

Ancestral acquisitions, gene flow and multiple evolutionary trajectories of the type three secretion system and effectors in *Xanthomonas* plant pathogens

Déborah Merda^a, Martial Briand^a, Eran Bosis^b, Céline Rousseau^a, Perrine Portier^a, Matthieu Barret^a, Marie-Agnès Jacques^{a§}, and Marion Fischer-Le Saux^{a§}

^aIRHS, Agrocampus-Ouest, INRA, Université d'Angers, SFR 4207 QuaSaV, 49071, Beaucouzé, France

^bDepartment of Biotechnology Engineering, ORT Braude College, Karmiel 2161002, Israel.

Table of contents of this file

Fig. S1: Comparisons of genomic environments of T3SS clusters in different *Xanthomonas* species.

Fig. S2: Comparisons of large genomic environments of the T3SS clusters in different *Xanthomonas* species using a window of 200 kb upstream and downstream of the cluster.

Fig. S3: Individual maximum likelihood phylogenies built for each *hrc/hrp* gene coding for the T3SS

Fig. S4: Genomic environments of *avrBs2* in *Xanthomonas arboricola* strains

Fig. S5: Genomic environments of 10 predicted type three effector (T3E) genes specific to *Xanthomonas arboricola* group A strains

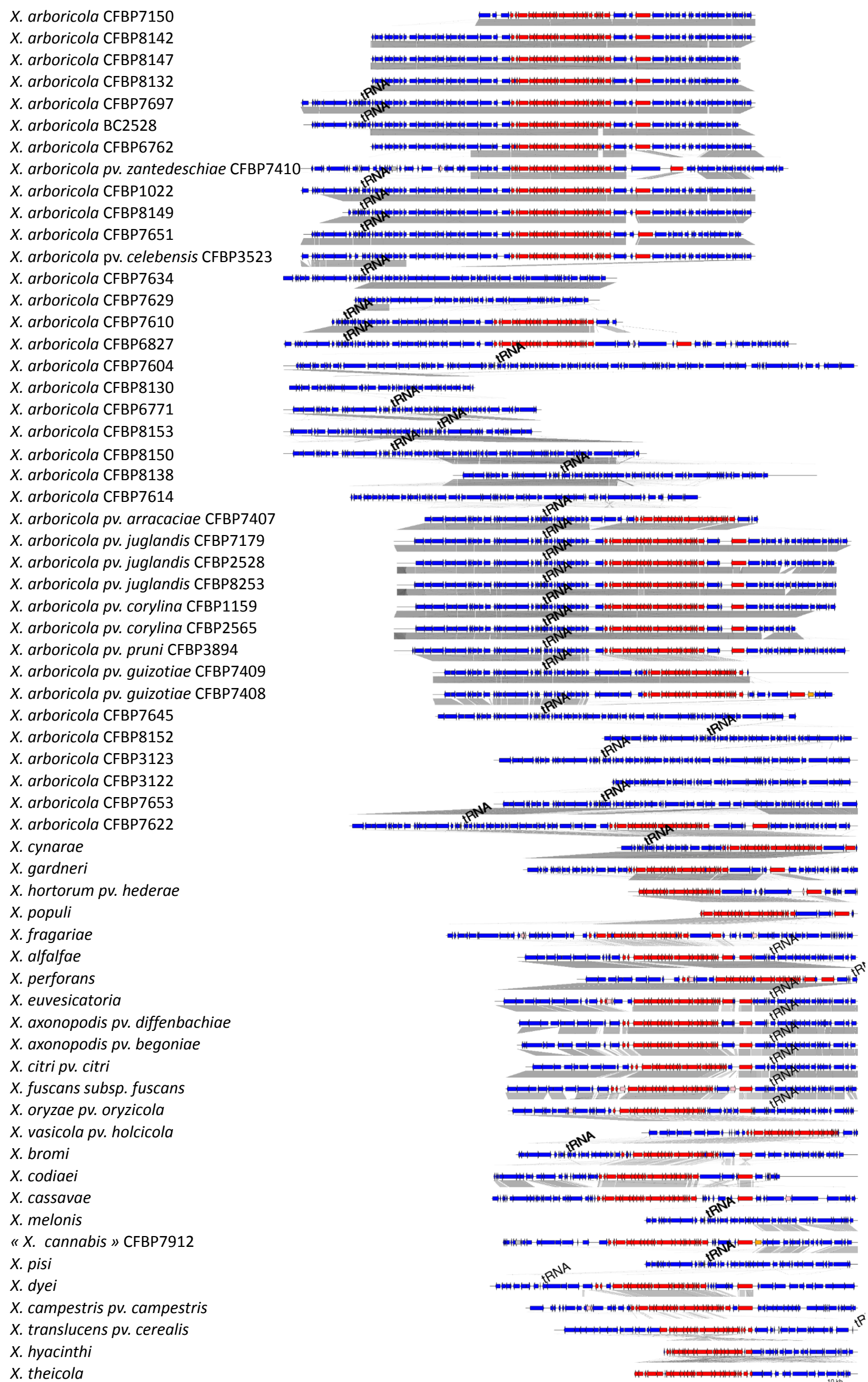
Fig.S6: Graphical circular representation of the draft genomes of *Xanthomonas arboricola* strains belonging to the three successful pathovars : pv. *juglandis* (CFBP 2528, CFBP 7179, CFBP 8253), pv. *pruni* (CFBP 3894) and pv. *corylina* (CFBP 2565, CFBP 1159).

Table S1: Results of Shimodaira-Hasegawa tests comparing phylogenies of T3SS coding genes in *Xanthomonas*

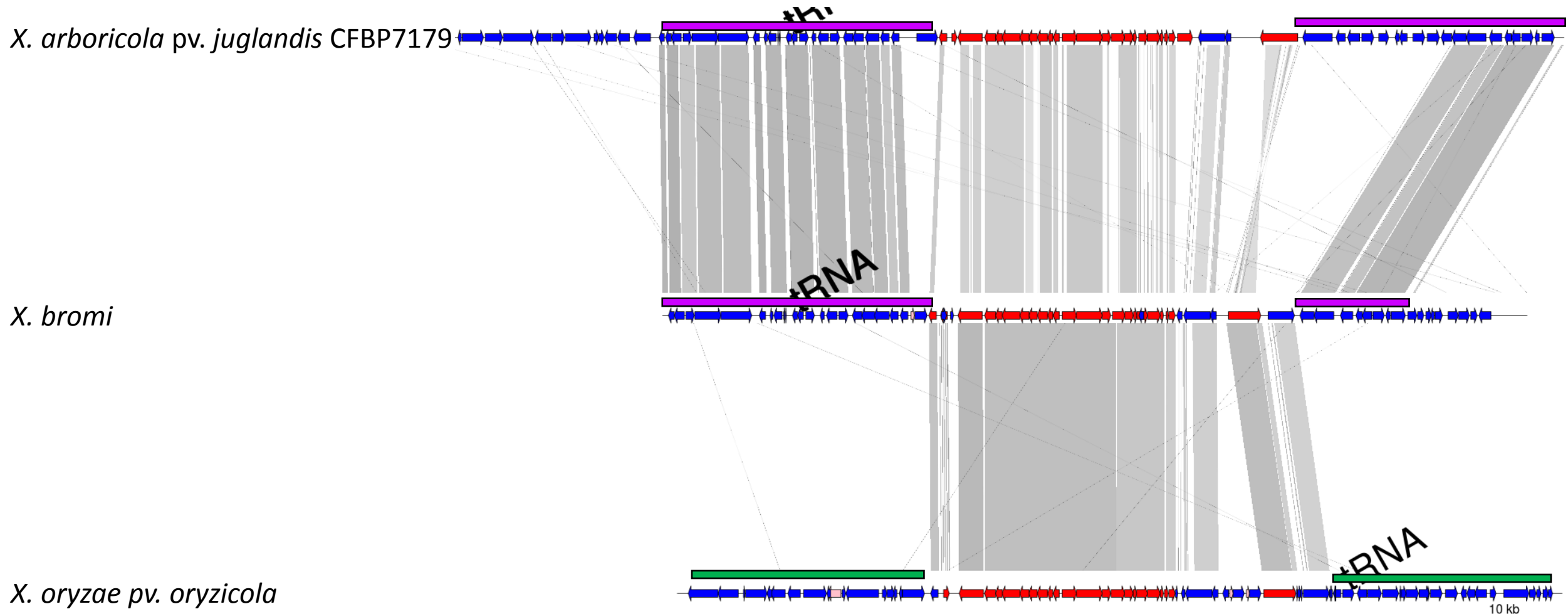
Supplementary Figure S1

FIGURE S1 Comparisons of genomic environments of T3SS clusters in different *Xanthomonas* species. Genomic environments (20 kb on each side) were compared using the R package Genoplots. The genes of the T3SS cluster are represented by red arrows. Pink arrows represent genes encoding transposases, orange arrows represent integrons, grey arrows represent phages. Other genes are represented by blue arrows. DNA fragments showing BLASTn similarities are connected with grey shading. (A) Comparison of strains representing all the phylogenetic clusters identified in the *Xanthomonas* genus. Strains are ordered according to their phylogenetic relationships. (B) to (F) : examples of comparisons showing similarities and differences between genomic environments of T3SS cluster. Similar genomic environments of the T3SS cluster were highlighted by bars of the same color placed above the schematic representation of the sequences. (B) Comparison between the strain CFBP 7179 of *X. arboricola* used as reference (representing the genomic environment retrieved in *X. arboricola* strains), *X. bromi* (CFBP 1976) and *X. oryzae* (BAI3). This comparison reveals that the genomic environment of *X. bromi* T3SS cluster shares similarities with the one of *X. arboricola* but not with the one of *X. oryzae* (this latter is shared by other clade B species). (C). Comparisons showing the diversity of genomic environments of T3SS cluster in *X. codiaei*, "*X. cannabis*", *X. cassavae*, and *X. dyei* and absence of similarities (except for *X. dyei*) with the genomic environment of clade B represented in green. (D). Comparisons showing the mosaic structure of the genomic environment of T3SS cluster in *X. arboricola* pv. *guizotiae* and in *X. dyei*. (E). Comparison showing that the genomic environment of T3SS cluster in the group 1 species *X. translucens* shares no similarity with the one retrieved in group 2 species from clades A and B. (F). Comparison showing similarities between the genomic environments of T3SS cluster in the three group 1 species *X. translucens*, *X. hyacinthi* and *X. theicola*.

A



B



X. oryzae pv. *oryzicola*

X. codiae

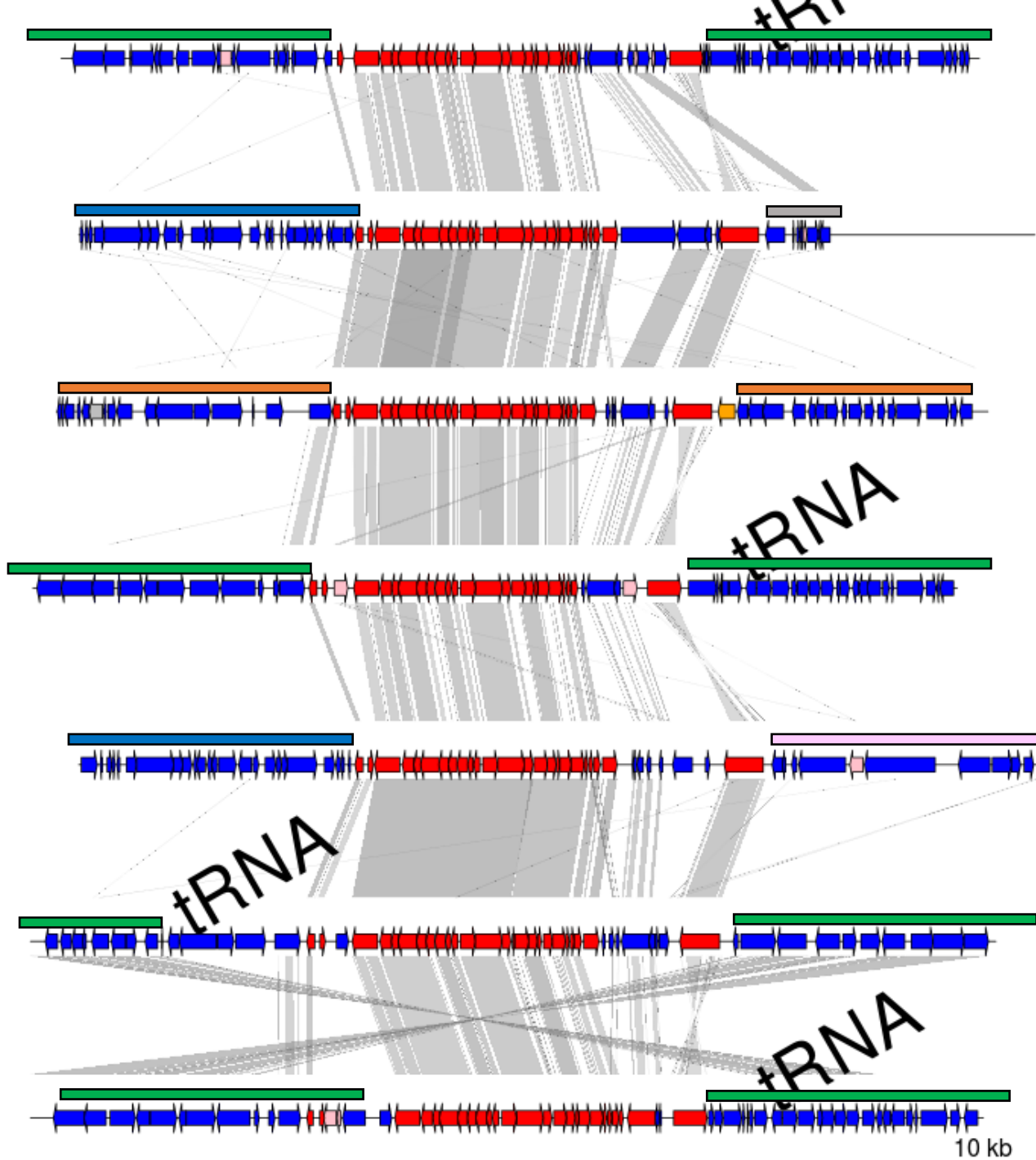
« *X. cannabis* » CFBP7912

X. fuscans subsp. *fuscans*

X. cassavae

X. dyei

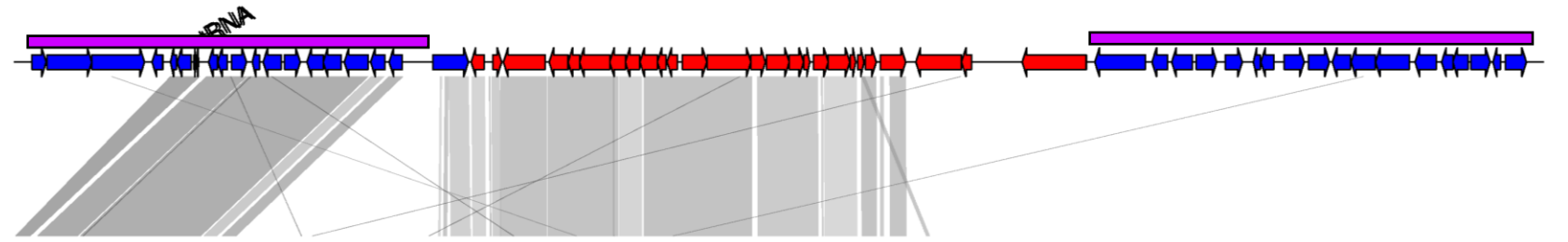
X. euvesicatoria



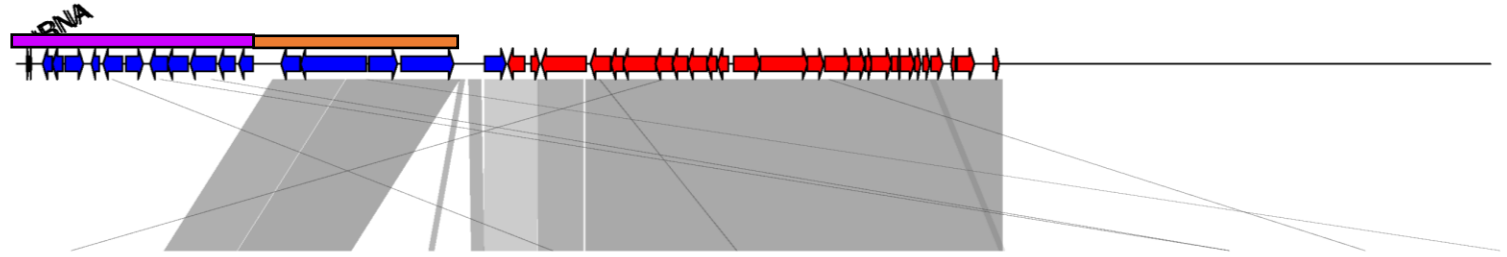
C

D

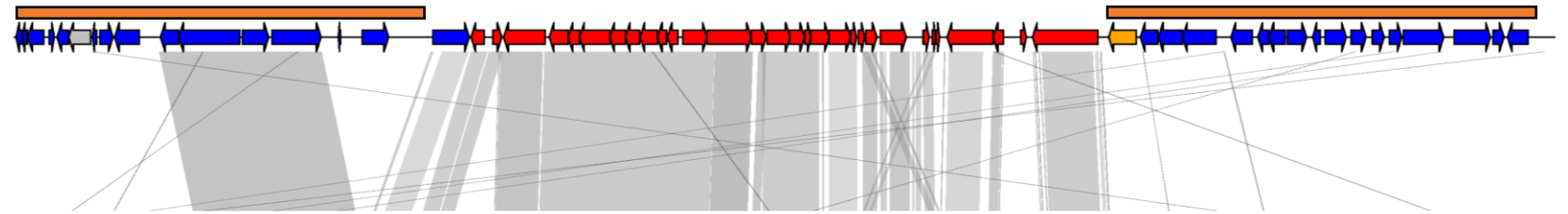
X. arboricola pv. *juglandis* CFBP2528



X. arboricola pv. *guizotiae* CFBP7408



X. « cannabis » CFBP7912

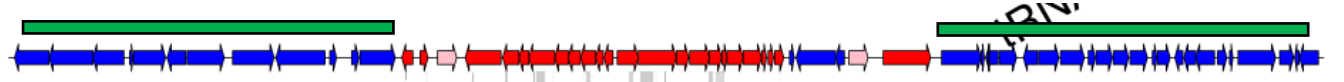


X. dyei

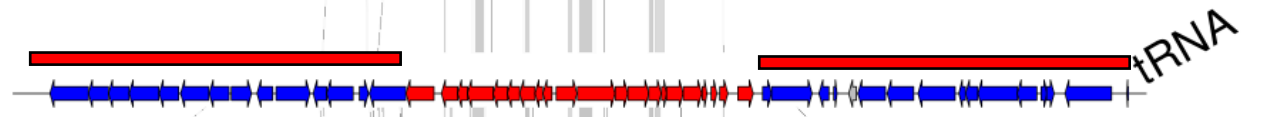


E

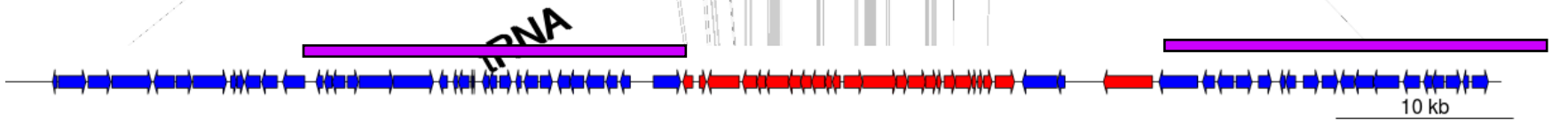
X. fuscans subsp. *fuscans*



X. translucens pv. *cerealis*

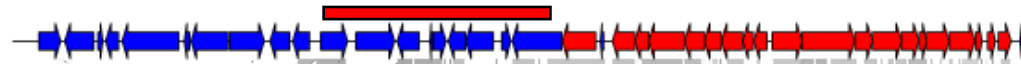


X. arboricola pv. *juglandis* CFBP7179

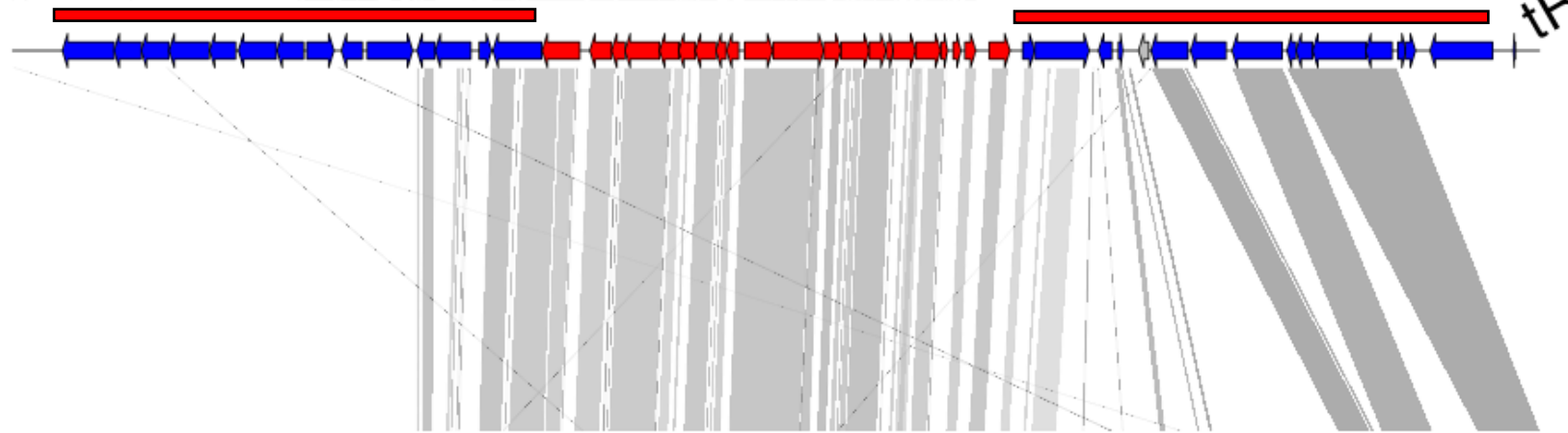


F

X. hyacinthi



X. translucens pv. *cerealis*



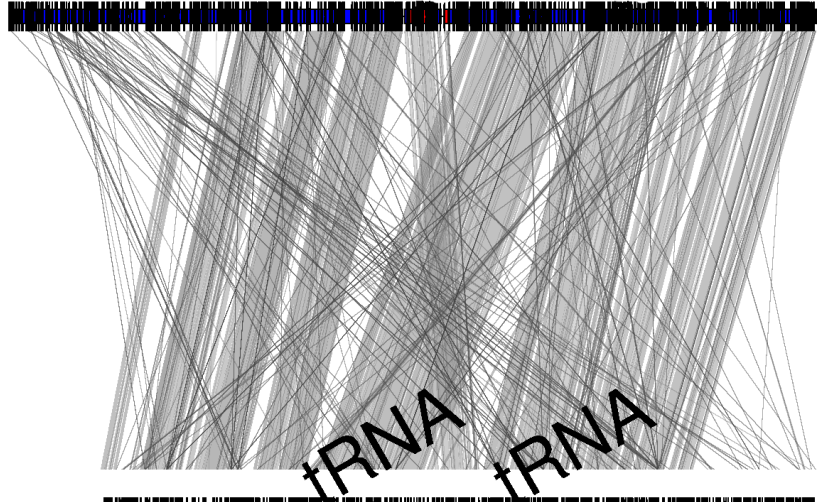
X. theicola



Supplementary Figure S2

Supplementary Figure S2: Comparisons of large genomic environments of the T3SS clusters in different *Xanthomonas* species using a window of 200 kb upstream and downstream of the cluster. Genomic environments were compared using the R package Genoplots. The genes of the T3SS cluster are represented in red; other genes are represented in blue. DNA fragments showing similarities are connected with grey shading.

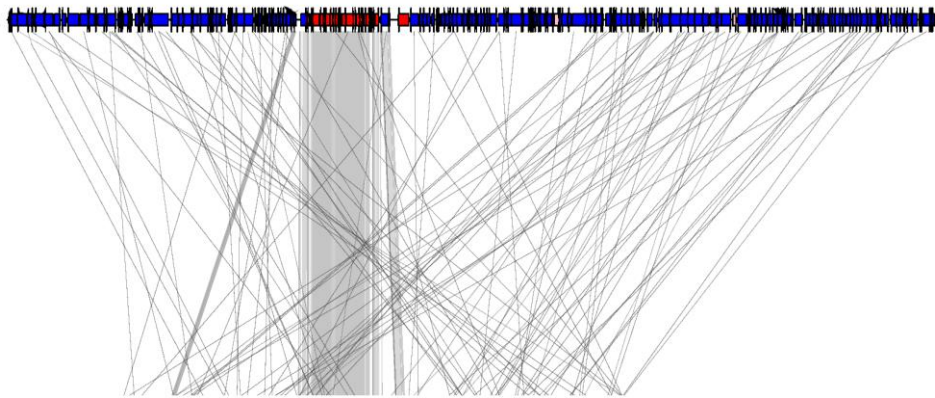
X. arboricola pv. *juglandis*



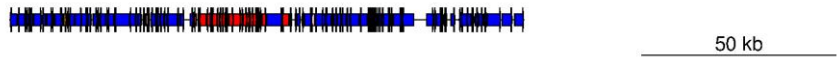
X. fuscans pv. *fuscans*



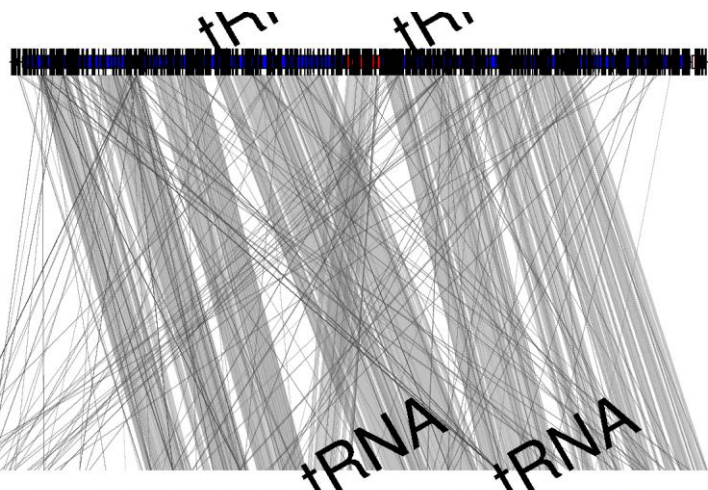
X. arboricola pv. *juglandis*



X. fragariae



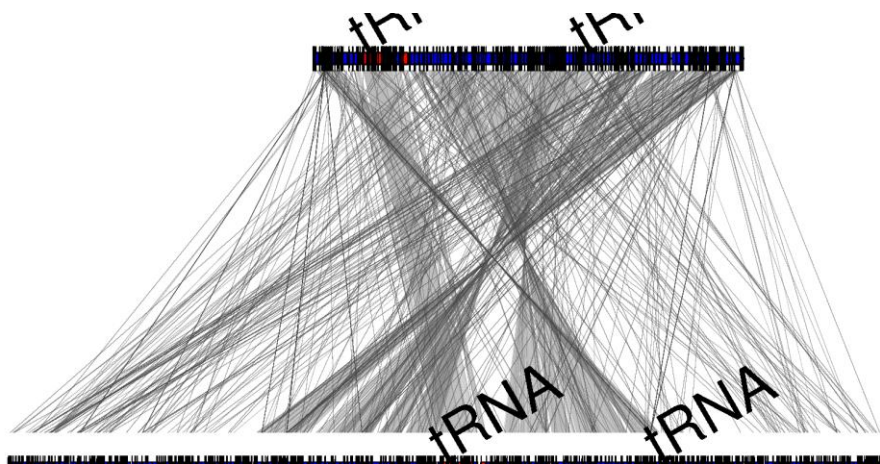
X. euvesicatoria



X. arboricola pv. *juglandis*



X. dyei

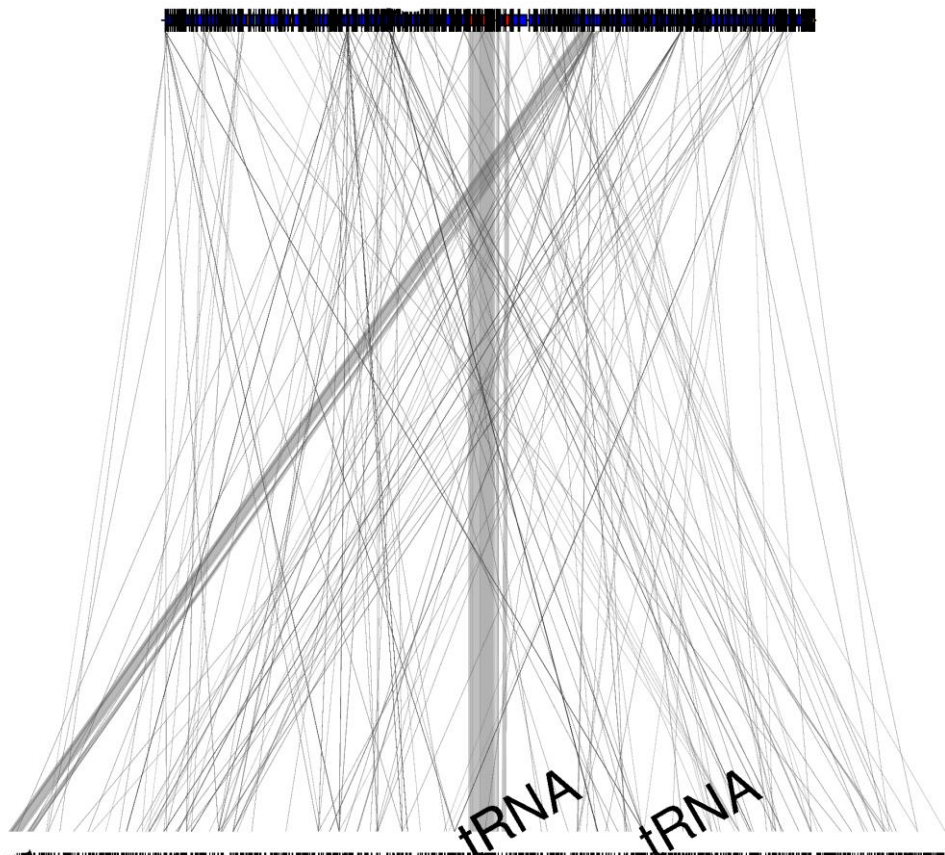


X. arboricola pv. juglandis



50 kb

X. cassavae

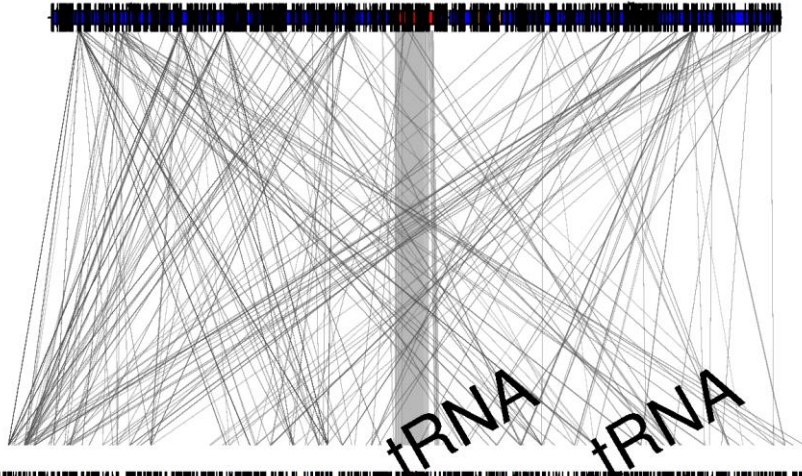


X. arboricola pv. juglandis



100 kb

X. campestris

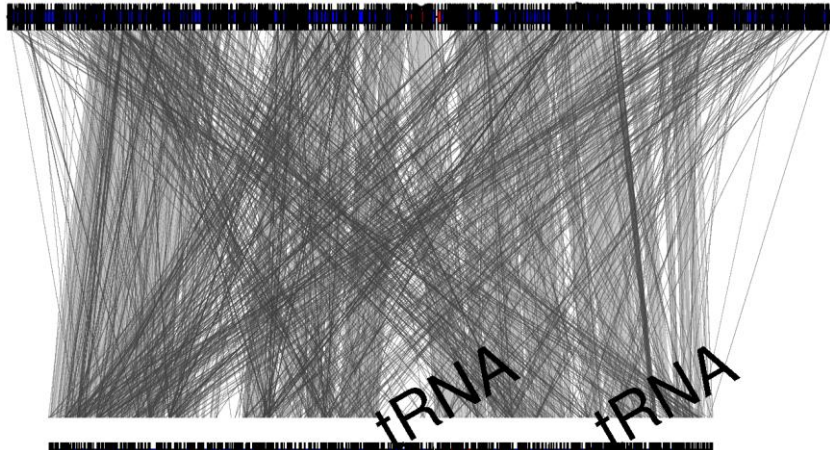


X. arboricola pv. *juglandis*



50 kb

X. arboricola pv. *juglandis*

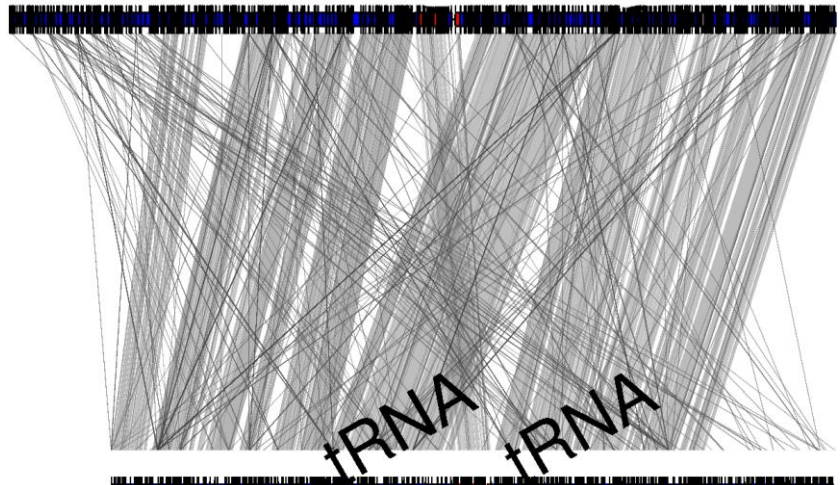


X. "cannabis" CFBP 7912



100 kb

X. arboricola pv. *juglandis*



X. alfalfae

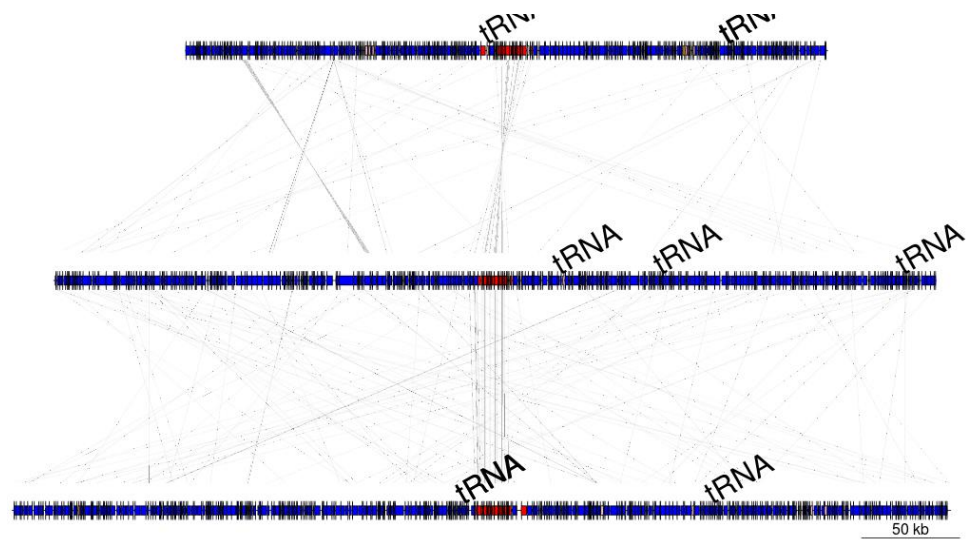


50 kb

X. fuscans subsp. *Fuscans*

X. translucens pv.
cerealis

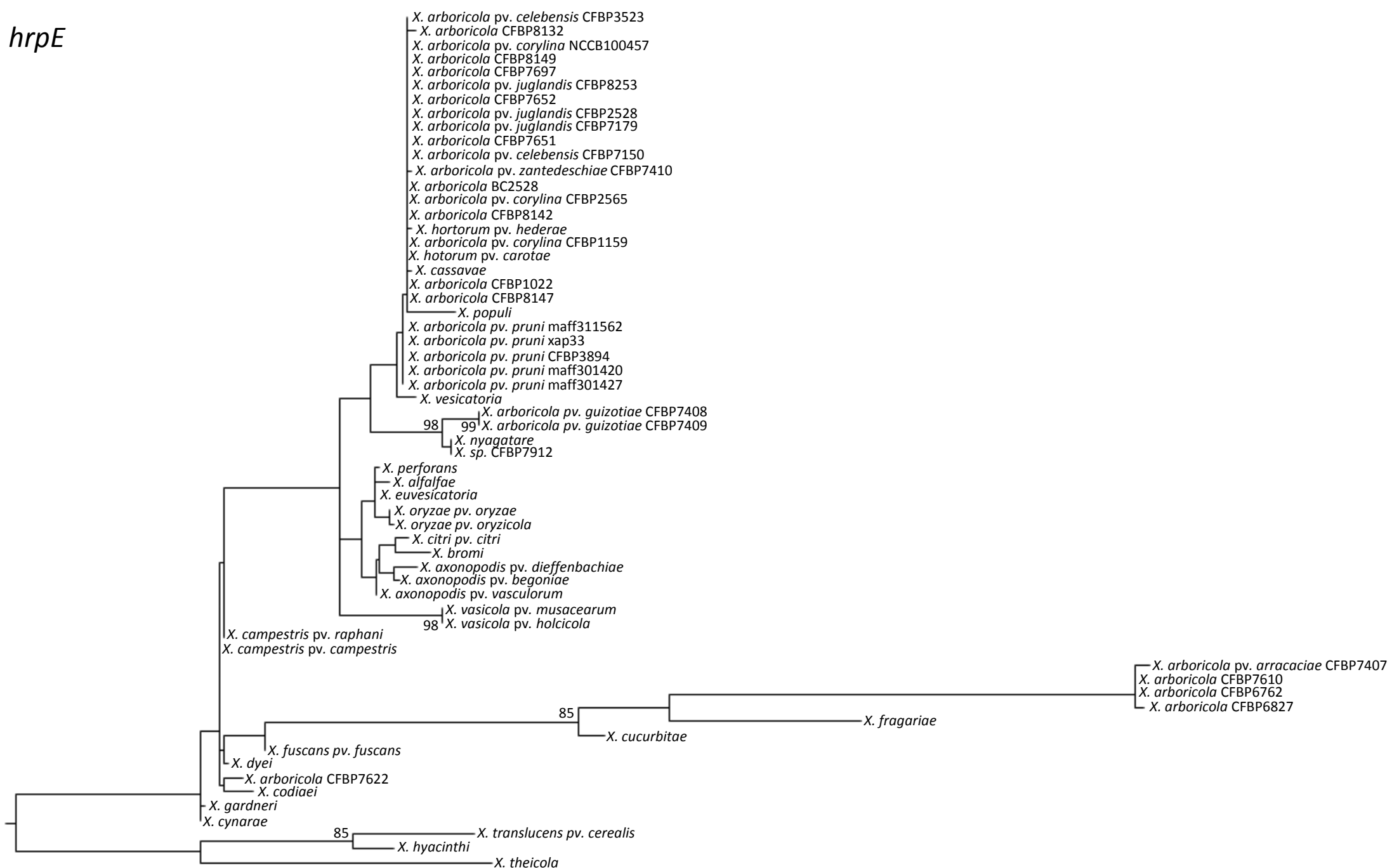
X. arboricola pv.
juglandis CFBP7179



Supplementary Figure S3

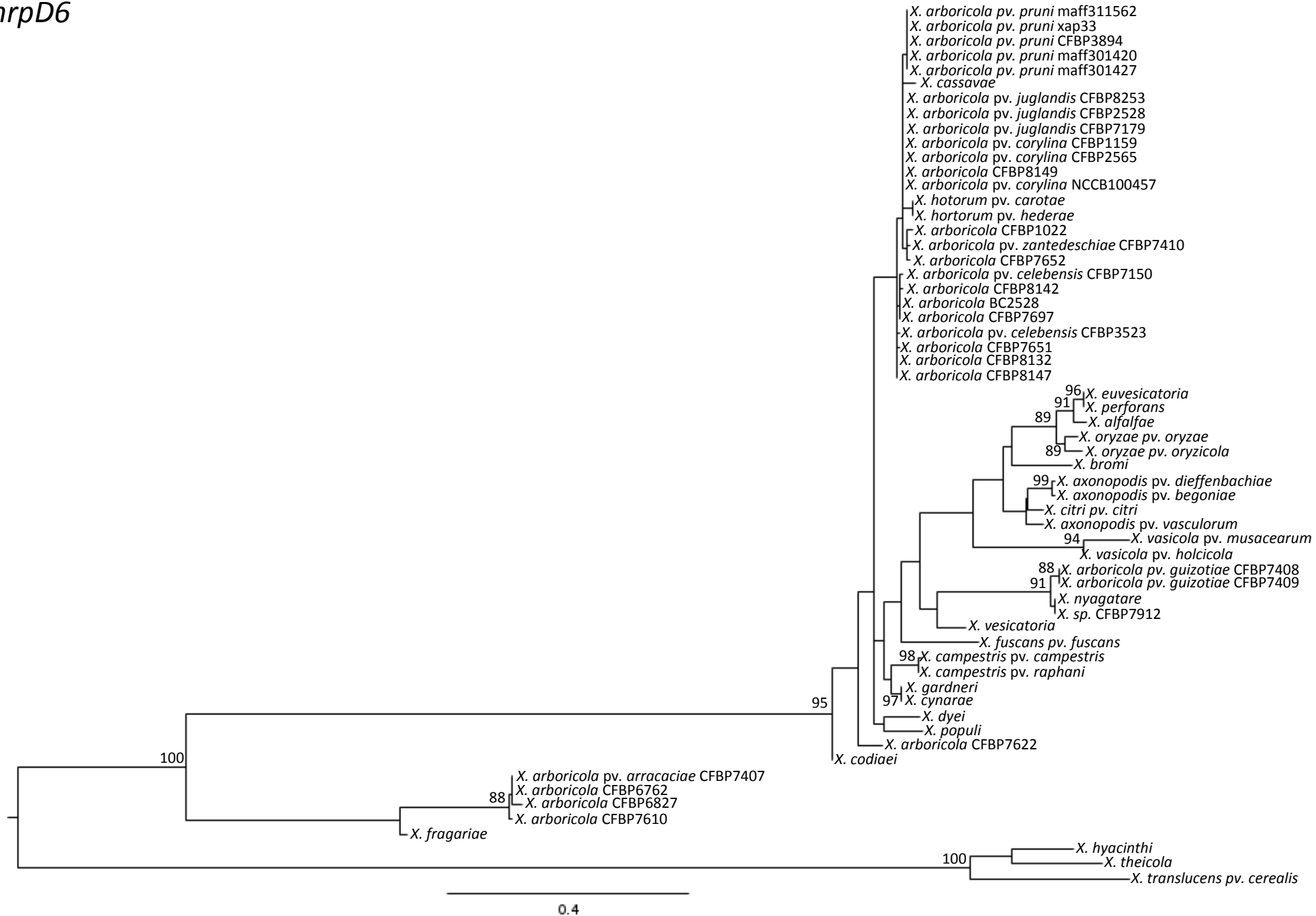
Supplementary Figure S3: individual maximum likelihood phylogenies built for each *hrc/hrp* gene coding for the T3SS. Bootstrap scores (1000 bootstraps) higher than 85% are displayed at each node.

hrpE

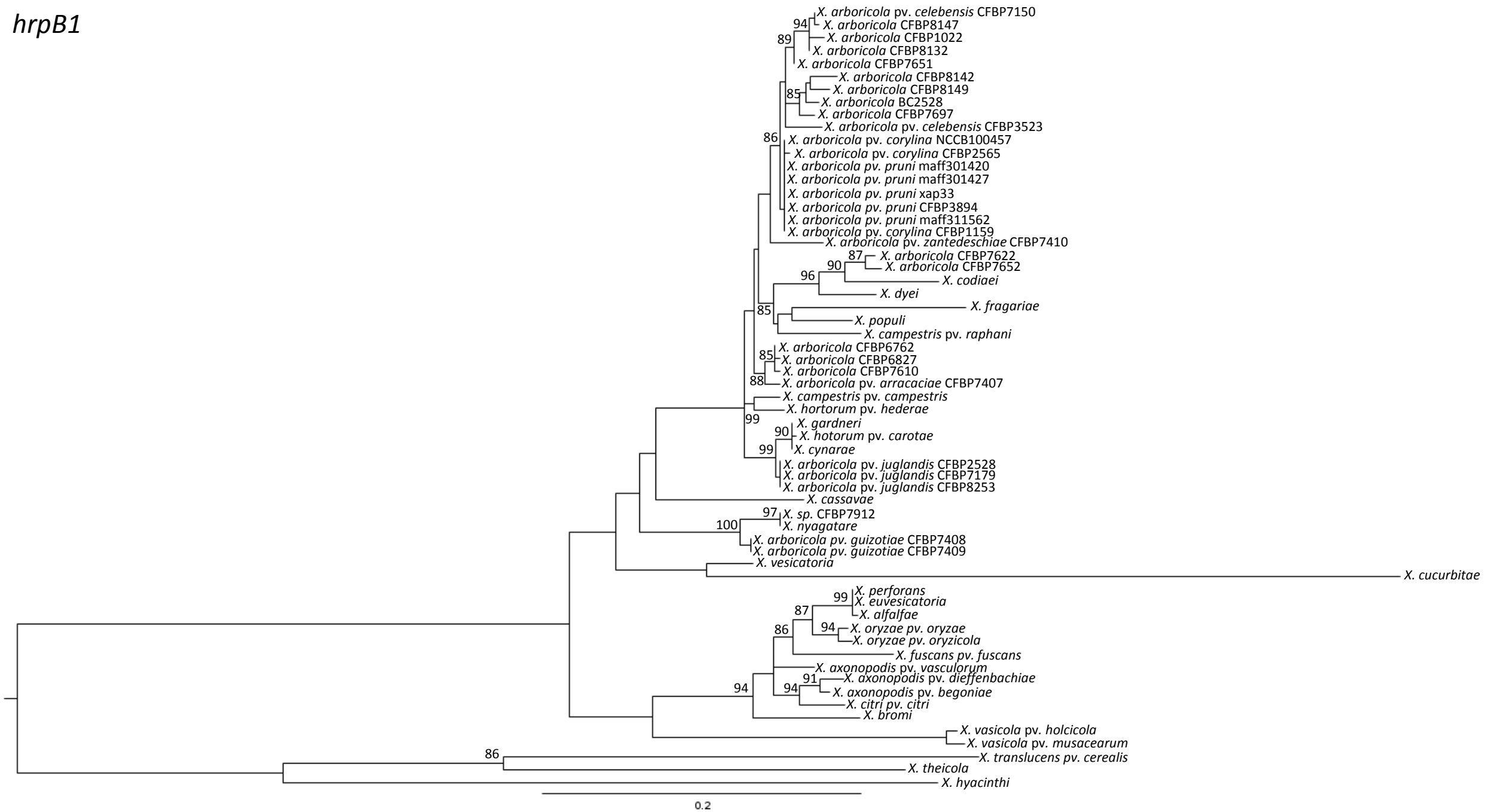


0.3

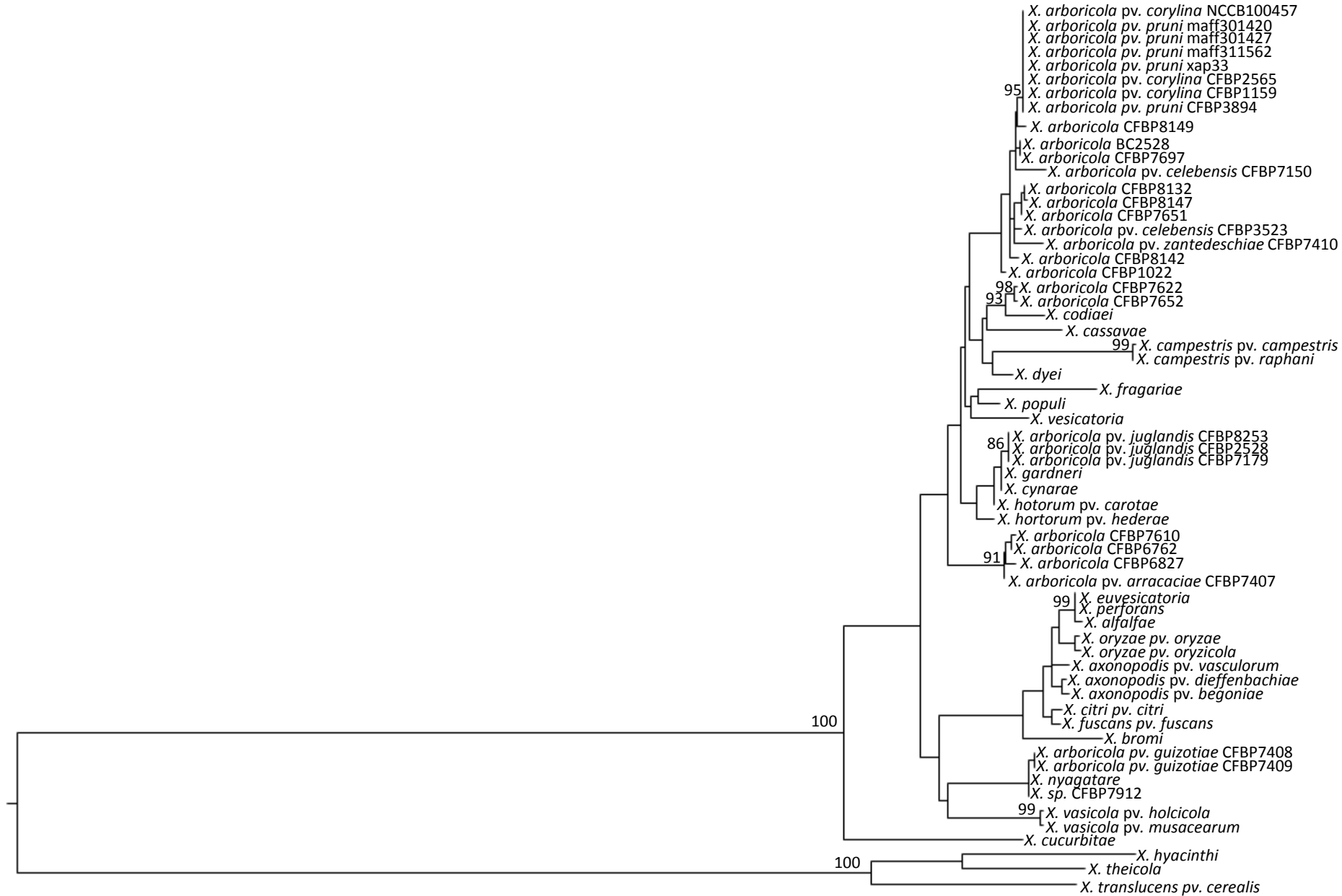
hrpD6



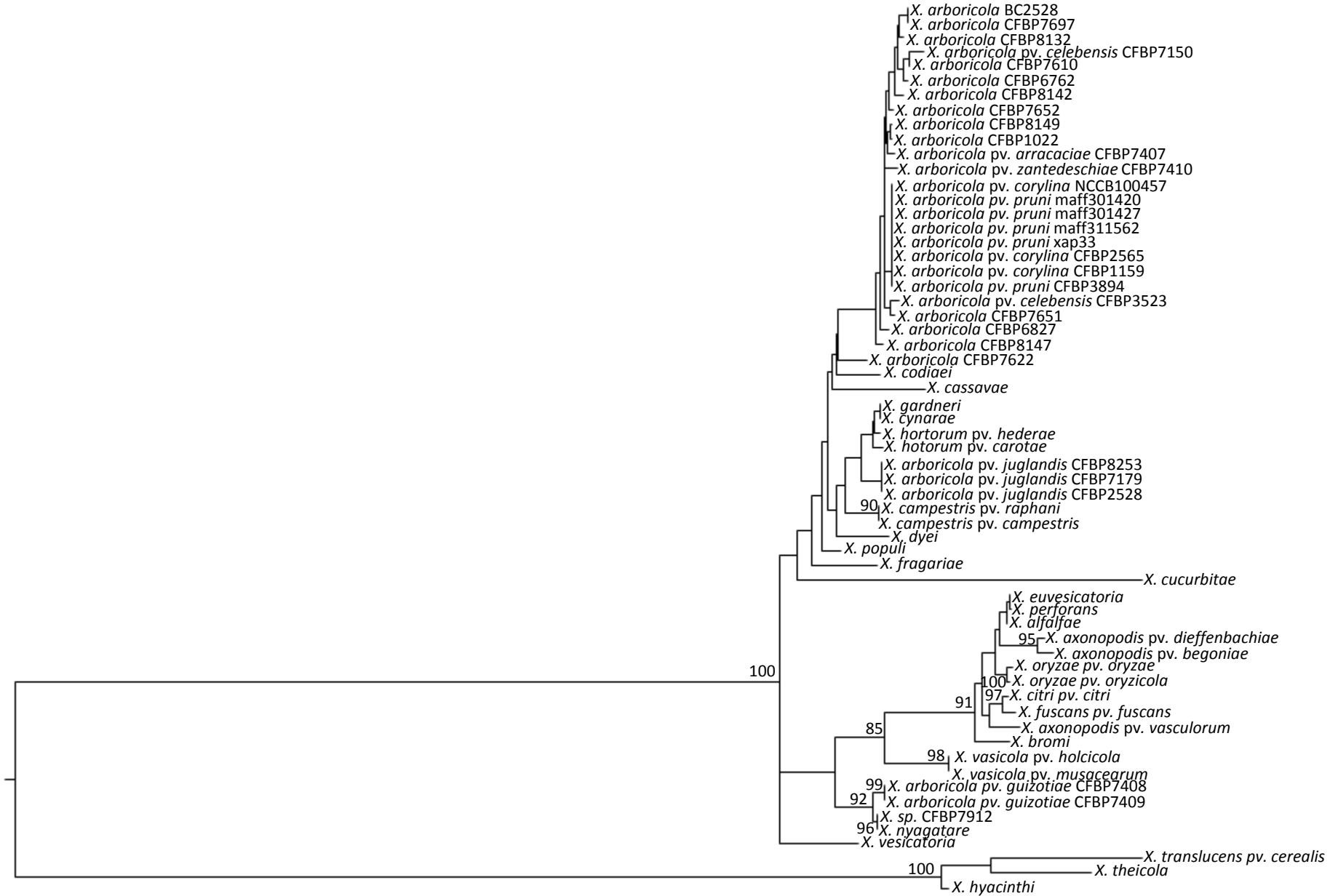
hrpB1



hrpB2

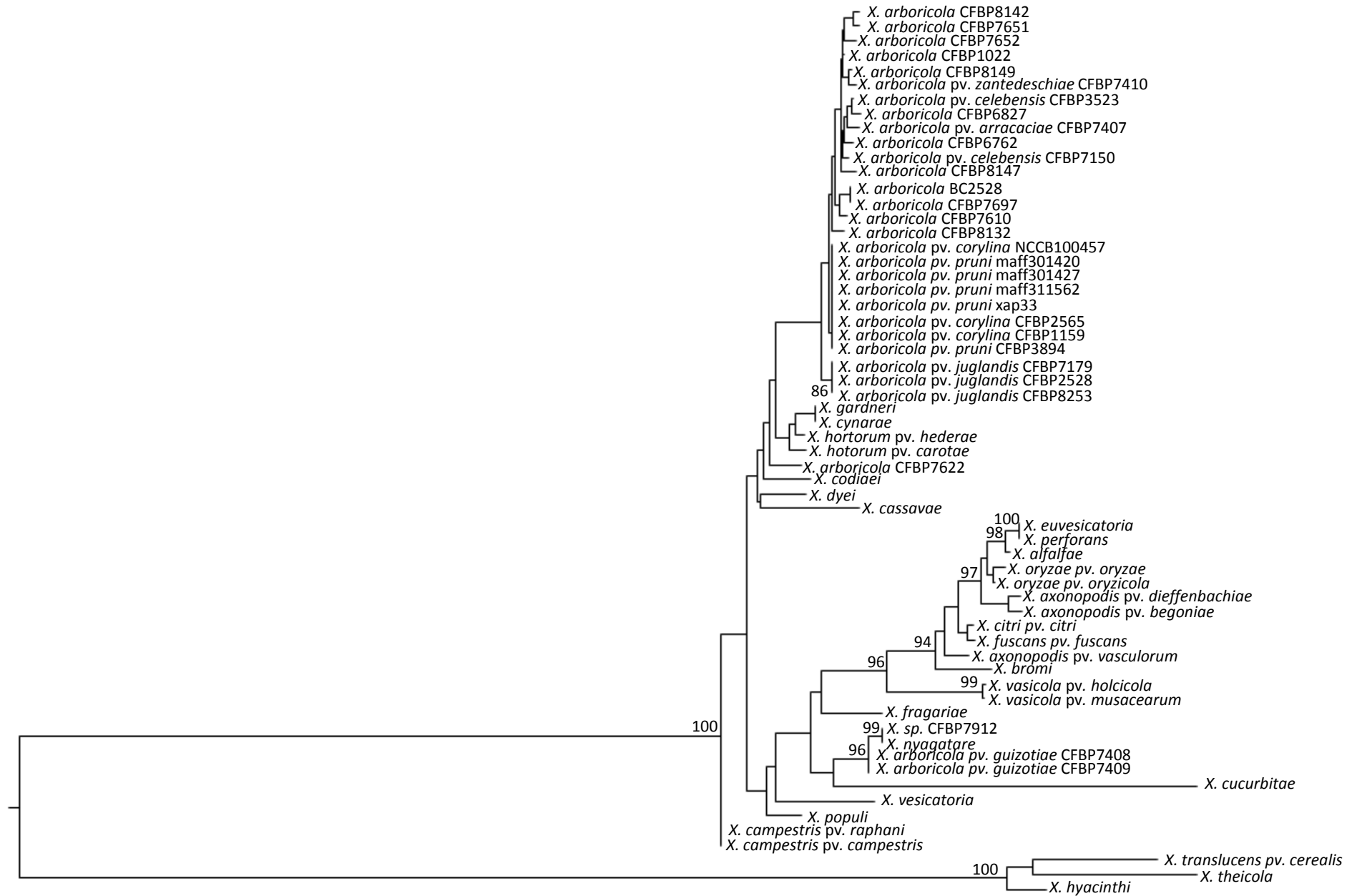


hrpB4



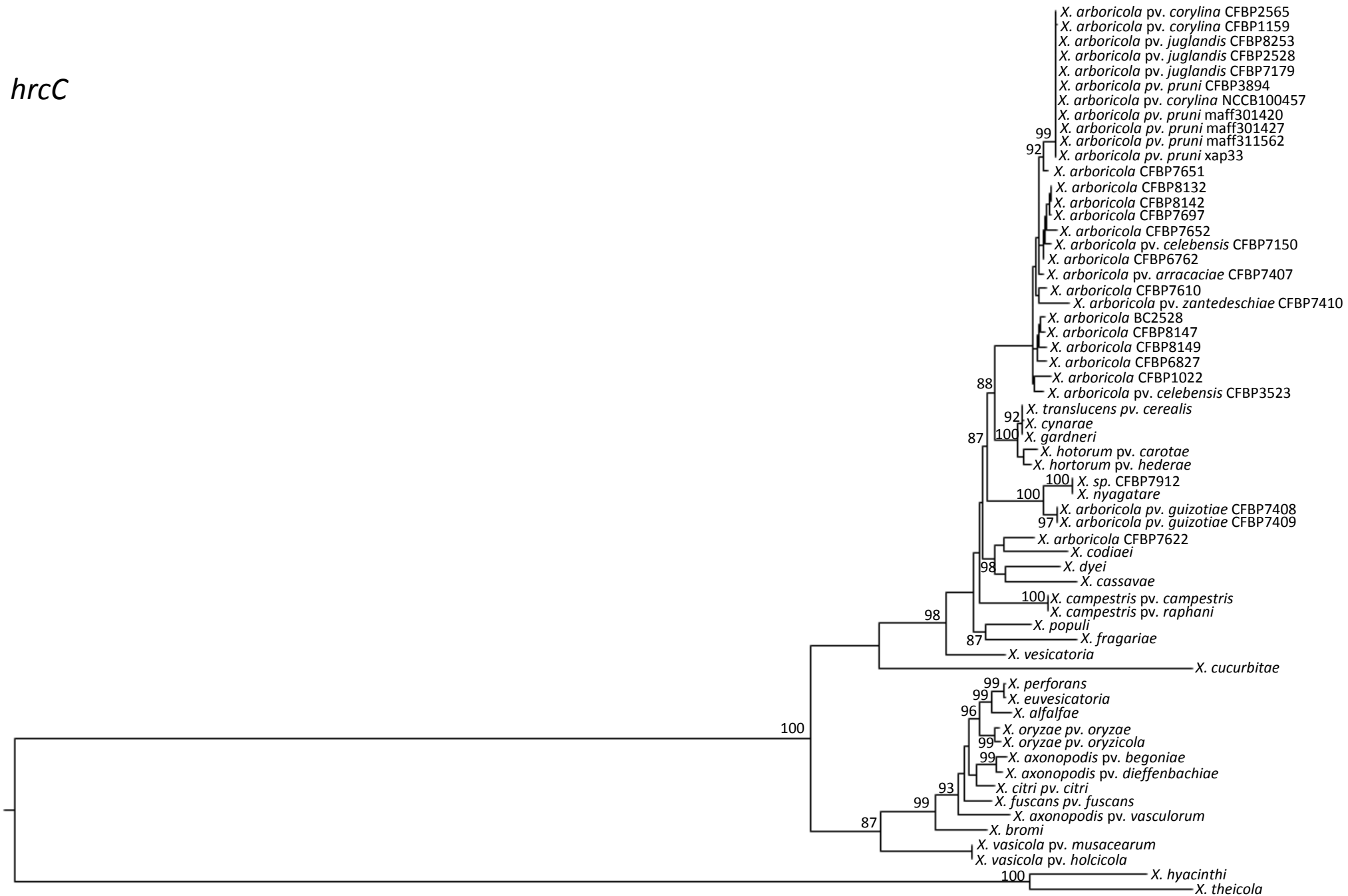
0.4

hrpB7



0.4

hrcC

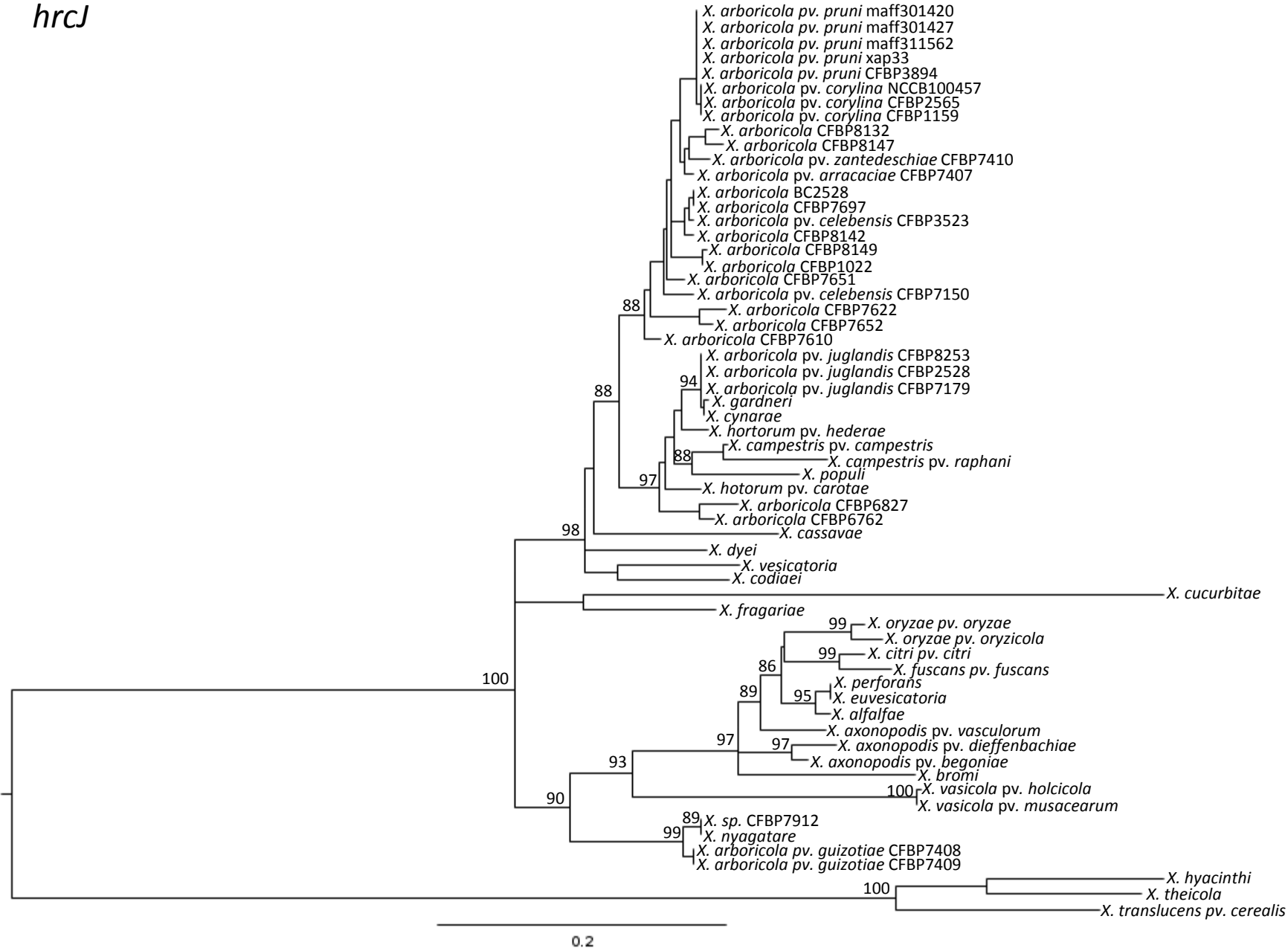


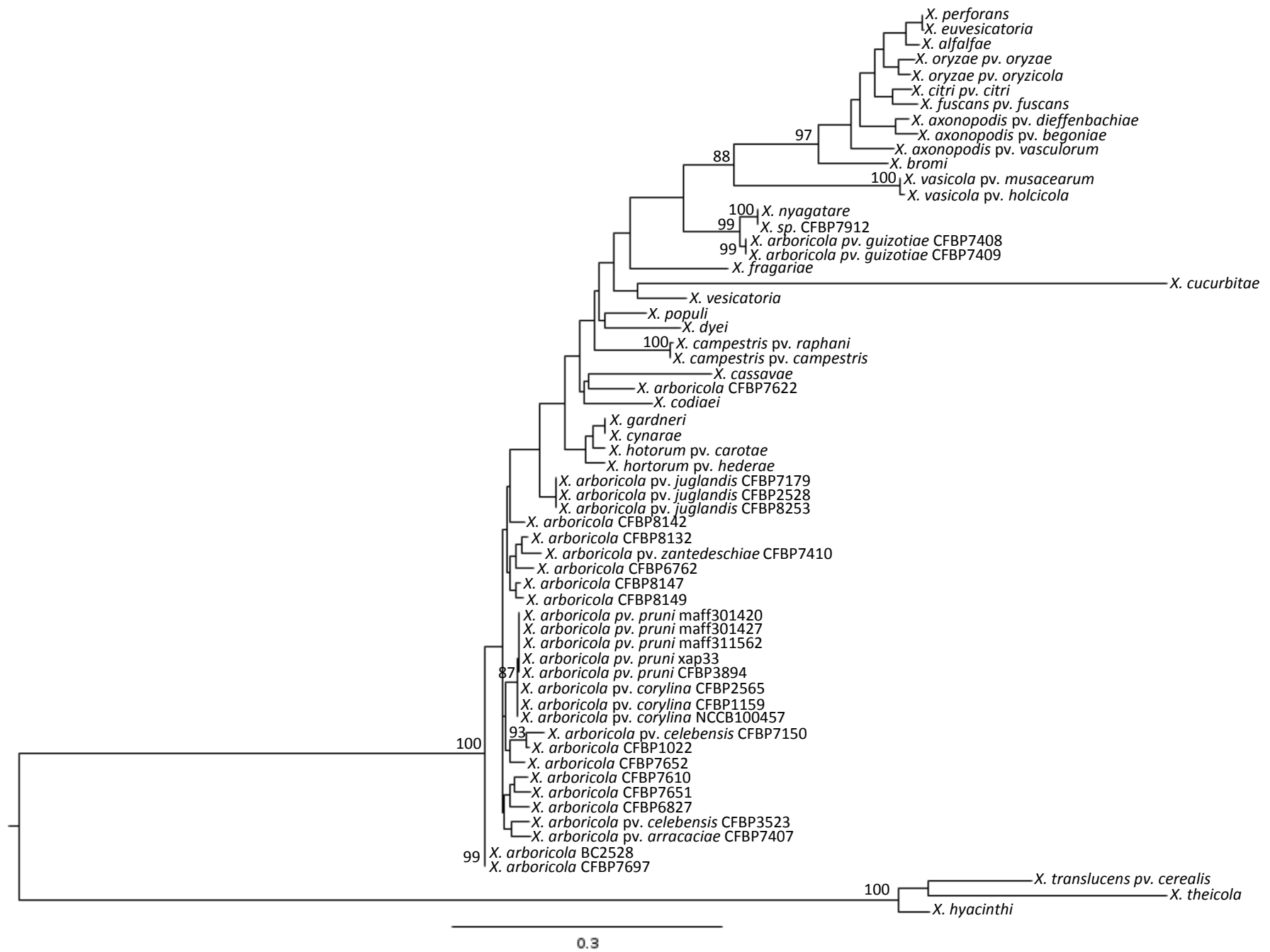
0.3

hrcD



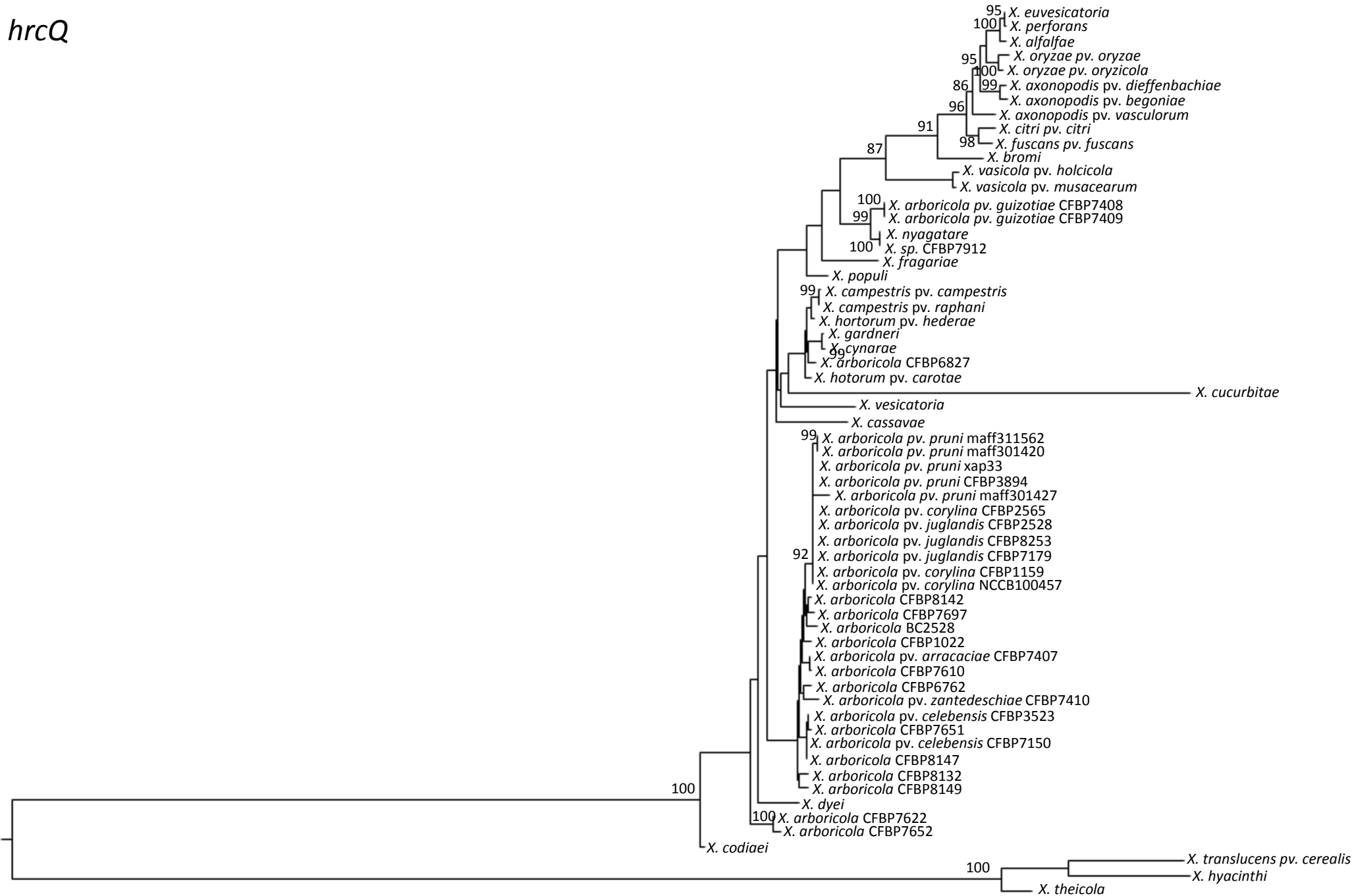
hrcJ





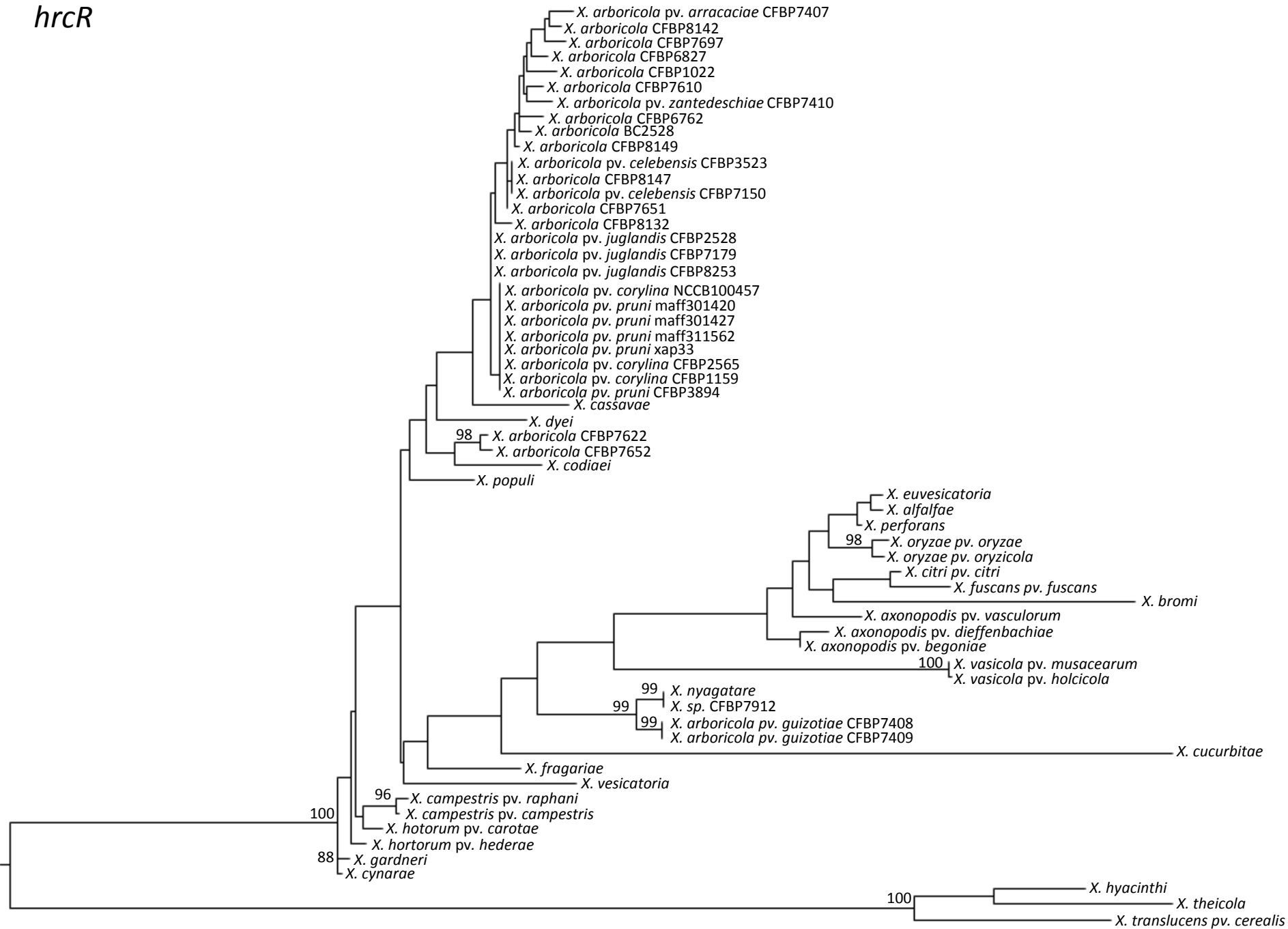
0.3

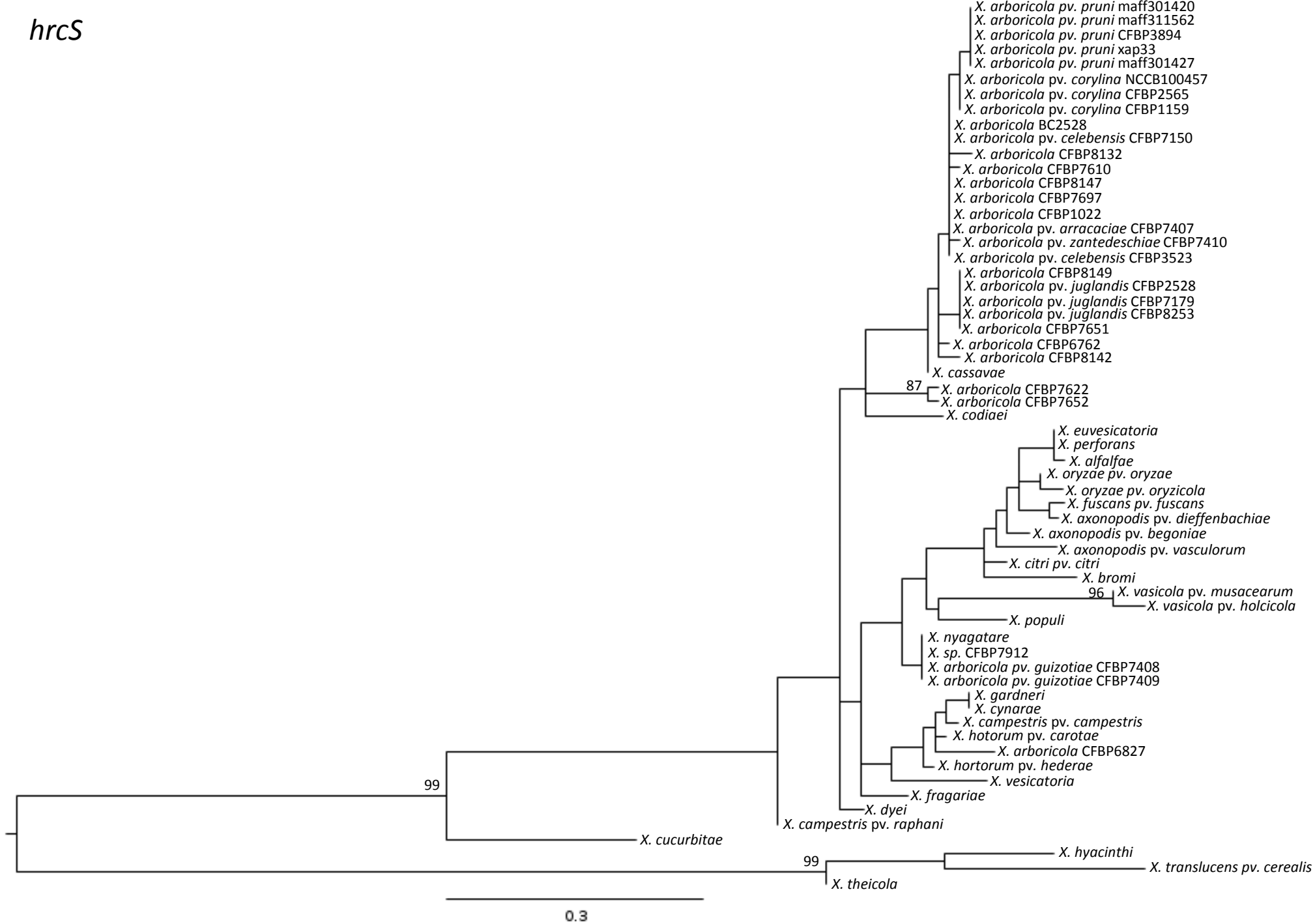
hrcQ



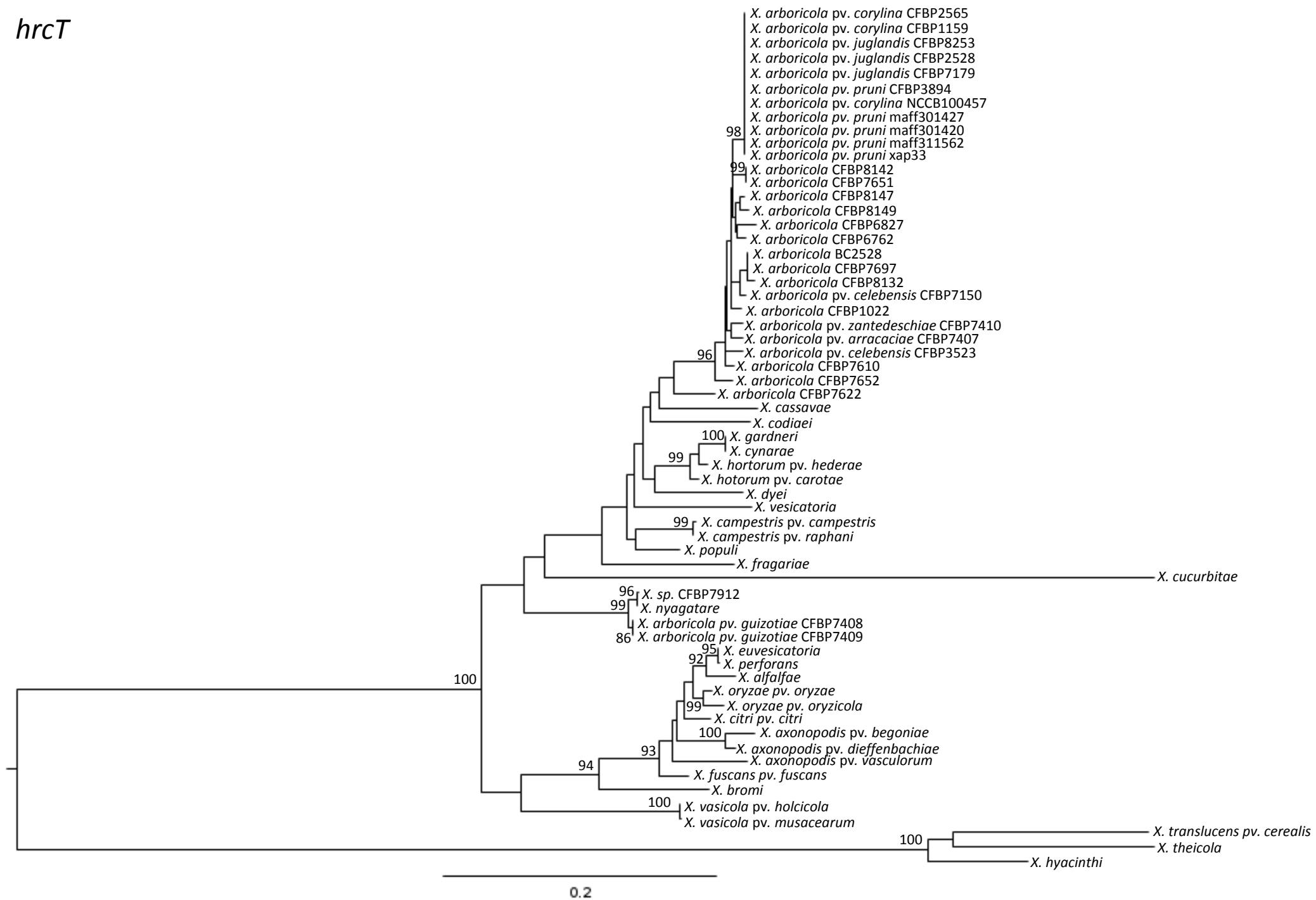
0.5

hrcR

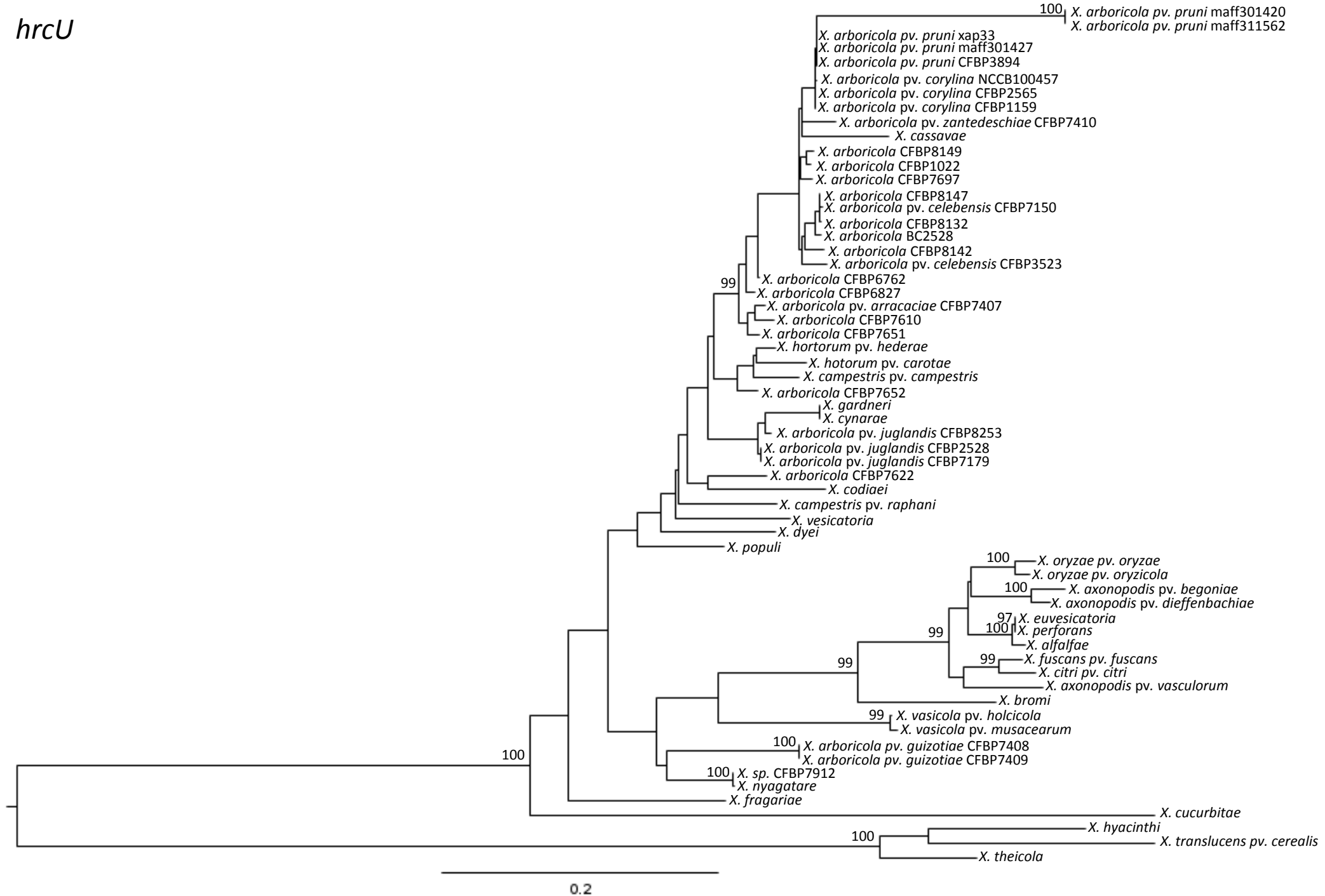




hrcT



hrcU

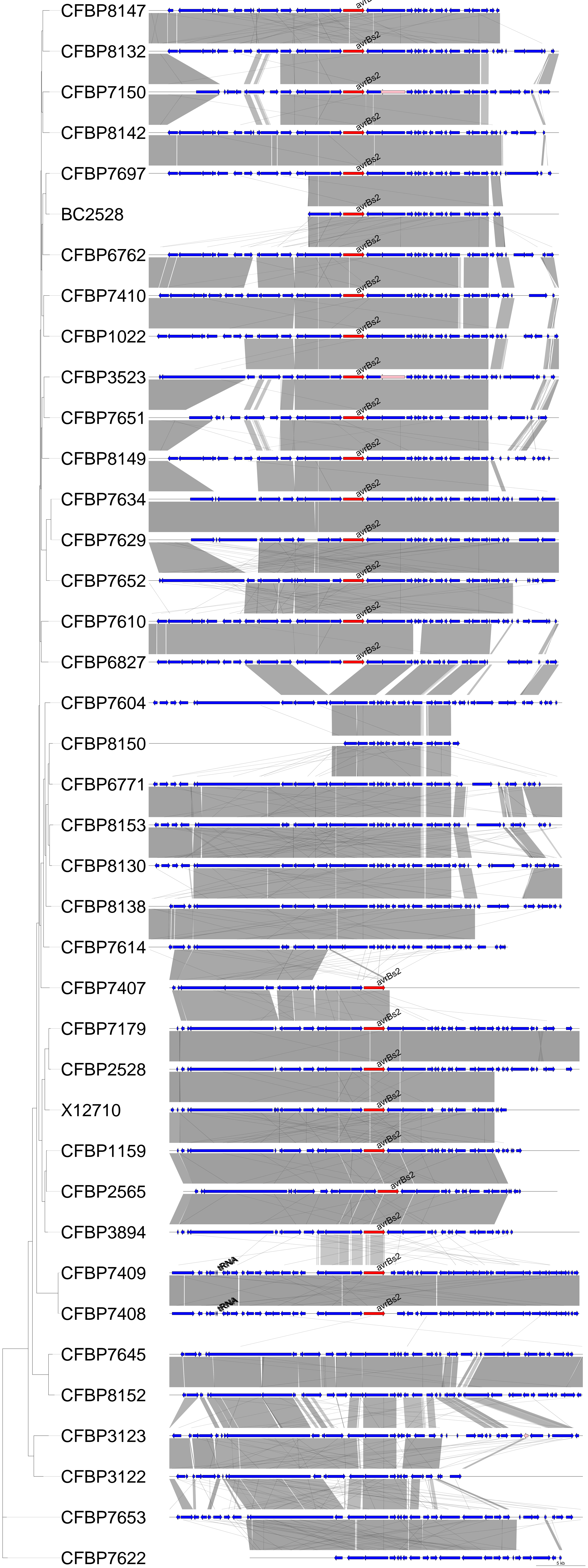


hrcV



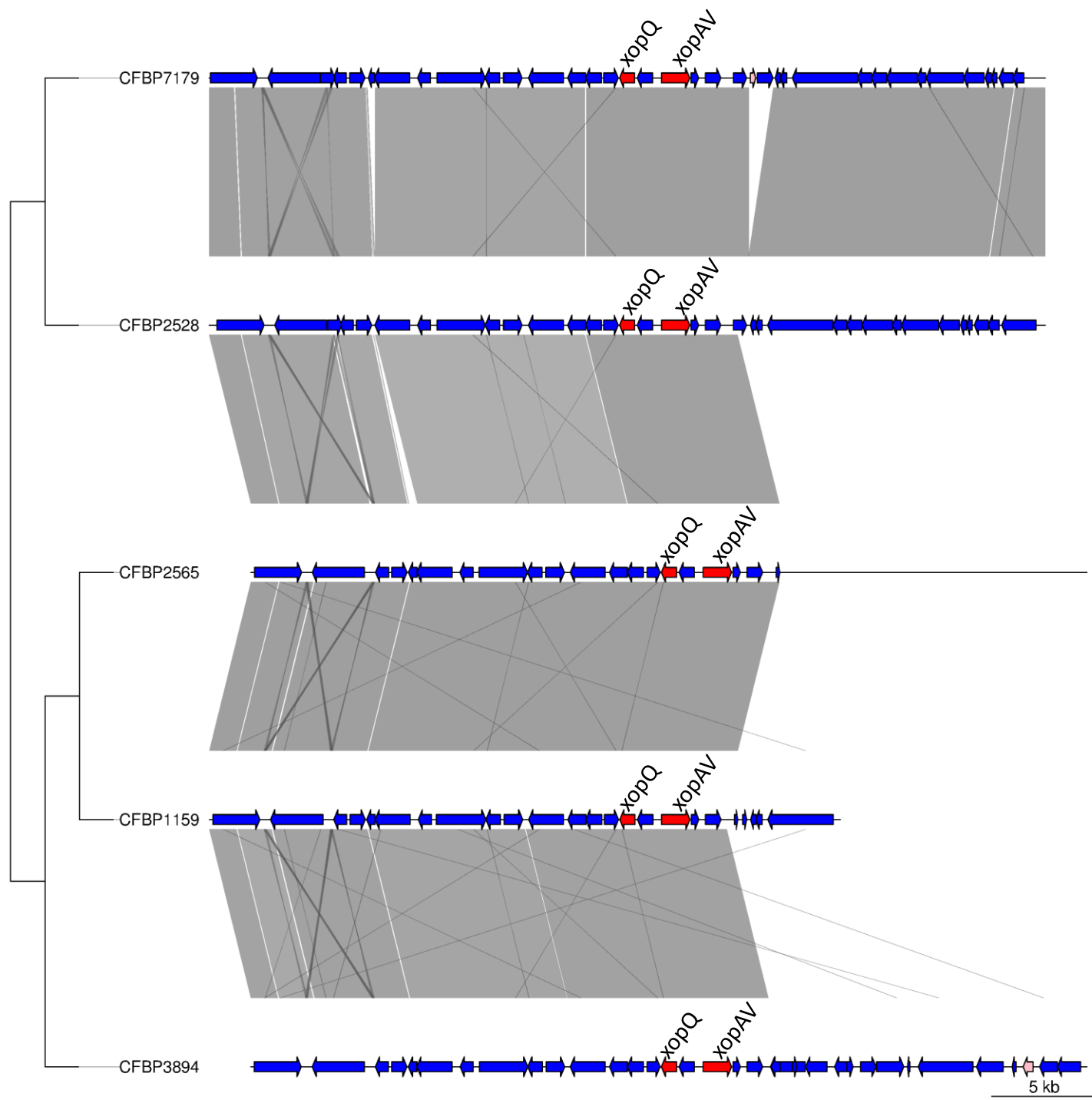
Supplementary Figure S4

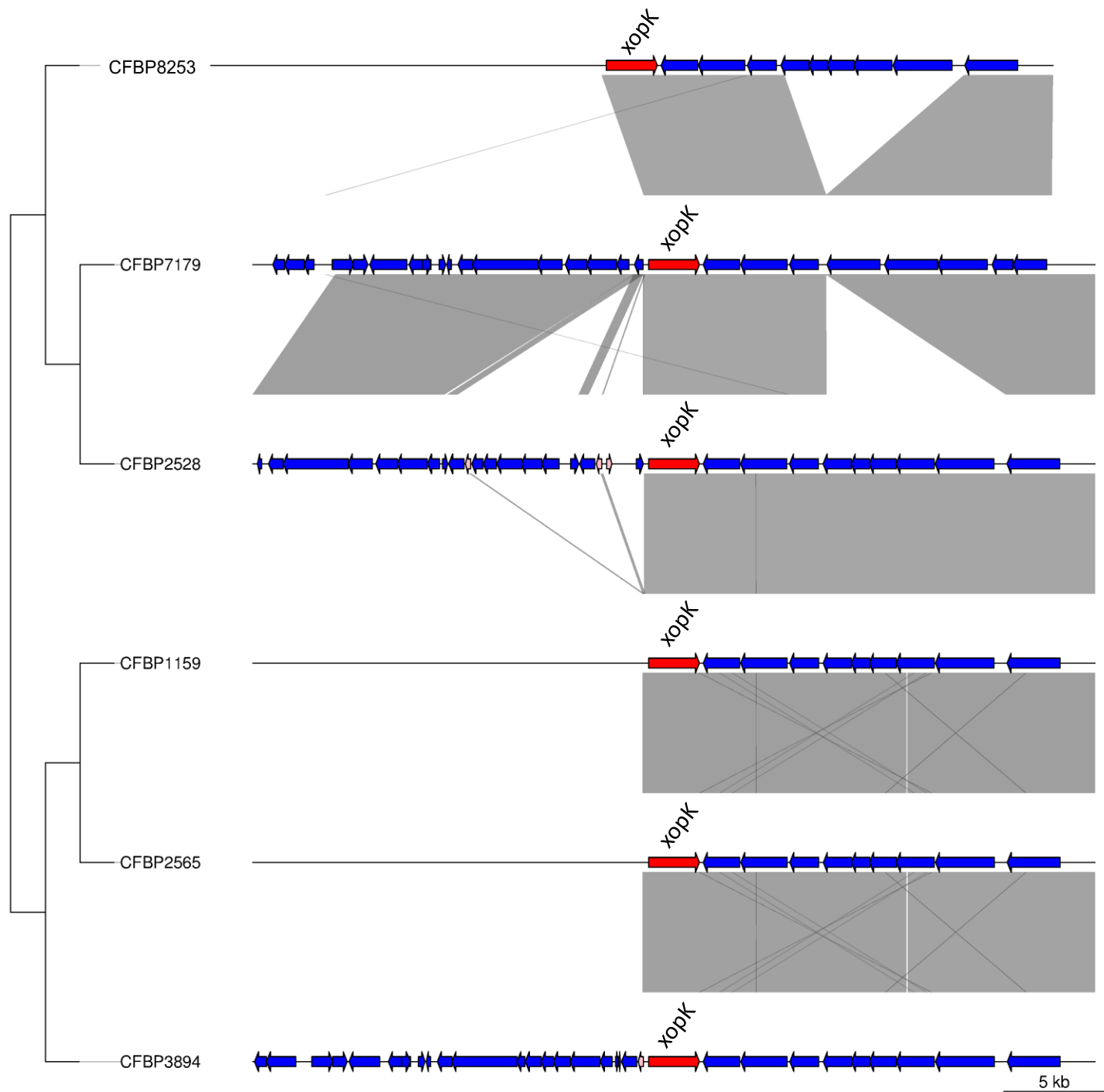
Supplementary Figure S4: Genomic environments of *avrBs2* in *Xanthomonas arboricola* strains. On the left, the dendrogram corresponds to the phylogenetic relationship between strains inferred from the core genome. On the right, genomic environments of *avrBs2* insertion site, within a window of 20 kb upstream and downstream of *avrBs2*. Red arrows represent *avrBs2*. Pink arrows represent genes encoding transposases. Blue arrows represent other genes within genomic environments. DNA fragments sharing similarities are connected with gray shading.

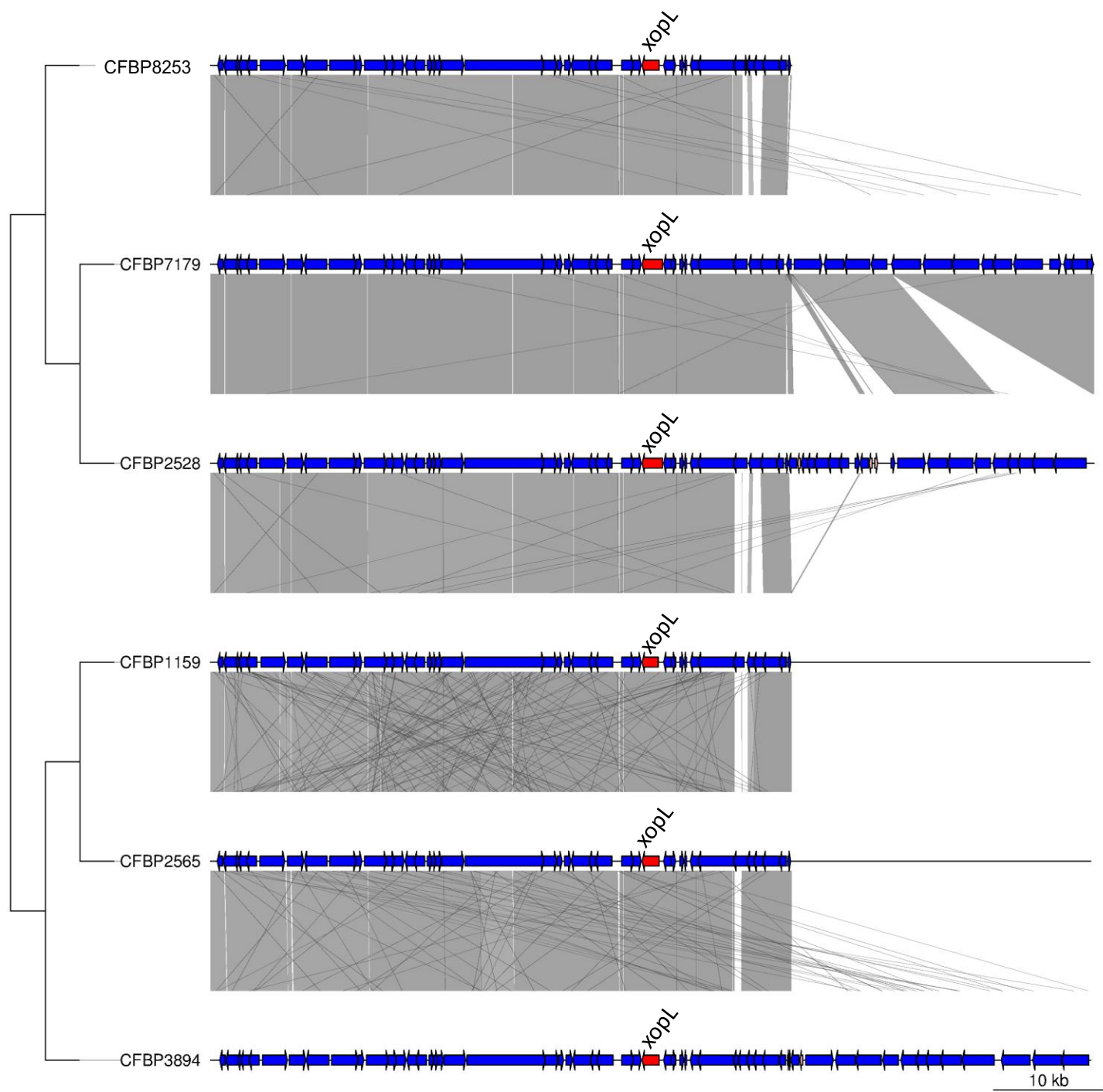


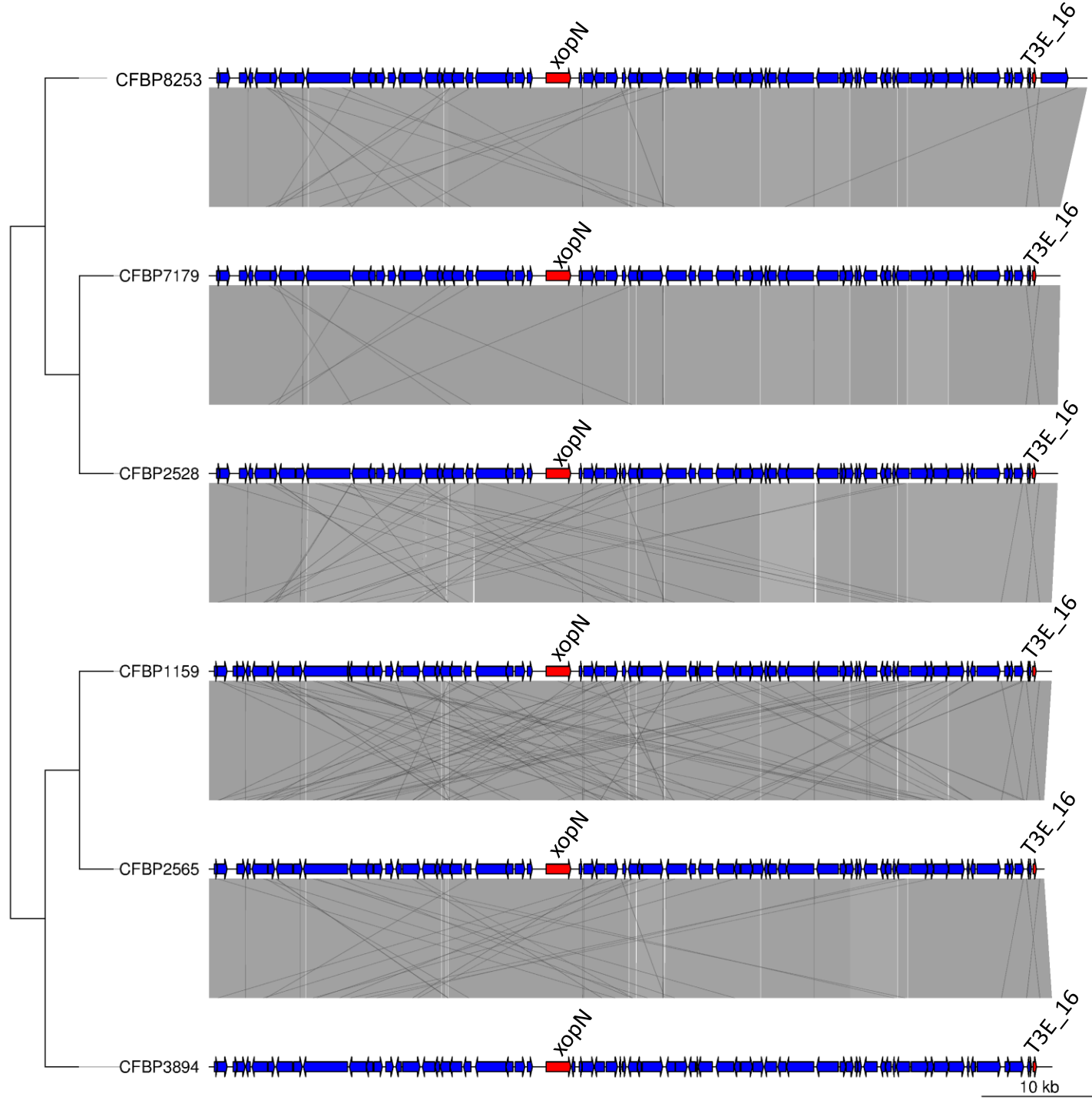
Supplementary Figure S5

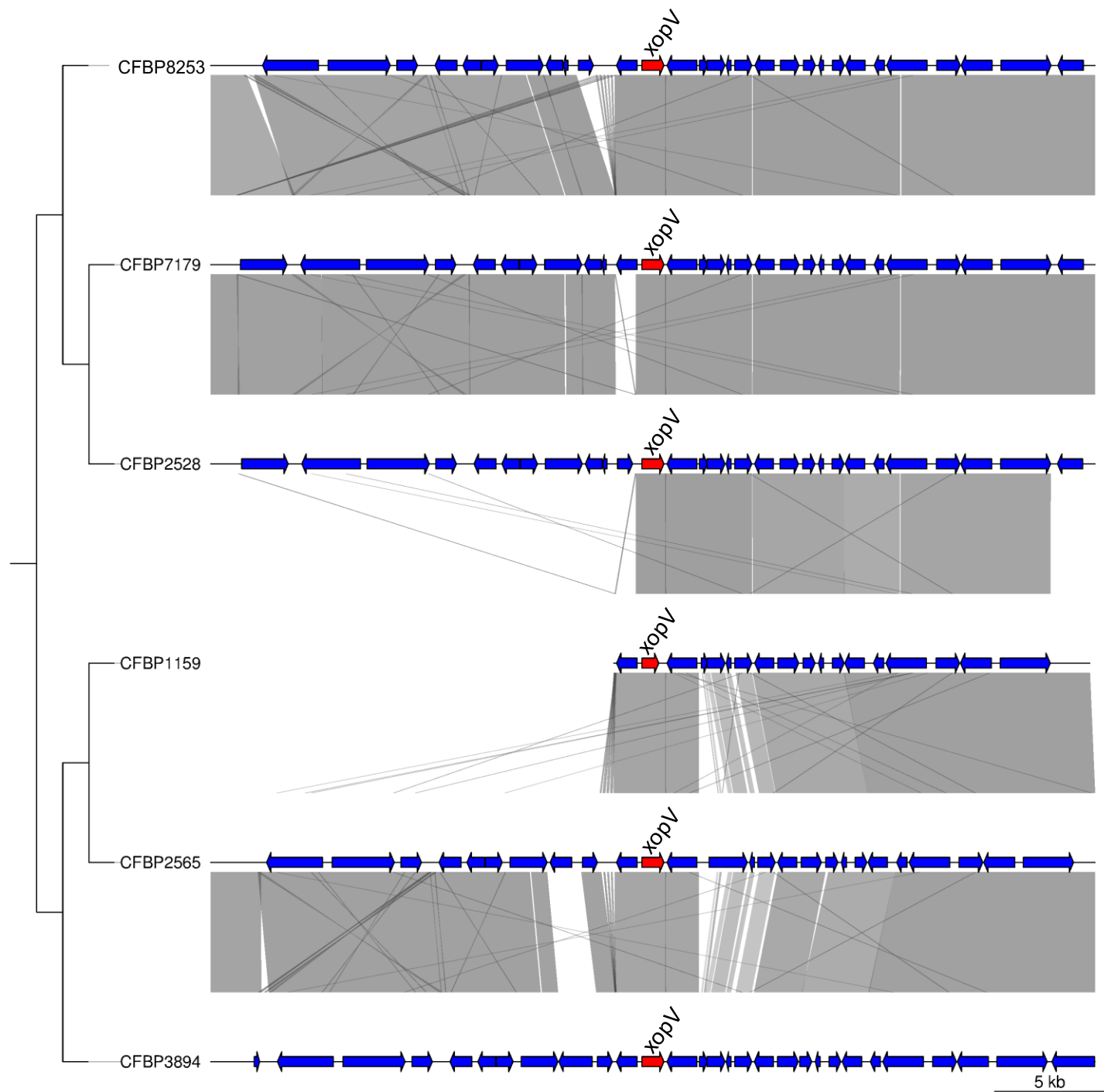
Supplementary Figure S5: Genomic environments of 10 predicted type three effector (T3E) genes specific to *Xanthomonas arboricola* group A strains. On the left, the dendrogram corresponds to the phylogenetic relationship between strains. On the right, genomic environments of predicted T3E genes, within a window of 20 kb upstream and downstream of the gene, are represented. Red arrows represent predicted T3E genes. Pink arrows represent genes encoding transposases. Green arrows represent phages. Blue arrows represent other genes within genomic environments. DNA fragments sharing similarities are connected with gray shading.

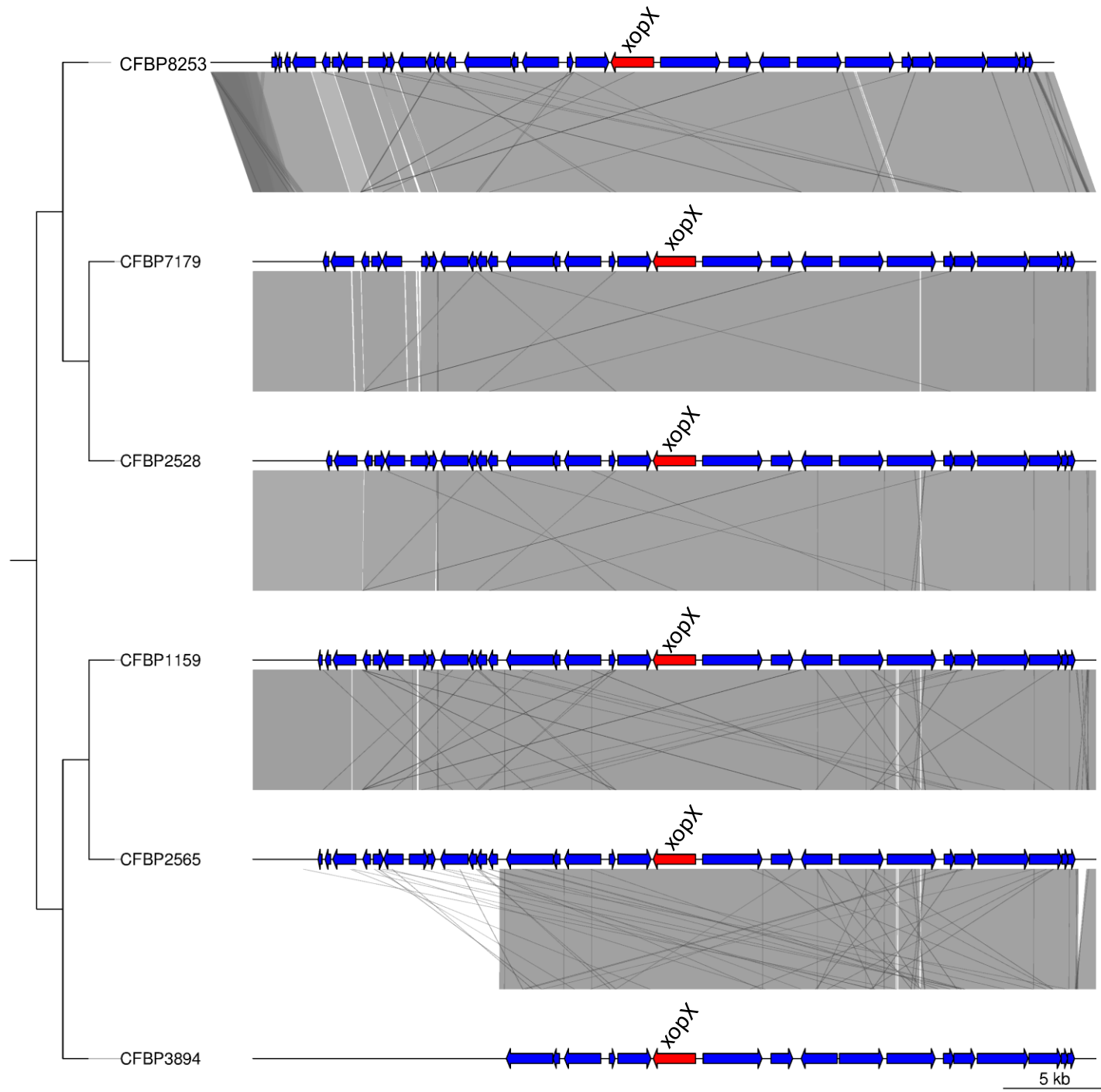


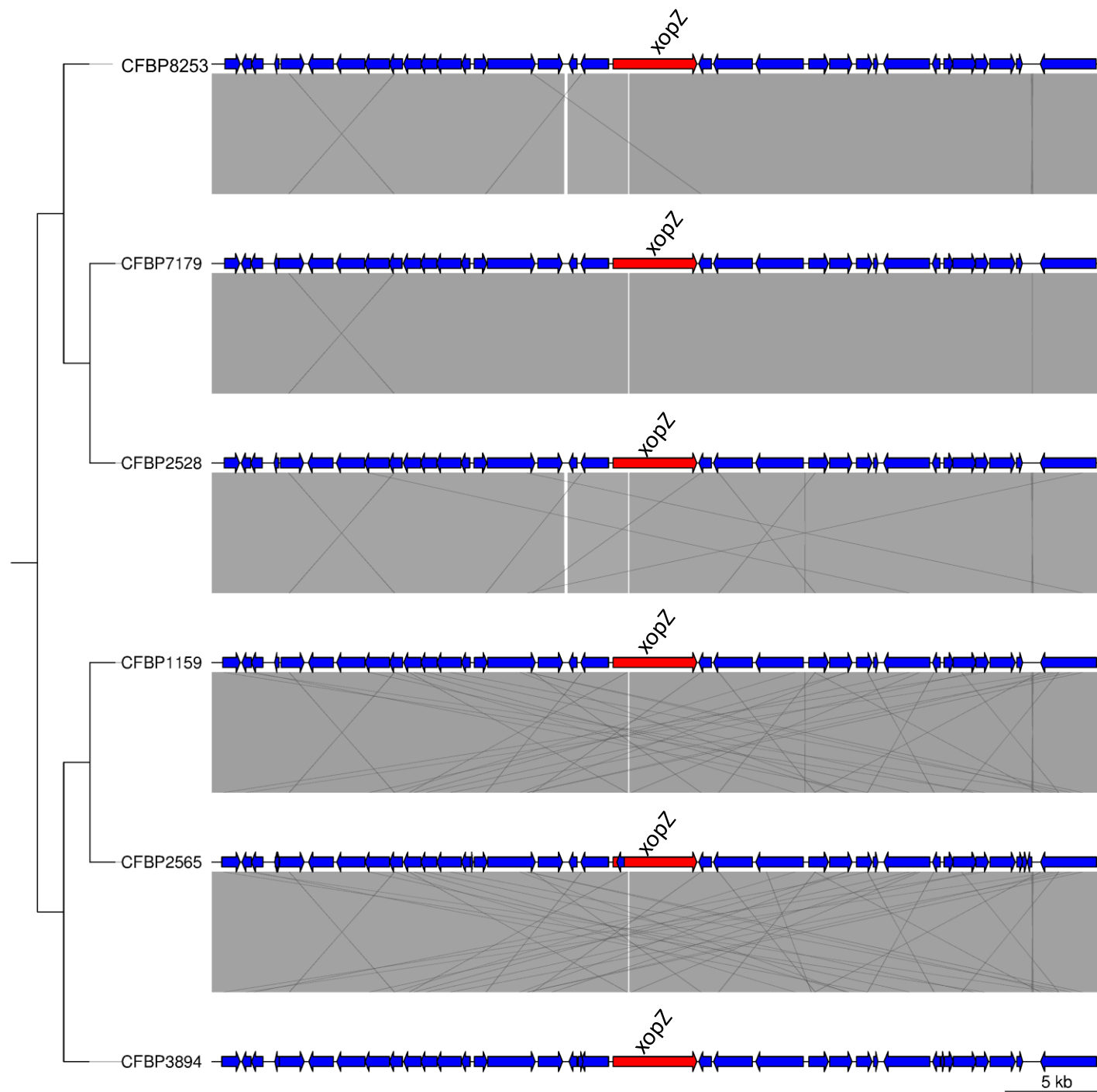












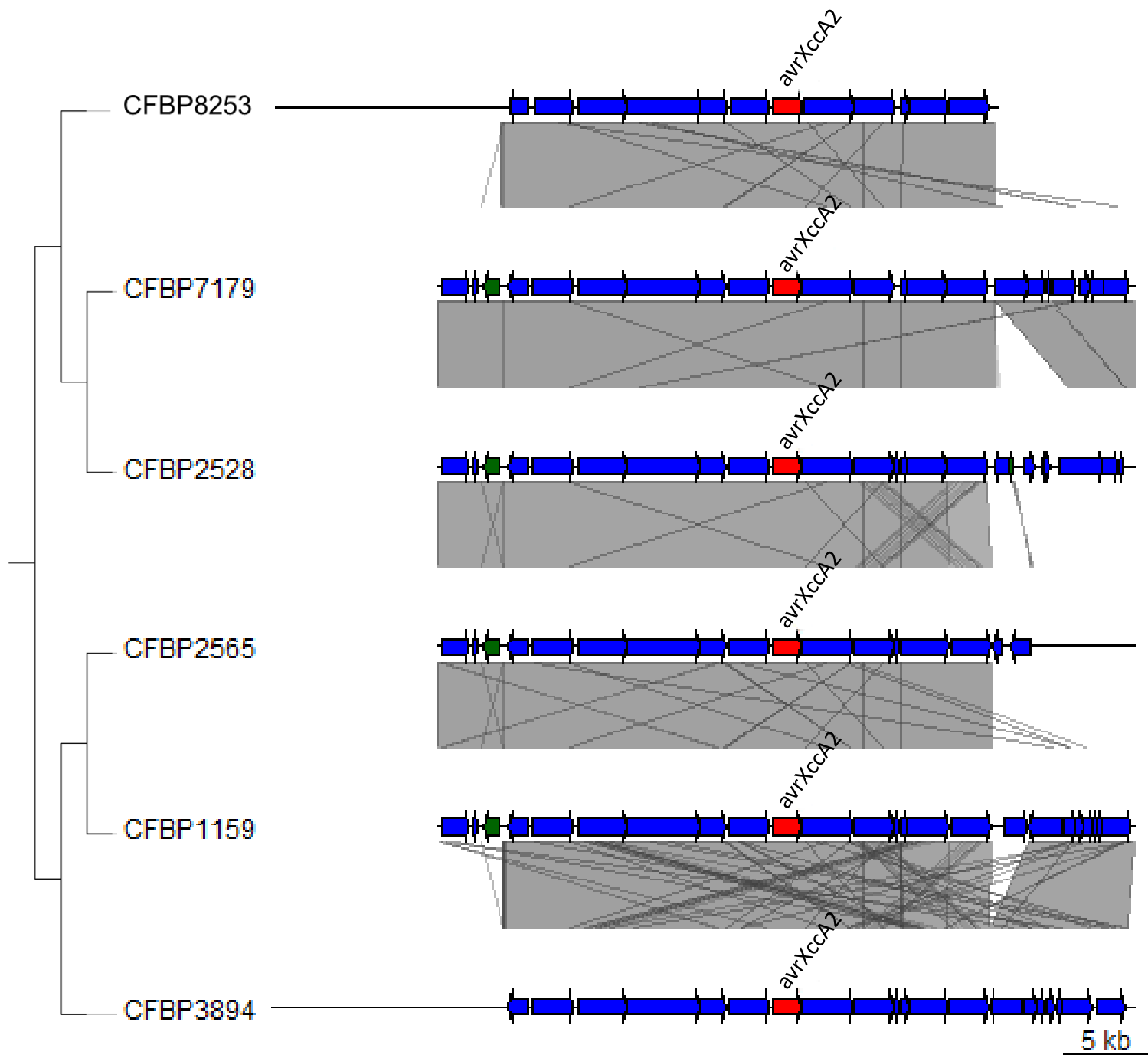


Table S1 : Results of Shimodaira-Hasegawa tests comparing phylogenies of T3SS coding genes in *Xanthomonas*

		H1 : topologies alternatives															
		<i>hrc_concat</i>	<i>hrcC</i>	<i>hrcD</i>	<i>hrcJ</i>	<i>hrcN</i>	<i>hrcQ</i>	<i>hrcR</i>	<i>hrcS</i>	<i>hrcT</i>	<i>hrcU</i>	<i>hrcV</i>	<i>hrpE</i>	<i>hrpB1</i>	<i>hrpB2</i>	<i>hrpB4</i>	<i>hrpB7</i>
HO: alignement	<i>hrc_concat</i>	0,96034	0	0,88536	0,04023	0	0,00083	0,16132	0,00578	0,20897	0,0206	0,00684	0,02441	0,03873	0,11108	0,3825	0,01626
	<i>hrcC</i>	0,37406	0,91395	0,34218	0,00832	0,52414	0,72117	0,65727	0,00259	0,26576	0,0002	0,46417	0	0,94543	0,89079	0,82633	0,7081
	<i>hrcD</i>	0,00025	0	0,94864	0,43053	0	0,04018	0,00005	0	0,00233	0	0,12352	0	0	0,41333	0,3802	0,00036
	<i>hrcJ</i>	0,05185	0	0	0,92988	0,07593	0	0,02137	0,04262	0	0,29572	0,43796	0	0,29101	0,73003	0,46092	0,02644
	<i>hrcN</i>	0	0,00001	0,00012	0,05916	0,91767	0,23817	0,03422	0,33393	0,0009	0,00322	0,27644	0	0,53221	0,42616	0,16445	0,09918
	<i>hrcQ</i>	0	0	0,85718	0,94507	0,9462	0,95422	0,69799	0,1154	0	0,08249	0,8301	0	0,76662	0,75029	0,90246	0,78262
	<i>hrcR</i>	0	0,01155	0,19194	0,35578	0,66565	0,28707	0,90366	0,27447	0,73229	0,08124	0,21551	0	0,32863	0,97158	0,02542	0,55015
	<i>hrcS</i>	0	0	0,00005	0,00029	0,00041	0,05747	0,0254	0,76005	0,0018	0,11881	0,01921	0	0,00016	0,00014	0,00008	0,0001
	<i>hrcT</i>	0	0	0,03912	0,00017	0,00778	0,06186	0,0903	0,17226	0,90986	0,30794	0,51963	0,00227	0,66663	0,19137	0,80863	0,72965
	<i>hrcU</i>	0	0	0,00026	0,03987	0,00043	0	0,00025	0,92947	0,00792	0,8608	0,03076	0,92051	0,62439	0,20338	0,0045	0,00021
	<i>hrcV</i>	0	0	0,46891	0,13923	0,06054	0,03686	0,04388	0,49159	0,00934	0,00053	0,94393	0,05171	0,02328	0,00002	0,91582	0,46595
	<i>hrpE</i>	0	0,00187	0,01405	0	0,02687	0,00386	0	0,00001	0,00011	0,02048	0	0,8257	0,01079	0,08082	0	0,00646
	<i>hrpB1</i>	0	0,00572	0,03533	0,48392	0,2558	0,04064	0,30653	0,00634	0,00264	0,19341	0,01187	0	0,92934	0,11034	0,50696	0,68073
	<i>hrpB2</i>	0	0	0,06997	0,956	0,00082	0,02319	0,70404	0,00021	0,00054	0,92429	0,97765	0,03508	0,00551	0,96611	0,9265	0,62853
	<i>hrpB4</i>	0	0	0	0,95477	0,28745	0,00553	0,03089	0	0	0,02795	0,00873	0	0,44886	0,0748	0,86887	0,29673
	<i>hrpB7</i>	0	0	0,00349	0,00318	0,38469	0,6936	0,81701	0,90497	0,02205	0,56477	0,00167	0	0,67116	0,94118	0,01916	0,9435