Repeated replacement of an intrabacterial symbiont in the tripartite nested mealybug symbiosis

1 Citrus mealybug *Planococcus citri* has two bacterial endosymbionts with an unusual nested arrange-
2 ment: the γ-proteobacterium *Moranella endobia* lives in the cytoplasm of the β-proteobacterium *Trem-*
3 *blaya princeps*. To test the stability of this three-way symbiosis, host and symbiont genomes for five diverse
4 mealybug species were sequenced. β-proteobacterial genomes from diverse mealybug species are *Trem-*
5 *blaya* with similar genome sizes, while γ-proteobacteria are from different clades with different genome
6 sizes. Therefore, it is inferred that *Tremblaya* is the result of a single infection in the ancestor of mealy-
7 bugs, while the γ-proteobacterial symbionts result from multiple replacements of inferred different ages
8 from related but distinct bacterial lineages.

9 Three scenario of the order and timing of the γ-proteobacterial infections are proposed. In idosyn-
10 cratic scenario, there was a single γ-proteobacterial acquisition in the ancestor of the Pseudococcinae that
11 has evolved idiosyncratically as mealybugs diversified over time, leading to seemingly unrelated genome
12 structures and coding capacities. In independent scenario, the γ-proteobacterial infections occurred in-
13 dependently, each establishing symbioses inside *Tremblaya* in completely unrelated and separate events.
14 In replacement scenario, there was a single γ-proteobacterial acquisition in the Pseudococcinae ancestor
15 that has been replaced in some mealybug lineages over time.

16 The idosyncratic scenario can be discarded as phylogenetics of γ-proteobacterial symbionts reveals
17 that they have originated from clearly distinct and well-supported bacterial lineages. The independent
18 and replacement scenarios are more difficult to tell apart. Under the independent scenario, *Tremblaya*
19 may experience two rounds of genome corruption: one after association with mealybug ancestor, and
20 one after infection of γ-proteobacteria. Therefore, one should expect diverse genome sizes in β- and γ-
21 proteobacteria. Conserved genome sizes of β-proteobacteria and diverse sizes of γ-proteobacterial genomes
22 favor the replacement scenario.

23 Two reasons why γ-proteobacteria end with living inside β-proteobacteria are proposed. The first is
24 that it was easier to use the established transport system between the insect cell and *Tremblaya* than to
25 evolve a new one. The second is that the insect immune system likely does not target *Tremblaya* cells,

26 and so the *Tremblaya* cytoplasm is an ideal hiding place for a newly arrived symbiont.