

Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis

1 The smallest reported bacterial genome belongs to *Tremblaya princeps*, a symbiont of *Planococcus*
2 *citri* mealybugs (PCIT). *Tremblaya* PCIT not only has a 139 kb genome, but possesses its own bacterial
3 endosymbiont, *Moranella endobia*. Genome and transcriptome sequencing, including genome sequencing
4 from *Tremblaya* symbiont of *Phenacoccus avenae* (PAVE), which lacks intracellular bacteria, reveals that
5 the extreme genomic degeneracy of *Tremblaya* PCIT likely resulted from acquiring *Moranella* as an
6 endosymbiont. In addition, at least 22 expressed horizontally transferred genes from multiple diverse
7 bacteria to the mealybug genome likely complement missing symbiont genes. However, none of these
8 horizontally transferred genes are from *Tremblaya*, showing that genome reduction in this symbiont has
9 not been enabled by gene transfer to the host nucleus.

10 Acquisition of *Moranella* symbiont may trigger extreme genome degeneracy in *Tremblaya* PCIT.
11 Genome reduction in *Tremblaya* PAVE occurs to a degree consistent with other previously reported
12 tiny symbiont genomes, and *Tremblaya* PCIT gene set is an almost perfect subset of *Tremblaya* PAVE.
13 These results suggest that much of the reductive genome evolution observed in *Tremblaya* (down to
14 approximately 170 kb) occurred before the acquisition of *Moranella* in the common ancestor of *Planococcus*
15 *citri* and *Phenacoccus avenae* and that the extreme genomic degeneracy observed in *Tremblaya* PCIT
16 (from 170 kb to 140 kb) was likely due to the acquisition of *Moranella* by *Tremblaya* at some point in the
17 lineage leading to *Planococcus citri*. This scenario is consistent with studies showing that massive and
18 rapid gene loss can occur in bacteria that transition to a symbiotic lifestyle (Mira *et al.*, 2001; Moran Mira,
19 2001; Nilsson *et al.*, 2005), after which gene loss slows, and gross genomic changes become infrequent,
20 even over hundreds of millions of years (McCutcheon and Moran, 2010; Tamas *et al.*, 2002; van Ham *et*
21 *al.*, 2003).

22 Pathways for translation, synthesis of essential amino acids, vitamins and peptidoglycan in *Tremblaya*
23 PCIT are complemented by *Moranella*, mealybug genes originated from bacteria-to-mealybug horizontal
24 gene transfers (HGTs) and mealybug genes of eukaryotic origin. In PCIT, ten HTGs group closely with
25 other alphaproteobacterial sequences in phylogenetic trees, and nine HTGs from Gammaproteobacteria,

26 two from Bacteroidetes, and one that is phylogenetically unresolved. The majority of these HGTs are not
27 present in *Tremblaya* and *Moranella* genomes.

28 The presence of a large number of HTGs involved in peptidoglycan production and recycling is consis-
29 tent with the hypothesis that cell lysis is the mechanism used to share gene products between *Moranella*
30 and *Tremblaya* PCIT (Koga *et al.*, 2013; McCutcheon and von Dohlen, 2011). This idea was initially sug-
31 gested based on a lack of transporters encoded on the *Moranella* genome combined with the large number
32 of gene products or metabolites involved in essential amino acid biosynthesis and translation that would
33 need to pass between *Moranella* and *Tremblaya* PCIT for the symbiosis to function (McCutcheon and
34 von Dohlen, 2011). Subsequent electron microscopy on mealybugs closely related to PCIT showed that
35 although most gammaproteobacterial cells infecting the *Tremblaya* cytoplasm were rod shaped, some were
36 amorphous blobs seemingly in a state of degeneration (Koga *et al.*, 2013). The results suggest a plausible
37 mechanism for how the insect host controls this process: by differentially controlling the expression of
38 the horizontally transferred genes, the host could regulate the cell wall stability of *Moranella*. Increasing
39 the expression of *murABCDE* genes would increase the integrity of *Moranella*'s cell wall, and increasing
40 the expression of *mltD/amiD* would tend to decrease *Moranella*'s cell wall strength. As *Tremblaya* PCIT
41 encodes no cell-envelope-related genes and likely uses host-derived membranes to define its cytoplasm, it
42 would be unaffected by changes in gene expression related to peptidoglycan biosynthesis. This hypothesis
43 is testable, because the levels of *Tremblaya* and *Moranella* are uncoupled in mealybugs closely related
44 to PCIT; in males in particular, *Moranella* levels drop to undetectable levels while *Tremblaya* persists
45 (Kono *et al.*, 2008). In situations where *Moranella* is reduced with respect to *Tremblaya*, low expression
46 of *murABCDEF* and increased expression of *mltD/amiD* would be expected.