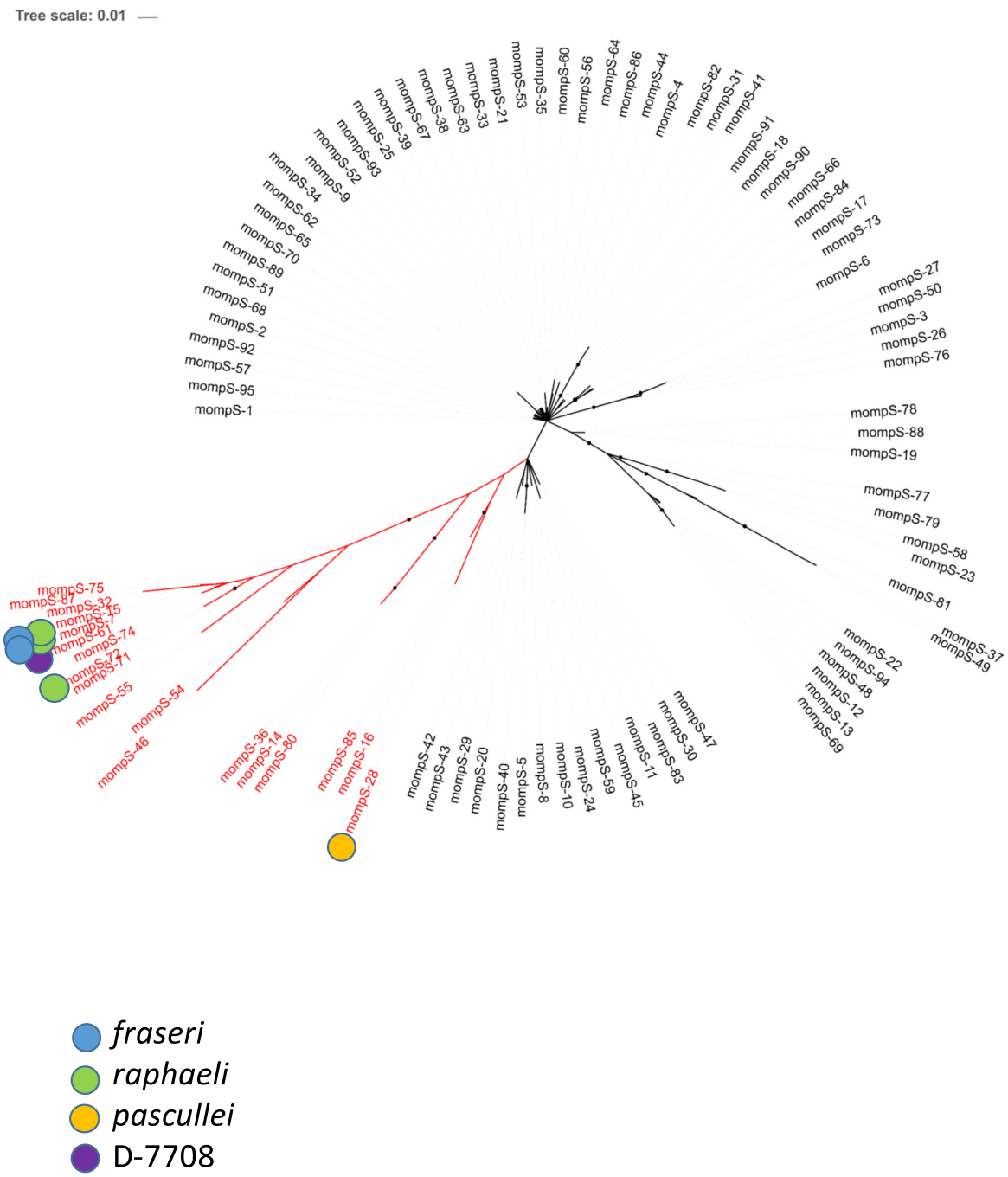

Figure S6

C: *mompS*

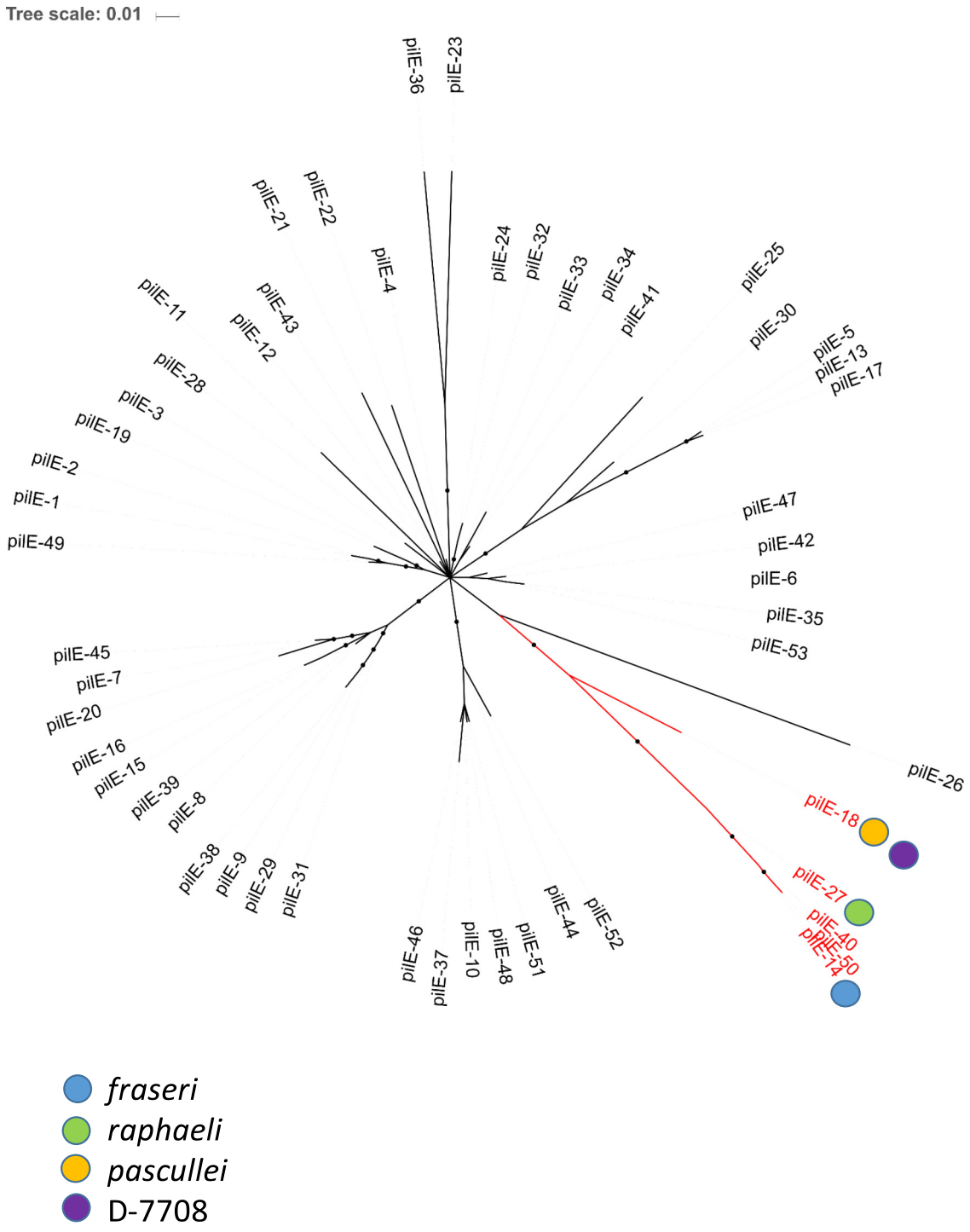

Figure S6
D: *pilE*

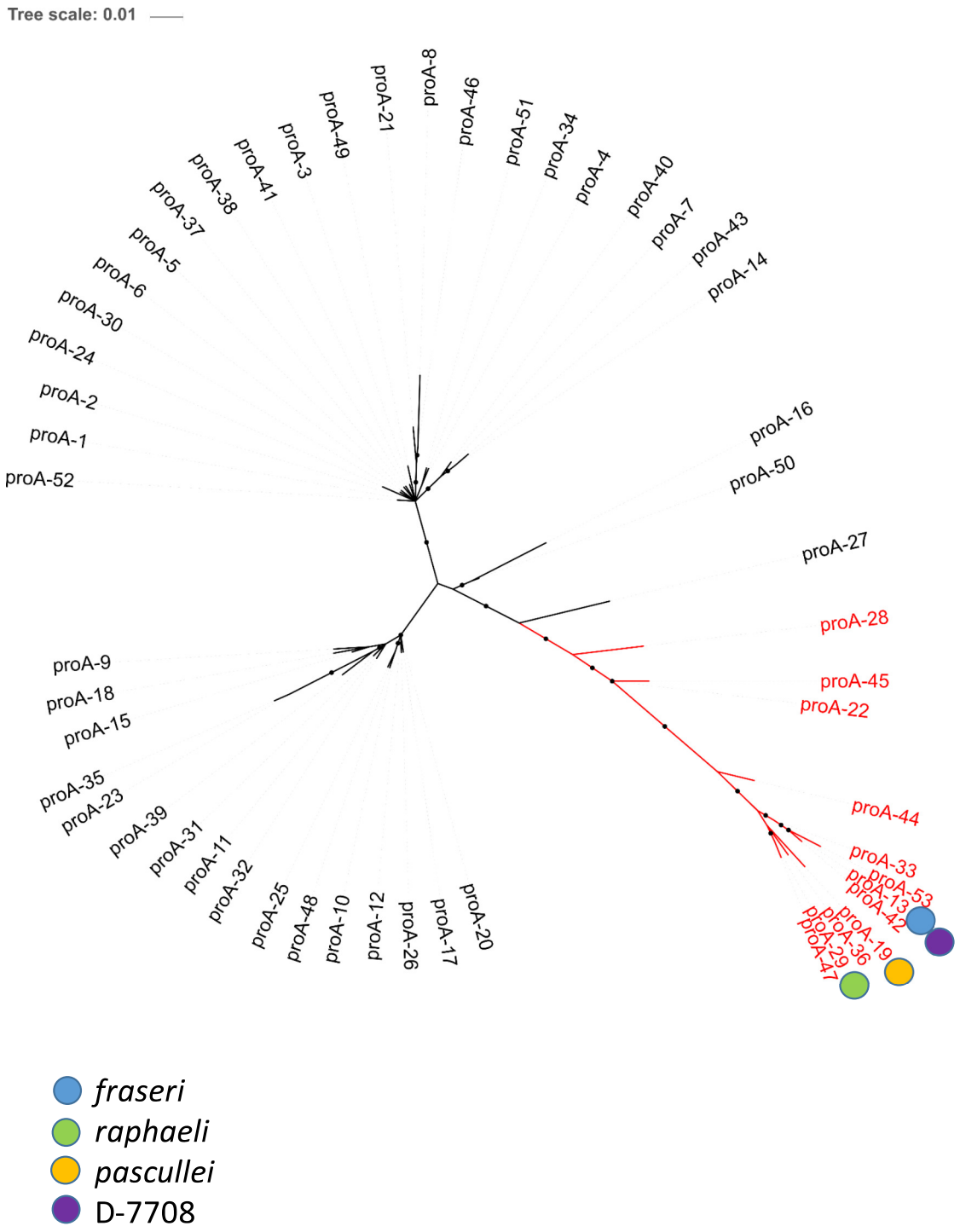

Figure S6

E: *proA*

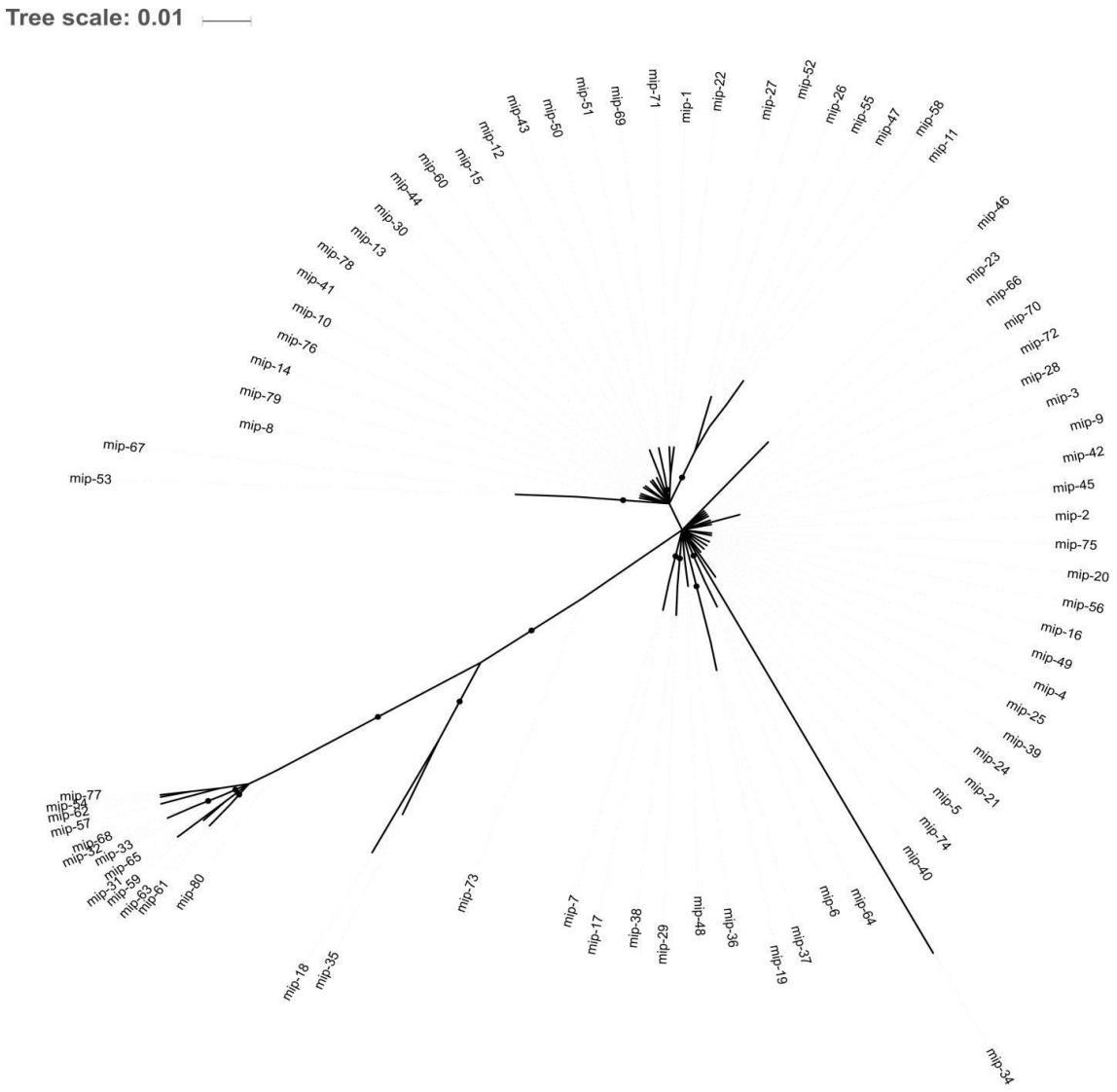

Figure S6

F: *mip*

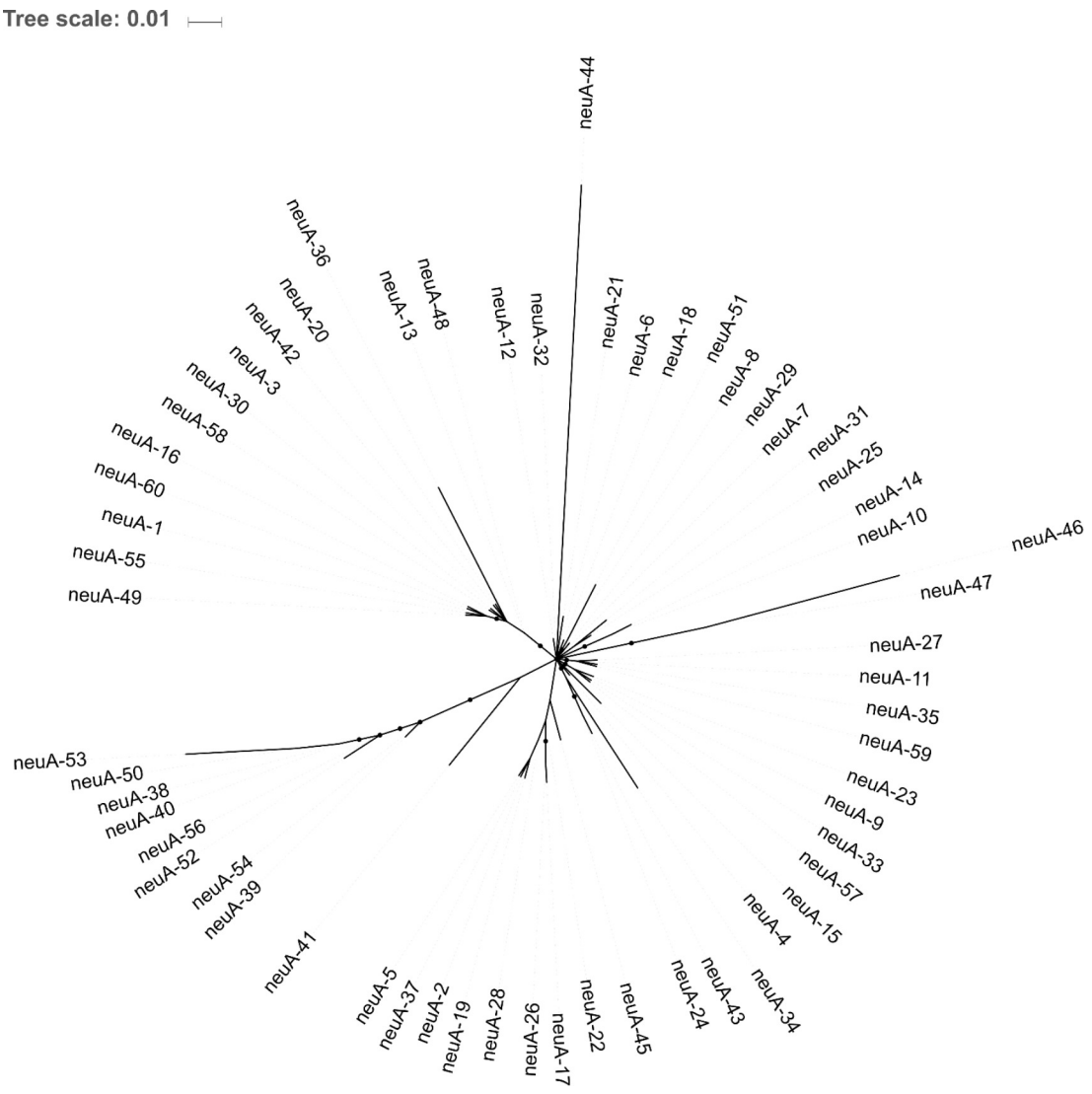

Figure S6

G: *neuA*

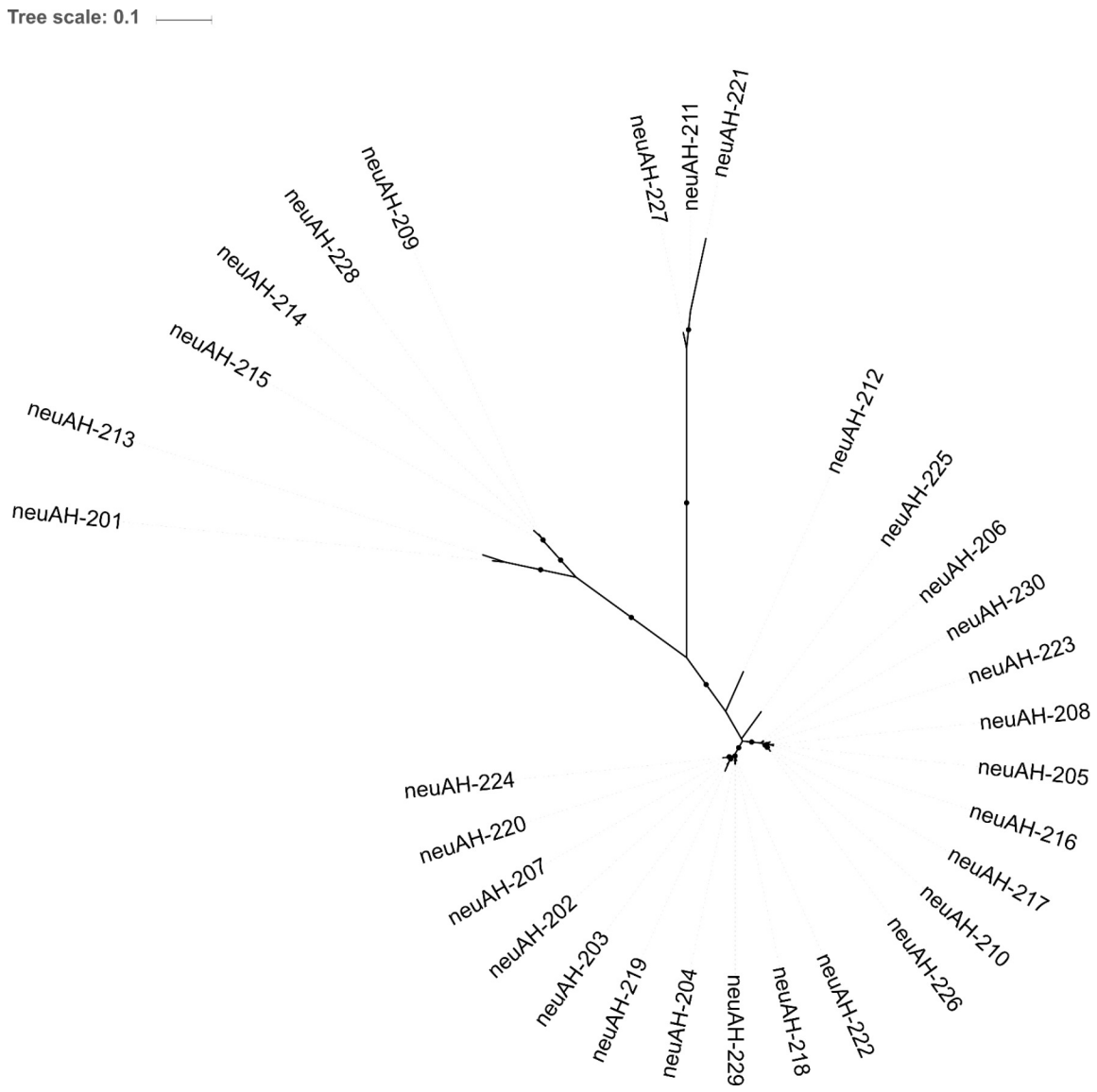

Figure S6

H: *neuAH*