

1 **Modeling trophic dependencies and exchanges among insects' bacterial**
2 **symbionts in a host-simulated environment**

3 Itai Opatovsky,^{ab} Diego Santos-Garcia,^c Tamar Lahav,^a Shani Ofaim,^a Laurence Mouton,^d
4 Valérie Barbe,^e Einat Zchori-Fein,^f Shiri Freilich,^{a#}

5

6 Institute of Plant Sciences, Newe Ya'ar Research Center, The Agricultural Research
7 Organization, Ramat Yishay, Israel^a; Regional Agricultural Research and Development
8 Center, Southern Branch (Besor), Israel^b; Department of Entomology, Hebrew University of
9 Jerusalem, Rehovot, Israel^c; Laboratoire de Biométrie et Biologie Evolutive, Université de
10 Lyon, 69622 Villeurbanne Cedex, France^d; CEA/DSV/IG/Genoscope, Evry, France^e; Newe
11 Ya'ar Research Center, The Agricultural Research Organization, Ramat Yishay, Israel^f

12

13 Running Head: A comparative model for studying symbiont interactions

14

15 # Address correspondence to Shiri Freilich, shiri@volcani.agri.gov.il.

16 I.O. and D.S.G. contributed equally to this work.

17 Abstract word count: 250

18 Text word count: 4707

19

20

21 **Abstract**

22 Individual organisms are linked to their communities and ecosystems via metabolic activities.
23 Metabolic exchanges and co-dependencies have long been suggested to have a pivotal role in
24 determining community structure. Metabolic interactions with bacteria have been key drivers
25 in the evolution of sap-feeding insects, enabling complementation of their deprived nutrition.
26 The sap-feeding whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae) harbors an obligatory
27 symbiotic bacterium, as well as varying combinations of facultative symbionts. We took
28 advantage of the well-defined bacterial community in *B. tabaci* as a case study for a
29 comprehensive and systematic survey of metabolic interactions within the bacterial
30 community and their associations with documented frequency of bacterial combinations. We
31 first reconstructed the metabolic networks of five common *B. tabaci* symbionts (*Portiera*,
32 *Rickettsia*, *Hamiltonella*, *Cardinium* and *Wolbachia*), and then used network analysis
33 approaches to predict: (1) species-specific metabolic capacities in a simulated bacteriocyte-
34 like environment; (2) metabolic capacities of the corresponding species' combinations, and
35 (3) dependencies of each species on different media components.
36 The automatic-based predictions for metabolic capacities of the symbionts in the host
37 environment were in general agreement with previously reported genome analyses, each
38 focused on the single-species level. The analysis suggested several previously un-reported
39 routes for complementary interactions. Highly abundant symbiont combinations were found
40 to have the potential to produce a diverse set of complementary metabolites, in comparison to
41 un-detected combinations. No clear association was detected between metabolic co-
42 dependencies and co-occurrence patterns. The findings indicate a potential key role for
43 metabolic exchanges as key determinants shaping community structure in this system.

44

45

46 **Importance**

47 This study harnesses the rapid advances in tools developed within the newly emerging field
48 of eco-systems biology to study a small, closed, well-defined micro-ecosystem of a bacterial
49 community, allowing a detailed description of its trophic networks. In addition to indicating
50 un-reported routes for complementary interactions between co-located symbionts of *Bemisia*
51 *tabaci*, this study provides a generic tool for creating testable predictions of metabolic
52 interactions in complex communities. Understanding the overall metabolic interactions in a
53 given system is of key importance in ecology and evolution and can provide a powerful tool
54 for expanding knowledge on inter-species bacterial interactions in various ecosystems.

55 **Introduction**

56 Metabolic interactions are one of the main factors shaping communities and ecosystems by
57 forming complex trophic networks. In bacterial communities, metabolic exchanges are
58 ubiquitous and play a pivotal role in determining community structure (1–8). Bacteria also
59 exchange metabolites with multicellular organisms, and such of mutualistic interactions have
60 been a key driver of evolution, enabling eukaryotic expansion into new ecological niches and
61 species diversification (9, 10). Among the most studied evolutionary radiations that has
62 depended on symbiosis are the sap-feeding insects such as whiteflies, aphids, psyllids,
63 cicadas and spittlebugs. All have intimate associations with maternally transmitted,
64 intracellular bacteria that provide essential nutrients (mainly essential amino acids) and
65 thereby enable dietary specialization on phloem or xylem sap of vascular plants (11–13) - a
66 poor environment composed mainly of simple sugars and non-essential amino acids (14). The
67 interaction with these inherited partners is obligatory for insect survival, and the bacteria are
68 thus located inside specialized insect cells termed bacteriocytes. In addition, insects may
69 harbor a diverse array of facultative, nonessential bacterial associates in the bacteriocytes or
70 other body tissues (15). Facultative symbionts are suggested to serve as a “horizontal gene
71 pool”, where variation in their combinations may have functional significance (16–19).
72 Notably, since the obligatory symbionts are exposed to an irreversible process of genome
73 reduction that can erode their metabolic potential (20), facultative symbionts can, in some
74 cases, complement or replace parts of the lost functions (21–23).

75 In recent years, metabolic approaches, based on genome-driven network constructions, have
76 been applied to predict the potential metabolic dependencies and metabolic exchanges
77 between bacterial species (4, 8, 24). Newly developed tools for genome-based metabolic
78 reconstruction enable predicting sets of interactions formed between species combinations ,
79 and the specific exchange of fluxes within multi-species systems (25, 26). Crossing such

80 predictions with corresponding co-occurrence patterns allows deciphering the importance and
81 meaning of variations in such bacterial assemblages (3, 27). To this end, multiple information
82 layers are required, including symbiont co-occurrence patterns, environmental conditions,
83 genetic background of both host and symbionts, and genome-driven predictions for
84 symbionts' potential activities. Here, based on the availability of both distribution patterns
85 and bacterial genome sequences, we focused on exploring the functional significance of
86 combinations of facultative symbionts in the sweetpotato whitefly *Bemisia tabaci*
87 (Hemiptera: Aleyrodidae) and their potential role in shaping alternative community
88 structures.

89 *Bemisia tabaci* is a major pest of several key crops worldwide (28) and is referred to as a
90 complex of species, consisting of at least 28 morphologically indistinguishable, genetically
91 delimited groups or species (29, 30). All whiteflies, including *B. tabaci*, harbor the primary
92 symbiont “*Candidatus Portiera aleyrodidarum*” (hereafter *Portiera*) (31), which has undergone
93 substantial genomic reduction as other obligatory symbionts (20), but is still able to produce
94 most of the essential amino acids (32, 33). In addition, *B. tabaci* has been reported to harbor
95 varying combinations of facultative symbionts, from bacterial genera *Rickettsia*,
96 *Hamiltonella*, *Wolbachia*, *Arsenophonus*, *Cardinium*, *Hemipteriphilus* and *Fritschea* (34).
97 The occurrence and frequencies of combinations of these bacterial symbionts were
98 investigated using a dataset of over 2,000 whiteflies, representing both the largest and the
99 most comprehensive meta-study of insects for which communities of facultative symbionts
100 have been described (34). MEAM1 and MED-Q1, the two most widespread genetic groups of
101 *B. tabaci*, were found to typically harbor the facultative symbiont “*Ca. Hamiltonella defensa*”
102 (hereafter *Hamiltonella*) in addition to the obligatory symbiont *Portiera*. A combination of
103 *Hamiltonella* and “*Ca. Rickettsia* sp.” (hereafter *Rickettsia*) seemed to be unique to MEAM1
104 individuals, while combinations of *Hamiltonella* with either “*Ca. Cardinium hertigii*” or “*Ca.*

105 *Wolbachia* sp.” (hereafter *Cardinium* and *Wolbachia* respectively) were unique to individuals
106 of the MED-Q1 genetic group. Because the analysis revealed no correlation between specific
107 facultative symbiont complexes and any of the environmental factors tested (34), we
108 hypothesized that metabolic interactions may be involved in shaping the bacterial community
109 structure. The recent release of the genome sequences of *Portiera*, *Rickettsia*, *Hamiltonella*,
110 and *Cardinium* (23, 32, 35–40) has promoted analyses of interactions between the obligatory
111 symbiont *Portiera* and its *B. tabaci* host (23, 33, 39, 41), the facultative symbionts and *B.*
112 *tabaci* (23, 35, 36, 39), and the obligatory and facultative symbionts (23, 33). At both trophic
113 levels, metabolic exchanges were suggested to be required for the completion of essential
114 metabolic pathways. Branched Chain Amino Acids (BCAs), for example, are synthesized
115 through *Portiera*–host complementary interaction (33, 39, 41) while lysine biosynthesis can
116 occur via *Portiera*–host or *Portiera*-*Hamiltonella* complementation (23, 39).

117 As metabolic cross talk is suggested to convey functional capacities associated with specific
118 species combinations, we conducted comparative-interaction analysis considering
119 interactions formed between pairwise combinations of residing symbionts. We first
120 reconstructed the metabolic networks of five symbionts (*Portiera*, *Rickettsia*, *Hamiltonella*,
121 *Cardinium* and *Wolbachia*), and then used network analysis approaches to predict: (1)
122 species-specific metabolic capacities in a simulated host’s bacteriocyte-like environment; (2)
123 metabolic capacities of species' combinations, and (3) the dependencies of each species on
124 the different media components.

125 **Results**

126 *Metabolic capacities of individual symbionts in the simulated bacteriocyte environment*

127 The complete genomes of *Portiera*, *Cardinium*, *Hamiltonella*, and *Rickettsia* from *B. tabaci*
128 MEAM1 and MED species were retrieved from public resources (Table 1) and the genome of
129 *Wolbachia* was assembled *de novo* (Supplemental material, Table S1). All genomes were

130 analyzed using a standard automated procedure followed by manual revision. For each
131 bacterium, a metabolic-network was reconstructed based on the identification of its genome-
132 derived enzyme content.

133 Beyond the static representation of data as a network, computational simulations allow
134 addressing the influence of environmental inputs (nutritional resources) on the network
135 structure and composition, *i.e.*, the metabolic capacities of a species in a given environment,
136 for example, in terms of its ability to produce essential metabolites. More specifically,
137 expansion algorithms generate the set of all possible metabolites that can be produced given a
138 set of starting compounds (source-metabolites) and a set of feasible reactions (42). We
139 defined the starting compounds as a compilation of nutrients provided by the host whitefly in
140 the bacteriocyte environment based on previous studies (33, 39, 41, 43). Our predicted
141 bacteriocyte environment was composed of 50 compounds including ATP, co-factors and
142 vitamins such as NAD⁺, heme and thiamine, six non-essential amino acids, and sugars (Table
143 S2).

144 For each of the symbionts we simulated metabolic activity in the bacteriocyte environment
145 and listed a sub-set of essential metabolites predicted to be produced (Table S3). It was found
146 that most of the secondary symbionts are capable of producing nucleic acids (Fig. S1),
147 whereas their ability to produce amino acids and co-factors varied (Fig. 1). *Portiera*, being an
148 obligatory symbiont that has undergone substantial genomic reduction, was the most limited
149 in its metabolic capacities. It was capable of synthesizing alanine and the essential amino
150 acids threonine, methionine, tryptophan and phenylalanine (Fig. 1), in accordance with
151 previous reports regarding its metabolic capacity and interaction with the whitefly host (32,
152 40). All of the facultative symbionts were capable of synthesizing the non-essential amino
153 acid glycine, which was not produced by *Portiera*. As previously reported alanine is only
154 produced by *Hamiltonella* and *Cardinium* (23, 35, 39). In addition, and in accordance with

155 previous results, asparagine could be produced by the facultative symbionts *Hamiltonella*,
156 *Wolbachia* and *Cardinium* (23, 39). Overall, the automatic-based predictions for metabolic
157 capacities of the symbionts in the host environment generated by the model were in general
158 agreement with previously reported genome analyses.

159 Complementary production of amino acids

160 The genome-specific differences in the production of amino acids (Fig. 1) suggested that
161 complementary metabolic interactions can potentially take place in the bacteriocyte eco-
162 system, increasing the total number of amino acids that can be synthesized by the residing
163 bacteria. This is supported by some established examples that demonstrate the co-
164 production of amino acids by bacterial combinations through complementation of metabolic
165 pathways in various ecological systems, including insect-symbiont interactions (3, 44–46). To
166 predict complementation patterns, we repeated co-growth simulations for pairwise
167 combinations in the exact same environment as for single-species simulations. A metabolite
168 was defined as "complementary" if its synthesis requires a combination of bacterial species
169 (*i.e.*, individual members of the combination cannot produce it). Overall, complementary
170 interactions for the co-synthesis of four essential amino acids were detected (Fig. 1): lysine
171 production by *Hamiltonella-Wolbachia* and *Portiera-Hamiltonella* combinations and
172 production of the three BCAs (leucine, valine and isoleucine) by the *Portiera-Rickettsia*
173 combination. While the complementation of *Hamiltonella-Wolbachia* for lysine production
174 has not been previously reported, our results are in agreement with the possible cooperation
175 of *Portiera* and *Hamiltonella* for its production (23, 39). The production of BCAs in the
176 bacteriocyte environment has been suggested to take place through a complementary
177 interaction between *Portiera* and *B. tabaci*. Our analysis suggested an alternative route for
178 the production of BCAs through an interaction between the obligatory symbiont *Portiera* and
179 the facultative symbiont *Rickettsia*. This previously unreported complementation is in

180 agreement with identification of the *ilvE* gene in *Rickettsia* from *B. tabaci*, carrying the final
181 reaction in the BCA-synthesis pathway (47).

182 Profiles of complementary metabolites

183 Beyond the complementary production of amino acids, we recorded, for each pairwise
184 bacterial combination, a vector describing the set of potential complementary metabolites
185 (Table S4). The interactions formed between the most frequent symbionts - the obligatory
186 symbiont *Portiera* and the partially fixated symbiont *Hamiltonella* - and the other symbionts,
187 produced a high number of complementary metabolites per interaction (average of ~12; Table
188 2). In comparison, the lowest number of complementary metabolites was predicted for
189 *Cardinium* (average of ~4, Table 2), the symbiont with the lowest number of appearances in
190 the surveyed populations (34). Overall, the interaction matrix included seven occurring
191 combinations (blue, Table 2) versus three non-occurring combinations (red), with an average
192 number of ~12 versus ~3 complementary metabolites.

193 Principle Component Analysis (PCA) of the complementary-metabolite vectors suggested
194 four key types of interaction-groups (Fig. 2): *Portiera* associated interactions (with
195 *Hamiltonella*, *Rickettsia* and *Wolbachia*), the two divergent *Hamiltonella*-associated
196 interactions (with *Wolbachia* and *Rickettsia*), and the non-occurring combinations *Cardinium*-
197 *Wolbachia*, and *Rickettsia*-*Wolbachia* and *Rickettsia*-*Cardinium* (red combinations in Table 2
198 and Fig. 2). *Cardinium*-*Portiera* combination is classified together with *Hamiltonella*-
199 *Wolbachia* and not with the other *Portiera* associated combinations. Metabolites common to
200 the *Portiera*-associated combinations included amino-acyl transferases and many primary
201 metabolites such as amino acids and co-factors. Complementary metabolites common to the
202 co-clustered *Portiera*-*Hamiltonella* and *Portiera*-*Wolbachia* combinations included potential
203 precursors of methionine and purine/thiamine (Table S4); all potential interactions have been
204 previously suggested for *Hamiltonella* (39), but not for *Wolbachia*.

205 The relatively divergent clustering pattern recorded for the combinations of facultative
206 symbionts *Hamiltonella-Wolbachia* and *Hamiltonella-Rickettsia* (Fig. 2) might be attributed
207 to the fact that most of these metabolites are not common but rather interaction-specific:
208 interactions between *Hamiltonella* and *Wolbachia* were mostly involved in the synthesis of
209 secondary metabolites, mainly terpenoids; interactions between *Hamiltonella* and *Rickettsia*
210 were mostly involved in butanoate and amino sugar metabolism (Table S4). Finally, non-
211 occurring combinations typically led to a low number of potential complementary
212 metabolites and were clustered.

213 Co-dependencies of symbionts on specific media components

214 Under the assumption that highly similar metabolic demands may hint at resource
215 competition and potentially lead to exclusion of the less fit competitor, the extent to which
216 symbiont combinations rely on common resources was assessed. Scores were evaluated using
217 NetCmpt, which provides predictions for the degree of effective metabolic overlap between
218 pairs of bacterial species, ranging between 0 (no overlap) and 1 (complete overlap) (26).
219 Scores are a-symmetrical whereas the effect of interactions on pair members is likely to differ
220 (*i.e.*, one of the species is likely to be more affected than its potential competitor). The score
221 is indicative of the effect of the column species over the row species. For example,
222 *Hamiltonella* was almost unaffected by *Portiera* and *Cardinum* and was more sensitive to the
223 presence of *Wolbachia* and *Rickettsia* (Table 2). Overall, pairwise scores were relatively low,
224 ranging between 0.03 (the effect of *Portiera* on *Hamiltonella*) and ~0.35 (the effect of
225 *Hamiltonella* on *Wolbachia* and *Rickettsia*). The observed average competition score, 0.18
226 (Table 2), was relatively low compared to an average of 0.36 calculated for other modeled
227 bacterial communities (4). Notably, no significant difference was observed in the level of
228 metabolic overlap between occurring versus non-occurring combinations (Table 2).

229 Since resource overlap is thought to determine community structure only under limited
230 carrying capacity of the habitat (48), we further simulated species-specific growth in the
231 bacteriocyte-like environment, rather than considering the generic optimal environment
232 assumed by the NetCmpt tool. We estimated the specific qualitative effect of each metabolite
233 on growth capacity following iterative removal of one component at a time. As expected,
234 *Portiera* exhibited the most differentiated dependency profile of all symbionts (Fig. 3). In the
235 specific bacteriocyte simulated environment, *Portiera* relied uniquely on D-ribose 5-
236 phosphate, D-erythrose 4-phosphate and phosphoenolpyruvate for tryptophan production, as
237 well as on L-homocysteine for methionine production. Metabolite dependencies that were
238 common to more than a single symbiont included dependencies on the amino acids L-
239 cysteine (*Wolbachia* and *Rickettsia*) and L-serine (*Cardinium*, *Hamiltonella* and *Wolbachia*).
240 Hence, co-dependency might lead to a mutually exclusive distribution pattern, as suggested
241 for *Wolbachia* and *Rickettsia* (34).

242 In addition, common dependencies on NAD⁺ (*Hamiltonella*, *Wolbachia* and *Rickettsia*) and
243 ATP (*Cardinium* and *Rickettsia*) reflected the energy production pathways of the
244 corresponding symbionts. NAD⁺ dependent bacteria all have a citrate cycle requiring NAD⁺
245 as a reducing force. *Rickettsia* and *Cardinum*, both missing glycolytic pathways, rely on the
246 host for ATP production. Though *Rickettsia* possesses a citrate-cycle, capable of producing
247 ATP, its activation requires thiamine diphosphate, which was not present in our bacteriocyte
248 environment. In our simulations, *Wolbachia* was the only symbiont that could produce
249 thiamine diphosphate from the thiamine provided through the activity of thiamine
250 diphosphokinase. Like *Cardinum*, *Portiera* does not possess either a citrate-cycle or
251 glycolysis pathway. However, at least to a minimal amount, ATP production can potentially
252 occur through the activity of ATP phosphoribosyltransferase in the histidine-metabolism

253 pathway requiring D-ribose 5-phosphate as input. In addition, *Portiera* can also obtain ATP
254 through carotenoid biosynthesis (49).

255 **Discussion**

256 We harnessed the rapidly advancing tools developed within the newly emerging field of eco-
257 system biology to study a small, closed, well-defined micro ecosystem of a bacterial
258 community. The focus on this unique community allowed exploring metabolic interactions
259 between all relevant pairwise combinations, providing a detailed description of the trophic
260 networks. Using simulation models to predict metabolic exchanges and co-dependencies we
261 aimed to shed light on the role played by symbiotic interactions in shaping host ecology and
262 how the ecology within the host can constrain community structure. The analysis was based
263 on several assumptions and limitations that should be acknowledged: (1) we assumed a free
264 flux of metabolites between the host and the symbionts and among the symbionts themselves.
265 Several descriptions of the frequent exchanges in microbial communities support this
266 assumptions (3, 50, 51). (2) The model is qualitative, only providing binary predictions for
267 the production or absence of a metabolite rather than quantitative estimates for metabolite
268 consumption/production as produced for stoichiometric networks using constraint based
269 modeling. Hence, metabolites that are common resources for several symbionts might not
270 induce competition, as they are not necessarily limiting. Similarly, the coproduction of
271 nutrients might take place in negligible amounts, (3) the model is limited to the identification
272 of metabolic interactions which are not likely to be the only factor affecting community
273 structure. However, despite the inherent limitations of the approach, the analysis successfully
274 captured previous genome-based predictions of metabolic complementations at host-
275 symbiont and symbiont-symbiont levels in the bacteriocyte (23, 32, 39). Such evidence
276 supports the relevance of our tool for the formulation of new, testable predictions of
277 metabolic exchanges in an automated manner. Moreover, our simulations take into account a

278 specific environment, hence reflecting the common notion that interactions are dynamic and
279 can vary with the addition or depletion of nutrients (4, 6, 44).

280 This study focused on diet-limited insects that rely on obligatory associations with bacteria
281 for complementation of their nutritional needs. The role of cooperative coevolution in
282 selecting for traits that enable and stabilize such symbioses has been thoroughly discussed in
283 the literature (10, 11, 13, 15, 16, 19, 52, 53). One of the most important negative
284 ramifications of symbiotic alliances is the genome-reduction process in the obligatory
285 symbionts that limits beneficial contributions (54). Consequently, a new symbiont may
286 replace or supplement the capabilities of a previous one. The dynamic acquisition and loss of
287 horizontally transmitted facultative symbionts enable the continuous persistence of many
288 species. Although the facultative symbiont's ability to colonize a new host is strongly
289 influenced by metabolic similarities between the new and old host (35, 55), it also relies, at
290 least to some extent, on the metabolic interactions that it forms with its new environment
291 (56). Accordingly, transient bacterial species are expected to co-occur less frequently than
292 expected by random chance if they are competing for limiting metabolic resources. Similarly,
293 if their metabolic pathways are complementary with respect to the production of a mutually
294 required resource, they are expected to co-occur more frequently than expected by random
295 chance. Such interactions can also suggest a possible gain that compensates for the fitness
296 cost of co-infections (56, 57).

297 Here, using automated tool rather than relying on genome-specific metabolic mappings (23,
298 33, 35, 39), we predicted four previously un-reported routes for transient complementary
299 interactions. These interactions can potentially increase the amount of the resulting amino
300 acids in the bacteriocyte by providing alternative synthesis routes. Examples include
301 complementation of the synthesis of BCAs is possible through the insect host (*B. tabaci*)
302 obligatory symbiont (*Portiera*) interaction but also, by a previously un-reported interaction

303 between *Portiera* and the facultative symbiont *Rickettsia*. Similarly, production of lysine as
304 well as of the co-factor 5-methyl- tetrahydrofolate, the predominant form of dietary folate
305 (58), occurs through the complementary *Portiera-Hamiltonella* interaction. The reported
306 *Portiera-Hamiltonella* complementation of lysine could indicate a more intimate relationship
307 between these symbionts, compromising the evolution of *Hamiltonella* toward a co-
308 obligatory symbiont in some *B. tabaci* species (23, 36, 39). Though some of the
309 complementary metabolites are redundant between co-existing interactions, they might
310 suggest alternative production routes, possibly compensating for the limited transcriptional
311 regulation of symbionts (59). Such complementation can be mutualistic, increasing the total
312 amount of essential nutritional sources for all community members. Alternatively, it might
313 only be beneficial for specific species and reflect a parasitic life style. For example,
314 complementary production of BCAs is possible through *Portiera-Rickettsia* interactions. The
315 *Rickettsia* from *B. tabaci* is part of the *R. bellii* group that includes many pathogenic
316 members (60, 61). The complementation might reflect the dependency of *Rickettsia* on the
317 BCA intermediates that it scavenges from the host-environment, bypassing the host's control
318 of BCA biosynthesis (47).

319 The model suggests several complementary pathways for metabolic co-production of
320 additional metabolites, typical of *Portiera* interactions with the facultative symbionts. All of
321 these interactions are involved in the production of metabolites compensating for the loss of
322 aminoacyl-tRNAs in the *Portiera* lineage (L-tryptophanyl, N-formylmethionyl, L-methionyl
323 and L-alanyl-tRNAs, Table S4) (33). Although these losses are assumed to reflect the
324 dependency of *Portiera* on its host (30,31,58), the analysis suggests alternative routes for
325 such complementation.

326 Complementary interactions also lead to the potential synthesis of secondary metabolites
327 regulating host-parasitoid interactions(62, 63). For example, dimethylallyl diphosphate, a

328 terpenoid, is involved in the metabolism of aphid's alarm pheromones(64); sialic acids have
329 diverse functions in host-bacteria interactions, including as signaling molecules and
330 nutritional sources (65) (Table S4).

331 Specific combinations of co-occurring symbionts have been shown to correlate with
332 delimited genetic groups of *B. tabaci* (34). Combinations of *Hamiltonella* with *Rickettsia* are
333 unique to individuals from MEAM1, whereas combinations of *Hamiltonella* with *Wolbachia*
334 are commonly found in individuals from MED-Q1. Notably, both combinations, which are
335 highly dominant in their corresponding genetic group (34), have the potential to co-produce a
336 diverse set of primary and secondary metabolites (14 and 18, respectively), which can
337 increase host fitness, favoring their maintenance on this species. Unlike the relatively
338 conserved profile of complementary metabolites produced through interactions between the
339 obligatory and facultative symbionts, the complementary profiles formed by *Hamiltonella*-
340 *Rickettsia* and *Hamiltonella*-*Wolbachia* are relatively diverse (Fig. 2), suggesting a biotype-
341 specific functional adaptation. *Hamiltonella*-*Cardinum* combination is mainly found in the
342 MED-Q1 group. This combination is less frequent (34), which could possibly be explained
343 by their low complementation potential (zero metabolites). Consistent with these specific
344 examples, we observed an overall trend of low complementary potential in non-occurring
345 combinations in comparison to occurring ones. However, the limited sample size precludes
346 significance of these observations.

347 While the analysis suggested an association between high-complementation and frequent co-
348 occurrence no such indication was detected for competitive interactions (Table 2). One
349 possible interpretation is that metabolic exchanges are more dominant in shaping bacterial
350 communities (66). Indeed, whereas according to classical ecology theory, inter-species
351 competition over common resources should lead to mutual-exclusion distribution patterns
352 (48), relevant examples are rarely identified based on potential metabolic screens (3, 4, 67). A

353 possibly explanation can be that only a narrow set of factors are quantitatively limited and
354 therefore relevant for competition and determining community structure. To identify such
355 potential limiting factors, we characterized metabolic co-dependencies between bacterial
356 pairs. Predicted co-shared metabolites included the amino-acids L-cysteine (*Wolbachia* and
357 *Rickettsia*) and L-serine (*Cardinium*, *Hamiltonella* and *Wolbachia*). Whereas *Hamiltonella*-
358 *Cardinium* and *Hamiltonella-Wolbachia* combinations are frequent, *Wolbachia-Rickettsia*
359 combinations are rare (34), indicating at cysteine as a potential limiting factor. Although
360 cysteine is a non-essential amino acid that can be supplied by the host and is found in the
361 phloem, it is the main sulfur source required for Fe-S protein biogenesis (68). In addition,
362 common dependencies in NAD⁺ and ATP which reflect the energy-production pathways of
363 the corresponding symbionts can have a strong influence on symbiont co-occurrences. For
364 example, *Rickettsia* and *Cardinium*, both missing the glycolytic pathways and relying on their
365 host for ATP production, are not found together in the host (34). In the *Rickettsia* genus, and
366 other intracellular parasites, ADP/ATP translocases are known to play a crucial role in the
367 exploitation of host ATP (60, 69). Interestingly, in *Cardinium* and related bacteria, ADP/ATP
368 translocases are also present, indicating to a parasitic past (33, 70, 71). In contrast, it seems
369 that *Wolbachia*, independent of its parasitic status, does not present (or has not acquired) the
370 ADP/ATP translocases, relying on its own machinery to produce ATP (72).

371 Despite its obvious limitations, this model provides a tool for generating predictions for
372 testable hypotheses of metabolic interactions in bacterial communities. Understanding the
373 overall metabolic interactions in a given system is of key importance in ecology and
374 evolution and can provide a powerful tool for expanding knowledge on inter-specific
375 bacterial interactions in various ecosystems. With respect to applied aspects, symbiotic
376 microorganisms have been shown to influence the success rates of various biological control
377 programs of agricultural pests (73, 74). Attempts to establish more efficient pest-management

378 strategies involve the removal of specific symbionts or the introduction of others, and our
379 proposed model is expected to contribute to the efficiency and productivity of such efforts.
380 The presented simple model system offers a level of tractability that is crucial for paving the
381 way to the simulation, prediction and management of microbial communities that can
382 expanded to more complex ecosystems, such as the guts of humans and livestock, water
383 resources and soils.

384 **Materials and Methods**

385 Genome assembly and annotation

386 Relevant genomes were collected from multiple public sources (Table 1), with the exception
387 of the *Wolbachia* genome which was assembled *de novo* using sequence data produced by a
388 Genoscope-funded project (<http://www.genoscope.cns.fr>). The sequence was deposited in the
389 European Nuclear Archive (<http://www.ebi.ac.uk/ena/data/view/>) under project number
390 PRJEB15492. The procedure is fully described in the supplemental data.

391 A standard protocol for annotation retrieval was applied for all genomes. Annotations were
392 carried out using several genome-annotation pipelines: IMG/M (75), Kbase (<http://kbase.us/>),
393 Rast (76), MG-rast (76). To estimate the accuracy and comprehensiveness of the predictions,
394 we benchmarked the EC (enzyme commission) predictions for the *Cardinium* genome,
395 retrieved from the four pipelines, with annotations derived from a detailed manual curation.

396 The IMG/G predictions were the most comprehensive and in highest agreement with the
397 manual curation (Fig. S2). Hence, for consistency, annotations for all genomes were retrieved
398 using the JGI platform. For *Portiera*, out of four published genomes (Table 1), annotations
399 for CP003835.1 were considered in the analysis, based on cross-genome comparative
400 analysis of the enzymatic sets and the annotation status (manually curated, Fig. S2).

401 Following annotation retrieval from JGI, reciprocal BLAST searches were carried out to

402 eliminate contaminated sequences between co-occurring symbionts. The phylogenetic origin
403 of highly similar sequences was determined according to BLAST best hits.

404 Putative pseudogenes for all re-annotated genomes were predicted using GenePrimp (77).

405 Manual inspection was performed for all candidate pseudogenes that had an assigned

406 metabolic function (EC number). In addition, previous annotations of *Cardinium* and

407 *Portiera* (32, 35) were used as supportive information for pseudogene cleaning in these

408 species. Finally, predicted pseudogenes with valid EC accessions were removed from the

409 predicted EC list before conducting follow-up analyses. The number of ECs annotated for

410 each genome is indicated in Table 1. The final EC lists are provided in Table S5.

411 Metabolic activity simulations

412 Metabolic activity simulations were carried using the Expansion algorithm (42) which allows

413 predicting the active metabolic network (expanded) given a pre-defined set of substrates and

414 reactions. The full expansion of the network reflects both the reaction repertoire of each

415 species/species-combination and the primary set of compounds, termed here "source-

416 metabolites". Briefly, the algorithm starts with a set of one or more biochemical compounds

417 acting as source metabolites for a feasible reaction, i.e., a reaction for which all required

418 substrates are available. This reaction is selected out of the reaction pool and added to the

419 network. In an iterative process, the products of the chosen reaction are turned into the new

420 substrates, and so on. Processing of the starting-point compounds by relevant reactions

421 increases the number of available compounds that can act as substrates for other, previously

422 in-activated reactions. The network stops expanding when there are no more feasible

423 reactions. Although, the closest organisms with a well-known and defined bacteriocyte

424 environment are aphids, we decided not to use the information generated for this organisms,

425 based on the long divergence time between aphids and whiteflies (more than 250 Mya) and

426 differences in their symbiotic communities and their mode of transmission (53, 78–80). Here,

427 we described the resources available in the whitefly bacteriocyte by compiling several such
428 pre-published lists that are based on genomic-driven analyses of the whitefly genome (23, 32,
429 39, 41). The list is composed of metabolites produced by the host only, though each symbiont
430 changes the environment by consuming/secreted unique set of metabolites. The limitation of
431 the environment to host secreted metabolites allows predicting potential pairwise interactions
432 that would otherwise be masked by alternative host-symbiont routes. These compounds were
433 termed "source metabolites" (detailed in Table S2) and were used as starting points for
434 unfolding a meta-network formed when considering all enzymes detected across all bacterial
435 genomes, leading to the construction of niche-specific networks.

436 Prediction of complementary interactions

437 Complementation was predicted through a three-stage model (1) constructing a combined set
438 of metabolic reactions (EC accessions) for each pairwise combination; (2) simulating co-
439 growth of both individual and combined bacterial genera in the predicted environment; (3)
440 comparing the set of metabolites produced by the combined genomes to those formed by the
441 individual genomes. Complementary/Synergistic metabolites were those formed by species
442 combinations but not by the individual species. A list of the complementary metabolites
443 produced in each interaction and their mapping to KEGG pathways is provided in Table S4.
444 PCA for the vectors of synergistic metabolites was carried out using R software (81).

445 Prediction of co-dependencies in source metabolites

446 The competition scores for each pair of symbionts were calculated by the network-based tool
447 NetCmpt (26). Beyond the quantitative estimates, NetCmpt was further extended to identify
448 dependencies on specific source metabolites. To this end, growth simulations were carried in
449 the bacteriocyte-like environment used throughout the analysis, rather than in the optimal
450 environment used for the generic NetCmpt calculations. Within each simulation, the number

451 of essential metabolites was determined (e.g., amino acids, nucleic acid and co-factors, Table
452 S3) (26). Iterative simulations were carried out while removing one source metabolite at a time. For
453 each iteration, the number of essential metabolites that could not be produced following the
454 removal of a source metabolite was recorded. The procedure is illustrated in Fig. S3.

455

456

457 **Funding information**

458 The study was supported by the Israel Science Foundation, grant no. 1481/13.

459 **Acknowledgments**

460 We thank Genoscope (<http://www.genoscope.cns.fr>) for providing the RAW data for
461 assembly of the *Wolbachia* genome.

462

463 **Tables**

464 **Table 1:** Genomes list of obligatory and facultative symbionts of *Bemisia tabaci*.

Symbiont	Host	Resource {GeneBank ID} (publication)	Number of ECs^a
<i>Portiera</i>	MEAM1	NCBI {CP003708.1} (40)	103
<i>Portiera</i>	MEAM1	NCBI {CP003868.1}	101
<i>Portiera</i>	MED-Q1	NCBI {CP003835.1}	101
<i>Portiera</i>	MED-Q2	NCBI {CP003867.1}	104
<i>Cardinium</i>	MED-Q1	NCBI {GCA_000689375.1}	112
<i>Hamiltonella</i>	MED-Q1	NCBI {GCA_000258345.1}	398
<i>Rickettsia sp.</i>	MEAM1	NCBI {GCA_000429565.1}	247
<i>Wolbachia sp.</i>	MED-Q2	ENA {PRJEB15492}	253

465 ^a Following annotation, filtering and manual curation. EC = enzyme commission.

466

467 **Table 2:** Predictions of pairwise interactions in the bacteriocyte system between occurring
 468 (blue) and non-occurring (red) pairwise combinations of symbionts. Occurrence versus non-
 469 occurrence was determined according to a detailed survey of symbiont occurrence from 2030
 470 whitefly individuals (34). The first value in each cell represents the number of
 471 complementary metabolites produced in each combination; the second value (in parentheses)
 472 represents the predictions of the competition values (Effective Metabolic Overlap); the third
 473 value (in square brackets) represents the number of source metabolites that induce co-
 474 dependency of both pair members. The primary endosymbiont is denoted in bold face.

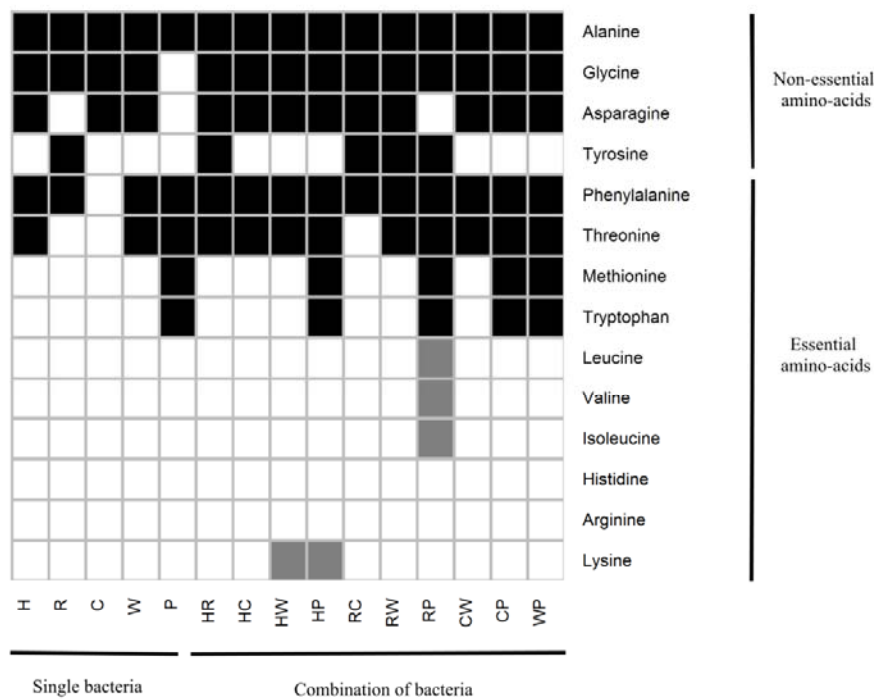
	<i>Hamiltonella</i>	<i>Rickettsia</i>	<i>Cardinium</i>	<i>Wolbachia</i>	<i>Portiera</i>
<i>Hamiltonella</i>		14 (0.2) [2]	0 (0.05) [0]	18 (0.2) [3]	15 (0.03) [0]
<i>Rickettsia</i>	14 (0.21) [2]		1 (0.12) [1]	8 (0.14) [3]	13 (0.07) [0]
<i>Cardinium</i>	0 (0.14) [0]	1 (0.12) [1]		1 (0.12) [0]	14 (0.14) [0]
<i>Wolbachia</i>	18 (0.36) [3]	8 (0.34) [3]	1 (0.14) [0]		8 (0.12) [0]
<i>Portiera</i>	15 (0.25) [0]	13 (0.25) [0]	14 (0.25) [0]	8 (0.25) [0]	

475

476 **Figures**

477 **Figure 1:** Predicted ability of single and pairwise species combinations to synthesize amino
478 acids in the predicted bacteriocyte environment. Amino acids available in the bacteriocyte
479 environment are not shown. Black/white/gray coloring of the cells – synthesis/no
480 synthesis/production of complementary metabolites, respectively. P, C, H, R, W represent
481 *Portiera*, *Cardinium*, *Hamiltonella*, *Rickettsia* and *Wolbachia*, respectively.

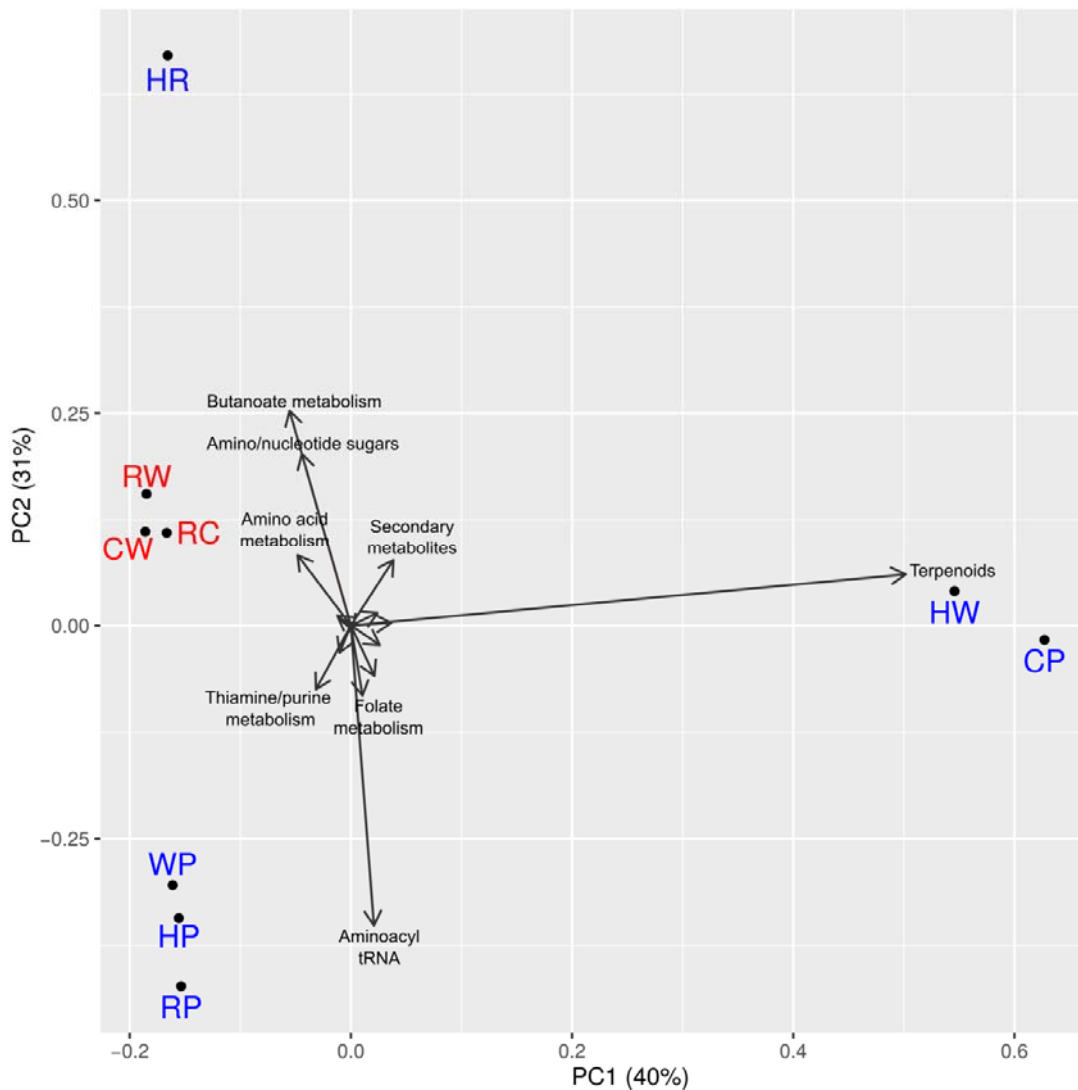
482



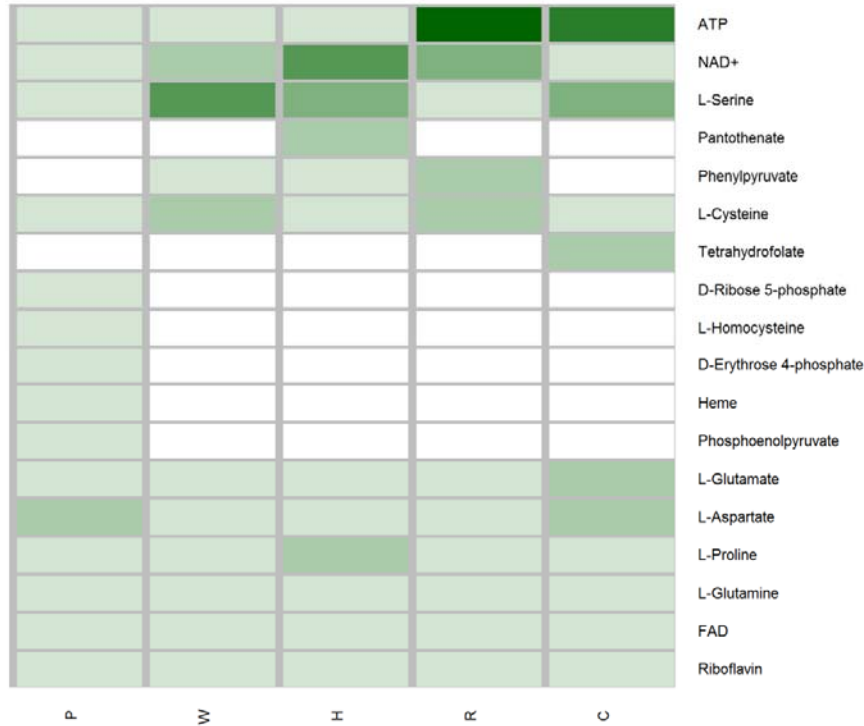
483

484

485 Figure 2: Principal Component Analysis (PCA) diagram of the synergistic metabolite profiles
486 produced through pairwise interactions (Table S4). Synergistic metabolites are those whose
487 synthesis requires the coexistence of both pair members and cannot be produced by either
488 member alone in the predefined environment in which the simulations were carried out. Blue,
489 co-occurring combinations; red, non-occurring combinations. P, C, H, R and W
490 represent *Portiera*, *Cardinium*, *Hamiltonella*, *Rickettsia* and *Wolbachia*, respectively. HC
491 combination has no synergistic metabolites and consequently is not represented.
492 Vectors names represent the metabolic pathway of each synergistic metabolite in Table S4.
493 For plotting reasons, only names of the most important vectors are displayed.



495 **Figure 3:** Reduction in symbiont's ability to produce essential metabolites following removal
496 of specific source metabolites (metabolites predicted to be available to the endosymbionts in
497 the bacteriocyte). Only source metabolites whose removal affected at least one species are
498 shown. P, C, H, R and W represent *Portiera*, *Cardinium*, *Hamiltonella*, *Rickettsia* and
499 *Wolbachia*, respectively.



500

501

502 **References**

- 503 1. **Marx CJ.** 2009. Microbiology. Getting in touch with your friends. *Science* **324**:1150–
504 1151.
- 505 2. **Fuhrman JA.** 2009. Microbial community structure and its functional implications.
506 *Nature* **459**:193–199.
- 507 3. **Freilich S, Zarecki R, Eilam O, Segal ES, Henry CS, Kupiec M, Gophna U,**
508 **Sharan R, Ruppin E.** 2011. Competitive and cooperative metabolic interactions in
509 bacterial communities. *Nat Commun* **2**:589–595.
- 510 4. **Freilich S, Kreimer A, Meilijson I, Gophna U, Sharan R, Ruppin E.** 2010. The
511 large-scale organization of the bacterial network of ecological co-occurrence
512 interactions. *Nucleic Acids Res* **38**:3857–3868.
- 513 5. **Klitgord N, Segrè D.** 2011. Ecosystems biology of microbial metabolism. *Curr Opin*
514 *Biotechnol* **22**:541-546.
- 515 6. **Klitgord N, Segrè D.** 2010. Environments that induce synthetic microbial ecosystems.
516 *PLoS Comput Biol* **6**:1–17.
- 517 7. **Großkopf T, Soyer OS.** 2014. Synthetic microbial communities. *Curr Opin Microbiol*
518 **18**:72-77.
- 519 8. **Levy R, Borenstein E.** 2013. Metabolic modeling of species interaction in the human
520 microbiome elucidates community-level assembly rules. *Proc Natl Acad Sci*
521 **110**:12804–12809.
- 522 9. **Zilber-Rosenberg I, Rosenberg E.** 2008. Role of microorganisms in the evolution of
523 animals and plants: The hologenome theory of evolution. *FEMS Microbiol Rev*
524 **32**:723-735.
- 525 10. **Moya A, Peretó J, Gil R, Latorre A.** 2008. Learning how to live together: genomic
526 insights into prokaryote–animal symbioses. *Nat Rev Genet* **9**:218–229.

- 527 11. **Buchner P.** 1965. Endosymbiosis of animals with plant microorganisms. New York,
528 John Wiley.
- 529 12. **Baumann P.** 2005. Biology bacteriocyte-associated endosymbionts of plant sap-
530 sucking insects. *Annu Rev Microbiol* **59**:155–189.
- 531 13. **Moran NA, McCutcheon JP, Nakabachi A.** 2008. Genomics and evolution of
532 heritable bacterial symbionts. *Ann Rev Genet* **42**:165–190.
- 533 14. **Douglas AE.** 2006. Phloem-sap feeding by animals: Problems and solutions. *J Exp Bot*
534 **57**:747–754.
- 535 15. **Zchori-Fein E, Kostas B.** 2011. Manipulative Tenants: Bacteria associated with
536 arthropods. CRC press.
- 537 16. **Moran NA.** 2007. Symbiosis as an adaptive process and source of phenotypic
538 complexity. *Proc Natl Acad Sci* **104**:8627–8633.
- 539 17. **Gueguen G, Vavre F, Gnankine O, Peterschmitt M, Charif D, Chiel E, Gottlieb Y,**
540 **Ghanim M, Zchori-Fein E, Fleury F.** 2010. Endosymbiont metacommunities,
541 mtDNA diversity and the evolution of the *Bemisia tabaci* (Hemiptera: Aleyrodidae)
542 species complex. *Mol Ecol* **19**:4365–4378.
- 543 18. **Jaenike J.** 2012. Population genetics of beneficial heritable symbionts. *Trends Ecol*
544 *Evol* **4**:226-232.
- 545 19. **Henry LM, Peccoud J, Simon JC, Hadfield JD, Maiden MJC, Ferrari J, Godfray**
546 **HCJ.** 2013. Horizontally transmitted symbionts and host colonization of ecological
547 niches. *Curr Biol* **23**:1713–1717.
- 548 20. **Bennett GM, Moran NA.** 2015. Heritable symbiosis: The advantages and perils of an
549 evolutionary rabbit hole. *Proc Natl Acad Sci* **112**:10169–10176.
- 550 21. **MacDonald SJ, Thomas GH, Douglas AE.** 2011. Genetic and metabolic
551 determinants of nutritional phenotype in an insect-bacterial symbiosis. *Mol Ecol*

- 552 **20:2073–2084.**
- 553 22. **McCutcheon JP, Von Dohlen CD.** 2011. An interdependent metabolic patchwork in
554 the nested symbiosis of mealybugs. *Curr Biol* **21**:1366–1372.
- 555 23. **Rao Q, Rollat-Farnier PA, Zhu DT, Santos-Garcia D, Silva FJ, Moya A, Latorre**
556 **A, Klein C, Vavre F, Sagot MF, Liu SS, Mouton L, Wang XW.** 2015. Genome
557 reduction and potential metabolic complementation of the dual endosymbionts in the
558 whitefly *Bemisia tabaci*. *BMC Genomics* **16**:226–238.
- 559 24. **Stolyar S, Van Dien S, Hillesland KL, Pinel N, Lie TJ, Leigh JA, Stahl DA.** 2007.
560 Metabolic modeling of a mutualistic microbial community. *Mol Syst Biol* **3**:92–104.
- 561 25. **Ebenhöh O, Handorf T, Heinrich R.** 2004. Structural analysis of expanding
562 metabolic networks. *Genome Inform* **15**:35–45.
- 563 26. **Kreimer A, Doron-Faigenboim A, Borenstein E, Freilich S.** 2012. NetCmpt: a
564 network-based tool for calculating the metabolic competition between bacterial
565 species. *Bioinformatics* **28**:2195–2197.
- 566 27. **Zelezniak A, Andrejev S, Ponomarova O, Mende DR, Bork P, Patil KR.** 2015.
567 Metabolic dependencies drive species co-occurrence in diverse microbial communities
568 **112:6449-6454.**
- 569 28. **Stansly PA, Naranjo SE.** 2010. *Bemisia*: Bionomics and management of a global pest.
570 Springer Netherlands.
- 571 29. **De Barro PJ, Liu S, Boykin L, Dinsdale AB.** 2011. *Bemisia tabaci*□: A statement of
572 species status. *Annu Rev Entomol* **56**:1–19.
- 573 30. **Hu J, Jiang Z, Nardi F, Liu Y, Luo X, Li H, Zhang Z.** 2014. Members of *Bemisia*
574 *tabaci* (Hemiptera□: Aleyrodidae) Cryptic Species and the Status of Two Invasive
575 Alien Species in the Yunnan Province (China). *J Insect Sci* **6**:1–8.
- 576 31. **Thao ML, Baumann P.** 2004. Evolutionary relationships of primary prokaryotic

- 577 endosymbionts of whiteflies and their hosts. *Appl Environ Microbiol* **70**:3401–3406.
- 578 32. **Santos-Garcia D, Farnier P-A, Beitia F, Zchori-Fein E, Vavre F, Mouton L, Moya**
579 **A, Latorre A, Silva FJ**. 2012. Complete genome sequence of “*Candidatus* Portiera
580 aleyrodidarum” BT-QVLC, an obligate symbiont that supplies amino acids and
581 carotenoids to *Bemisia tabaci*. *J Bacteriol* **194**:6654–6655.
- 582 33. **Santos-Garcia D, Vargas-Chavez C, Moya A, Latorre A, Silva FJ**. 2015. Genome
583 evolution in the primary endosymbiont of whiteflies sheds light on their divergence.
584 *Genome Biol Evol* **7**:873–888.
- 585 34. **Zchori-Fein E, Lahav T, Freilich S**. 2014. Variations in the identity and complexity
586 of endosymbiont combinations in whitefly hosts. *Front Ecol Environ* **5**:1–8.
- 587 35. **Santos-Garcia D, Rollat-Farnier PA, Beitia F, Zchori-Fein E, Vavre F, Mouton L,**
588 **Moya A, Latorre A, Silva FJ**. 2014. The genome of *Cardinium* cBtQ1 provides
589 insights into genome reduction, symbiont motility, and its settlement in *Bemisia tabaci*.
590 *Genome Biol Evol* **6**:1013–1030.
- 591 36. **Rollat-Farnier PA, Santos-Garcia D, Rao Q, Sagot MF, Silva FJ, Henri H, Zchori-**
592 **Fein E, Latorre A, Moya A, Barbe V, Liu SS, Wang XW, Vavre F, Mouton L.**
593 2015. Two host clades, two bacterial arsenals: Evolution through gene losses in
594 facultative endosymbionts. *Genome Biol Evol* **7**:839–855.
- 595 37. **Rao Q, Wang S, Su YL, Bing XL, Liu SS, Wang XW**. 2012. Draft genome sequence
596 of “*Candidatus* Hamiltonella defensa,” an endosymbiont of the whitefly *Bemisia*
597 *tabaci*. *J Bacteriol* **194**:3558-3558.
- 598 38. **Rao Q, Wang S, Zhu DT, Wang XW, Liu SS**. 2012. Draft genome sequence of
599 *Rickettsia* sp. strain MEAM1, isolated from the whitefly *Bemisia tabaci*. *J Bacteriol*
600 **194**:4741-4742.
- 601 39. **Luan JB, Chen W, Hasegawa DK, Simmons AM, Wintermantel WM, Ling KS,**

- 602 **Fei Z, Liu SS, Douglas AE.** 2015. Metabolic coevolution in the bacterial symbiosis of
603 whiteflies and related plant sap-feeding insects. *Genome Biol Evol* **7**:2635–2647.
- 604 40. **Sloan DB, Moran NA.** 2012. Endosymbiotic bacteria as a source of carotenoids in
605 whiteflies. *Biol Lett* **8**:986–989.
- 606 41. **Upadhyay SK, Sharma S, Singh H, Dixit S, Kumar J, Verma PC, Chandrashekar**
607 **K.** 2015. Whitefly genome expression reveals host-symbiont interaction in amino acid
608 biosynthesis. *PLoS One* **10**:1–16.
- 609 42. **Ebenhöh O, Handorf T, Heinrich R.** 2004. Structural analysis of expanding
610 metabolic networks. *Genome Inform* **15**:35–45.
- 611 43. **Rao Q, Rollat-Farnier PA, Zhu DT, Santos-Garcia D, Silva FJ, Moya A, Latorre**
612 **A, Klein CC, Vavre F, Sagot MF, Liu SS, Mouton L, Wang XW.** 2015. Genome
613 reduction and potential metabolic complementation of the dual endosymbionts in the
614 whitefly *Bemisia tabaci*. *BMC Genomics* **16**:226.
- 615 44. **Wintermute EH, Silver PA.** 2010. Emergent cooperation in microbial metabolism.
616 *Mol Syst Biol* **6**:407–413.
- 617 45. **Osterman A, Overbeek R.** 2003. Missing genes in metabolic pathways: A
618 comparative genomics approach. *Curr Opin Chem Biol* **7**:238-251.
- 619 46. **Gosalbes MJ, Lamelas A, Moya A, Latorre A.** 2008. The striking case of tryptophan
620 provision in the cedar aphid *Cinara cedri*. *J Bacteriol* **190**:6026–6029.
- 621 47. **Zhu DT, Xia WQ, Rao Q, Liu SS, Ghanim M, Wang XW.** 2016. Sequencing and
622 comparison of the *Rickettsia* genomes from the whitefly *Bemisia tabaci* Middle East
623 Asia Minor I. *Insect Sci* **23**:531–542.
- 624 48. **den Boer PJ.** 1986. The present status of the competitive exclusion principle. *Trends*
625 *Ecol Evol* **1**:25–28.
- 626 49. **Calle-Espinosa J, Ponce-de-Leon M, Santos-Garcia D, Silva FJ, Montero F,**

- 627 **Pereto J.** 2016. Nature lessons: The whitefly bacterial endosymbiont is a minimal
628 amino acid factory with unusual energetics. *J Theor Biol* **407**:303-317.
- 629 50. **Seth EC, Taga ME.** 2014. Nutrient cross-feeding in the microbial world. *Front*
630 *Microbiol* **5**:1–6.
- 631 51. **Paczia N, Nilgen A, Lehmann T, Gätgens J, Wiechert W, Noack S.** 2012. Extensive
632 exometabolome analysis reveals extended overflow metabolism in various
633 microorganisms. *Microb Cell Fact* **11**:122–135.
- 634 52. **Mccutcheon JP, Moran NA.** 2011. Extreme genome reduction in symbiotic bacteria.
635 *Nat Rev Microbiol* **10**:13–26.
- 636 53. **Hansen AK, Moran NA.** 2014. The impact of microbial symbionts on host plant
637 utilization by herbivorous insects. *Mol Ecol* **23**:1473–1496.
- 638 54. **Bennett GM, Moran NA.** 2013. Small , Smaller , Smallest□: The origins and
639 evolution of ancient dual symbioses in a phloem-feeding insect. *Genome Biol Evol*
640 **5**:1675–1688.
- 641 55. **Lukasik P, Guo H, van Asch M, Henry LM, Godfray HCJ, Ferrari J.** 2015.
642 Horizontal transfer of facultative endosymbionts is limited by host relatedness.
643 *Evolution* **69**:2757–2766.
- 644 56. **Ferrari J, Vavre F.** 2011. Bacterial symbionts in insects or the story of communities
645 affecting communities. *Philos Trans R Soc B-Biological Sci* **366**:1389–1400.
- 646 57. **Oliver KM, Moran NA, Hunter MS.** 2006. Costs and benefits of a superinfection of
647 facultative symbionts in aphids. *Proc Biol Sci* **273**:1273–80.
- 648 58. **Pietrzik K, Bailey L, Shane B.** 2010. Folic acid and l-5-methyltetrahydrofolate:
649 Comparison of clinical pharmacokinetics and pharmacodynamics. *Clin Pharmacokinet*
650 **49**:535-548.
- 651 59. **Russell CW, Poliakov A, Haribal M, Jander G, van Wijk KJ, Douglas AE.** 2014.

- 652 Matching the supply of bacterial nutrients to the nutritional demand of the animal host.
653 Proc R Soc B Biol Sci **281**:20141163–20141163.
- 654 60. **Blanc G, Ogata H, Robert C, Audic S, Suhre K, Vestris G, Claverie JM, Raoult D.**
655 2007. Reductive genome evolution from the mother of *Rickettsia*. PLoS Genet **3**:0103–
656 0114.
- 657 61. **Weinert LA, Werren JH, Aebi A, Stone GN, Jiggins FM.** 2009. Evolution and
658 diversity of Rickettsia bacteria. BMC Biol **7**:6.
- 659 62. **Leroy PD, Sabri A, Heuskin S, Thonart P, Lognay G, Verheggen FJ, Francis F,**
660 **Brostaux Y, Felton GW, Haubruge E.** 2011. Microorganisms from aphid honeydew
661 attract and enhance the efficacy of natural enemies. Nat Commun **2**:348.
- 662 63. **Oliver KM, Noge K, Huang EM, Campos JM, Becerra JX, Hunter MS.** 2012.
663 Parasitic wasp responses to symbiont-based defense in aphids. BMC Biol **10**:11–21.
- 664 64. **Vandermoten S, Mescher MC, Francis F, Haubruge E, Verheggen FJ.** 2012. Aphid
665 alarm pheromone: An overview of current knowledge on biosynthesis and functions.
666 Insect Biochem Mol Biol **42**:155-163.
- 667 65. **Vimr ER, Kalivoda KA, Deszo EL, Steenbergen SM.** 2004. Diversity of microbial
668 sialic acid metabolism. Microbiol Mol Biol Rev **68**:132–153.
- 669 66. **Tilman D.** 1982. Resource competition and community structure. Princeton University
670 Press.
- 671 67. **Chaffron S, Rehrauer H, Pernthaler J, von Mering C.** 2010. A global network of
672 coexisting microbes from environmental and whole-genome sequence data. Genome
673 Res **20**:947–959.
- 674 68. **Py B, Barras F.** 2010. Building Fe-S proteins: bacterial strategies. Nat Rev Microbiol
675 **8**:436–446.
- 676 69. **Schmitz-Esser S, Haferkamp I, Knab S, Penz T, Ast M, Kohl C, Wagner M, Horn**

- 677 **M.** 2008. *Lawsonia intracellularis* contains a gene encoding a functional rickettsia-like
678 ATP/ADP translocase for host exploitation. *J Bacteriol* **190**:5746–5752.
- 679 70. **Bertelli C, Greub G.** 2012. Lateral gene exchanges shape the genomes of amoeba-
680 resisting microorganisms. *Front Cell Infect Microbiol* **2**:110.
- 681 71. **Schmitz-Esser S, Tischler P, Arnold R, Montanaro J, Wagner M, Rattei T, Horn**
682 **M.** 2010. The genome of the amoeba symbiont “*Candidatus Amoebophilus asiaticus*”
683 reveals common mechanisms for host cell interaction among amoeba-associated
684 bacteria. *J Bacteriol* **192**:1045–1057.
- 685 72. **Wu M, Sun L V, Vamathevan J, Riegler M, Deboy R, Brownlie JC, McGraw EA,**
686 **Martin W, Esser C, Ahmadinejad N, Wiegand C, Madupu R, Beanan MJ,**
687 **Brinkac LM, Daugherty SC, Durkin AS, Kolonay JF, Nelson WC, Mohamoud Y,**
688 **Lee P, Berry K, Young MB, Utterback T, Weidman J, Nierman WC, Paulsen IT,**
689 **Nelson KE, Tettelin H, O ’neill SL, Eisen JA.** Phylogenomics of the Reproductive
690 Parasite *Wolbachia pipientis* wMel: A Streamlined Genome Overrun by Mobile
691 Genetic Elements. *PLoS Biol* **2**:327-341.
- 692 73. **Zindel R, Gottlieb Y, Aebi A.** 2011. Arthropod symbioses: A neglected parameter in
693 pest- and disease-control programmes. *J Appl Ecol* **48**:864-872.
- 694 74. **Gebiola M, White JA, Cass BN, Kozuch A, Harris LR, Kelly SE, Karimi J,**
695 **Giorgini M, Perlman SJ, Hunter MS.** 2016. Cryptic diversity, reproductive isolation
696 and cytoplasmic incompatibility in a classic biological control success story. *Linn Soc*
697 London **117**:217–230.
- 698 75. **Markowitz VM, Chen I-MA, Palaniappan K, Chu K, Szeto E, Pillay M, Ratner A,**
699 **Huang J, Woyke T, Huntemann M, Anderson I, Billis K, Varghese N, Mavromatis**
700 **K, Pati A, Ivanova NN, Kyrpides NC.** 2014. IMG 4 version of the integrated
701 microbial genomes comparative analysis system. *Nucleic Acids Res* **42**:D560–7.

- 702 76. **Meyer F, Paarmann D, D'souza M, Olson R, Glass E, Kubal M, Paczian T,**
703 **Rodriguez A, Stevens R, Wilke A, Wilkening J, Edwards R.** 2008. The
704 metagenomics RAST server – a public resource for the automatic phylogenetic and
705 functional analysis of metagenomes. *BMC Bioinformatics* **9**:386–384.
- 706 77. **Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A,**
707 **Kyrpides NC.** 2010. GenePRIMP: a gene prediction improvement pipeline for
708 prokaryotic genomes. *Nat Methods* **7**:455–457.
- 709 78. **Moran NA, Bennett GM.** 2014. The Tiniest Tiny Genomes. *Annu Rev Microbiol*
710 **68**:195–215.
- 711 79. **Misof B, Liu S, Meusemann K, Peters RS, Al. E.** 2014. Phylogenomics resolves the
712 timing and pattern of insect evolution. *Science* **346**:763–767.
- 713 80. **Szklarczyk T, Moskal A.** 2001. Ultrastructure, distribution, and transmission of
714 endosymbionts in the whitefly *Aleurochiton aceris* Modeer (Insecta, Hemiptera,
715 Aleyrodinea). *Protoplasma* **218**:45–53.
- 716 81. **R Core Team.** 2012. R: A language and environment for statistical computing. R
717 Foundation for Statistical Computing, Vienna, Austria.
- 718
- 719