

1 **Supporting information**

2 **Table S1.** Metadata for all isolates included in this study.

3 **Table S2.** Assembly statistics for newly sequenced *X. fastidiosa* subsp. *fastidiosa* isolates from  
4 Site1 and Site2.

5 **Table S3.** Genes unique to Southeast USA PD-I and PD-III cluster showing all isolates in the  
6 source and descendant populations.

7 **Table S4.** EggNOG-mapper and interproscan5 results for unique genes listed in Table 1.

8 **Table S5.** Recombinant genes within Southeast USA/Taiwan and California/Spain.

9 **Table S6.** Fst and MKT among PD-causing populations.

10

11 **Figure S1. Non-recombinant Maximum Likelihood (ML) tree showing phylogenetic**  
12 **grouping of worldwide PD-causing subsp. *fastidiosa* isolates.** Color represents isolates from  
13 the same geographical location: California (Red), Texas (Pink), Georgia (Green), North Carolina  
14 (Dark green), Florida (Yellow), Spain (Light blue), and Taiwan (Dark blue). PD-causing strains  
15 have been divided into three phylogenetically supported clades (PD-I, PD-II, PD-III). **b.**

16 Maximum likelihood (ML) tree of PD-causing subsp. *fastidiosa* isolates. Tree was built using the  
17 core genome alignment after removing recombinant segments detected in fastGEAR. Bootstrap  
18 values mark branch support. Arrows point towards the base of PD-causing clades (-I to -III).

19

20 **Figure S2. Maximum Likelihood trees (ML) showing the stochastic probability of gene**  
21 **gain/loss per tree branch within Southeast USA/Taiwan and California/Spain.** Branch

22 lengths are proportional to the total number of gain and lost events. Estimated number of gene  
23 gain (up, bolded) and loss (down, italics) events are shown for each tree branch. Color represents  
24 isolates from the same geographical location: California (Red), Texas (Pink), Georgia (Green),  
25 North Carolina (Dark green), Florida (Yellow), Spain (Light blue), and Taiwan (Dark blue).  
26 Trees were built using the core genome alignment without removing recombinant segments for  
27 the California/Spain and Southeast USA/Taiwan populations. Bootstrap values mark branch  
28 support. **a.** ML tree showing number of branch-associated gene gain/loss events within the  
29 California/Spain population; **b.** ML tree showing number of branch-associated gene gain/loss  
30 events within the Southeast USA/Taiwan population.

31

32 **Figure S3. Circle plot showing strain-specific recombination events within**

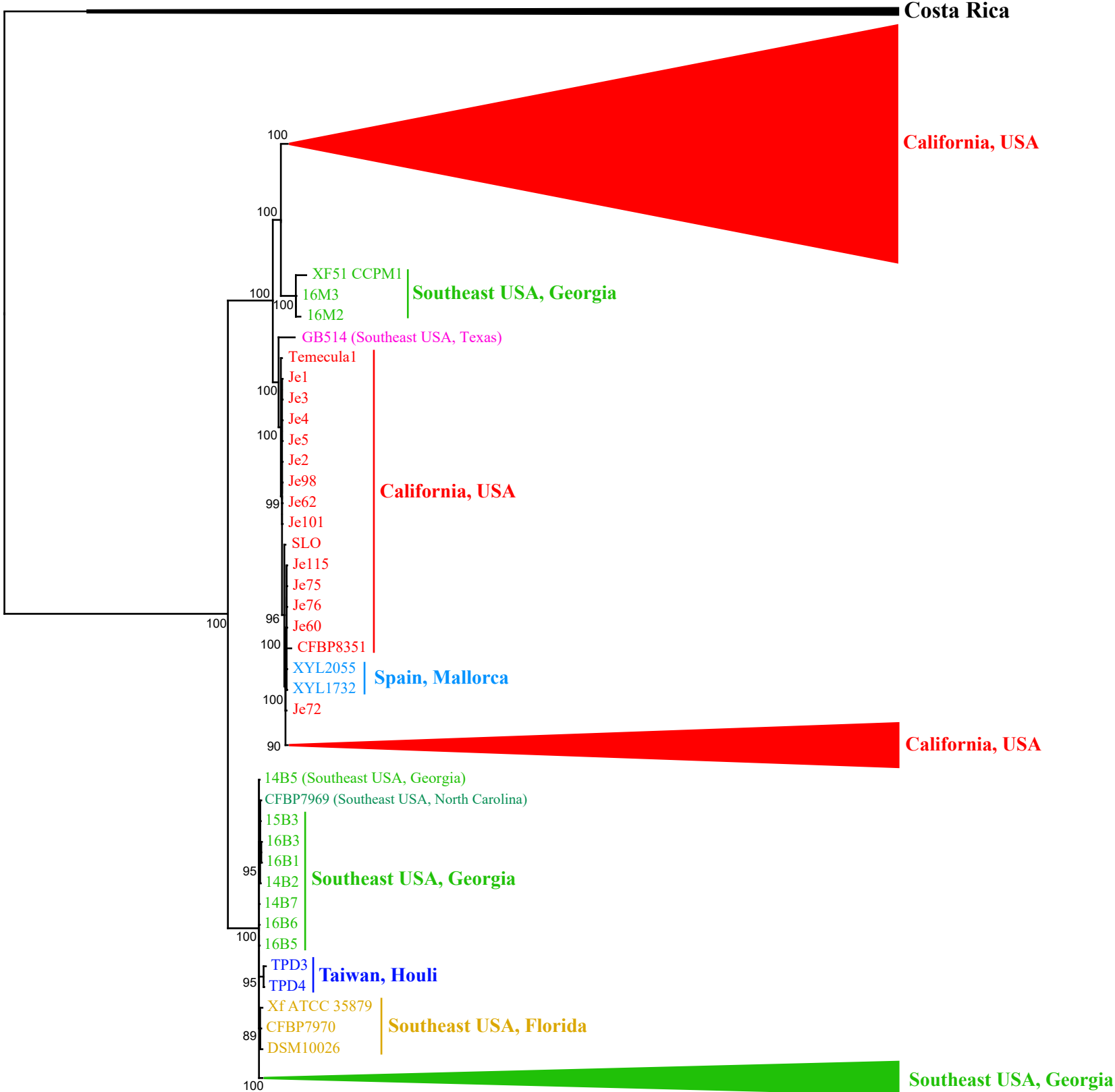
33 **California/Spain.** Each line represents a recombinant event, with the width and placement of the  
34 line indicating recombinant segment size and core genome alignment position. Colors  
35 correspond to lineages identified during fastGEAR analysis.

36

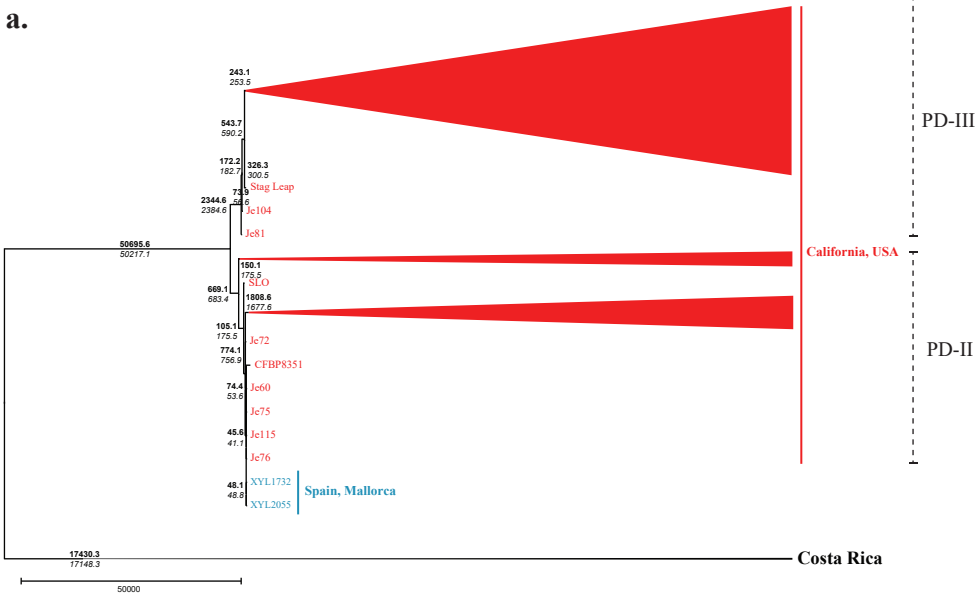
37 **Figure S4. Circle plot showing strain-specific recombination events within Southeast**

38 **USA/Taiwan.** Each line represents a recombinant event, with the width and placement of the  
39 line indicating recombinant segment size and core genome alignment position. Colors  
40 correspond to lineages identified during fastGEAR analysis.

41



0.0020

**a.****b.**