Characterization of *Serratia proteamaculans* strain AGR96X encoding an anti-feeding prophage (tailocin) with activity against grass grub (*Costelytra giveni*) and manuka beetle (*Pyronota* spp.) larvae

Mark RH Hurst[1,2]; Amy Beattie[1]; Sandra A Jones[1], Aurelie Laugraud [3], Chikako van Koten [3] and Lincoln Harper[4]

[1]Forage Science, AgResearch, Lincoln Research Centre, Private Bag 4749, Christchurch 8140, New Zealand

[2]Bio-Protection Research Centre, Lincoln, Christchurch, New Zealand

[3]Knowledge & Analytics, AgResearch, Lincoln Research Centre, Private Bag 4749, Christchurch 8140, New Zealand

[4]Curtin University; Department of Environment and Agriculture; Centre for Crop and Disease Management; Kent St, Bentley, WA 6102 Australia

Table S1 AGR96X genome assembly data

Assembly	A5/AGR96X_final.scaffolds.csv
Number of scaffolds	19
Total size of scaffolds	5471622
Total scaffold length as percentage of known genome size	4.9
Longest scaffold	2436133
Shortest scaffold	366
Number of scaffolds > 500 nt	17
Percentage of scaffolds > 500 nt	89.5
N50 scaffold length	861236
L50 scaffold count	2
N50 scaffold - NG50 scaffold length difference	861236
Percentage of assembly in scaffolded contigs	0
Percentage of assembly in unscaffolded contigs	100
Average number of contigs per scaffold	1
Average length of break (>25 Ns) between contigs in scaffold	0.789474
Number of contigs	19
Number of contigs in scaffolds	0
Number of contigs not in scaffolds	19
Total size of contigs	5471622
Longest contig	2436133
Shortest contig	366
Number of contigs > 500 nt	17
N50 contig length	861236
L50 contig count	2
N50 contig - NG50 contig length difference	861236
Scaffold number | size	**X coverage**
scaffold1|size 2436133	676.0289
scaffold2|size 861236	693.3952
scaffold3|size 452547	680.6453
scaffold4|size 309239	711.3371
scaffold5|size 287972	693.3683
scaffold6|size 264846	697.1971
scaffold7|size 214164	687.5617
scaffold8|size 208714	689.7185
scaffold9|size 122773	677.5747
scaffold10|size 114992	382.5988
scaffold11|size 954207	707.2675
scaffold12|size 64368	723.2397
scaffold13|size 36883	680.0028
scaffold14|size 1620	479.8568
scaffold15|size 1562	311.815
scaffold19|size 366	32.92896
Genome size 5471622	677.9828

```
Se : AAAGGCTCGGGTATGGGAAATAGCAAAATTTACAACAAACGTGAGTAGATGATTTGCTCCTTCGGCACCGCAAATTCCTCCAGGCTGAAGTCTGGTACCGCAGGAAAGTATGTGTTCGCAAGCG :  124
Sp : AAAGGCTCGGGTATGGGAAATAGCAAAAGTTACAACAAACGTGAGTAGATGATTTGCTCCTTCGGCACCGCAAATTCCTCCAGGCTGAAGTCTGGTACCGCAGGAAAGTATGTGTTCGCAAGCG :  124

Se : AGTTTGGTCCAGTTCCAGGTGATATCGCCATGTTGTTGAGAAACCGGCACTAGCCATTGTCA--TGCTGGTGCCGCAACTTTGCATGTAGTG-AATATATATGATGCAGACGGCATTGATGTGA :  245
Sp : AGTTTGGTCCAGTTCCAGGTGATATCGCCATGTTGTTGAGAAACCGGCACTAGCCATTGTCATATGTTGGTGCCGCAACTTTGCATGTAGTGGAATATATATGATGCCGACGGCATTGATGTAA :  248

Se : TGCGGGGTCTTAAAGACGGTTGAGGCAGCAATACTGCCGTGCGCGCTA---------------CCTCGACTGCATTACAGGTGAATTATATAATTATTATTAATTGATTTCTATTGATATTTGT :  354
Sp : TGCGGGGTCTTAAAGACGGTTGAGGCAGCAATACTGCCGTGCGCGCTATGTAGGGGTCGCTACCTTCGACTGCATTACAGGTGAATTATATAATTTTTATTTATTGATTTCTATTGATATTTGT :  372

Se : TTTTGTTTTGGTGTTTAAATCTTACTTTTTAAGATTTATTGGTTTTAAATTTTAATTTATTTTGTTTTGTTGCTTTTTTACTATTATTTGGTCTGTTGATGCTGCTGAATTAAAGGGTGGTGAT :  478
Sp : TTTTGTTTTGGTGTTTAAATCTTACTTTTTAAGATTTATTGGTTTTAAATTTTAATTTATTTTGTTTTGTTGTTTTTTTACTCTTATTTACTCTGTTGATGCTGCTGAATTAAAGGATGGTGAT :  496

Se : GATCTATTAAACTTACATTGTAAACTTTAATGATATAATATTCTAGTTTTTATTTATTTTCTGGTTTATTGTTTTGTTATTTTATTTTTTTATCGATTGTTGAGTGGTGTTGATTGTTGATGTT :  602
Sp : GATCTATTAAACTGACATTTTAAGCTTTAATGATATAACATTCTATTTTTTATTTATTTTTTGGTTTATTGTTTTGTTATTTTATTTCTTTATCGATTGTTTAGTTGTGTTGATTGTTGATGTT :  620

                                                                                                                    anfA1
Se : CATGAATTATCCTGATCATGTTGCTTAAGAATATACATTCTCTTTGTTGAGTTTTTGGCGGAGCTTCGCCTTACACCAGTTTCTTATAGTTTAATTATTGGTGAATCGTAATGAAAAAAAAATG :  726
Sp : AATGAATTATCCTGATCATGTTGCTTAAGAATATACATTCTCTTTGTTGAGTTTTTGGCGGAGCTTCGCCTTACACCAGTTTCTTATAATTTAATTATTGGTGAATCGTAATGAAAAAAAAATG :  744

Se : GTATTTGATTAGATGTAATTACGGGAAAATTGACTTAGGGCAAAAAAGCTTGGAGGCTATGGGTGTTGAGTGTTTCTGCCCAAGATTAAGTCCGGCTGAGTTGAAGGTTTTATTTGACAATAAA :  850
Sp : GTATTTGATTAGATGTAATTACGGGAAAATTGACTTAGGGCAAAAAAGCTTGGAGGCTATGGGTGTTGAGTGTTTCTGCCCAAGATTAAGCCCGGCTGAGTTGAAGATTTTATTTGACAATAAA :  868

Se : AAAAATTTTTCAAAAGGATGGTGTATTTTCCCTCCTTACCTTTTTGTGAGGTTTGACATTGATCACATTCCAATATCAAAAATTCAACATGCCATCGGTGTAAAAGGATTTGTCAGGTTTGGTG :  974
Sp : AAAAATTTTTCAAAAGGATGGTGTATTTTCCCTCCTTACCTTTTTGTAAGGTTTGACATTGATCACATTCCAATATCAAAAATTCAACATGCCATCGGTGTAAAAGGATTTGTCAGGTTTGGTG :  992

Se : GGGGGATAAATTCAATACCCGATATTGTTATCGGAAAAATAATGAAAAATGATTACTTGTCAGGGATTGATAAAGATGTAATTAAATTAATAAGATGTGAAAATAAAGTCTTGCGGGTTGATAC : 1098
Sp : GGGGGATAAATTCAATACCCGATATTGTTATCGGAAAAATAATGAAAAATGATTACTTGTCAGGGATTGATAAAGATGTAATTAAATTAATAAGATGTGAAAATAAAGTTTTGCGGGTTGATAC : 1116

Se : TTTATTGTCAATGATAGAGAGCGAGCTAGGGAAAACAATCATTAGGAATAAGCTTGTATAAAACGTTAATAATCCTTGAGTTAATGTGATTTAACAGATTGTAAAAAACACATTGCCATCCATT : 1222
Sp : TTTATTTTCAATGATAGAGAGCGAGCTAGGGAAAAAAATCATTGGGAATAATCTTATATAAAATGTTAATAATCCTTGAGTTAATGTGATTTAACATATTGTAAAAAACACATTGCCATCCATT : 1240

                                                                                        anfA2
Se : TTTTGTTAATTCAGGGGTTTTAATTGCGTCAATTTTAATTGGCTATGGGGACATTGTGAACCATTATTCTGGCTTTCTTGATATGACTATAGACTGGTTAAAATTATATATATCTTCGGTCTAT : 1346
Sp : TTTT---AATGCAGGGGTTTTAATTGCGTCAATTTTAATTGGCTATGGGGGCATTGTGAACCATTATTCTGGCTTTCTTGATATGACTATAGACTGGTTAAAATTATATATATCTTCGGTATAT : 1361

Se : GCTGCTTTAAGTGCACTGAGTGTTTCTGCATTGACGAGTTTGAAAGATAATAAAGAAATTAAAGAGGTTGTTACTTCAATGCTCGCGTGTGGTATTCTTGCCCTGTCTTTTTCTGGCATGTTAG : 1470
Sp : GCTGCTTTAAGTGCACTGAGTGTTTCTGCATTGACGAGTTTGAAAGATAATAAAGAAATTAAAGAGGTTGTTACTTCAATGCTCGCGTGTGGTATTCTTGCCCTGTCTTTTTCTGGCATGTTAG : 1485

Se : AGCATTTTGGACTACCTCCGAACAGTGCTGGCATCGTCGGTGGTGTTATTGGGCTACTTGGCGTGGATAGAATCCGATATATGGTTTCTTCTAAATTCAACTCCAAGACAAAGAAAGAATGAAA : 1594
Sp : AGCATTTTGGACTACCTCCGAACAGTGCTGGCATCGTCGGTGGTGTTATTGGGCTACTTGGCGTGGATAGAATACGATATATGGTTTCTTCTAAATTCAACTCCAAGACAAAGAA----TGAAA : 1605

Se : AAAGTTGATGTTATCATCTAATGAATAGGATTAAAATATTCATTTATTTAAATTGTTTTGTTCCGGATGTGTTTGATCTTAAAGCTTTACCTTCCGTACTGCTTAATCGTTCGATAGTTTGGAA : 1718
Sp : AAAATTAATGTTATTATCTAATGAATAGGATTAAAATATTCATTTATTTAAATTGTTTTGTTCCGGATGTGTTTGATCTTAAAGCTTTACCTTCCGTGCTGCTTAATCGTTCGATAGTTTGGAA : 1729
```

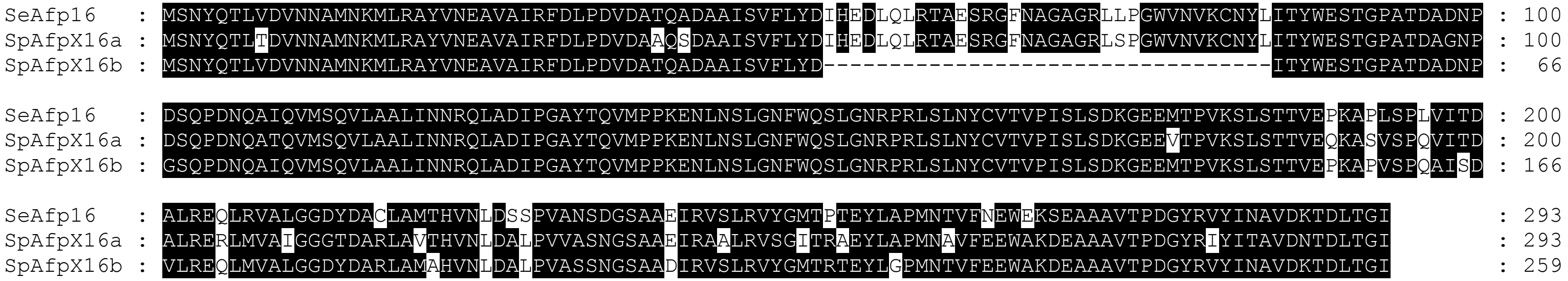
```
Se : ACTCAATTTCCCTTTTTTTAGCTTTCTAACCTTGTACAGTAGCGGATTACAATCCCGGATTGCGGGGTATAGCGATAGGTTGTGTGTAAAAATACTGACTTTGACTTGTTTTTAATCTAAAACA : 1842
Sp : ACTCAATTTCCCTTTTTTGAACTTTCTAACCTTGTACAGTAGCGGATTACAATCCCGGATTGCGGGGTATAGCGAGAGGTTGTGTGTAAAAATACTGACTTTGACTTGTTTTTAATCTAAAACA : 1853

Se : CCTTTTTCAGCTACATTGAGCCAATGCTGGCATCTGGAGCCATCGTTTTTT-GCCAGTTAGTGAAGTCTGTCAGACAACCGTTTTATGGTGTGCTCGCCCGCCTGAGCGCTGGCAAACGCAACC : 1965
Sp : CCTTTTTCAGCTACATTTAGCCAATGCTGGCATCTGGAGCCATCGTTTTTTTGCCAGTTAGTGAAGTCTGTCAGACAACCGTTTTATGGTGTGCTCGCCCGCCTGAGCGCTGGCAAACGCAACC : 1977

Se : GATTGCCACACAAGATACATTGGTACGGATCTGTGCCCAGGAAGCCTTTCATCAGCACCGCAAACCCTGCCCATTTGGGGTTCTCTCGAACCGTCATCTACCGCAGTTAGTAGACCTTCGGTAG : 2089
Sp : GATTGCCACACAAGATACATTGGTACGGATCTGTGCCCAGGAAGCCTTTTATCAGCACCGCAAACCCTGCCCATTTGGGTTTCTCTCGAACCGTCATCTACCGCAGCTAGTAGACCTTCGGTAG : 2101
                                  sea23
Se : TAGCGTGCCGCAGTAACGCACCATCTTAAAATGTCAGGCTGGGATATGGCTGACATAGCGCTGCAGTATGTCCTCTTGATGGATTTTTTGACGTTTGTGCTGTTGTGCCCTGTGATCGTAATAC : 2213
Sp : TAGCGTGCCGCAGTAACGCACCATCTTAAAATGTCAGGCTGGGATATGGCTGAAATAGCGCTGCAGCATTTCCTCTTGATGGATTTTTTGACGTTTGTGCTGTTGTGCCCTGTGATCGTAATAC : 2225

Se : GGATGGACCATGGCGCTGCCGCGATAATGTCGTAGCTGCAAGACCGCGACAGACGGGCTCGGATAAAACAGGGAAACAGGTCGTGGAGTAGCGGCCTGTTGTTGTTGACGAACGGCCAGAGGTA : 2337
Sp : GGATGGACCATGGCGCCGCCGCGATAATGTCGTAGATGCGAGACCGCGACAGACGGGCTCGGATAAAACAGGGAAACAGGTCGTGGAGTAGCGGCCTGTTGTTGTTGACGAACGGCCAGAGGTA : 2349

Se : GCGGTCATGGGTACGCAGTCCTATCAGGCGTAGCCGTAGCGGCGTGCGCCTATGAAGCAGGTGCCGCAGGTAATAATGCGCTCGTTGCGAGTTTGGTCCAGTCGTTGAGGGTATCGCCATGTTT : 2461
Sp : GCGGTCATGGGTACGCAGTCCTATCAGGCGTAGCCGTAGCGGCGTGCACCTATGAAGCAGGTGCCGCAGGTAATAATGCGCTCGTTGCGAGTTTGGTCCAGTCGTTGAGGGTATCGCCATGTTT : 2473
                                                                                                                           Product5
Se : ATTGTCATGCTGGAACAGCAGCTTTGCTGGGGCGGGTAATATCCATGAGGCAAAGTCTAATGACAGGAGACAGGAAAAGGGCGTGGCGCGAGCCGGCTCACGTCATAGAGGGGT-TAATTTATG : 2584
Sp : ATTGTCATGCTGGAACAGCTGCATTGCTGGGTCGGGTAATATCCATGAGGCAAAGTCTAATGACAGGAGACAGGAAAAGGGCGTGGCGCGACCCGGCTCACGTCATAGAGGGGCATAATTTATG : 2597

Se : CGCGGGGGTGTGGATGGCGCCACAGCAAGTGCGGTAGGATCATCAACTGGCGCTGGCGTACTTGCACTCCATCTGGTGTAGGTTCCGCTTGCTGACCAAACGTACGGCTAACGATGTCTGCTGC : 2708
Sp : CGCGGGGGTGTGGATGGCGCCACAGCAAGTGCGGTAGGATCATCAACTGGCGCTGGCGTACTTGCACTCCATCTGGTGCAGGTTCCGCT-GCTGACCAAACGTATGGCTAAAGATGTCTGCTGC : 2720

Se : TTTACCCCCAGCGCATCAATCGAGATACATGGCCCAACGTTAGTGCGATTTAGGCCGCAGAACAACCGTGCTTGTATCGATGTATATAATATGCATATTTTAATAACCATGATCACTACCATTT : 2832
Sp : TTTACCGCCAGCGCATCAATCGAGATACATGGCCCAACGTTAGTGCGATTTAGGCCGCAGAACAACCGTGCTTGTATCGATGTATATAATATGCATATTTTAATAACCATGATCACTATCATGC : 2844

Se : TAATTGAAAGCTTTTCCCGATTGTTATTGTTTTTTTGATAGCCCATAACTATGAATTGGATTGGTAACGATCGTTACTAACGATTTCTATTTGTTCTGAATTGATCCTATAGTTAATTCAGACA : 2956
Sp : TCATTGAAAGTTTTTTCCGATCGTTATTGTTTTTTTGATAGATTTAACTATGCATTAGACTGATAACGATCGTTACTAACGATTTCTATTTGTTCTGAATTGATTCTATAGTTAATTCAGACA : 2968

Se : AACAAGAGAGCGAACCTGAAAGCAGTCGTTTAAGATTTTTGATGAAAAATTGATTAATAGTCGCTAGGCGTGTTTTTAATACTTCCTATTATGTCAAT-------------------------- : 3054
Sp : AACAAGAGAGCGAACCTGAAAGCAGTCGTTTAAGATTTTTGATGAAAAATTGATTAATAGTCGCCAGGCGTATTTTTAATACGTCCTATTATGTCAATCATATCTATAATTAGATGAGTCTGTT : 3092
                                                                                                                          enp1
Se : -----------TATCTCAATATTTCTTACTTTCTATTTTATTAATCGTGATGTTATATTTGGTTTTTTATTATGTTATCACTTCAATAGAGAAGAATGTTCGAAGGGGTTATTTATAAATGCGA : 3167
Sp : TTTATATCAATTATCTCAAGGTTTCTTACTTTTTATTTTCTTAATCGTGATGTTATATTTGGTTTTTTATTATGTTATCACTTCAATAGAGAAAAATGATCGAAGGGGTTATTTATAAATGCGA : 3216

Se : ATAGCTATGTTGTGCGGAATCATCGTCGTAATTGCAGCTTTTTCTTGGTCAACCCGATATTATTATCAACAGTCAATCATCTGGAAGGAAAAATCGAAGGAAGCAAACGCATCAGTCAGGAAAT : 3291
Sp : ATAGCTATGTTGGGCGGGATTATCGTCGTGATTGCAGCTTTTTCTTGGTCAACCCGATATTACTATAAACAGTCAATCATCTGGCAGGAAAAAGCGAAGGAAGCAAACGCATCAGTCAGGAAAT : 3340

Se : ATGAAAATCAAATTGAATTGATACAAAGTCAGCATCGCAAACTAGCTGAATTAGATGAACACCACACGGAGAAACTGAATGAGGCTGAACAAGAGAATACGGTCCTGCGCACTCAGCTTGCCCG : 3415
Sp : ATGAAAATCAAATTGAATTGATACAAAGTCAGCATCGCAAACTAGCTGAATTAGATGAGCACCACACGGAGAAACTGAATGAGGCTGAACAAGAGAATACGGTCCTGCGCACTCAGCTTGCCCG : 3464

Se : TGGCCATCGCAGGATGCTCGTCTCTGGACAAATATCCTGTCCAGACAGAACATCCCCCTCAACCCGCAGCTTGGGCTATGACGGCCCCTTCGAACTCTCTGCAGATACTGGACAGCGTATTCTT : 3539
Sp : TGGCCATCGCAGGATGCTCGTCTCTGGACAAATATCCTGTCCAGACAGAACATCCCCCTCAACCCGCAGCTTGGGCTATGACGGCCCCTTCGAACTCTCTGCAGATACTGGACAGCGTATTCTT : 3588

Se : TCTATCCGGGAAGGCATCATCAGAGACCAACAAAAACTGATGTATCTTCAGAGTTATATTCGGCAGTTCTGTCTTAAAGATTAAAAATATATATATTTTTAATATATGGTACGTGCTAATTATT : 3663
Sp : TCTATCCGGGAAGGCATCATCAGAGACCAACAAAAACTGATGTATCTTCAGAGTTATATTCGGCAGTTCTGTCTTAAAGATTAAAAATATATATATTTTAAATATATGGTACGTGCTCATTATT : 3712

Se : ACGAGCCGATTAACGTGAGTTTATAATATTAACTAAAATTGGAGTTCTAATGATTATTTTTTCGTTACCAAAACCGCTTCCTGATAGTGACATGGTATTCTCTTGGTTGATTGTAGGAGGATTC : 3787
```

```
Sp : ACGAACCGATTAACGTGAGTTTATAATGTTAACTAAAATTGGAGTTCTAATGATGATTTTTCCGTTACCAAAACCGCTTCCTGATAGTGACATCGTATTCTCTTGGTTGATTGTAGGAGGATTC : 3836
                                   hol1
Se : TCCGCTTGAGGCGGTATCGTTAGATATTTAATGGAAAACAAAGAGTCTGGGGAGAAAATTTCATGGCTTGAGGTTTTCAAGCAAGTCGTTACTTCCGGATTTTCGGGGCTTTTAGCGGGGATCT : 3911
Sp : TCCGCTTGGGGCGGTATCGTTCGATATTTAATGGAAAACAAAGAGTCTGGGGAGAAATTTTCATGGCTTGAGGTTTTCAAGCAAGTCGTTATTTCTGGATTTACGGGGTTTTTAGCGGGGATCT : 3960

Se : ATGGCTATGAACAAGGATACAGCGAGTTTATAACCATGGCCTTCTCCGGTTTGAGTTGGGCATTAGGTGGTCACTTATTGGACTTATTATGGAAGCGATTAACTAATTAATTAGAAAAAGAAAA : 4035
Sp : ATGGCTATGAACAAGGATACAGCGAGTTTATAACCATGGCCTTCTCCGGTTTGGGTGGGGCATTAGGTGGTCACTTATTGGACTTGTTATGGAAGCGATTAACTAATTCATTAGAAAAAGAAAA : 4084

Se : TTATTCTAAAAACTAGCTGTGGCGACAAGTTGGCTAGCACTGCTAGTGTTTCAGTGATATATATTTAAATTTAACCAGATAAATAAAAATGTAATTTATAATAAATTGGTTGATAACTACAAAA : 4159
Sp : TCATTCTAAACACTAGCTGTGGCGACAAGTTGCCTAGCCCTGCTAGTGTTTCGGTGATATATATTTAAATTTAACCAGACAAATAAAA-TGTAATTTATAATAAATTGGTTGATAACTACAAAA : 4207

Se : ACATCATTTTATGGCACCCTATATAGATGCGGATTTGTTTTTACAGGGTGTCAGTATGTCTTAGTATCTTTTGATGCGTAATACCTTACAGATGTGGGCGCCTTTTCCTTTTTGGATTTAAATT : 4283
Sp : GCATTATTTTATGGAACCCTATATAGATGCGAATTTGTTTTTACAGGGTGTCAGTAT---------TTTTTTGATGCGTAATACCTTACAGATGTGGGTGCCTTTTCCTTTTTGGATTTAAATC : 4322

Se : GATCCGCTCAATAGTGAGTTTTTAAAATATGGCGGGGATTAATTTCAAACGTCAAGATCCTTTTTATGAAGTGATCCGGTTTAGAGAGTAGATGTCGCCGTCGTTGTAATATTTTATATTCATG : 4407
Sp : GAGTCACTCAATAGTGAATTTTTAAAATAGGACGGGGATTAATTTAAAATATTAAGTTCCTTTTTATGAAGTGATCCGGTTTAGAGAGTGGCTGGCGCCGTCGTTGTAATGTTTTATATTAATG : 4446

Se : ATTTCAAATCACATATCCTGTAAATACTTTCAATGTCAATTTTGTCGATATTTTCATTTATCTTTCTGTATTGAATGCATACAT-TTTCGAATATGGATGGATGCCTGTGGCGATATTTCAGTG : 4530
Sp : ATTTCACATCACATATCCGGTAAATACTTGCAATGTCAATTTTGTCGATATTTTCATTTATCTTGCTGTATTGAATGCATACATGTTTAGAATATAGATGGATGCCTGTGGCGATATTTCAGTG : 4570
                                                          mur1
Se : ATGTAGTTTTATCTCTGCGATCAGCTGTCAGATAAACCTGTTGGTACGGAGGTTACAATGATAATCGATGTTAATGGACTTAAACTCATTAAACATTTTGAGGGGCTAAGGCTCCGTGCTTATC : 4654
Sp : ATGTAGTTTTATCTCTGCCATCAGCTGTCAGATAAACCTGTTGGTACGGAGGTTACAATGATAATAGATGTTAATGGACTTAAACTCATTAAACATTTTGAGGGGCTAAGGCTCCGTGCTTATC : 4694

Se : AATGCTCGGCCAACGTATGGTCCATTGGATATGGTCATACTGCCGGCGTCGGTCCGGATGATGTTATTACCGAAGGGCAGGCTATTTCCTTCTTGCGTCAGGATGTAGCGGAGAGTGAACGGGC : 4778
Sp : AATGCTCGGCCAACGTATGGACCATTGGATATGGTCATACCGCCGGCGTCGGTCCGGATGATGTTATCACCGAAGAGCAGGCTATTTCCTTCTTGCGTCAGGATGTAGCGGAGAGTGAACGGGC : 4818

Se : AGTAAATCAGTATGTGCATGTTCCGCTTACGCAAAATCAGTTTGATGCCCTGGTTTCATTTGTTTTCAACCTGGGCGTCGGGAATTTCCGAACCTCAACGCTGCTGAAAAAACTCAATGCCGGT : 4902
Sp : AGTAAATCAGTATGTGCATGTTCCGCTTACGCAAAATCAGTTTGATGCCCTGGTTTCCTTTGTTTTCAACCTGGGCGTCGGGAATTTCCGAACCTCAACGCTGCTGAAAAAACTCAATGCCGGT : 4942

Se : GATTATGACGGCGCGGCCCAGGAATTTGGGCGTTGGATCCATGCCGGCGGTAAAGCGCTGCCCGGCCTGGTACGCAGAAGAGAGGCCGAAAGCGCACTCTTTCTGAAATAAGTAACACAGTTAG : 5026
Sp : GATTATGACGGCGCGGCCCAGGAATTTGGGCGTTGGATCCATGCCGGCGGTAAAGCGCTGCCCGGCCTGGTACGCAGGAGAGAGGCCGAAAGCGCACTCTTTCTGAAATAAGTAACACAGTTAG : 5066
                                                     ops
Se : ACGTATTTTTTATTTATAAGCATGAAAAAACAACTGGTGTAAATCAGGGGCGGTAGCATGCTTTTTTTCAGGGGAAACGGAGTTAATTAAATTGTTTTTCATGCAGATAATTTAAAAGGATTAT : 5150
Sp : ACGTATTTTTTATTTATAAGCATGAAAAAACAACTGGTGTAAATCAGGGGCGGTAGCATGCTTTTTTTCAGGGGAAACGGAGTTAATTAAATTGTTTTTCATGCAGATAATTTAAAAGGATTAT : 5190
                                            afp1
Se : TTGCACCCACTATTGATTGTAAAAAGAAGGATTTAACTCATGGCTATTACCGCAGACGACATTGCAGTGCAATATCCAATCCCTACTTATCGCTTTATTGTGACCCTGGGCGATGAGCAGGTGC : 5274
Sp : TTGCACCCACTATTGATTGTAAAAAGAAGGATTTAACTCATGGCTATTACCGCAGACGACATTGCAGTGCAATATCCAATCCCTACTTATCGCTTTATTGTGACCCTGGGCGATGAGCAGGTGC : 5314
```

Fig. S1 Nucleotide alignment of the region upstream of *S. entomophila* (Se) *afp1* and *S. proteamaculans* (Sp) AGR96X *afpX1*. The regulatory loci *anf*A1 and *anf*A2 and the phage lysis cassette (*mur*1, *enp*1, and *hol*1) are labelled above their respective initiation codons. The *ops*-associated element (bold underscore), located between *enp*1 and *afpX*1, is indicated. Non-shaded regions denote nucleotide dissimilarity. Italics denote the ORFs.

```
                                                           afp17
SeAfp17 : AATATTAATCAGTGGTGAGTTTGCATTTACACCCCGTCATTTATTTAAAAGAGGTATATTTATGCCGACTAAAACACCACAGTTACAGCTGGCGATCGAAGAATTTAATAAAGCAATATTACGG :
SpAfp17 : AATATTAATCAGTGGTGAGTTTGCATTTACACCCCGTCATTTATTTAAAAGAGGTATATTTATGCCGACTAACACACCACAGTTACAGCAGGCGATTGAAGAATTTAATAAAGCAATATTACGG :

SeAfp17 : GCTGCAGAATATGAATGGAATATTGAAAATAATATTTACGATAAAGTTCACACACTTTTCCCAGCAGCAGATAAATCTCTAAAAGAGAGTGTGGTCGCGCAGAGCGTTGAAAAAATATCTGAGG :
SpAfp17 : GCTGCAGAATATGAAGGGAATATTGAAAATATTATTTACGATAAAGTTCACACACTTTTCCCAGCAGCAGATAAATCTCTAAAAGAGAGTGTGGTCGCGCAGAGCGTTGAAAAAATATCTGAGG :

SeAfp17 : AAAACCAGGCGTTTGAGTTTAACCACTTGGGTAACGATCTCCTTGAATTTCCGGCAGTAGAATATATTGCTCGTTCTTTTTTACGCAGAATACTGGAAAAACTGGCTATTTATCAAGCTAGAGA :
SpAfp17 : CAAACCAAGCGTTTGAGTTTAACCACTTGGGTAACGACCTCCTTGAATTTCCGGCAGTAGAATATATTGCTCGTTCTTTTTTACGCAGAATACTGGAAAAACTGGCTATTTATCAAGCTAGAGA :

SeAfp17 : TCCATCATTTACTATAACGAAAACCAATCAGGATGCCGGGATAGGTTTCTTGGAAGGGTGGGAAAGCGATAACGCGAAATGGGCAAAGCAAATGAACGCGCTGCTGAATGAAAGCAAGGTTAAC :
SpAfp17 : TCCATCATTTACTATAACGAAAACCAATCAGGATGCCGGGATAGGTTTCTTGGAAGGGTGGAGCAGCGATAACGCGGAATGGGCAAAACAAATGAACGCGCTGCTGAATGAAAGCAAGGTTAAC :

SeAfp17   GCCCGCGTTTTGACCAGCCTGGCATACAGCTATATCCGCTTTTTTCGCCCGCTGAAGACATTGCGCAAGCTGGGAGGGGAGGCTAATGAGTCTTTGCTGAACAGCAAAGAAGCCAACCCGCGTC :
SpAfp17 : GCCCGCGTTTTGACCAGCCTGGCATACAGCTATATCCGCTTTTTTCGCACGCTGAAGACATTGCGCAAGCTGGGAGGGGAGGCTAATGAGTCTTTGCTGAACAGCAAAGAAGCCAACCCGCGTC :

SeAfp17 : TGGATATTTTTGCCAGCTCAACGGCGGGAGCGCCGCCGGTCATGTCGAAGCCGATCACCACTCAGCCTCACGGCTGGGTGTGGCTAAAATACGCCGACTGGAACGCCGCGGAAGCGTTGGGTAT :
SpAfp17 : TGGATATTTTTGCCATCTCAACGGCGGGATCGCCGCCGGTCATGTCGAAGCCGATCACCACTCAGCCTCACGGCTTGCCGTGGCTAAAATACGCCGACTGGAACGCCGCGGGAGCGTTGGGTAA :

SeAfp17 : CAAAACGTCCACGGATTTAATTGGCAGCCGGCCAAGCAGCGACCAGCATAAAGGTACAGGCTTCAAGTTTAAGGGCATGCTGCAGGCGATAAAGCCGCCGCTGTATCGACGCCATTCATTTTCC :
SpAfp17 : CAAAACGTCCACGGATTTAATTGGCAGCCGGCCAAGCAGCGACCCGCATAAAGGTACAGGCTTCAAGTTTAAGGGCATGCTGCAGGCGATAAAGCCGCCGCTGTATCGACGCCATTCATTTTCC :

SeAfp17 : CTTAGGGAACGCGACGAACATGAATCGGGGTACGGCGGTTTCGTTATCAACGATAAGTCGATGACCTCAATAGGGGCAGAAAAACCGTTGGTGGTGAATTGGCGTCAGCAGAACATGGATAAAA :
SpAfp17 : CTTAGGGAACGCGACGAACATGAATCGGGGTACGGCGGTTTCGTTATCAACGATAAGTCGATGACCTCAATAGGGGCAGAAAAACCGTTGGTGGTGAATTGGCGTCAGCAGAACATGGATAAAA :

SeAfp17 : AGCTGCCGATGTTTTCCGGCCCGTCCAGCACCACTTCCTATATGTATGAAATTGCCCGTTTGCTGAATCTGCCTGCTGCTGAAACACAGGTATCTACGTTTTCGATAATTCGGCTCTGTCAAGT :
SpAfp17 : AGCTGCCGATGTTTTCCGGCCCGTCCAGCACCACTTCCTATATGTATGAAATTGCCCGTTTGCTGAATCTGCCTGCTGCTGAAGCACAGGCATTCCGGTTGCTTCTGTTGGGCTGGATGATCCA :

SeAfp17 : ATTTAGGTCAGAGTGTAACTAATAGTTTTGTCACATTAGCGTAGGTCAGATACCAGCGAATACACTGAATATGCAGCCACA :
SpAfp17 : GGCGCGCGATCACAGTTTTACCGAGATTATGGGGGCGCTGG-ATGCCTATGCCATGGAGTTTGACGAGCGGGGCCCCACGT :
```

Fig. S2 Nucleotide alignment of *S. entomophila* (Se) *afp17* and *S. proteamaculans* (Sp) AGR96X *afp*X17. Non-shaded regions denote nucleotide dissimilarity. Italics denote the ORF.



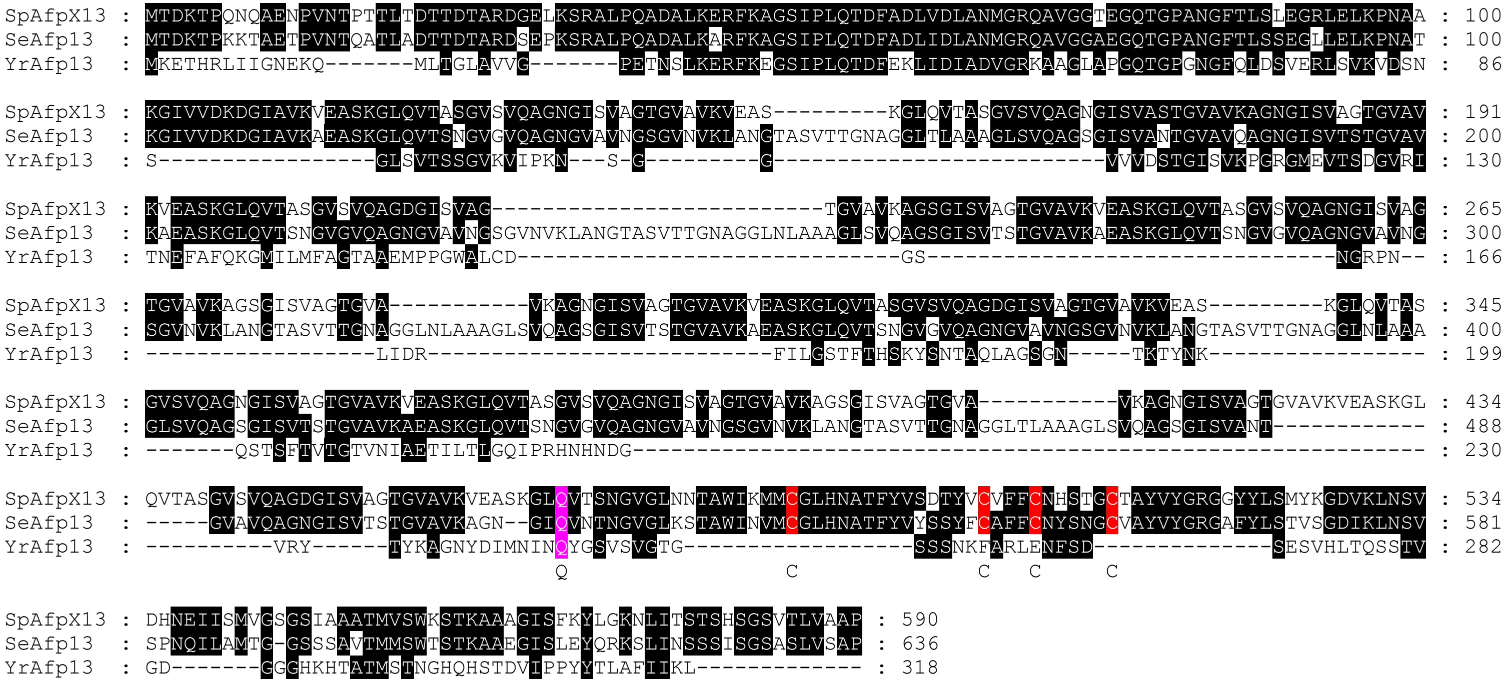
```
SeAfp16   : MSNYQTLVDVNNAMNKMLRAYVNEAVAIRFDLPDVDATQADAAISVFLYDIHEDLQLRTAESRGFNAGAGRLLPGWVNVKCNYLITYWESTGPATDADNP : 100
SpAfpX16a : MSNYQTLTDVNNAMNKMLRAYVNEAVAIRFDLPDVDAAQSDAAISVFLYDIHEDLQLRTAESRGFNAGAGRLSPGWVNVKCNYLITYWESTGPATDAGNP : 100
SpAfpX16b : MSNYQTLVDVNNAMNKMLRAYVNEAVAIRFDLPDVDATQADAAISVFLYD----------------------------------ITYWESTGPATDADNP :  66

SeAfp16   : DSQPDNQAIQVMSQVLAALINNRQLADIPGAYTQVMPPKENLNSLGNFWQSLGNRPRLSLNYCVTVPISLSDKGEEMTPVKSLSTTVEPKAPLSPLVITD : 200
SpAfpX16a : DSQPDNQATQVMSQVLAALINNRQLADIPGAYTQVMPPKENLNSLGNFWQSLGNRPRLSLNYCVTVPISLSDKGEEVTPVKSLSTTVEQKASVSPQVITD : 200
SpAfpX16b : GSQPDNQAIQVMSQVLAALINNRQLADIPGAYTQVMPPKENLNSLGNFWQSLGNRPRLSLNYCVTVPISLSDKGEEMTPVKSLSTTVEPKAPVSPQAISD : 166

SeAfp16   : ALREQLRVALGGDYDACLAMTHVNLDSSPVANSDGSAAEIRVSLRVYGMTPTEYLAPMNTVFNEWEKSEAAAVTPDGYRVYINAVDKTDLTGI         : 293
SpAfpX16a : ALRERLMVAIGGGTDARLAVTHVNLDALPVVASNGSAAEIRAALRVSGITRAEYLAPMNAVFEEWAKDEAAAVTPDGYRIYITAVDNTDLTGI         : 293
SpAfpX16b : VLREQLMVALGGDYDARLAMAHVNLDALPVASSNGSAADIRVSLRVYGMTRTEYLGPMNTVFEEWAKDEAAAVTPDGYRVYINAVDKTDLTGI         : 259
```

Fig. S3 Amino acid alignments of *S. proteamaculans* (Sp) AGR96X, AfpX16a, and AfpX16b against the *S. entomophila* (Se) Afp16 amino acid sequence.

```
YrAfp18   : MPYFNKSKK-NEIRPEKSKEEVGGVLFDDSAIHENIDHNME-P-QTGDSVATFPDNS-DEVVGGDLAALRARLQATIGD-YLPEFYEFQQTGRNILYPKPVDIKALQARLNSM------- :  108
SpAfpX18  : MPYASELKKLDK-KP--TENE--NFL-TNS---EN-E-T-ELPAEI-PS-SFFTE-SLDEEIEADLAELRAHLQAIISD-ALPGFYEFQQAGYNLLYPKKVDIEALQARVNTMQI----- :   99
SeAfp18   : MPYSSESKE---------K-E------THSKETERDN---ADPVFQRVSQGGVGVSPPDEGGDLSAGSNNDHLFAFIRETHLEEAQEFRLKGYNSQVPEVDEGLLTQLRDDLAAAKSRQY :  101

YrAfp18   : -------PITAWIQPLQQQ-LEQAGNAFLDRFRLFRKEKDWHNNALEFVKFRLAVKETGWDEQQTRDYNEALKHANRLSGYLSGTLDLIQHL-NDNPERQTSWEAVRKTSYDMLLNELLP :  219
SpAfpX18  : -------PVL--IQP-QHKLLEQAANAFLDRFRRFRKESDLHNNALEFIKFRLEVNKAGWDEPQNRDYHQALKHANRLSGYLSGTLDLMS-LINENVERQTPWETTRQANYDVLLNELLP :  208
SeAfp18   : VQASVGGPLGSRWLLEQIGKIERLTDKFSDEVIKRWVKHELP---------RKNVGEAAGTELGSR--RRAQSLARRLESTMQTLIRVVHTLHAMDDYSTP----E--KRHDMLVAELTG :  204

YrAfp18   : VDAE-NNGNQYTAKIRGNLT--------KNHPDNIKFNTIVNSVTRDERMASEALSAETKGLSARLLSG--------------------------------IALTRNSAEKIDKKSETSG :  298
SpAfp18X  : VNADAHN-NQYAAKTRGNLTG-------K-QPDNIKFNAIVYSVTLDERIAREALSAEIERLSARLLRG--------------------------------ITLTRNYAEKPDETGETSG :  287
SeAfp18   : AFSSSDKNTVIQALGRMTAETSGHAAWASLMQARTAAWEAGKSATVDAMIATSQHEAEKANKKARELQGDSQAFFSHVAAYLQSLSSDLAKVSINMGQSTSSSTIRNEAEKASELQ---- :  320

YrAfp18   : ITRDIVALCQVLQGVIKTIKRRGEYIQPLTDWQPSEHKLPGQKEELTKAETVKRELVKSR-KVMMQKVQGAKTILG-VLSDKTNKNRHRITQTLPYS-ANPDTNTVSRKTNEAVFRAGIM :  415
SpAfpX18  : IKRDIDFLCQVLQSVIHDVPYAGQQIQPET-YQP----L-GQKEAHTEIETVKHEL-KNIWKTARQKVQGAKAS-GQALSDKFNKNKHRIVHAFPGSTAS-DKNGDSHNINKAVFQVGIR :  398
SeAfp18   : --GDS-------Q---TFVSHVAAYLQSLPSDLAKVPINMWQNMDFVKTKLQAQTAVAGMKVHAHRLVRIAQHG----HSTKTPAK----DPEQ------VVTDSIIRSILWLWQQPAIK :  414

YrAfp18   : LLDKIQQTTSGIHKASL-ASRPLQHAVTQYSAL-EQMSSGMPSNRILDAKLRSESGRWQEKAEKSKKQ-------LQHILREITALAE--EHLKHKFMSALRDELKRA--PETLASNNII :  522
SpAfpX18  : LLDKIQQTTSDMHKA-IQTSRPLQQAVTHYSELGE-MLSGMPHNSILDAKLRAESDRWRQKAEESKEQ-------LQQLLGTITTLSD--ESMKQRYLSALREELNSVTNP--LASNNII :  505
SeAfp18   : IQCASTALLSKVGELKKIEGILASYAVTDELGN-EQERRNPPGEEDLDAQLR-EWVSASLKQENPENQRAVKVVTLERLLGGDIASARNLVERLGKAEEGIQNLLRRQQVAVLKMSVERL :  532

YrAfp18   : SNFDTRAKIVVEGLASIEIGMNQSVVRLAGHGEAGRK------DLDEQVVAWLQRLERIKGELKTDITQATGQSINNFSRQGMLARRMGEWNEAEKQRYLAALSTEDRAVAEMQYNTLFF :  636
SpAfpX18  : RDFDTQIKATVEGLSDIQKALCQALLRLSEHGQAGGK------ELDKHTVSWLQQLKGIKDNLKTGITQATGQSINNFSREGMLARWMAEWSEAEKQRYLGTLSAEGRAVTEKHYDTVFF :  619
SeAfp18   : PAAVSSLKAVDKLLPDLAGDLTASIAALSNALQAAEYPARDFSEAKIQAGYAQLWATKVKESLSAASARLTERPLDEHSRG---ARLAKHWANLAKERSVGNYPPPD------AQQVLAS :  643

YrAfp18   : EVIQHYLPLLSKETDPQGERLLQRLRLEVSNAAEGTTVYPATMAEILAGMKSTEQAIRDWSGR-----------KLLRVVFLAACLEGVKLMPKLAALPLRVAIKFVITGAKVAWATHKG :  745
SpAfpX18  : EVIQHYLPLLSKESDPQGEGLLQRLRLEVGNAAKGNTLYPATMADILAGMKSREQAIRDWSER-----------KLIRGAFLAVCLGGVKLLPNLAALPLRLPIKFAITGAKVAWGAHKG :  728
SeAfp18   : LKEQGLLAGTLSTGDPAGYLFATRLAGELENAGNGELRLPMSPEQYVALEKGLVEYIVKWGQRRISQGVTRIVIELSFEQALDTVSFNVSKLFRLPYKVLKASIKIPYNVNKVNNYTMPG :  763

YrAfp18   : QQGIRGGEGDVDDEIGEYAKRSFKTASVKVVLSLPPGLATMLGVASIALDVY--------------------------------------------EGGLKGAGGKIAKNIVGDAPWRAL :  821
SpAfpX18  : QQGIRGGEGDISDEIAEYTKQSYKTAAIKIVLSLPPGLATTLGIASIVWDVY--------------------------------------------EDGLKGAGEKVAKHIIGEAPWRGL :  804
SeAfp18   : HDKPYKAIYGLLGKKLKQLGFNLLTAPVPGVMKLVAGTGITAGAGVHNLRVENREKTFSAVYQHVVEGKQSEKIKMDSVKSMIFDSVLDTATTAAFKGGRRAWQSGRNKNDIEINEVTLS :  883

YrAfp18   : NQGSKIAAEAY-TTASMNAALKEGGANPVSHSSTLQQQMDAKPFFDNSDQDADQP-RVRRKREMTDEIMLSDGSSRSEAKALPDENELDTDQSKSRPESALAVETLQSEHFDFDRGIRYQ :  939
SpAfpX18  : DTGTRTVAEAYATTL-IEAAITEEKITSATHSSPLQPQTDVMPLSDEHDPGINQPHRIR-KRAVADT-ALRP-APR---RS-P---QWH-D-N-I-P-SD-A--TLQSGHFDFDRDIRYQ :  905
SeAfp18   : QAQQRQVAANNVDTAWGNAAANERKVEPVSHSSPLQQQTDVRPFSDKSDQDAAQP-RVR-KRAVADMAPPPP-------------HQWHNN--------IPANTTLQSEHFDFDRDIRYQ :  980

YrAfp18   : DFSDEQKKQTYLHGIKFVLLQIENDGHFAQNIRNNAYLARIGAKLAVPVDIYR-YKLNNT-FLLPDEIDSKSGVLIRLDSEIPYYYVSEGKDLLENIAWAMPYNAAN-RGPLKF-S-LD- : 1053
SpAfpX18  : DFSGEQKKQTYLHGIQFVLFQIENDGRFSQQIRNHAYLARIGDKLLVPVDI-KGHKLNNTIFL-PDSRGAKSGVLIRLDSDIPYYYVDEGKDLLENVEWAMPHN-ANKREP-RIVTRLTR : 1021
SeAfp18   : GFSDEKKKQTYLHGIQLVLLQIQNDGRFAQKIRNNAYLASIGVSNLVPVDIEG-YKLKNT-FLLPDSPSAKSGVIVRLDSENPYYYVGKGSDLLEDIKWAMPHDANEQKYEYMPDS---- : 1094

YrAfp18   : PGEVTS--SH-SGVDILNNIRSERFKF-ETYFNYNAPEAMSIESLSAQLANTIEADYKFKNTSPTNKILISRAIVGAHI--PDPGVRATQGE-YHIEFDS-DELAPAKYLRSFARPFSTL : 1165
SpAfpX18  : VGRMRGFISFPSGVDILNQIRSGKARFGEN-FNANNPDPMDIASLSARLADTMEADYKLKGQSITNKLLISRAIAGAHISAPDSDITFTE-EKYHLEF-TWDNLTPAEYLRSFSRPFSTL : 1138
SeAfp18   : -DGNAHFYRLRNGFDIFNDIRMGNSSF-EYFFNYNNPEPISITRLSSYLANTIEGDYKNGSSPPKNKLLISRAITGAHIP--DPGVTATAAG-YHIEF-TWDNLTSAEYLRSFSRPFSTL : 1208

YrAfp18   : SGEMQLISSSIKGETIQETELHVHQAEYIGSWVDATAGAIISFTPEGWFLNTAQSAAEITADLTEGKDPDPLAVAGLLVGIIPGGKIAAKVGKFTRIGGKTVKYGLILGNKSVDLAIVGK : 1285
SpAfp18X  : SGEMQLVTSSIKGETIQETELHVHQAEYIGSWVDATVGAIISFTPAGWVLNTAQSAADIAADLTEGKDPDSLAVAGLVVGCIPGERIAAKVGKFTRIGGKVVKYGVMLGNKAVDLAIVGR : 1258
SeAfp18   : SGEMQLVSSSINGETIQETELHVHQAEYIGSWVDATVGAITSFTPEGWVLNTTLSAADIAADLTEDKDPDPMAVAGLVVGCIPGGRIAAKVGKFTRIGGKVVKYGLMLGNKAVDLAIVGK : 1328

YrAfp18   : SIKTAVDTGEPLAIYQAFLASGMSVKNSYDIAKNMSSELKISKKIEESARLRKLKALQKNTY-KYSMNSKMPV-RKFRVGQTDLLGKIHKGEIKISRN-NGT------TWEKGNQLHLLA : 1396
SpAfpX18  : SIKTAVDAGEPLAIYQAFLASGMSVKNSYDMARNMSSRLKINRRMEESASLEELEALQ-NDRPEYA-TS-VPVIRTFRVGLTEMLGRINNGDIEIF-NEN-TS-----TWEKGSKLHLLA : 1368
SeAfp18   : SVKIAVDTGDPLAIYKALLISGIGTTNSYDMAKNMSSKLKISKKLEDSVQLDKPLQNLR-----GG--SKRPE--EGESTSQKRSQPDESPENSQEAVYSGPSGRARVTFDQTTEHHIML : 1439

YrAfp18   : YRLQNAGGGRLLPDVFRDKIVIGEYSFKRVKYNQ-----KKLNEMMRIAKMYTPTSNSTERIAKLQQNYKTGKEMSHAPQYDTYNDLSLGEKLDLFINS-NTDATTRGVLAGKINESITN : 1510
SpAfpX18  : YRLQNAGGGRRLAALFSNKIVIGEHTFKRVKYSQ-----DKFNEMMRIAKTYTPTSNSTERIAKIQQDYRAGKEMSHAPQHDHYNNLSLDEKLNLF-NKPDTDAITRGVLAGKINESIAN : 1482
SeAfp18   : NRFSRLVLAR---SFAEKRIQSFEHSSSKFKSSDEFKVDKSIPLDAQNASLNDMASKTDNLINQLEQFTGDKKESISANLKN--IQLMIGDIMENRNNRNDNLNFHTYVLVGKNTTYDEI : 1554
```

```
YrAfp18   : INLYETAKGVDAWKTSANKATDVVLAPQNIFLKGRAGECLPESILMGWALQSGQDTKLAKKLMNIHSSSNIAANPLYKSLVELHSDGNASRFGESVISDINMKMLSG-AESKLFPTENSS : 1629
SpAfpX18  : INLYETAKAADDWKASANKATEVVLAPQNIFLKGRAGECLPESVLMGWALQSGQDAKLAKKFMSIYSSSNVADDPLYKSLVELHADGNASKFSAAAISDVKMNAL-GDAESRLFPKENAS : 1601
SeAfp18   : SNYYGFAEFKFSINNGN--LTLNNLIGHPYVVINKYPEFRDYLIRNGVITQPELEPYNIKNVARYLGSKAVSSEIGYYESIPATKVKILSFTGANPITQKLGNTLQS--ISDLFSKKRFS : 1670

YrAfp18   : ----VRVDIPEHTMLISKVNKDGKIKYVFYDPNYGMAYFNKYKEMIS-FFKKKIKGYDTPKRSTS-FRQLDYSHLSDIKIKGKNLN---EIIDGEIPQIFRQEDVNLEGITPQDGIYRML : 1740
SpAfpX18  : ----VRVDIPEHTMLLSKVNQEGKVKYVFYDPNYGLAYFDKYKD-LSAFFKKKLEAYDTPENSTNLY-QLDYSHLPDVKIKGRNLN---EIINGEIPALYKQERVNLQGISPHEGVYKMP : 1712
SeAfp18   : DEPNDAITQDELFTRQTDLYTYDEMINHLKTIVRTTFTPVKPIEEIS----------------SE--------KT-DAAIKGIMTNGDAEIISENIIESK-------------------- : 1745

YrAfp18   : GTHQQENNTYIKSQN-NIYQVEWDQTTNTWRVFDSANTNRSRITVPIKRDTDGEWFKYTETGLKGGGLFDEIKNNWLQRKRFKNLQDFNDIVDFEENKWPSEPINKDIHMIWVGTRNISE : 1859
SpAfpX18  : G----DKN-YIK-VNDNVYQVEWDQTMSTWRVFDPANTNRSRVTVPVKRDANGEWFRHLDTGLKGGGLFDEIKNNWLSGKKFKRLQDFTEVVELEANKWPPEPVNKNIHMIWIGTKKISE : 1826
SeAfp18   : --------------N-------SLDHD----------TSYFSHEVISAVDRAYEKVNSVKSLFERAESQPEIKSQLKS------------------------------------------ : 1792
                                                                                                               Glycoslyltransferase domain
                                                                                                                              DXD
YrAfp18   : KNIGLSLETARKNSDYNTTIIYDSGIAGYESAKDFMTEKFKDSKVTLVDFRKKS---------YFHQLQQEPSFPYYEQAIRDKK---------YAQASDILRLLVLKYEGGIYKDIDDV : 1961
SpAfpX18  : KNIGLSIETAQKNPDYNTTIIYDSGIPGHESAKSFMAEKFNDSNVTLVDFRNKI---------YFHQLQKEPAFSYYEQAIRDKK---------YAQASDILRLLVLKYEGGIYKDIDDV : 1928
SeAfp18   : ----LINEATG----LNDEVIIDIDKVDVSSISDMAYNRFKDNINTLHDYLSKQKSGGYTSFVLFKYRQFNPSKPDAFAMPYDSKKRIFLAMLPENQRHGGLVDTVTHEASHNSSYTLDH : 1904

YrAfp18   : QVKAFGSLAFPKGIGVMREYAPEAGKTTAFPNTPIAATKNNPVVNKTLELAVENYHRGETNVLKLAGPDVFTESLYQEIPGMRPQVLGAQLDQFELAKRQALGMRLEKPKGFADEKLTLQ : 2081
SpAfpX18  : QVKGFGSLAFPKGIGVMREYAPEAGKATAFPNTPIAATKNNSVVNRALELAVENYRRGETNVLKLAGPDVFTEALYQEMPGMRPQVLGQQLEQFELAKRQALGMPLEKPKSFADEQLTLQ : 2048
SeAfp18   : TYIGN-----------------ARSNTGSFPSSFEPPTRDS-------RYFYENER-----VGKLSAKYALDLPENANITQNQLAQADILLRNSRLTKADTLLNAAEYN--AYMIDVLSR : 1993

YrAfp18   : EKAKI-RQPYEAIRGLSGYVDNGADHSWVTD-MPGNSTQSSGLS---------------------------------------------------------------------------- : 2123
SpAfpX18  : EKSRISR-PYEAIRGLSGYVENGADHSWISDIPGTSALGNSQSEVSATFEPAIFKQALTEFVSQQAAQGPGDPAYTIYQSAFNSLPESQEITLQQVNDLKRALSASGNDLVAQAKYQYAL : 2167
SeAfp18   : VKINGSHIEFEENSSRSKRSVSGGNHKLDDIIILASMNATLPRREGGEFESSTFKKTLTEFINKGTTSSQVDPVYASYLPAFNTIVDSEGITLQQVKDLKSALPATGNDLRAQARYNYAV : 2113

SpAfpX18  : LAYLSGPRGIQLPIDKNEIR--------------------------------------------------NRIVQEIEHADYVYGWKVKEREQFQAQLNKAIERKNEISSRYSVTLRPKM : 2237
SeAfp18   : LQYLSGSAGIKLPIDTDVIRNSINDDLNIASCVDNCKKAASQRLELLKAVPEYLSGPEAIHAQIDTNLMINRIKNTLKHENYTYDLKVNERD----YLLLQLEQKYIASSKPPKVGAQKW : 2229

SpAfpX18  : HANADRKVVNLRNEVTERNRLMNEWASKKNQDSLRLDMLNDSWHESYD------------------------------------------------------------------------ : 2285
SeAfp18   : SADINKEINYLCDEIEEKKREVVCWENKKNESSRRLDLWSTATKYLSDPKGIKPPIDTDVILNNINDVLNHAGYVYNYKVREKDNLEKKLSRTPLTSQIEVNRLQTEIAKRYRLIDYWGK : 2349

SeAfp18   : KADETLLMFILWRNVFG : 2366
```

Fig. S4 Amino acid alignment of *Y. ruckeri* (Yr) Afp18, *S. proteamaculans* (Sp) AGR96X AfpX18, and *S. entomophila* (Se) Afp18 sequences. The conserved glycosyltransferase DxD, which is essential for sugar donor substrate binding, is indicated. Red (conserved) and yellow (shared) coloring indicates amino acid similarity between AfpX18 and the other two sequences, as outlined in Jank (1).

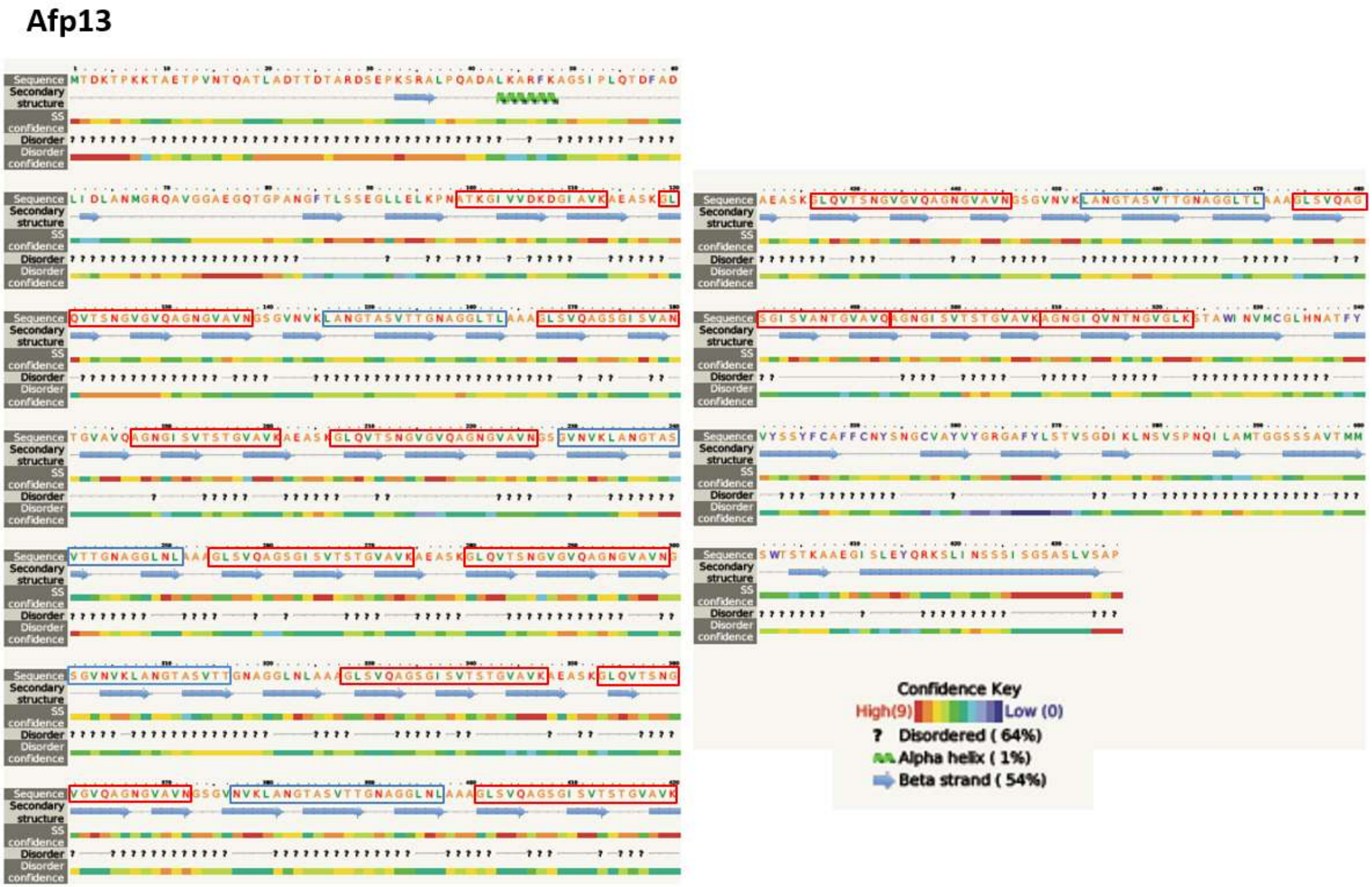


```
SpAfpX13 : MTDKTPQNQAENPVNTPTTLTDTTDTARDGELKSRALPQADALKERFKAGSIPLQTDFADLVDLANMGRQAVGGTEGQTGPANGFTLSLEGRLELKPNAA : 100
SeAfp13  : MTDKTPKKTAETPVNTQATLADTTDTARDSEPKSRALPQADALKARFKAGSIPLQTDFADLIDLANMGRQAVGGAEGQTGPANGFTLSSEGLLELKPNAT : 100
YrAfp13  : MKETHRLIIGNEKQ-------MLTGLAVVG-------PETNSLKERFKEGSIPLQTDFEKLIDIADVGRKAAGLAPGQTGPGNGFQLDSVERLSVKVDSN :  86

SpAfpX13 : KGIVVDKDGIAVKVEASKGLQVTASGVSVQAGNGISVAGTGVAVKVEAS---------KGLQVTASGVSVQAGNGISVASTGVAVKAGNGISVAGTGVAV : 191
SeAfp13  : KGIVVDKDGIAVKAEASKGLQVTSNGVGVQAGNGVAVNGSGVNVKLANGTASVTTGNAGGLTLAAAGLSVQAGSGISVANTGVAVQAGNGISVTSTGVAV : 200
YrAfp13  : S-----------------GLSVTSSGVKVIPKN---S-G---------G--------------------------VVVDSTGISVKPGRGMEVTSDGVRI : 130

SpAfpX13 : KVEASKGLQVTASGVSVQAGDGISVAG--------------------------TGVAVKAGSGISVAGTGVAVKVEASKGLQVTASGVSVQAGNGISVAG : 265
SeAfp13  : KAEASKGLQVTSNGVGVQAGNGVAVNGSGVNVKLANGTASVTTGNAGGLNLAAAGLSVQAGSGISVTSTGVAVKAEASKGLQVTSNGVGVQAGNGVAVNG : 300
YrAfp13  : TNEFAFQKGMILMFAGTAAEMPPGWALCD------------------------------GS--------------------------------NGRPN-- : 166

SpAfpX13 : TGVAVKAGSGISVAGTGVA-----------VKAGNGISVAGTGVAVKVEASKGLQVTASGVSVQAGDGISVAGTGVAVKVEAS---------KGLQVTAS : 345
SeAfp13  : SGVNVKLANGTASVTTGNAGGLNLAAAGLSVQAGSGISVTSTGVAVKAEASKGLQVTSNGVGVQAGNGVAVNGSGVNVKLANGTASVTTGNAGGLNLAAA : 400
YrAfp13  : ------------------LIDR---------------------------FILGSTFTHSKYSNTAQLAGSGN-----TKTYNK----------------- : 199

SpAfpX13 : GVSVQAGNGISVAGTGVAVKVEASKGLQVTASGVSVQAGNGISVAGTGVAVKAGSGISVAGTGVA-----------VKAGNGISVAGTGVAVKVEASKGL : 434
SeAfp13  : GLSVQAGSGISVTSTGVAVKAEASKGLQVTSNGVGVQAGNGVAVNGSGVNVKLANGTASVTTGNAGGLTLAAAGLSVQAGSGISVANT------------ : 488
YrAfp13  : -------QSTSFTVTGTVNIAETILTLGQIPRHNHNDG-------------------------------------------------------------- : 230

SpAfpX13 : QVTASGVSVQAGDGISVAGTGVAVKVEASKGLQVTSNGVGLNNTAWIKMMCGLHNATFYVSDTYVCVFFCNHSTGCTAYVYGRGGYYLSMYKGDVKLNSV : 534
SeAfp13  : -----GVAVQAGNGISVTSTGVAVKAGN--GIQVNTNGVGLKSTAWINVMCGLHNATFYVYSSYFCAFFCNYSNGCVAYVYGRGAFYLSTVSGDIKLNSV : 581
YrAfp13  : ----------VRY------TYKAGNYDIMNINQYGSVSVGTG------------------SSSNKFARLENFSD--------------SESVHLTQSSTV : 282
                                           Q                 C              C   C     C

SpAfpX13 : DHNEIISMVGSGSIAAATMVSWKSTKAAAGISFKYLGKNLITSTSHSGSVTLVAAP : 590
SeAfp13  : SPNQILAMTG-GSSSAVTMMSWTSTKAAEGISLEYQRKSLINSSSISGSASLVSAP : 636
YrAfp13  : GD-------GGGHKHTATMSTNGHQHSTDVIPPYYTLAFIIKL------------- : 318
```

Fig. S5 Amino acid alignment of *S. entomophila* (Se) Afp13, *S. proteamaculans* (Sp) AGR96X AfpX13, and *Y. ruckeri* (Yr) Afp13. Similar amino acid residues are black-shaded. Conserved tryptophan (Q) and cysteine residues (C) are indicated. Refer to Supplemental Fig. S6 for an alignment of the tail fiber amino acid repeats.

AfpX13


Sequence
Secondary structure
SS confidence
Disorder
Disorder confidence
Confidence Key
High(9)
Low (0)
? Disordered (64%)
Alpha helix (1%)
Beta strand (54%)



Afp13

Fig. S6 Secondary structure and disorder prediction derived from the Phyre^2 server (2). Red boxes denotes amino acid repeats present in *S. proteamaculans* AfpX13 and *S. entomophila* Afp13. Blue boxes indicate the less-conserved *S. entomophila* Afp13 amino acid repeat.

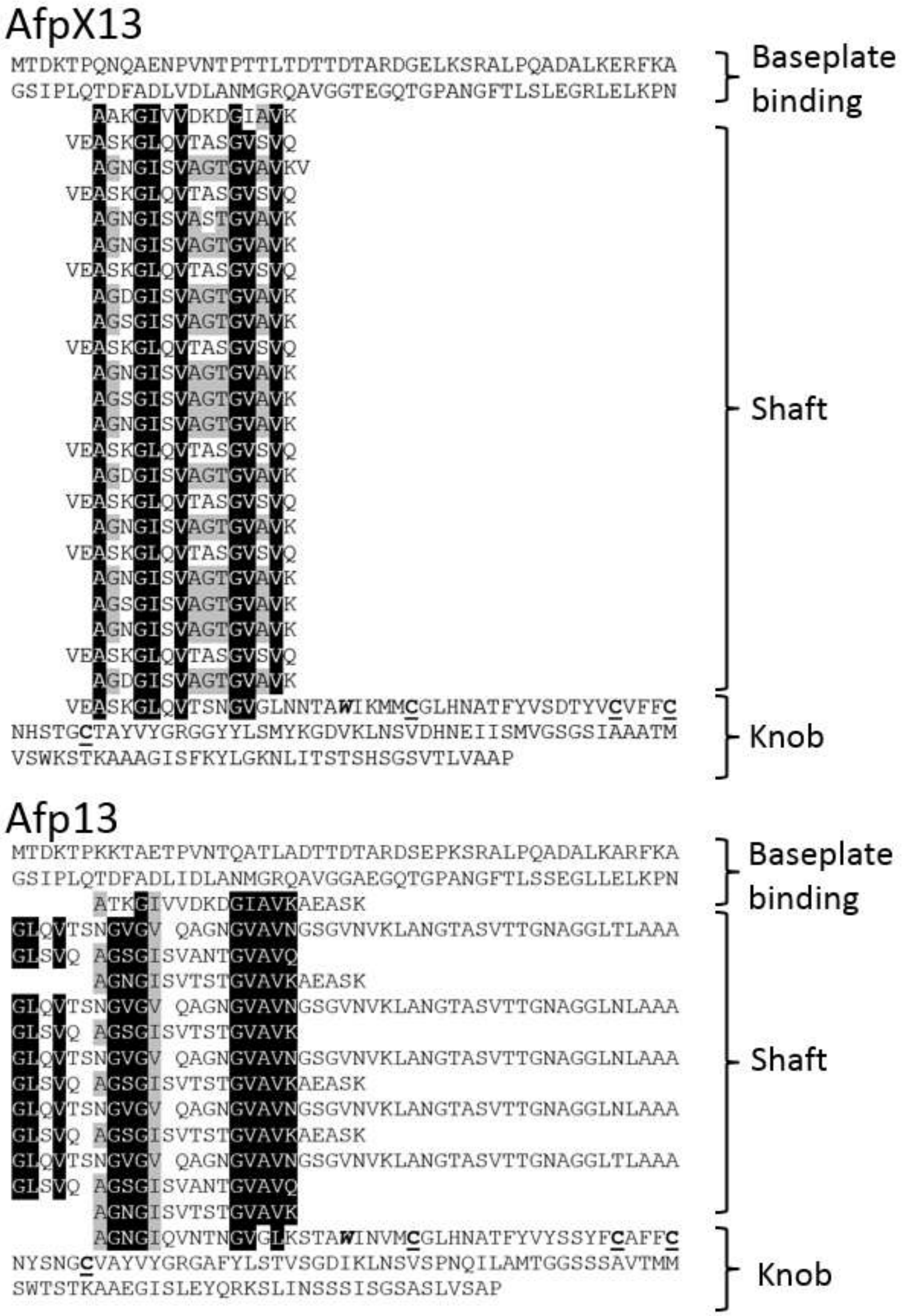


Fig. S7 Amino acid alignment of the predicted *S. proteamaculans* AfpX13 (24 repeats) and *S. entomophila* Afp13 (14 repeats) tail fiber shaft repeats. Putative Afp baseplate, shaft, and knob regions are indicated. A repeat consisting of a tryptophan residue (***W***) followed by a series of four cysteine residues (<u>**C**</u>) was located within five residues of the C-terminus. Predicted baseplate binding, shaft, and knob regions are indicated.

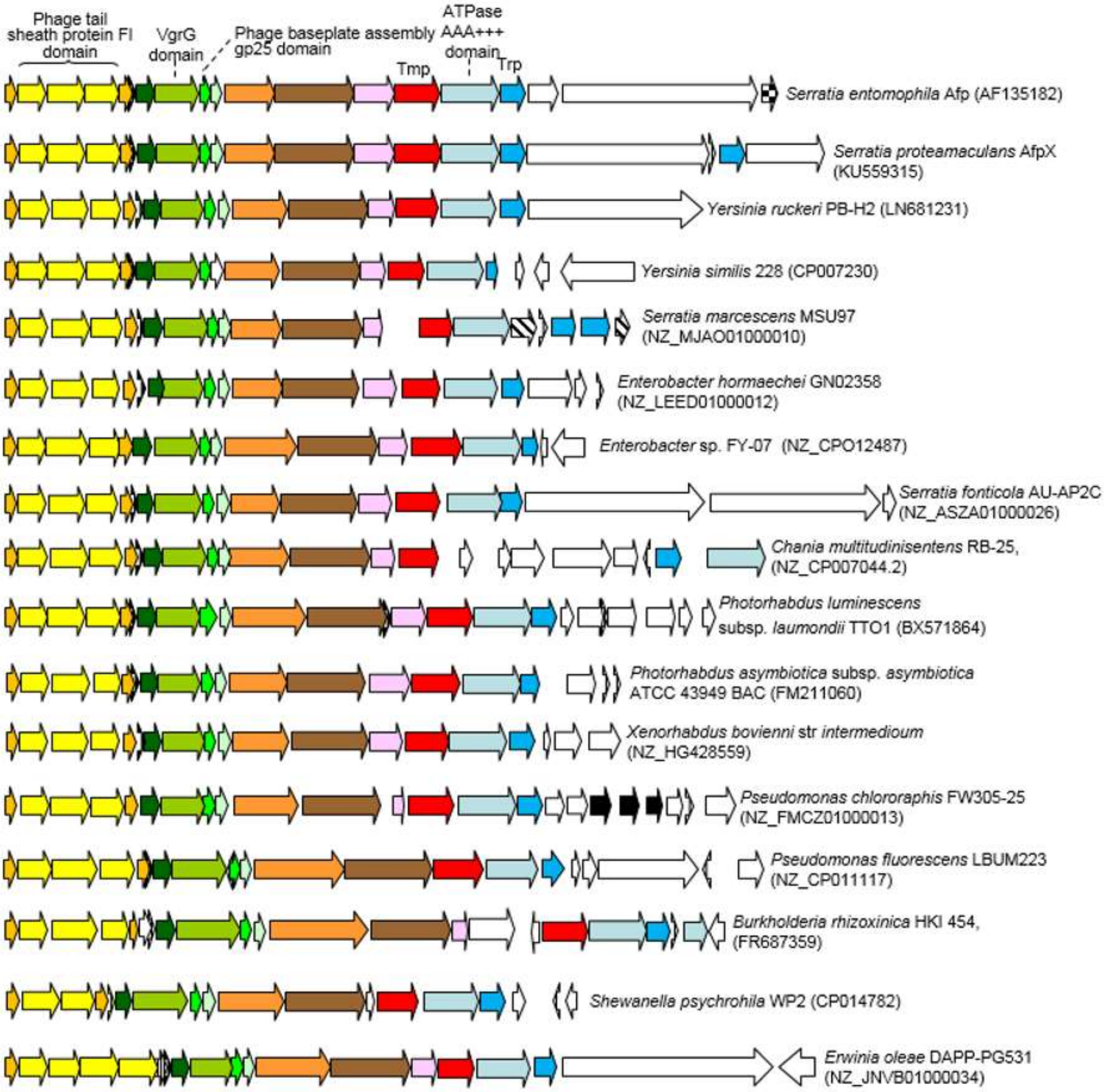


Fig. S8 Afp variants found in the draft genome sequences of various bacterial species sourced from the NCBI nrBlast database. ORFs are denoted by arrows. Amino acid regions with high similarity are indicated by similar coloring. Predicted and ascribed protein domains are listed above the *S. entomophila afp* gene cluster. Forward diagonal lines within the ORFs of *Serratia marcescens* MSU97 denote orthologs of the type three secretion system effector proteins WP_073532279 and WP_073532282. Refer to Supplemental Fig. S9 for sequences used and associated references.

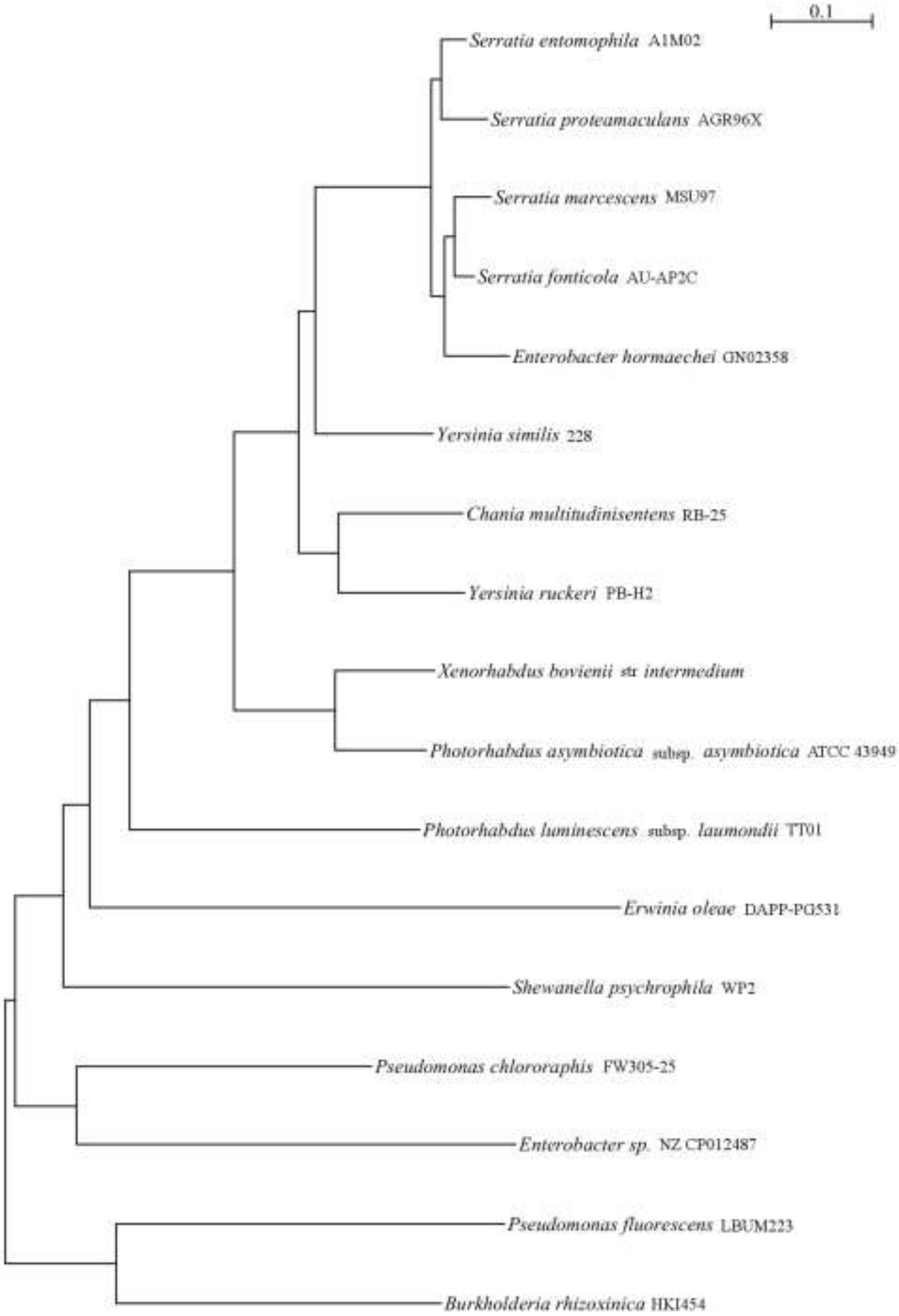


Fig. S9 Phylogenetic assessment of AfpX15 of selected Afp variants found in the draft genome sequences of other bacterial species sourced from the NCBI nrBlast database. Sequences included: *Serratia entomophila* Afp (AF135182) (3); *Serratia proteamaculans* AfpX (KU559315); *Yersinia ruckeri* PB-H2 (LN681231) (4); *Yersinia similis* 228 (CP007230) (5); *Serratia marcescens* MSU97 contig00010 (NZ_MJAO01000010) (6); *Enterobacter hormaechei* GN02358 (NZ_LEED01000012) (7); *Enterobacter* sp. FY-07 (NZ_CPO12487) (8); *Serratia fonticola* AU-AP2C contig26 (NZ_ASZA01000026) (9); *Chania multitudinisentens* RB-

25, (NZ_CP007044.2) (10); *Photorhabdus luminescens* subsp. *laumondii* TTO1 (BX571864) (11); *Photorhabdus asymbiotica* subsp. *asymbiotica* ATCC 43949, BAC clone 20-21 (FM211060) (12); *Xenorhabdus bovienii* str. Intermedium (NZ_HG428559) (13); *Pseudomonas chlororaphis* FW305-25 (NZ_FMCZ01000013) (14); *Pseudomonas fluorescens* LBUM223 (NZ_CP011117) (15); *Burkholderia rhizoxinica* HKI 454, (FR687359) (16); *Shewanella psychrophila* strain WP2 (CP014782) (17); *Erwinia oleae* strain DAPP-PG531 (NZ_JNVB01000034) (18).

REFERENCES

1. Jank T; Eckerle S, Steinemann M, Trillhaase C, Schimpl M, Wiese S, van Aalten DM, Driever W, Aktories K. 2015. Tyrosine glycosylation of Rho by *Yersinia t*oxin impairs blastomere cell behaviour in zebrafish embryos. Nat Commun 6:7807.
2. Kelley LA, Sternberg MJE. 2009. Protein structure prediction on the web: a case study using the Phyre serve. Nat Prot 4:363-371.
3. Hurst MR, Glare TR, Jackson TA. 2004. Cloning *Serratia entomophila* antifeeding genes--a putative defective prophage active against the grass grub *Costelytra zealandica*. J Bacteriol 186:5116-5128.
4. Nelson MC, LaPatra SE, Welch TJ, Graf J. 2015. Complete genome sequence of *Yersinia ruckeri* strain CSF007-82, etiologic agent of red mouth disease in salmonid fish. Genome Announc 3 (1): e01491-14.
5. Sprague LD. and Neubauer H. 2014. Genome sequence of *Yersinia similis* Y228T, a member of the *Yersinia pseudotuberculosis* complex, Genome Announc 2 (2): e00216-14. (CP007230).
6. Matilla MA, Udaondo Z, Krell T, Salmond GP. 2017. Genome sequence of *Serratia marcescens* MSU97, a plant-associated bacterium that makes multiple antibiotics. Genome Announc 5 (9): e01752-16.
7. Adams M, Sutton G, Nelson K, Bonomo R, Mc Corrison J, Sanka R, Brinkac L, Nierman W. 2015. *Enterobacter hormaechei* GN02358 Deposited (01-JUN-2015) JCVI, 4120 Capricorn lane, La Jolla, CA 92037, USA (NZ_LEED01000012).
8. Ji K, Wang W, Zeng B, Chen S, Zhao Q, Chen Y, Li G, Ma T. 2016. Bacterial cellulose synthesis mechanism of facultative anaerobe *Enterobacter* sp. FY-07. Sci Rep 6: 21863.
9. Devi U, Khatri I, Kumar N, Sharma D, Subramanian, S, Saini AK. 2013. Draft genome sequence of plant-growth-promoting rhizobacterium *Serratia fonticola* strain AU-AP2C, isolated from the pea rhizosphere. Genome Announc 1 (6): e01022-13.

10. Robson EHJ. 2014. Unpublished, complete genome. *Chania multitudinisentens* RB-25, deposited (09-JAN-2014) Genetics and molecular biology, University of Malaya, Kuala Lumpur 50603, Malaysia (NZ_CP007044.2).
11. Duchaud E, Rusniok C, Frangeul L, BuchrieserC, Givaudan A, Taourit S, Bocs S, Boursaux-Eude C, Chandler M, Charles JF, Dassa E, Derose R, Derzelle S, Freyssinet G, Gaudriault S, Medigue C, Lanois A, Powell K, Siguier P, Vincent R, Wingate V, Zouine M, Glaser P, Boemare N, Danchin A, Kunst F. 2003. The genome sequence of the entomopathogenic bacterium *Photorhabdus luminescens*. Nat Biotechnol 21:1307-1313.
12. Waterfield NR, Sanchez-Contreras M, Eleftherianos I, Dowling A, Yang G, Wilkinson P, Parkhill J, Thomson N, Reynolds SE, Bode HB, Dorus S, ffrench-Constant RH. 2009. Rapid Virulence Annotation (RVA): identification of virulence factors using a bacterial genome library and multiple invertebrate hosts. Proc Natl Acad Sci USA 105:15967-15972.
13. Murfin K, Klassen J, Lee M, Forst S, Stock P, Goodrich-Blair H. 2013. Unpublished, Deposited (12-SEP-2013) Bacteriology Department, University of Wisconsin-Madison, Madison, WI 53706, United States. (*Xenorhabdus bovienni* str *intermedioum* (NZ_HG428559)
14. Crable, B, unpublished, *Pseudomonas chlororaphis* FW305-25, whole genome shotgun sequence deposited (11-AUG-2016) Oak ridge national laboratory, one bethel valley road, Oak Ridge, TN 37831, USA. (NZ_FMCZ01000013).
15. Roquigny R, Arseneault T, Gadkar VJ, Novinscak A, Joly DL, Filion M. 2015. Complete genome sequence of biocontrol strain *Pseudomonas fluorescens* LBUM223. Genome Announc 3(3): e00443-15.
16. Lackner G, Moebius N, Partida-Martinez L, Hertweck C. 2011. Complete genome sequence of *Burkholderia rhizoxinica*, an endosymbiont of *Rhizopus microspores*. J Bacteriol 193:783-784.

17. Xu G, Jian H, Xiao X, Wang F. 2017. Complete genome sequence of *Shewanella psychrophila* WP2, a deep-sea bacterium isolated from west Pacific sediment. Marine Genomics 35:19-21.
18. Moretti C, Cortese C, da Silva DP, Venturi V, Firrao G, Buonaurio R. 2014. Draft genome sequence of *Erwinia oleae*, a bacterium associated with olive knots caused by *Pseudomonas savastanoi* pv. *savastanoi.* Genome Announc 2(6): e01308-14.