1 **Microbial ecology–based methods to characterize the bacterial communities of non-model**
2 **insects**

3 Erica M. Prosdocimi[a], Francesca Mapelli[a], Elena Gonella[b], Sara Borin[a], and Elena Crotti[a*]

4

5 [a]Dipartimento di Scienze per gli Alimenti, la Nutrizione e l'Ambiente (DeFENS), Università degli
6 Studi di Milano, Milano, Italy
7 [b]Dipartimento di Scienze Agrarie, Forestali e Alimentari (DISAFA), Università degli Studi di Torino,
8 Grugliasco, Italy

9

10

11 *corresponding author: Tel. +39.02.50319122; Fax +39.02.50319238.

12

13 E-mail addresses: erica.prosdocimi@unimi.it (Erica M. Prosdocimi); francesca.mapelli@unimi.it
14 (Francesca Mapelli); elena.gonella@unito.it (Elena Gonella); sara.borin@unimi.it (Sara Borin);
15 elena.crotti@unimi.it (Elena Crotti).

16

17

18

19 **Abstract**

20 Among the animals of the Kingdom Animalia, insects are unparalleled for their widespread diffusion,
21 diversity and number of occupied ecological niches. In recent years they have raised researcher
22 interest not only because of their importance as human and agricultural pests, disease vectors and as
23 useful breeding species (e.g. honeybee and silkworm), but also because of their suitability as animal
24 models. It is now fully recognized that microorganisms form symbiotic relationships with insects,
25 influencing their survival, fitness, development, mating habits and the immune system and other
26 aspects of the biology and ecology of the insect host. Thus, any research aimed at deepening the
27 knowledge of any given insect species (perhaps species of applied interest or species emerging as
28 novel pests or vectors) must consider the characterization of the associated microbiome. The present
29 review critically examines the microbiology and molecular ecology techniques that can be applied to
30 the taxonomical and functional analysis of the microbiome of non-model insects. Our goal is to
31 provide an overview of current approaches and methods addressing the ecology and functions of
32 microorganisms and microbiomes associated with insects. Our focus is on operational details, aiming
33 to provide a concise guide to currently available advanced techniques, in an effort to extend insect
34 microbiome research beyond simple descriptions of microbial communities.

35

36 **Keywords**

37 Insect microbiome, bacterial symbiont cultivation, molecular methods, non-model insects, bacterial
38 community, microbial ecology

39

40

41 ## 1. Introduction

42 There is increasing awareness of the importance of microbial symbionts for animal physiology, and
43 the number of studies focusing on non-pathogenic bacteria hosted by a wide range of organisms is
44 also on the increase. However, for decades, research has been focused primarily on pathogenic
45 bacteria, where most of the animal-associated microorganisms are most likely harmless, variable and
46 acquired from the environment (Turnbaugh et al., 2007). In any case, a number of beneficial bacteria
47 have been found, and among these is a group of essential, vertically transmitted endosymbionts that,
48 with their host, form an inseparable holobiont (Moran et al., 2008).
49 Insects have been investigated by microbiologists not only as models for human-microbiome
50 interactions, but also for their importance as pests and disease vectors. The most studied bacteria-
51 insect associations fall roughly into two groups, i.e. heritable symbionts (including primary and
52 secondary symbionts) (Moran et al., 2008), and gut symbionts. Primary (P) symbionts are bacteria
53 necessary for insect survival and/or reproduction, and they inhabit highly specialized cells –
54 bacteriocytes - that lie dispersed in the gut epithelium or be grouped within specialized organs called
55 bacteriomes. They typically have a reduced genome, and share a long evolutionary history with the
56 host as they propagate only through maternal transmission. This kind of symbiosis is common in
57 insects with poor diets, such as aphids, which feed exclusively on phloem sap and host a
58 gammaproteobacterium of the genus *Buchnera* for the synthesis of the amino acids lacking in their
59 diet (Akman Gündüz and Douglas, 2009). Instead secondary (S) symbionts are not essential for host
60 survival, although they can improve host fitness. They colonize various cells and organs, including
61 hemolymph, and are able to infect new hosts, establishing stable associations with them through
62 maternal transmission (Favia et al., 2007). This category also includes reproductive manipulators with
63 maternal transmission that spread throughout the population, promoting the reproduction of infected
64 females through daughters. This is accomplished through cytoplasmic incompatibility,
65 parthenogenesis, male feminization and son-killing (Stouthamer et al., 1999).

Gut symbionts have been extensively reviewed by Engel and Moran (Engel and Moran, 2013), and it has been found that most are commensals that reside in the gut, being neither clearly harmful nor beneficial to the host (Dillon and Dillon, 2004). Our use of the term "commensals" refers to a broad range of microorganisms that vary greatly, even among members of the same species. These bacteria are generally acquired from both the environment and the diet, their selection depending on the chemical and physical conditions inside the gut, such as pH, oxygen availability, and retention time of the food bolus. Furthermore, the host immune system plays an active role in bacteria selection as it is elicited by specific bacterial features, e.g. the excretion of uracil (Lee et al., 2013). Despite the extreme variability of this type of microbial consortium, there is increasing evidence that commensals can critically affect the host physiology, acting on the immune system (Lee et al., 2013), on larval development (Shin et al., 2011) and even on the choice of mate (Sharon et al., 2010). However, these effects, though sometimes important, cannot clearly classify a microorganism as "mutualistic" or "pathogenic" (Dillon and Dillon, 2004). For example, in many cases just the presence of commensal microflora can prevent pathogen colonization (Ryu et al., 2008): in fact, any alteration in the bacterial community due to pathogen colonization can lead to "dysbiosis", which is detrimental for the host (Hamdi et al., 2011). Moreover, there are known cases of specialized gut symbionts where the relationship with the host resembles primary symbiosis (genome shrinkage, strict heritability) (Hosokawa et al., 2006). Indeed, in Hemiptera, the vertical transmission of gut symbionts smeared on the eggs or encased in symbiont capsules, is well studied (Fukatsu and Hosokawa, 2002; Kikuchi et al., 2011; Prado et al., 2006; Sudakaran et al., 2012).

Considering the crucial roles of the above listed symbionts, a descriptive study of any insect cannot disregard the characterization of the bacterial community associated with it. Thus, research on emerging pests, disease vectors, and beneficial insects (e.g. honeybees and silkworms) should include microbiological screening. Apart from achieving a more complete image of the physiology of the insect itself, such a survey could lead to the development of useful products and applications, such as probiotics and pest and disease control strategies (Bourtzis et al., 2014; Crotti et al., 2012; Jurkevitch,

92 2011). With an eye towards potential technological developments, culture-dependent assays must not
93 be underestimated. Indeed, though far less powerful than molecular methods in describing bacterial
94 communities, such assays allow the manipulation of single cultivable strains. The aim of this review
95 is to give researchers approaching this kind of study an overview of the microbiological methods
96 available (figure 1), together with information on the practical considerations necessary for their
97 application to insects and for a correct interpretation of the results. In addressing scientists new to this
98 field, the paper focuses mainly on the best-known and most feasible research techniques.
99 Nevertheless, some advanced genetic analysis techniques will also be presented in order to promote
100 possible progress studies beyond the simple taxonomic characterization of bacterial communities.
101

102 ## 2. Sampling

103 The gathering of suitable insect specimens for microbial community analysis can be challenging,
104 especially for insects not reared in the laboratory. Depending on the target species and its habitat, the
105 field season is often restricted to a short period of the year and the collection and identification of
106 samples requires time and experience. Therefore, it is important to start with a precise plan of the
107 number and type of individuals required.

108 ### 2.1 Specimen choice

109 The microbiome composition of insects of the same species is largely dependent on ecological
110 parameters such as temperature, geographical location and, above all, diet. Indeed, the qualitative and
111 quantitative differences in the gut community composition of insects reared on different diets has
112 been reported in different papers over the years (Broderick et al., 2004; Chandler et al., 2011; Colman
113 et al., 2012; Kane and Breznak, 1991; Montagna et al., 2015; Santo Domingo et al., 1998).. In some
114 cases, insects of highly divergent taxonomical groups that feed on the same substrate can harbour
115 very similar microbial communities, regardless of the host taxonomy that normally shapes the
116 microbiome; this can be seen in some species that live on decaying wood (Colman et al., 2012). Also
117 to be noted is that the microbiome found in the gut of wild insects is generally more diverse than that

118 of reared insects, though studies on reared *Drosophila melanogaster* show that also laboratory
119 populations host different bacterial species that can exert important and similar physiological roles
120 (Douglas, 2011; Shin et al., 2011; Storelli et al., 2011). However, wild *Drosophila* gut microbiomes
121 are much more complex, the most diverse communities being present in flower and mushroom feeders
122 (Chandler et al., 2011). Finally, it is important to note that most standard diets used in insect stocks
123 contain preservatives and/or antibiotics; these avoid the proliferation of microorganisms in the diet
124 itself but obviously have an impact on the natural microbiome of the insects. Even the distribution of
125 secondary endosymbionts such as *Wolbachia*, *Spiroplasma*, and *Rickettsia* can depend on ecological
126 factors such as diet, temperature and snowfall, or even on the co-occurrence of other symbionts (Toju
127 and Fukatsu, 2011). The reproduction-manipulation strategy of endosymbionts like *Wolbachia* can
128 rapidly spread the bacterium across a population, resulting in a patchy distribution. Moreover, this
129 behaviour can influence the frequency of other vertically-transmitted symbionts through a hitchhiking
130 effect (Toju and Fukatsu, 2011). On the contrary, primary symbionts are, by definition, present in all
131 the specimens of one species as they establish a tight interdependence with the host.
132 For all these reasons, it may be necessary, according to the aim of the research, to sample insects of
133 different populations. If the primary goal is to describe the gut bacterial community of an insect it is
134 better to analyse wild samples, rather than rely on laboratory stocks. Rearing insects on an artificial
135 diet, even for short periods, can lead to important biases (Hammer et al., 2014). Moreover, to explain
136 possible differences in the specimens' gut bacterial composition it is advisable to monitor ecological
137 parameters at the collection site, at least of the substrate on which the insects are feeding.
138 Instead, to survey the presence/infestation rate of endosymbionts in a species, the sampling sites
139 should be as geographically diverse as possible. If the infestation rate is 100% across genetically and
140 spatially separate populations, the hypothesis of a primary symbiosis should be verified.

141 2.2 Target organs

142 When choosing target organs, specimen size plays an important role as dissection can be challenging:
143 while addressing single organs in smaller insects is very difficult, when studying larger insects the

144 analysis of single organs is a better choice. In fact, hard cuticles and high quantities of biomass can
145 impair DNA extraction. Moreover, the bacterial cells in the gut usually outnumber those in any other
146 body part. Thus, examining the insect as a whole can give results similar to sampling only the gut
147 (Sudakaran et al., 2012). Therefore, at least in the case of "dissectible" insects, it is useful to know
148 which organs are more likely to host symbionts (figure 2).

149 *2.2.1 Gut.* The gut of insects can be divided into at least three tracts, the foregut, midgut and hindgut,
150 that can host different bacterial communities. In the foregut and hindgut, which are of ectodermal
151 origin, a cuticle layer of chitin separates the epithelium from the lumen, impeding nutrient absorption,
152 making the midgut the main site where nutrient absorption takes place. Here, in many insects, a
153 peritrophic matrix is secreted by the epithelial cells. This forms a permeable envelope separating the
154 bolus and the gut bacteria from the epithelium, allowing nutrient transit. This matrix is in a continuous
155 state of being shed and replaced. In the anterior hindgut the Malpighian tubules (insect excretory
156 organs) deliver waste products, such as uric acid, collected from the body cavity into which they
157 extend. Therefore, this area in the insect gut environment, where there is a mixture of nitrogen and
158 food waste, differs from that of vertebrates where nitrogen waste is separated out. Depending on the
159 species and life stage, these three tracts can assume very different shapes and comprise diverticula or
160 extremely specialized compartments in which specific bacteria are hosted, as in the case of termites
161 (Köhler et al., 2012). Microsensor measurements in the beetle *Pachnoda ephippiata* show that the
162 redox condition and the pH can vary sharply along the gut: the bacterial community composition
163 varies accordingly (Lemke et al., 2003). Therefore, knowledge of the differences between the
164 bacterial communities in the three tracts can be critical in understanding the gut physiology, and such
165 knowledge can be achieved through dissection. Nevertheless, dissecting the gut prior to analysis is a
166 delicate step: special attention is needed to avoid the leakage, or the mixing, of the gut content of the
167 different parts.

168 *2.2.2 Bacteriomes.* Should the target of the research be primary symbionts, the presence of
169 bacteriomes must be investigated. In these organs, bacterial symbionts are hosted in specialized cells

170 named bacteriocytes. Bacteriomes are often associated with the external midgut ((tsetse flies
171 (Balmand et al., 2013), louse (Perotti et al., 2007)), appearing as whitish round shaped bodies; their
172 first observation can be traced back to the 17th century (Hooke, 1667). Bacteriocytes can also be
173 located in the fat body, where they can be identified by microscopy because of their cytoplasm, which
174 is densely populated with bacteria (Toenshoff et al., 2014), or intercalated in midgut tissue (Stoll et
175 al., 2010). Separating bacteriomes from gut tissue without breaking the gut epithelium can be difficult
176 in smaller animals, however, the detection of endosymbionts by molecular means, such as diagnostic
177 PCR, allows the analysis of the whole insect, thus avoiding this step (see section 4.5).
178 *2.2.3 Gonads.* Reproductive manipulators, as well as maternally transmitted mutualists, colonize
179 ovaries where they infect trophocytes and oocytes, through which they are transmitted to the progeny
180 (Stouthamer et al., 1999; Veneti et al., 2004, 2003). Moreover, in the arthropod *Ixodes ricinus*, a hard
181 tick, an intra-mitochondrial endosymbiont that almost exclusively inhabits the ovaries, is harboured
182 (Sassera et al., 2006). Testes can also be colonized, enabling paternal transmission to the mating
183 female and thus to the offspring, as in the case of *Asaia* and *Anopheles* mosquitoes (Crotti et al.,
184 2009).
185 *2.2.4 Salivary glands.* The presence of endosymbionts has been investigated in the salivary glands
186 of different arthropods, such as blood-sucking insects and ticks. *Asaia* colonizes the salivary glands
187 of mosquitoes, co-localizing with the malaria vector *Plasmodium* spp., if present (Favia et al., 2007).
188 The presence of *Midichloria mitochondrii* has been demonstrated in the salivary glands of *Ixodes*
189 *ricinus* (a tick), which inoculate humans in sufficient amounts to stimulate antibody production
190 (Mariconti et al., 2012). These tiny organs have been mainly examined by Fluorescent In Situ
191 Hybridization (Gonella et al., 2011) (FISH, a technique that it will be discussed in section 6.1).
192 *2.2.5 Cuticles and Antennae.* A peculiar type of symbiotic relationship exists between ants and
193 bacteria of the genus *Streptomyces*. The bacterium occupies specific locations on the bacterial cuticle
194 and produces an antifungal compound that suppresses the growth of parasitic *Escovopsis*, which
195 threatens the fungal gardens grown by the ants (Currie et al., 1999). Another bacterium of the same

196 genus plays a similar role in a symbiotic relationship with beewolves of the genus *Philantus*
197 (Kaltenpoth et al., 2005). Adult wasps carry symbionts in specialized glands in the antennae, and
198 females secrete it in the brood chamber in which they lay eggs. Thanks to the antibiotic activity of
199 the *Streptomyces* strain, this behaviour in the wasp protects the larva and cocoon from bacterial
200 pathogens, which could grow on the prey that is larval food (Kaltenpoth et al., 2005).

201 *2.2.6 Other locations. Wigglesworthia glossinidia*, symbiont of the tse-tse fly, is transmitted to
202 offspring through the milk gland. The secretion of this gland feeds larvae that develop inside the
203 mother's uterus, until they reach the third instar (Balmand et al., 2013).

204 **2.3 Life stages**

205 Insects are characterized by a unique development, thanks to which they can change dramatically in
206 morphology and lifestyle, according to the life stage. Within the largest insect class, the Pterigota,
207 there are two main types of postembryonic development, as discussed in the following paragraphs.

208 *2.3.1 Exopterigota (heterometabolous insects).* In the exopterigota, the nymph hatched from the egg
209 is similar in morphology to the adult. To reach the adult stage, it undergoes a series of moults,
210 increasing its body mass without changing substantially in shape. Within this group, modification of
211 the gut bacterial community during development has been studied mostly in Hemiptera. In the
212 European firebug, *Pyrrhocoris apterus,* the gut microbiome appears to be stable throughout several
213 moulting steps, the presence of the most represented bacterial taxa already being within the egg
214 (Sudakaran et al., 2012). However, as confirmed also in another member of the same order, *Riptortus*
215 *pedestris,* the colonization of the insect by a specific bacterial community is not completed before the
216 third nymphal instar (Kikuchi et al., 2011). During the moults, an increase in the production of
217 antimicrobial peptides causes a drop in the bacterial load of *R. pedestris* midgut (Kim et al., 2014). A
218 possible explanation is the vulnerability of the moulting insect to injury and pathogens, which leads
219 to an up-regulation of the immune system.

220 Regarding primary endosymbionts, it has been demonstrated in aphids that the development of the
221 bacteriome occurs in the early embryonic stages, even in the absence of symbiotic bacteria

222 (aposymbiotic insects) (Braendle et al., 2003). This particularly intimate symbiotic relationship is not
223 affected in any way by insect development as the two organisms can be regarded as a single holobiont
224 that develops as a whole.

225 *2.3.2 Endopterigota (holometabolous insects).* Insects belonging to the endopterigota are born as
226 larvae, and are subject to a complete metamorphosis prior to the adult stage. The extraordinary
227 evolutionary success of holometabolous insects has been attributed to the differentiation of the food
228 sources utilized by larvae and adults, which avoids competition between juvenile and mature
229 conspecifics. This differentiation is reflected in the composition of the gut community. Generally, the
230 pupation phase implies a simplification and reduction of the gut microbiome, which, nevertheless, is
231 not completely erased. After that, as the emerging adult starts to feed, bacteria start to grow again,
232 comprising on one hand the species that survived the pupation and, on the other, a new species that,
233 according to the difference in diet, may or may not be similar to the ones in the larvae (Arias-Cordero
234 et al., 2012; Hammer et al., 2014). A particular case is the honeybee, *Apis mellifera*, in which, due to
235 the cleaning behaviour of nutrices, the larvae are regarded as an almost-sterile environment, and the
236 worker bee acquires a characteristic microflora as it emerges from the brood cell through two
237 mechanisms: trophallaxis and contact with the hive (Martinson et al., 2012).

238 The maternal transmission of endosymbionts, present from the first larval stages to the reproducing
239 adults, indicates that metamorphosis does not affect the presence of the bacteria in the bacteriomes
240 throughout the life cycle. Nevertheless, bacterial load can change with time. The dynamics of
241 bacteriocytes during metamorphosis has been studied in the carpenter ant *Camponotus floridanus*,
242 hosting the mutualistic symbiont *Blochmannia floridanus*. In these ants there is a surprising increase
243 in both bacterial load and number of bacteriocytes in the midgut epithelium during metamorphosis,
244 starting in the last larval instar (Stoll et al., 2010). During metamorphosis *Blochmannia* appears to
245 also colonize non-bacteriocyte cells, while in adults it gradually decreases.

246 In conclusion, if surveying the presence of maternally-transmitted symbionts, the life stage of the
247 specimens can be neglected. However, if dealing with gut bacteria it is necessary to pay attention to
248 this feature, especially in the case of holometabolous insects.
249

250 3. Sample storage and dissection

251 3.1 Storage

252 The fate of each collected specimen must be determined at the very beginning of sampling. In fact,
253 different techniques require different methods of sample storage.
254 Specimens for molecular analyses are usually stored in ethanol and kept at -20°C. When cooperating
255 with unqualified personnel or volunteers during sampling campaigns, it is especially important to
256 underline that denatured ethanol is not suitable for sample preservation. To ensure the permeation of
257 the body of larger insects with ethanol, one or more legs can be removed. If dissection is required, it
258 is better to perform it before the addition of ethanol and store separate organs. In fact, tissue
259 dehydration caused by ethanol can hamper dissection, even after rehydration in saline.
260 Specimens dedicated to bacterial isolation should be analysed immediately. It must be noted that
261 feeding the insects on artificial diets, even for a few days, can dramatically change their gut
262 communities due to the fast growth of specific bacterial contaminants on diet material (Montagna et
263 al., 2015). For this reason, the best option when immediate culture-based analysis is not feasible is to
264 collect live insects together with their native feeding substrate, and to analyze them as soon as
265 possible.

266 3.2 Dissection

267 Prior to dissection, the insects should be killed, or at least anesthetized. Although studies on
268 nociception in *Drosophila* are still in the early stages, there is no doubt that insects show a pain
269 response to mechanical, thermal and chemical noxious stimuli (Im and Galko, 2012). A well-
270 established entomologic technique for killing insects is to saturate the atmosphere in their cage with
271 chloroform or ether. This technique is suitable for molecular studies, as it is not likely to damage the

272 nucleic acids. However, to the best of our knowledge there is no knowledge of the possible effect of
273 these gases on living bacterial symbionts inside the insect; therefore, for specimens dedicated to
274 bacterial isolation, some authors prefer to use anesthetisation (Arias-Cordero et al., 2012) or sacrifice
275 (Sudakaran et al., 2012) by exposure to low temperatures (ice or refrigerator). To avoid contamination
276 by bacteria attached to the cuticle, which could be influenced by the environment and manipulation,
277 it is recommended to sterilise the surface as a first step before the dissection, or before the smashing
278 for bacterial isolation purposes. In the washing procedure, the specimen is rinsed a few times in water
279 or ethanol (Arias-Cordero et al., 2012; Zouache et al., 2011); detergents, like SDS solutions, can also
280 be used (Sudakaran et al., 2012). If the ethanol and detergents permeate the insect body it could affect
281 the viability of the microbial symbionts. For this reason, the timing of the exposure to these chemicals
282 should be evaluated in relation to the insect's body mass: for big beetles with thick cuticles this
283 observation is irrelevant, while it can be important for small flies and mosquitoes. In any case, insect
284 ingestion of chemicals should be avoided.
285 The dissection of larger insects usually starts with the removal of the cuticle from the abdomen. At
286 this point, with some experience, it is possible to visualize the native arrangement of the gut and other
287 organs in the abdominal cavity (Ceja-Navarro et al., 2014). Organs of interest can then be removed.
288 For larvae, the cuticle is usually cut along the side of the body, and all the fat tissue is gradually
289 removed to uncover the gut. From small insects, like mosquitoes or *Drosophila* flies, it is still possible
290 to obtain entire guts by gently pulling the head with pincers until the gut slides out of the body,
291 attached to the mouth parts.
292 The dissection of the gut of wild animals can be far more challenging that the dissection of specimens
293 reared on artificial diets. In fact, the latter usually have larger amounts of good quality food available,
294 resulting in a bigger and more swollen gut that is clearly distinguishable from the surrounding tissues.
295 It is important to note that it is impossible to surface-sterilize single organs. Therefore, these organs
296 can be contaminated by bacteria from the haemocoel and abdominal cavity.
297

298 **4. Culture-independent methods for community characterization**

299 The biggest advantage of applying molecular methods is the detection of non-cultivable species,
300 which outnumber cultivable ones in almost every environment. Indeed, insect primary endosymbionts
301 are often intrinsically not cultivable, making the molecular approach the only available choice.
302 Furthermore, samples for molecular analyses can be collected throughout the field season and stored
303 in ethanol without affecting the analysis, whereas the cultivation-based approach requires the
304 immediate processing of the collected insects, restricting the time available for microbiome study. As
305 is known, PCR-based methods are affected by multiple factors that can influence the reliability of
306 microbial community characterization, as reported below.

307 **4.1 DNA extraction.**

308 The first critical step is the extraction of DNA. In soil, different extraction methods result in markedly
309 different community profiles (de Lipthay et al., 2004; Feinstein et al., 2009). Although soil is
310 considered one of the most "difficult" substrates for DNA extraction, even for insects some critical
311 factors must be taken into account.

312 Extraction protocols start with lysis, for which a combination of mechanical, enzymatic and chemical
313 treatments is used. In the first step, for entire insects, especially those with hard cuticles, liquid
314 nitrogen freezing and ceramic pestle-smashing is the most popular option (Sudakaran et al., 2012),
315 while single organs (e.g. gut) can be smashed in saline using a small, tube-fitted pestle (Crotti et al.,
316 2009). Bead beating or sonication is also reported (Broderick et al., 2004). However, aggressive
317 beating is not generally required to detach bacteria from insect tissues, and even a single brief
318 centrifugation can be sufficient to separate most bacteria from gut cells and other debris (Santo
319 Domingo et al., 1998). Instead, to achieve enzymatic lysis, proteases, lysozymes or other lytic
320 enzymes are added. Indeed, it is advisable to carry out lysozyme treatment as it promotes the release
321 of nucleic acids from the gram-positive, which, due to their harder cell wall, can have a lower
322 extraction yield. The use of specific enzymes to extract DNA from particular gram-positives is also
323 reported: for example, Nikodinovich et al. (Nikodinovic et al., 2003) recommend a chromopeptidase

324 to improve extraction from Streptomycetes. Finally, chemical lysis disrupts cell membranes and
325 releases DNA: this is particularly important for endosymbionts, which are enclosed in bacteriocytes.
326 For this purpose, the sample is treated with detergents such as CTAB (cetyltrimethylammonium
327 bromide) or SDS (sodium dodecyl sulfate). Sometimes, EDTA (ethylenediaminetetraacetic acid) is
328 added during mechanical or chemical lysis to impair DNAses naturally present in the sample.
329 Enzymatic lysis must be performed before the addition of detergents because these latter denature
330 protein, preventing enzymatic activity.
331 After lysis and the release of nucleic acids from the cells, the DNA must be purified. Two main
332 strategies are used to accomplish this step: i) chloroform-isoamyl alcohol extraction followed by
333 isopropanol precipitation of DNA, ii) binding of the DNA to a silica membrane in the presence of
334 chaotropic salts: the most commercial extraction kits rely on this method (e.g. MoBio Power Soil,
335 Qiagen Blood&tissue). The first protocol, described in Zouache et al. (2011), is usually cheaper but
336 is labour-intensive. When dealing with insects from soil, humic acids may also be co-extracted with
337 this procedure. With regard to the second method, there are a number of kits and protocols available.
338 For example, the popular Qiagen Blood & Tissue kit (Qiagen GmbH, Germany) comprises a protocol
339 that is specific for insects. In this case, it is important to point out that the goal of this kit is to extract
340 DNA from the insect itself, rather than from the bacteria, therefore at least one lysozyme pre-
341 treatment should be added. If humic acid inhibition is a problem (as can be the case for insects that
342 live or feed on ground substrates), a good choice could be the commercial kits developed for DNA
343 extraction from soil. These kits have also been used for insects that live above ground (Hammer et
344 al., 2014). If using such a commercial kit, careful attention should be paid to cuticle grinding or
345 removal as cuticle can easily clog the microcolumns.
346 It is advisable to test the presence and quality of bacterial DNA in the extraction product by direct
347 PCR amplification of the 16S rRNA gene, using the universal primer for bacteria rather than by
348 measuring the DNA concentration spectrophotometrically or by gel electrophoresis. Moreover, in this
349 case, a high DNA yield from the extraction would not reflect a high amount of bacterial DNA as the

350 insect genome is always co-extracted. Co-extraction of the host genome and the use of the mixed
351 DNA template in PCR reactions implies two pitfalls that will be discussed in detail below: the
352 unspecific amplification of the host 18S rRNA gene (paragraph 4.2) and the amplification of bacterial
353 genes that are encoded in the host genome due to horizontal gene transfer (paragraph 4.5).

354 **4.2 Primer choice**

355 When amplifying the 16S rRNA gene, the critical points of a good primer choice are three: i)
356 universality, ii) variability of the amplified region, and iii) specificity for bacteria.

357 4.2.1 *Universality*. Although the literature describes a wide range of primers on the 16S rRNA gene,
358 none is truly universal, at least not without degeneracy. Several studies have tested, for their
359 universality, primer pairs in silico (Baker et al., 2003; Wang and Qian, 2009), proposing new couples.
360 Since it is impossible to identify the perfect couple, it is still important to ensure that the chosen
361 primer pairs are suitable for the given research purpose. Thus, the primer pairs should be tested in-
362 silico for their universality using online tools such as the RDP ProbeMatch
363 (http://rdp.cme.msu.edu/probematch/search.jsp) or Silva TestPrime (http://www.arb-
364 silva.de/search/testprime/).

365 4.2.2 *Variability of the amplified region*. Given that many molecular analyses, including pyrotag (see
366 paragraph 4.5) and DGGE (4.3), are performed on amplicons of a maximum of 500 base pairs, the
367 analysis must be restricted to one or two variable regions of the 16S rRNA gene. For this purpose,
368 the most common choice is to select the V1-V3 or V5-V6 regions as major variability lies in the V1,
369 V3 and V6 regions (Andersson et al., 2008). Yarza and colleagues (Yarza et al., 2014) carried out an
370 in-depth analysis, and showed that only by combining two or more variable regions can the full
371 recovery of all the taxonomic ranks present in the sample be achieved.

372 *4.2.3 Specificity for bacteria*. When working with mixed eukaryotic-prokaryotic DNA, the risk of
373 unspecific amplification should not be underestimated. In any DNA solution extracted from an insect
374 there is a mix of bacterial DNA and DNA from the insect and its food source (e.g. from the prey, or
375 from plant material). Some of the priming regions of 16S rRNA are conserved in plastid DNA or in

376 eukaryotic 18S rRNA (Huys et al., 2008; Prosdocimi et al., 2013). In some cases, the size of the
377 amplicon from eukaryotic DNA is different from the size of the prokaryotic amplicon, allowing
378 separation by means of gel electrophoresis. When the sizes are similar, the problem of unspecific
379 amplification is difficult to detect. However, due to the higher proportion of eukaryotic DNA in the
380 source sample, the bacterial amplicons will be underrepresented in the PCR product. To check primer
381 pairs for this problem, PrimerBlast is a useful tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast/)
382 although it does not support primers with degeneracy. If it is not possible to find a suitable primer
383 pair, one option is the nested-PCR approach: first, PCR amplification is performed with the primers
384 27F – 1492R, which are highly specific for bacteria and amplify the whole 16S rRNA gene; the result
385 of this amplification is used as the template for the second PCR with the 16S primers of choice. To
386 reduce the bias that comes from the double amplification (see following paragraph) the number of
387 cycles in the first PCR should be reduced as much as possible. Another option is the design of a
388 specific blocking primer (Vestheim and Jarman, 2008), a strategy that can require great effort for
389 optimization.

390 **4.3 Denaturing Gradient Gel Electrophoresis (DGGE).**

391 The DGGE method is based on the separation of the different amplicons generated by a 16S rRNA
392 PCR on a denaturing gradient polyacrylamide gel (Muyzer et al., 1993). DGGE band profiles are
393 immediately and visually comparable, giving the researcher an idea of the complexity and variability
394 of the bacterial community in many insect samples at the same time. This is useful in many cases: i)
395 to directly compare individuals across time-scales, or different breeding conditions, or to test if the
396 examined factor influences the bacterial community, and ii) to test for the presence of recognisable
397 patterns of bands common to one group of samples. Compared to pyrotag, this method requires more
398 bench work and optimization, but less bioinformatics analysis. However, it is possible to analyse the
399 gel data both qualitatively and quantitatively with both proprietary software (e.g. Bio Rad Quantity
400 One, Bio Rad Laboratories Inc., Hercules, California) and open-source software (ImageJ) if the
401 appropriate markers are included.

402 If we exclude special cases, such as the highly bio-diverse gut of termites, the DGGE profile obtained
403 from an insect consists of less than 20 bands (Marzorati et al., 2006; Raddadi et al., 2011; Rizzi et al.,
404 2013). For the first trial, the denaturing gradient can be set to 40%-60% of denaturing agents. Bands
405 originating from endosymbionts, whose genomes are usually rich in A and T, are mainly located on
406 the top of the gel. Thus, when looking for endosymbionts, it is advisable to choose a denaturing
407 gradient, starting from low denaturing agent percentages (<40%) to separate the bands better. In the
408 case of plant-feeding insects, even if nested PCR has been performed to improve amplification
409 specificity, bands corresponding to chloroplast can appear in the middle part of the gel. If *Wolbachia*
410 sequences are retrieved, it is important to note that some insects harbour long stretches of this
411 bacterium DNA in their genomes, due to horizontal gene transfer (HGT) events (Brelsfoard et al.,
412 2014; Doudoumis et al., 2012; Dunning Hotopp et al., 2007, see paragraph 4.5). These HGT events
413 have not yet been studied extensively, and could involve other bacteria besides *Wolbachia* (Dunning
414 Hotopp et al., 2007).

415 **4.4 Diagnostic PCR.**

416 This technique consists in directly testing the presence of a specific symbiont within an insect
417 individual, performing the PCR using specific primers targeting the symbiont. Research to date has
418 produced a long list of insect endosymbionts, many of which are restricted to certain taxa (especially
419 P-symbionts), while others are widespread through different orders (reproductive manipulators). This
420 knowledge has been exploited to design a number of specific primers that can be used to detect
421 symbiotic bacteria. This is the easiest kind of screening to assess the presence of known primary
422 symbionts in newly studied insects. Moreover, it is useful to assess the distribution of S-symbionts
423 across species, sampling sites and populations (Russell et al., 2012). Specific primers can also be used
424 to set up qPCR assays, which are the most reliable method to quantify the presence of uncultivable
425 symbionts, whereas other molecular methods give only an approximation of the contribution of a
426 single species to the overall bacterial community (Amend et al., 2010). The literature reports a number
427 of diagnostic primers for the best known symbionts: e.g. the list in Russell et al. (Russell et al., 2012).

428 As an alternative, it is possible to directly design specific primers from the 16S rRNA gene sequence
429 of the target bacterium. Primer specificity must be validated in silico, as discussed in paragraph 4.2,
430 and in control PCR reactions with both positive and negative control template DNA extracts.
431 The diagnostic PCR method has allowed researchers to shed light on the phenomenon of HGT from
432 bacteria, particularly *Wolbachia*, to their insect hosts. Antibiotic-treated insects, tse–tse flies
433 (*Glossina morsitans*), resulted positive for *Wolbachia,* and extensive insertions of *Wolbachia* genome
434 were found in the insect chromosomes (Brelsfoard et al., 2014; Doudoumis et al., 2012). Insertions
435 of the *Wolbachia* genome are also present in the genome of *Drosophila ananassae*, of wasps of the
436 genus *Nasonia*, and in the nematodes *Brugia malayi* and *Dirofilaria immitis* (Dunning Hotopp et al.,
437 2007). These insertions appear to be more common than previously thought, and could invalidate the
438 results of diagnostic PCR screenings, as well as other sequence-based surveys, giving false positives.

439 **4.5 Next Generation Sequencing of the 16S rRNA gene (pyrotag).**

440 In recent years, methods based on Next Generation Sequencing (NGS) have become the main tool to
441 investigate microbial communities, and are becoming increasingly popular. In fact, one of the
442 principal applications of NGS platforms is the sequencing of 16S rRNA amplicons, which is possible
443 from almost all kinds of environmental sample.
444 *4.5.1 Platform choice.* For sequencing, different devices can be chosen; the most common are
445 Roche 454FLX, the first pyrosequencing platform available, or Illumina (HiSeq or MiSeq), which is
446 rapidly overtaking 454FLX for its lower costs and higher sequencing depth. Indeed, the trend is
447 towards increasing usage of Illumina, and there are already papers investigating insect communities
448 with this method (Hammer et al., 2014). In this context, Illumina technology can be successfully
449 applied to compare bacterial communities from many different insects, and assess microbiome
450 modifications that occur with change in environmental parameters, populations or species. However,
451 454FLX is more suitable for in-depth studies of the composition of a single insect's microflora,
452 eliciting information about function as it produces longer readouts than Illumina technology. For this
453 purpose, the bioinformatics tool PICRUSt can be used (Langille et al., 2013; Montagna et al., 2015).

454 *4.5.2 Number of sequences required. Pilot studies.* The number of sequences per sample is largely
455 dependent on the microbial community therein. Hence, the most accurate way to determine the
456 minimum number of sequences is to perform a pilot study, and inspect the rarefaction curve that
457 correlates the number of taxa found to the sequencing effort. Different types of rarefaction curves can
458 be obtained easily by following the Qiime alpha-diversity pipeline (for Qiime see paragraph 4.5.3).
459 A pilot study to assess community diversity can also be performed by low resolution methods such
460 as DGGE. In fact, in some species the bacterial composition of different individuals can vary greatly.
461 For example, when analysing whole insects, it should be noted that sometimes primary and secondary
462 symbionts, particularly *Wolbachia*, can be present in very high loads, compared to other bacteria, thus
463 "shadowing" the remaining diversity. For the same reason, this kind of preliminary analysis can be
464 useful to decide if it is worth pooling small insects together or not. In this respect, it must be
465 remembered that pooling always implies a loss of information on individual variability, which can
466 lead to data misinterpretation.
467 *4.5.3 Data analysis.* One of the most popular platforms for data analysis is Qiime (Quantitative
468 Insights in Microbial Ecology, www.qiime.org) (Caporaso et al., 2010), which requires a Linux
469 operating system and basic knowledge on related command line writing. The Qiime workflow
470 comprises all the basic operations to de-noise and filter data, and to cluster the sequences into OTUs
471 with a taxonomic assignation. Moreover, a script included in the pipeline (ChimeraSlayer) excludes
472 the chimeric OTUS from the analysis (chimeras are DNA sequences composed of mixed DNA from
473 two or more parents, produced as artefacts of the PCR reaction). Finally, it estimates diversity within
474 samples (alpha) and between samples (beta). The Qiime analysis pipeline can be run as either an
475 almost totally automated way (using default parameters) or highly customized. The software is open-
476 source and can be modified by anyone experienced in python programming: detailed tutorials are
477 available on the website. An important note for insect related data regards the use of databases: Silva
478 databases comprise eukaryotic 18S sequences, allowing the detection of possible insect sequences in
479 the sample (originated by poor primer specificity, as seen before); instead Greengenes databases are

480 exclusively prokaryotic, eukaryotic sequences not being recognised and thus shown as unassigned.
481 An alternative to Qiime is Mothur (http://www.mothur.org). Furthermore, a conceptually different
482 method for the analysis of sequence data, based on the idea of "phylogenetic entropy", was recently
483 applied to the analysis of the bacterial community of a mite, with the web tool PhyloH (Sandionigi et
484 al., 2014).

485 A valuable instrument to reconstruct phylogenies, and identify the species from sequences, is ARB
486 (Ludwig et al., 2004). This software package is available on the Silva website (http://www.arb-
487 silva.de) and requires a Linux operating system, although it is not run by command line like Qiime.

488 **4.6 Phylogenetic micro arrays**

489 The product of a PCR amplification of the 16S rRNA gene from an insect sample can also be analysed
490 through the hybridization of a phylogenetic array. This application is reviewed, among other
491 metagenomics approaches, by Nikolaki and Tsiamis (Nikolaki and Tsiamis, 2013). Like other arrays
492 designed for gene expression, phylogenetic arrays contain a wide panel of DNA probes; in this case,
493 the probe design is based on the 16S rRNA genes of known bacterial families. After hybridization, it
494 is possible to recognise the taxonomic groups within the sample by checking which probes effectively
495 bind to matching amplicons in the PCR product. The most comprehensive phylogenetic array
496 platform is the PhyloChip (Affymetrix, Santa Clara, California), which has been used in a variety of
497 microbial ecology studies of different environments (Hazen et al., 2010; Kellogg et al., 2012;
498 Sagaram et al., 2009; Tsiamis et al., 2012). PhyloChip is based on the 16S database "Greengenes"
499 and is constantly updated. To the best of our knowledge, this technology has not yet been applied to
500 the study of the bacterial communities of insects. Nevertheless, it is a sensitive and effective method
501 that will prove useful also in this field. Its main drawback is the impossibility of discovering new taxa
502 that are not represented in the probes.

503 Based on the use of microarrays, a comparative genomic hybridization approach has been employed
504 to examine the genomes of several *Wolbachia* strains, revealing the presence in *Wolbachia* genomes

505 of mosaic features, and underlying the importance that lateral gene transfer events have had in
506 *Wolbachia* evolution and diversity (Ishmael et al., 2009).

507 **4.7 Sequencing metagenomes and metatranscriptomes.**

508 The term "metagenomes" and "metatranscriptomes" used here refers to the shotgun sequencing of
509 the whole DNA or cDNA from a given environmental sample. Analysis of the resulting sequence
510 dataset leads to a better understanding of both the taxonomical composition and the metabolic
511 potential of a given bacterial community, through comparison with known sequenced gene encoding
512 enzymes. To the best of our knowledge, metagenomic/metatranscriptomic sequencing has, to date,
513 been applied to insect symbionts by only a few authors with a specific interest in wood degrading
514 systems (Liu et al., 2013; Scully et al., 2013a) or in honeybee gut physiology (Engel et al., 2012; Lee
515 et al., 2014). More scientific literature is available on the transcriptomic profiles of the insect
516 themselves instead, sometimes connected to bacterial colonisation (Cornman et al., 2013). Scully and
517 colleagues (Scully et al., 2013a, 2013b) published two research papers describing the metagenome
518 and metatrascriptome of the gut of the cerambycid beetle *Anoplophora glabripennis*, casting light on
519 the contribution of both the host and its bacterial community to the digestion of recalcitrant polymers.
520 In planning this kind of experiment one major decision must be made: to sequence only the bacterial
521 DNA (RNA) or to sequence the nucleic acids from the symbionts and the host together. In the former
522 case, the gut bacteria can be separated from the surrounding insect tissue mechanically, or by
523 centrifugation (Liu et al., 2013; Scully et al., 2013a), while in the case of endosymbionts no separation
524 is possible. In the latter case, it must be remembered that the DNA and RNA content of eukaryotic
525 cells is much higher than that of prokaryotes; thus, depending on the sequencing depth, the
526 information from bacteria could be insufficient for a proper analysis. Therefore, a higher number of
527 sequences is required. Moreover, the separation between co-sequenced eukaryotic and prokaryotic
528 DNA can be tricky due to HGT events (see paragraph 4.5). Dunning Hotopp and colleagues speculate
529 that the common exclusion of prokaryotic sequences from eukaryote genome sequencing projects
530 leads to an underestimation of the HGT phenomenon (Dunning Hotopp et al., 2007).

531 “Information mining” within great amounts of sequence data depends strictly on the specific
532 questions of the researcher. An informative overview of the available techniques, for both sample
533 preparation and data analysis, can be found in the review by di Bella and co-workers (Di Bella et al.,
534 2013). For a detailed idea of the technical issues involved in transcriptomic experiments we suggest
535 the review by Westermann and colleagues (Westermann et al., 2012). Although speaking about
536 pathogens rather than symbionts, it gives a detailed discussion of the questions that arise from an
537 experimental design involving, at the same time, one eukaryotic host and one or more bacterial
538 residents. A final interesting example regarding the application of these methods to insect gut
539 symbiosis can be found in the paper by Xie and colleagues (Xie et al., 2012), who detected a number
540 of genes for resistance to xenobiotics in the *Bemisia tabaci* microbiome.
541
542 **5. Bacterial isolation and cultivation**
543 Although there is no agreement concerning the estimation of the proportion of uncultivable bacterial
544 species, it is common knowledge that they greatly outnumber cultivable ones. Therefore, a culture-
545 based survey of a bacterial community cannot claim to be complete. Most endosymbionts are not yet
546 cultivable (for one exception, see (Matthew et al., 2005)), therefore culture efforts should focus on
547 gut microflora. Moreover, handling a collection of bacterial strains gives the opportunity to test
548 directly, *in vivo* and *in vitro*, hypotheses on their role in insect physiology, an aspect that is lacking
549 in many studies about microbial characterization in insects. Thus, this approach could be the first step
550 to fill the gap in knowledge. Finally yet importantly, isolates can be used in a range of
551 biotechnological applications, beyond the pure investigation of natural environments (Crotti et al.,
552 2012). Bacteria have been proposed as probiotics to improve the health of useful insects like
553 honeybees (Crotti et al., 2013; Hamdi et al., 2011), or to improve the effectiveness of the sterile insect
554 technique (i.e. release of sterile males to control a pest population) (Augustinos et al., 2015; Gavriel
555 et al., 2011). On the other hand, paratransgenesis is a promising technique against vector-borne
556 diseases (Hurwitz et al., 2011a).

557 A drawback of culture-dependent methods is the need for fresh insect specimens, which cannot be
558 preserved after sacrifice. Bacterial isolation is typically performed by smashing insect gut (or the
559 whole animal) with a sterile pestle in saline, and plating or inoculating dilutions of the homogenate
560 on various culture media. By directly plating serial dilutions of the homogenate onto a rich culture
561 medium it is possible to isolate those bacteria that are present at higher concentration and that grow
562 faster than others. In insect gut, a number of Enterobacteriales usually fall in this category.
563 To maximize the proportion of cultivated strains, the best practice for media preparation is to mimic,
564 as closely as possible, the natural environment of the bacteria. To obtain information about how the
565 gut environment looks, in terms of physio-chemical conditions, micro-sensors have been used. For
566 example, Lemke et al. (Lemke et al., 2003) provided a detailed description of important parameters
567 such as pH, redox potential, and gas exchange along the entire gut tract of *Pachnoda ephippiata* larvae
568 (Coleoptera: Scarabeidae). These parameters can vary both spatially (in different gut tracts) and
569 temporally (in different life stages). Oxygen availability in the gut varies greatly among insects, but
570 it is usually higher in smaller insects, while in bigger ones the gut can be almost totally anoxic (Ceja-
571 Navarro et al., 2014) or divided into anoxic compartments. In such cases, oxygen usually enters the
572 foregut with food, decreasing as the food reaches the midgut and hindgut. Anoxic conditions for the
573 isolation of bacteria can be obtained using an anaerobic chamber or, if not available, using sealed jars
574 with oxygen subtractors such as Anaerocult A (Merck Millipore).
575 In the case of particular activities (e.g. degradation of specific substrates) or certain species,
576 enrichment cultures can be obtained in liquid culture media prior to plating. For example, for cellulose
577 degrading bacteria the literature reports a range of enrichment media (Gupta et al., 2012), containing
578 filter paper or carboxymethyl-cellulose, while the isolation of acetic acid bacteria, which are
579 widespread symbionts of sugar-feeding insects, is carried out on acid media containing high
580 proportions of sugar or ethanol (Kounatidis et al., 2009; Lisdiyanti et al., 2001). Several enrichment
581 steps can be performed by inoculating fresh enrichment medium with an aliquot of the enriched
582 culture.

583 Diversity in the strain collection can be assessed using fingerprinting techniques; the most popular
584 being Internal Transcribed Spacer (ITS)-PCR. Indeed, PCR amplification of the ITS regions produces
585 a characteristic profile of amplicons of different sizes, usually visualized on an agarose gel,
586 demonstrated to be species or subspecies-specific (Daffonchio et al., 1998) and widely applied for
587 de-replication of environmental bacterial collections (Mapelli et al., 2013; Marasco et al., 2012).
588 Identification of the strains grouped by ITS fingerprinting is then achieved by sequencing, only in
589 polymorphic strains, the 16S rRNA gene, followed by BLAST search or using other online tools such
590 as RDP classifier (Wang et al., 2007), SINA aligner (Pruesse et al., 2012) or EZTaxon (Kim et al.,
591 2012). Sequences can also be used to reconstruct phylogenies, together with sequences from type and
592 non-type strains, using ARB (Ludwig et al., 2004).

593 **5.1 Isolate screening.**

594 Some studies infer metabolic potential of the gut microbiome from metagenomic and
595 metatranscriptomic data (Engel et al., 2012; Scully et al., 2013a, 2013b). Direct screening of the
596 isolates for specific activities is another straightforward and simple way of assessing the possible
597 contribution of the bacteria to the host metabolism. Many *in vitro* tests can be performed quite easily,
598 and others can be created or re-adapted from the literature to fit the individual case study. For
599 example, after finding pectate-lyases during a metagenomic survey of the microbiome of honeybee
600 gut, Engel et al. (Engel et al., 2012) verified the pectin degradation ability of isolates from honeybee
601 gut. Interestingly, only some strains within the *Gilliamella* genus possess this activity that could
602 contribute to pollen digestion.

603 *5.1.1 Carbon metabolism.* A number of insects feed on recalcitrant plant polymers, and although
604 many of them secrete enzymes to digest such substances (Sugimura et al., 2003), the digestion of
605 complex polysaccharide matrices most likely results from cooperation between the insect and its
606 bacterial and fungal microbiome. Indeed, the termite has the best known wood-degrading microbial
607 consortia (Ohkuma, 2003). In order to assess the degrading capabilities of bacterial isolates, a number
608 of plate tests can be easily performed (Table 1). Isolates can also be grown on a mineral medium in

609 the presence of sugar monomers and dimers, originated by complex sugar hydrolysis, as the sole
610 carbon source, e.g., arabinose and xylose, the major components of xylan, or cellobiose, a cellulose
611 dimer (two glucose molecules joined by a beta bond). To this end, a plate reader spectrophotometer
612 can be used for O.D. (optical density) measurements, allowing multiple assays at a time.

613 *5.1.2 Nitrogen metabolism.* Herbivorous diets are generally poor in nitrogen, e.g., in wood the
614 carbon to nitrogen ratio can be as high as 1000/1. However bacterial symbionts can directly fix
615 atmospheric nitrogen, thus helping their host to cope with nitrogen scarcity (Behar et al., 2005; Ceja-
616 Navarro et al., 2014; Morales-Jiménez et al., 2013). Alternatively, the nitrogen waste can be recycled
617 from the insect itself: nitrogen excretion in insects passes through the Malpighian tubules, collecting
618 the uric acid in the body cavity and then conferring it the anterior hindgut. Other nitrogen sources in
619 the gut of herbivorous insects are the urea produced by the gut microbiome itself, and cell-wall
620 proteins. Easy biochemical tests reveal the presence of urease or uricase activity, or the secretion of
621 proteases. Another option to investigate nitrogen fixing ability is the detection of the nitrogenase gene
622 by PCR (Gaby and Buckley, 2012), and, finally, nitrogenase activity can be measured by acetylene
623 reduction assay (Hardy et al., 1968).

624 *5.1.3 Screening for other nutritional activities.* Bacteria can be involved in the detoxification of diet
625 compounds. It has also been demonstrated that they can even confer resistance to pesticides (Kikuchi
626 et al., 2012). In order to protect themselves many herbivorous insects specialize in feeding on plants
627 that synthesize toxic metabolites. A common plant defence against insects is the production of
628 tannins, which reduce protein availability in the insect diet. Several bacteria have been shown to
629 produce tannases, which could breakdown these toxic metabolites in insect gut (Aguilar et al., 2007).
630 The survival of isolated bacteria on toxic compounds can be tested directly on liquid or agar cultures.
631

632 **6. Fluorescent microscopy methods**

633 Microscopy techniques are an invaluable resource to obtain information on symbiont localization,
634 and to speculate on function as the presence of bacteria in different body parts is indicative of their

635 role and route of transmission. Bacteria often inhabit very specialized sub-niches in the gut or are
636 localized in specific positions in the bacteriocytes and reproductive organs, hence microscopy is the
637 method of choice to observe the situation in detail, e.g., *Candidatus* Midichloria mitochondrii were
638 found, as suggested by the name, to inhabit the mitochondria of the ovarian cells of ticks (Beninati et
639 al., 2004; Sassera et al., 2006). The most popular method of detecting a bacterial species in a sample
640 is *in situ* hybridization (ISH), based on an oligonucleotide probe complementary to a small fragment
641 of 16S rRNA gene. The visualization of the probe is done by immunohistochemistry or, more often,
642 fluorescence microscopy. In the former, the sample is treated with a DNA probe joined to the
643 digoxygenin antigen, which in turn is bound by an antibody conjugated with an enzyme; the activity
644 of this enzyme is then detected by adding a chromogenic substrate (Beninati et al., 2004; Favia et al.,
645 2007) that is visualized with light microscopy. The method is quite labour-intensive and has been
646 replaced in most cases by Fluorescence *In Situ* Hybridization (FISH), in which a DNA probe is
647 conjugated with a fluorescent label.
648 Immunocytochemistry is also used to stain different cells in host tissues, especially for an accurate
649 description of the interaction between symbionts and host organs (Frydman et al., 2006).
650 **6.1 FISH**.
651 Fluorescence *in situ* hybridization is based on the hybridisation of a DNA oligonucleotide conjugated
652 with a fluorochrome to the bacterial DNA. This technique is a powerful method to detect bacteria
653 directly in field-collected samples (Gonella et al., 2011). Tissues can be preserved by fixation, thus
654 enabling researchers to choose the most suitable timing for their experiments. The first step to perform
655 FISH is to design, or select from the literature, a specific probe. If the aim is to detect a particular
656 species, the probe should hybridize on the 16S rRNA gene of the target species. Indeed, specific
657 probes can be designed using bacterial 16S rRNA gene sequences from available databases.
658 Furthermore, the high copy number of ribosomal RNA genes inside bacterial cells provides a strong
659 fluorescence signal. Particular features of bacterial communities can be investigated by designing
660 FISH probes for functional genes. To design probes for the 16S rRNA gene, the PROBE_DESIGN

661 tool within the ARB software package can be used (Ludwig et al., 2004). The SILVA website is
662 another useful source of information about how to perform FISH (http://www.arb-silva.de/fish-
663 probes/). The probes can be verified in-silico also using the online tool "probe match" of RDP
664 (http://rdp.cme.msu.edu/index.jsp, (Cole et al., 2014)) If possible, FISH probes should then be tested
665 on bacterial isolates or known positive samples to assess the specificity and adjust the stringency
666 conditions. This step is also necessary for probes available in the literature as different working
667 conditions lead to different specificity features. Moreover, different substrates can include, or not,
668 species closely related to the target – a situation that requires a higher degree of stringency in the
669 assay. Thus, it is important to verify that the probes can detect the target species, and are not cross-
670 hybridizing with other bacterial DNA in the sample. FISH can be performed directly on bacterial
671 cultures of the isolates or, if the target bacterium is uncultivable, on an *E. coli* clone from a library,
672 carrying the 16S rRNA of the target bacterium (Schramm et al., 2002). Usually, a probe targeting all
673 bacteria is included in the assay. A second preparatory step for the FISH assay consists in the
674 observation of the target insect tissue by an epifluorescence microscope. Indeed, insect tissues and
675 even the plant based diet of the insect can have a level of autofluorescence. Thus, in accordance with
676 the endogenous level of fluorescence of the sample, and the area in which bacteria are expected to
677 be, the researcher can choose the appropriate fluorescent dyes conjugated to the probes.
678 Additional nuclear staining (e.g. DAPI (4',6-diamidino-2-phenylindole) Propidium Iodide) is
679 recommended to better localize bacteria within tissue, using the position of the eukaryotic nuclei as
680 the reference. For this, actin filaments in the host cells can be marked using fluorescent phalloidin.
681 For small insects, such as mosquitoes or *Drosophila* flies, whole organs can be stained in a 1.5 ml
682 plastic tube, mounting the glass slide as the final step. According to their thickness, these thin samples
683 can be observed in an epifluorescence microscope or, better still, in a confocal fluorescence
684 microscope, enabling the observation of multiple focal planes at different depths. For larger insects,
685 the organs (particularly gut) can often be too thick to observe a suitable focal plane on the Z-axis, or
686 to allow light penetration to the desired plane. In this case, good images require micro-sections. One

of the most feasible techniques is cryo-sectioning. This method preserves tissue morphology, without the need for histological preparations, such as resin infiltration, that require time and equipment. In this case, DAPI staining and hybridization with probes can be performed on tissue slices on the glass slides. Unfortunately, gut sections are especially fragile because the gut content is easily detached from tissue during the washing and staining procedures. For this reason it is also possible to perform DAPI staining and hybridization on whole organs before freezing, and to observe the slides immediately after tissue cutting. The success of this second option depends on the penetration rate of the probe into the insect tissue, which is in turn linked to the size of the organ, and can be improved by pepsin treatment. However, preliminary experiments are usually required to choose the best visualization options for each case study.

Examples of the application of FISH to the study of insect symbiont localization can be found in Balmand et al., 2013; Gonella et al., 2011; Hayashi et al., 2015; Matsuura et al., 2015; and Stoll et al., 2010.

6.2 Fluorescent protein tagged strains

If an interesting strain is found among the isolates, the first step to characterize its interaction with the host is its transformation using a fluorescent protein. The fluorescent clone can subsequently be administered to the host to investigate i) colonization routes and timing, ii) persistence over time and across different life stages, iii) transmission routes (maternal, paternal, environmental) iv) cross-colonization of different species (Damiani et al., 2008; Earl et al., 2015; Gonella et al., 2012; Kikuchi and Fukatsu, 2013).

This approach requires not only a bacterial strain that is easily cultivated and transformed, but also the management of stable insect breeding and rearing.

6.2.1 Marking systems for bacteria. Several marking systems are described in the literature to transform bacteria with genes encoding different fluorescent proteins: and it is from among them that the researcher has to choose, according to the target species and experimental constraints. The following, an overview of the principal features of these systems, will help researchers new to this

713 field to better understand the available options. Depending on the final position of the marker gene
714 (usually encoding a fluorescent protein, e.g. GFP) in the host strain, the marking can be chromosomal
715 if the marker gene is located on the bacterial chromosome, or episomal if it is on a plasmid.
716 The genes for fluorescent proteins are usually encoded in specific plasmids to enable them to be
717 transferred to the target strain, and coupled with an antibiotic resistance gene to select cells that
718 acquire them. If the marking is chromosomal, the plasmid lacks the replication origin that allows it
719 to replicate in the host. Thus, as it cannot persist in the host cytoplasm it integrates in the genome
720 through molecular mechanisms derived by transposons or phages. In some cases the insertion occurs
721 in a specific, neutral site, such as attTn7 in tn7-based systems (Choi and Schweizer, 2006; Lambertsen
722 et al., 2004), while in others the insertion is random, raising the issue of the influence of insertion on
723 gene expression. Chromosomal marking is usually more stable, and allows the cultivation of the
724 bacterium without selective pressure (i.e. antibiotics). As a consequence, insects receiving the
725 bacterial inoculum do not need antibiotic treatment, thus closely resembling the natural environment.
726 Conversely, for episomal marking, a vector carrying a replication origin recognized by the host is
727 used, the vector thus continuously replicating in its cytosol. In this case, continuous antibiotic
728 selection is needed to avoid plasmid loss through generations; however, the fluorescence can be
729 brighter due to the multiple plasmid copies present in a single cell at the same time. However, the
730 level of expression of the fluorescent protein is largely dependent on the bacterial strain.
731 Another important difference among marking systems concerns the delivery mode of the marking
732 gene to the bacterial cell: by transformation or by conjugation. In the former, only two strains are
733 involved: the target and the donor (usually *E. coli*). The plasmid is purified from the donor, usually
734 using a commercial kit, while the target is grown in liquid culture and washed to obtain competent
735 cells (i.e. cells capable of acquiring foreign DNA). The purified DNA is then introduced directly into
736 the competent cells by heat-shock transformation or electroporation. The first technique works well
737 with *E. coli*, but is generally regarded as less effective in terms of number of transformants. For this
738 reason, electrotransformation is advisable in the case of environmental isolates. The second method

739 requires electroporator equipment to subject the competent cells, mixed with plasmid DNA, to a
740 voltage (Mostafa et al., 2002). After transformation, the bacteria are usually grown for a short period
741 of time, one or a few hours in a rich liquid medium without antibiotics, and then plated on appropriate
742 plates with added antibiotic(s), to select the transformed cells.
743 The conjugation method exploits the natural horizontal DNA transfer mechanism between bacteria
744 of different species. To transfer the DNA by conjugation complex molecular machinery is required;
745 thus, unless the donor strain retains all the required features, more than two strains can be involved:
746 the donor strain, the target bacterium and one or two helper strains. The strains are grown separately
747 in liquid cultures, washed, mixed together in appropriate ratios and plated or spotted on filter paper.
748 Only after the conjugation time (usually overnight, or longer) the exconjugants are exposed to
749 antibiotic(s) to select target bacteria that acquired the marker gene.
750 To plan a transformation experiment, a marking system has to be chosen, according, primarily, to the
751 target bacterial species. Various marking systems have been developed, but they differ in many
752 features such as origin of replication of the plasmids and delivery mode; the systems must, primarily,
753 be suitable for different taxa. There are many taxon-specific characteristics that affect transformation
754 efficiency: recognized replication origin, methylation of the DNA, type of cell wall (i.e. Gram
755 positives/negatives), codon usage.
756 Prior to starting an experiment it is necessary to verify the antibiotic resistance profile of the target
757 strain, which should be sensitive to the antibiotics used in the selection of transformed cells. If using
758 a conjugation-based system, the target strain should have one antibiotic resistance in the genome, or
759 should grow in special selective conditions (e.g. high salinity, low pH) to allow separation from the
760 donor strain. For this purpose it is sometimes possible to select naturally occurring mutants that are
761 resistant to rifampicin. Furthermore, to optimize the visualization, the fluorescent protein suitable for
762 each study must be selected according to the natural fluorescence of the target insect organs. Also the
763 bacteria themselves can have natural autofluorescence that should be checked. Table 2 shows a short
764 list of marking systems that have been used in experiments involving insect colonisations.

765

766 ## 7. In vivo experiments

767 The recolonization of an insect with marked symbionts can have several purposes: i) to assess the
768 beneficial effect of the symbionts on insect growth or fitness, ii) to verify the transmission routes of
769 the bacterium iii) to assess the localization of the bacterium within the insect. All three aspects are
770 discussed by Kikuchi and colleagues in an elegant study on *Riptortus pedestris* (Kikuchi and Fukatsu,
771 2013). Conversely, it could be useful to deprive the insect of a symbiont to investigate whether the
772 symbiont is necessary, beneficial or neutral to the host (Chouaia et al., 2012).

773 ### 7.1 Symbiont deprivation and mono-associated insects.

774 The creation of "aposymbiotic" insects is the best method to assess the importance of a symbiont for
775 host survival and fitness. Moreover, it is the only applicable strategy when the target bacterium is an
776 uncultivable endosymbiont, of which the effect can only be proved in a negative manner (by
777 subtracting it rather than adding). Aposymbiotic insects have been studied since the fifties (Brooks
778 and Glenn Richards, 1955) and can be obtained by various methods, reviewed by Wilkinson
779 (Wilkinson, 1998), among which the most popular and effective is antibiotic treatment. This strategy
780 permits, for example, an assessment of the dependency of pea aphids on *Buchnera* symbionts for the
781 production of some essential amino acids (Wilkinson and Ishikawa, 2000). Germ-free *Drosophila*
782 can be produced also by bleaching the embryos and rearing them on sterile medium, without the use
783 of antibiotics. However, sterile conditions are very difficult to maintain; thus, germ-free stocks are
784 usually reared on diets supplemented with a mix of antibiotics (Storelli et al., 2011). The comparison
785 of germ-free (GF) and wild-type (WT) animals is a powerful method, but is not sufficient for most
786 scientific purposes. Indeed, aposymbionts are deprived of their entire microbiome, which is, in most
787 cases, composed by more than one species and it is thus impossible to attribute the better performance
788 observed in WT animals to a single bacterial species. Furthermore, it is difficult to prove that the
789 antibiotic itself is harmless for the insect. For these reasons, many studies use aposymbiotic insects
790 as a starting point to produce mono-associated insects, which are inoculated with a particular strain

791 usually provided with the feeding medium. In *Drosophila,* this strategy has allowed the screening of
792 an entire library of mutated strains to assess which bacterial genes were responsible for the beneficial
793 properties of the tested bacterium (Shin et al., 2011). In this case, the mono-associated insects were
794 compared to both WT and GF animals. In another study, authored by Chouaia and colleagues
795 (Chouaia et al., 2012), a rifampicin-resistant *Asaia* strain was used to recolonize rifampicin-treated
796 mosquitoes. The development of rifampicin-treated specimens was delayed, while the colonization
797 by *Asaia* restored the normal development rate, showing that the longer growing time was due to the
798 absence of the symbionts rather than to rifampicin itself.
799 Many gut symbionts are commensals acquired from the environment, and are selected in the gut by
800 the complex habitat created by the host and the resident microbial community. It is reasonable to
801 assume that these strains could persist and be beneficial to the insect through interaction with other
802 members of the gut community. For these reasons, the colonization of insects with isolates could, in
803 some cases, be performed directly on WT insects (Mitraka et al., 2013).
804 One critical aspect of these experiments is the mode and delivery time of the bacterium. Since most
805 of the cultivable symbionts are located primarily in the gut, bacteria are usually mixed with diet
806 material. To ensure bacteria acquisition, insects can be subjected to fasting before being exposed to
807 the inoculated food. Alternatively, the inoculated food is provided for sufficient time to ensure that
808 the insects have had at least one meal. The timing is chosen according to the size and feeding habits
809 of the animal. For example, Kikuchi and colleagues (Kikuchi and Fukatsu, 2013) left *Riptortus*
810 *pedestris* nymphs without water for one night, and the animals immediately drank the inoculated
811 water given to them the following morning; Storelli and colleagues (Storelli et al., 2011) allowed the
812 eggs of *Drosophila* flies to hatch on inoculated medium; and in another study mosquitoes were fed
813 an inoculated sugar diet for two hours (Damiani et al., 2008). In some cases, symbionts were acquired
814 in a specific temporal window during development, and this should be considered when planning
815 experiments (Kikuchi et al., 2011).

816 *7.1.1 Fitness measurements.* To quantify the effect of a symbiont on host fitness is not a trivial
817 matter. In many cases the beneficial effects become visible only in situations of stress or nutrient
818 scarcity, but it must be remembered that laboratory reared insects usually live in controlled
819 conditions, with plenty of food and optimal temperature and light/dark cycles. For example, the
820 influence of *Acetobacter pomorum* on *Drosophila* larvae development is visible only when the
821 amount of yeast extract in the diet is reduced (Shin et al., 2011), while in an applicative study on the
822 fruit fly *Ceratitis capitata* the symbiont *Klebsiella oxytoca* is used to restore mating competitiveness
823 of γ-irradiated males (Gavriel et al., 2011). Furthermore, the measurement of "fitness" can be
824 performed in different ways and at different life stages, according to the expected effect of the
825 bacterium. In most cases, growth rate or development time is measured at larval stages. In fact,
826 holometabolous insects grow mainly as larvae, and different larval stages can be monitored observing
827 moults, thus registering development time (Chouaia et al., 2012; Shin et al., 2011). Insects can also
828 be measured and weighed, or mated to assess the mean number of eggs laid as an indicator of
829 reproductive success (Kikuchi and Fukatsu, 2013). To specifically address the problem of male
830 mating success of *C. capitata*, Gavriel and colleagues (2011) observed the proportion of individuals,
831 from among sterile males inoculated with live and dead bacteria, that mated with WT females,
832 demonstrating the contribution of probiotic treatment to increase the efficacy of the SIT (sterile insect
833 technique) approach. For a similar SIT application, Augustinos and colleagues demonstrated the
834 efficacy of their symbiont-based strategy to improve the mass rearing of medfly sterile males with a
835 broad panel of measurements: pupal weight, mating competitiveness, survival to food and water
836 starvation, flight ability (Augustinos et al., 2015).
837 *7.1.2 Measurements of gene expression.* In addition to fitness measurements, the response of the
838 insect to the administration of a bacterium can be evaluated by measuring the expression of specific
839 genes, such as insect anti-microbial peptides (AMPs), using real-time PCR assays. Generally
840 speaking, AMPs are important and well-characterized immune effectors of insects and invertebrates
841 that lack an adaptive immune system (Evans and Lopez, 2004).

842 *7.1.3 Localization and transmission routes.* Having access to a marked clone of the target bacterium
843 allows the researcher to localize it in colonized insects. To do this, it is possible to take advantage of
844 the same microscopy techniques reviewed in the FISH paragraph. Fluorescent proteins are not
845 affected by tissue fixation or freezing. According to the colonization pattern observed, it is possible
846 to make and test hypotheses about the bacterial transmission. This approach was used to localize the
847 symbiont *Asaia* in the gut, salivary glands and gonads of *Anopheles stephensi*, then demonstrating its
848 maternal, paternal and environmental transmission, besides a capability to colonize different
849 mosquito species (Crotti et al., 2009; Damiani et al., 2008; Favia et al., 2007).
850

851 **8. Genetic manipulation techniques for advanced *in-vivo* studies**

852 A further goal in symbiont research is an understanding of the genetic mechanisms that allow
853 interaction between an insect and its bacterial hosts. In addition to metatranscriptomic studies, some
854 authors have started to apply advanced genetic techniques primarily set up for medical research.
855 These methods provide elegant demonstrations in hypothesis-based experiments, as well as the
856 possibility of applying wide screenings.
857 However, the conditions for setting up this kind of experiment are severe. It is necessary to have well-
858 established breeding of the target insect, with appropriate and fast screening systems to evaluate the
859 fitness of animals colonized by different bacteria. Furthermore, the bacterium of interest should be
860 easily cultivable and transformable. Finally, the availability of its sequenced genome is a great help.

861 **8.1 Screenings based on insect phenotype**

862 *8.1.1 Site-specific mutations.* Particular hypotheses on the interaction between bacteria and insects
863 can be tested by impairing a specific gene on the bacterial genome and assessing the effect of the
864 mutant strain on mono-associated insects. In the research group of Elke Genersch, this method has
865 been used to validate the discovery of new virulence factors of the honeybee pathogen *Paenibacillus*
866 *larvae* (Fünfhaus et al., 2013; Garcia-Gonzalez et al., 2014). With the constant increase in genomic
867 and transcriptomic data on bacterial symbionts, it will be possible to formulate, and test, many

868 specific hypotheses. As an alternative, a group of different mutants can be tested at the same time by
869 transposon mutagenesis.
870 *8.1.2 Transposon mutagenesis.* Shin and colleagues (Shin et al., 2011) created a library of
871 *Acetobacter pomorum* in which each strain carried an insertion of a Tn5 transposon in one random
872 position on the genome. The library was used to create cohorts of mono-associated *Drosophila*, which
873 were compared according to weight and growth rate. The strains that produced slow-growing flies
874 were subsequently analysed to understand which genes had been impaired by the insertion of the
875 transposon. The authors found that 11 genes involved in the periplasmic pyrroloquinolinequinone–
876 dependent alcohol dehydrogenase (PQQ-ADH)–dependent oxidative respiratory chain were damaged
877 in the selected strains. This strategy allowed them to understand that the production of acetic acid by
878 *A. pomorum*, mediated by the PQQ-ADH pathway, was important to enhance larval growth on a sub-
879 optimal diet.

880 **8.2 Screenings based on bacterial survival in the host**

881 IVET (In Vivo Expression Technology) and STM (Signature Tagged Mutagenesis), reviewed and
882 compared in Chiang et al. (Chiang et al., 1999), have been developed to discover *In Vivo* Induced
883 (IVI) genes. Reasonably, pathogenic organisms encounter, as they enter the host, environmental
884 conditions that they do not experience outside. Thus, they react to the mutated situation by changing
885 the gene expression profile, enhancing the transcription of the genes involved in colonization and
886 pathogenesis. Therefore, many IVI genes are expected to be virulence factors. The same reasoning
887 can be applied to symbionts. In the mutated conditions of the host body, they do not activate virulence
888 factors, but they should be able to modify their gene expression in order to escape the immune system,
889 to colonize their specific niche in the body and to exploit the food resources therein without damaging
890 the insect, even possibly giving it an advantage. For this reason, the application of these techniques
891 to the field of insect symbiosis appears, in our opinion, very promising. Both these techniques are
892 based on a genetic selection that is independent of host fitness.

893 *8.2.1 STM.* Signature-tagged mutagenesis (STM) consists of a large-scale screening to detect the
894 bacterial genes necessary for the survival of symbionts inside the host body. Like in the method
895 described in paragraph 8.1.2, a transposon is used to impair random genes. In this case, a number of
896 different transposon plasmids, each containing a sequence tag, is used to produce a set of libraries of
897 randomly mutated strains of the same bacterium. Each library is originated by a transposon with a
898 different tag. The tags are flanked by two invariant regions, to easily amplify them by PCR with the
899 same primers. The libraries are used to create Input Pools (IP), combining one strain from each
900 library. Each IP is administered to a different group of animals. After the experimental time, bacteria
901 are recovered from the animals, generating Output Pools (OP). Strains that are mutated in genes
902 fundamental to *in-vivo* survival of the bacteria will not overcome the phase inside the host, and will
903 not be retrieved in the OP. Note that IP and OP can be compared by dot blot, after tag amplification.
904 Strains present in the IP but not in the OP are thus analysed to detect their mutated genes, which are
905 assumed necessary for bacterium permanence inside the host. To ensure that each gene in the bacterial
906 genome is mutated in at least one strain, it is necessary to repeat the experiment with many input
907 pools. STM has been successfully applied to human enteric pathogens (Cummins and Gahan, n.d.) as
908 well as to the plant symbiont *Sinorhizobium meliloti* (Pobigaylo et al., 2006), while no report has yet
909 been published on insect symbionts.
910 *8.2.2 IVET.* A more refined genetic manipulation is required for In Vivo Expression Technology
911 (IVET) (Chiang et al., 1999). The aim of this technique is to select bacterial promoters that are active
912 only *in vivo*, i.e. when the bacterium is inside the host. As a first step, the genomic DNA of the target
913 bacterium is randomly fragmented through partial enzymatic digestion. The fragments are inserted
914 upstream from a promoterless antibiotic resistance gene into a specific plasmid that is subsequently
915 used to transform the target bacterium. In this last, the plasmids are unable to replicate, but they can
916 integrate in the genome through a homologous recombination event, between the cloned genomic
917 sequence and the corresponding region on the bacterial chromosome. Thus, a library of mutants is
918 produced. As a result of the homologous recombination process, each of the mutants carries two

919 tandem copies of a specific fragment of the genome. Downstream from one of the copies lies the
920 promoterless antibiotic-resistance gene. The library is subjected to a first screening step on antibiotic
921 medium to eliminate clones resistant to the antibiotic *in-vitro*, with a replica-plating-like process.
922 Indeed, only those bacteria in which a promoter in the cloned genomic fragments drives the
923 expression of the antibiotic resistance gene will survive on the antibiotic plates. These clones harbour
924 a cloned copy of a promoter that is probably constitutive as it is active in optimal laboratory
925 conditions, without any stress stimulus. Thus, these clones should be excluded from further
926 screenings. The remaining clones of the library of mutants are introduced in the eukaryotic host as a
927 whole, and, as in STM, are retrieved after some time. Differently from STM, the animal host is
928 subjected to an antibiotic therapy for the duration of the experiment. In these conditions, only bacteria
929 able to express the antibiotic resistance gene, due to the presence of a cloned *in vivo* activated
930 promoter upstream, will survive. The bacteria in which the cloned portion of the genome coincides
931 with inactive promoters or regions with other functions will die, or at least remain quiescent, inside
932 the host. To further refine the selection, the retrieved pool of clones can be used for a second *in vivo*
933 inoculation. Finally, the output pool will contain only bacteria in which a second copy of an *in vivo*
934 activated promoter is present. Further analyses are needed to sequence it, and determine its native
935 function in the genome, in order to understand which are the *in vivo* transcribed genes in its operon.
936 To this extent, knowledge of the genome of the target bacterium is very useful.
937 Many variants of this general idea have been developed. The screening can be based on the
938 complementation of auxotrophy rather than on antibiotic resistance, avoiding the need to treat the
939 hosts with an antibiotic (Gal et al., 2003). Finally, different methods have been developed to select
940 against the constitutive promoters. An advancement of IVET is called RIVET – Recombinase-based
941 IVET – and involves a recombinase that, when expressed due to promoter activation, permanently
942 cleaves a marker gene elsewhere in the genome. RIVET is much more sensitive and can detect
943 promoters activated at either low levels or transiently. Both IVET and RIVET are reviewed by Merrell
944 and Camilli (Merrell and Camilli, 2000). IVET has also been applied to research on *Bacillus cereus*

945 as a pathogen of insects, using the honeycomb moth *Galleria mellonella* as the model (Fedhila et al.,
946 2006).
947
948 **9. Conclusion**
949 In recent years, literature data on insect bacterial symbionts have greatly increased, together with the
950 fast development of sequencing techniques. However, a large number of works do not go beyond the
951 description of taxa associated with insects. In our opinion, this great amount of data is an important
952 resource that, so far, has not been exploited to any great degree. Indeed, in most cases the function of
953 insect symbionts is unknown. Conversely, metagenomic studies show that RNA transcripts from
954 symbionts can exert a variety of functions, but they give only a rough idea of which bacterium is
955 actually producing them. Connecting taxonomy to function and, possibly, to symbionts obtained in
956 culture, constitutes a great advance for science and biotechnology. The extraordinary evolutionary
957 success of insects is at least partially due to their symbiotic relationships. For example, insect gut is
958 a complex ecosystem that efficiently exploits a variety of food sources, including poor, unbalanced,
959 toxic and difficult to digest substrates. Assessing the mechanisms through which this happens would
960 be precious for research on waste management, including the possible reuse and exploitation of waste
961 matter, as well as for the detoxification of contaminated substrates. In this context, the management
962 of isolated strains would be an added high value, for both deepening knowledge and developing
963 applications. On the other hand, knowing the processes underlying the survival, and resistance to
964 treatment, of many agricultural pests and disease vectors could greatly contribute to research on
965 animal and plant illnesses. In insect vectors, symbionts and pathogens coexist, but their interactions
966 are far from fully elucidated. Instead, the interaction of gut symbionts with the host immune-system
967 is well-documented (Lee et al., 2013; Ryu et al., 2008). Several management strategies based on
968 symbionts, (reviewed in Crotti et al., 2012) have proved to be effective to control insect pests and
969 disease vectors. For examples see paratransgenesis (Durvasula et al., 1997; Hurwitz et al., 2011b),
970 symbiont-based improvements of the sterile insect technique (Augustinos et al., 2015; Gavriel et al.,

971 2011), and *Ceratitis capitata* control through the *Wolbachia* infection of males (Zabalou et al., 2004).
972 For these reasons, insect symbionts are a promising field, one to be investigated using all the
973 techniques available. Some are summarized in this review, the aim being to not only describe
974 microbial diversity but also to link it to function and to potentiate research to find biotechnological
975 applications.
976

977 **Acknowledgements**

978 F.M. was supported by Università degli Studi di Milano, DeFENS, European Social Found (FSE) and
979 Regione Lombardia (contract "Dote Ricerca").

980 **References**

981 Aguilar, C.N., Rodríguez, R., Gutiérrez-Sánchez, G., Augur, C., Favela-Torres, E., Prado-Barragan, 982 L. a, Ramírez-Coronel, A., Contreras-Esquivel, J.C., 2007. Microbial tannases: advances and 983 perspectives. Appl. Microbiol. Biotechnol. 76, 47–59. doi:10.1007/s00253-007-1000-2

984 Akman Gündüz, E., Douglas, A.E., 2009. Symbiotic bacteria enable insect to use a nutritionally 985 inadequate diet. Proc. Biol. Sci. 276, 987–91. doi:10.1098/rspb.2008.1476

986 Amend, A.S., Seifert, K. a, Bruns, T.D., 2010. Quantifying microbial communities with 454 987 pyrosequencing: does read abundance count? Mol. Ecol. 19, 5555–65. doi:10.1111/j.1365-988 294X.2010.04898.x

989 Andersson, A.F., Lindberg, M., Jakobsson, H., Bäckhed, F., Nyrén, P., Engstrand, L., 2008. 990 Comparative analysis of human gut microbiota by barcoded pyrosequencing. PLoS One 3, 991 e2836. doi:10.1371/journal.pone.0002836

992 Arias-Cordero, E., Ping, L., Reichwald, K., Delb, H., Platzer, M., Boland, W., 2012. Comparative 993 evaluation of the gut microbiota associated with the below- and above-ground life stages (larvae 994 and beetles) of the forest cockchafer, Melolontha hippocastani. PLoS One 7, e51557. 995 doi:10.1371/journal.pone.0051557

996 Arioli, S., Monnet, C., Guglielmetti, S., Parini, C., De Noni, I., Hogenboom, J., Halami, P.M., Mora, 997 D., 2007. Aspartate biosynthesis is essential for the growth of Streptococcus thermophilus in 998 milk, and aspartate availability modulates the level of urease activity. Appl. Environ. Microbiol. 999 73, 5789–96. doi:10.1128/AEM.00533-07

1000 Augustinos, A. a., Kyritsis, G. a., Papadopoulos, N.T., Abd-Alla, A.M.M., Cáceres, C., Bourtzis, K., 1001 2015. Exploitation of the Medfly Gut Microbiota for the Enhancement of Sterile Insect 1002 Technique: Use of Enterobacter sp. in Larval Diet-Based Probiotic Applications. PLoS One 10, 1003 e0136459. doi:10.1371/journal.pone.0136459

1004 Baker, G.C., Smith, J.J., Cowan, D.A., 2003. Review and re-analysis of domain-specific 16S primers. 1005 J. Microbiol. Methods 55, 541–555. doi:10.1016/j.mimet.2003.08.009

1006 Balmand, S., Lohs, C., Aksoy, S., Heddi, A., 2013. Tissue distribution and transmission routes for 1007 the tsetse fly endosymbionts. J. Invertebr. Pathol. 112, S116–22. doi:10.1016/j.jip.2012.04.002

1008 Behar, A., Yuval, B., Jurkevitch, E., 2005. Enterobacteria-mediated nitrogen fixation in natural 1009 populations of the fruit fly Ceratitis capitata. Mol. Ecol. 14, 2637–43. doi:10.1111/j.1365-1010 294X.2005.02615.x

1011 Beninati, T., Lo, N., Sacchi, L., Genchi, C., Noda, H., Bandi, C., 2004. A novel alpha-1012 Proteobacterium resides in the mitochondria of ovarian cells of the tick Ixodes ricinus. Appl. 1013 Environ. Microbiol. 70, 2596–602. doi:10.1128/AEM.70.5.2596-2602.2004

1014 Bourtzis, K., Dobson, S.L., Xi, Z., Rasgon, J.L., Calvitti, M., Moreira, L. a., Bossin, H.C., Moretti, 1015 R., Baton, L.A., Hughes, G.L., Mavingui, P., Gilles, J.R.L., 2014. Harnessing mosquito-1016 Wolbachia symbiosis for vector and disease control. Acta Trop. 132, S150–S163. 1017 doi:10.1016/j.actatropica.2013.11.004

1018 Braendle, C., Miura, T., Bickel, R., Shingleton, A.W., Kambhampati, S., Stern, D.L., 2003.
1019 Developmental origin and evolution of bacteriocytes in the aphid-Buchnera symbiosis. PLoS
1020 Biol. 1, E21. doi:10.1371/journal.pbio.0000021

1021 Brelsfoard, C., Tsiamis, G., Falchetto, M., Gomulski, L.M., Telleria, E., Alam, U., Doudoumis, V.,
1022 Scolari, F., Benoit, J.B., Swain, M., Takac, P., Malacrida, A.R., Bourtzis, K., Aksoy, S., 2014.
1023 Presence of Extensive Wolbachia Symbiont Insertions Discovered in the Genome of Its Host
1024 Glossina morsitans morsitans. PLoS Negl. Trop. Dis. 8. doi:10.1371/journal.pntd.0002728

1025 Broderick, N.A., Raffa, K.F., Goodman, R.M., Handelsman, J., 2004. Census of the bacterial
1026 community of the gypsy moth larval midgut by using culturing and culture-independent
1027 methods. Appl. Environ. Microbiol. 70, 293–300. doi:10.1128/AEM.70.1.293-300.2004

1028 Brooks, M.A., Glenn Richards, A., 1955. Intracellular symbiosis in cockroaches. I. Production of
1029 aposymbiotic cockroaches. [WWW Document]. Biol. Bull. Mar Biol Lab. doi:10.2307/1538656

1030 Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer,
1031 N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig,
1032 J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J.,
1033 Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J.,
1034 Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. Nat.
1035 Methods 7, 335–6. doi:10.1038/nmeth.f.303

1036 Cappuccino, J., Sherman, N., 1992. Microbiology: A Laboratory Manual, 3rd ed.
1037 Benjamin/Cummings Pub. Co., New York.

1038 Ceja-Navarro, J. a, Nguyen, N.H., Karaoz, U., Gross, S.R., Herman, D.J., Andersen, G.L., Bruns,
1039 T.D., Pett-Ridge, J., Blackwell, M., Brodie, E.L., 2014. Compartmentalized microbial
1040 composition, oxygen gradients and nitrogen fixation in the gut of Odontotaenius disjunctus.
1041 ISME J. 8, 6–18. doi:10.1038/ismej.2013.134

1042 Chandler, J.A., Lang, J.M., Bhatnagar, S., Eisen, J.A., Kopp, A., 2011. Bacterial communities of
1043 diverse Drosophila species: ecological context of a host-microbe model system. PLoS Genet. 7,
1044 e1002272. doi:10.1371/journal.pgen.1002272

1045 Chiang, S., Mekalanos, J.J., Holden, D.W., 1999. In vivo genetic analysis of bacterial virulence.
1046 Annu. Rev. Microbiol. 129–154. doi:10.1146/annurev.micro.53.1.129

1047 Choi, K.-H., Schweizer, H.P., 2006. mini-Tn7 insertion in bacteria with single attTn7 sites: example
1048 Pseudomonas aeruginosa. Nat. Protoc. 1, 153–61. doi:10.1038/nprot.2006.24

1049 Chouaia, B., Rossi, P., Epis, S., Mosca, M., Ricci, I., Damiani, C., Ulissi, U., Crotti, E., Daffonchio,
1050 D., Bandi, C., Favia, G., 2012. Delayed larval development in Anopheles mosquitoes deprived
1051 of Asaia bacterial symbionts. BMC Microbiol. 12 Suppl 1, S2. doi:10.1186/1471-2180-12-S1-
1052 S2

1053 Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., Brown, C.T., Porras-Alfaro, A.,
1054 Kuske, C.R., Tiedje, J.M., 2014. Ribosomal Database Project: data and tools for high throughput
1055 rRNA analysis. Nucleic Acids Res. 42, D633–42. doi:10.1093/nar/gkt1244

1056 Colman, D.R., Toolson, E.C., Takacs-Vesbach, C.D., 2012. Do diet and taxonomy influence insect
1057 gut bacterial communities? Mol. Ecol. 21, 5124–37. doi:10.1111/j.1365-294X.2012.05752.x

1058 Compant, S., Reiter, B., Sessitsch, A., Nowak, J., Clément, C., Barka, E.A., 2005. Endophytic
1059 Colonization of Vitis vinifera L. by Plant Growth-Promoting Bacterium Burkholderia sp. Strain
1060 PsJN. Appl. Environ. Microbiol. 71, 1685–1693. doi:10.1128/AEM.71.4.1685-1693.2005

1061 Cornman, R.S., Lopez, D., Evans, J.D., 2013. Transcriptional response of honey bee larvae infected
1062 with the bacterial pathogen Paenibacillus larvae. PLoS One 8, e65424.
1063 doi:10.1371/journal.pone.0065424

1064 Crotti, E., Balloi, A., Hamdi, C., Sansonno, L., Marzorati, M., Gonella, E., Favia, G., Cherif, A.,
1065 Bandi, C., Alma, A., Daffonchio, D., 2012. Microbial symbionts: a resource for the management
1066 of insect-related problems. Microb. Biotechnol. 5, 307–17. doi:10.1111/j.1751-
1067 7915.2011.00312.x

1068 Crotti, E., Damiani, C., Pajoro, M., Gonella, E., Rizzi, A., Ricci, I., Negri, I., Scuppa, P., Rossi, P.,
1069 Ballarini, P., Raddadi, N., Marzorati, M., Sacchi, L., Clementi, E., Genchi, M., Mandrioli, M.,
1070 Bandi, C., Favia, G., Alma, A., Daffonchio, D., 2009. Asaia, a versatile acetic acid bacterial
1071 symbiont, capable of cross-colonizing insects of phylogenetically distant genera and orders.
1072 Environ. Microbiol. 11, 3252–64. doi:10.1111/j.1462-2920.2009.02048.x

1073 Crotti, E., Sansonno, L., Prosdocimi, E.M., Vacchini, V., Hamdi, C., Cherif, A., Gonella, E.,
1074 Marzorati, M., Balloi, A., 2013. Microbial symbionts of honeybees: a promising tool to improve
1075 honeybee health. N. Biotechnol. 30, 716–22. doi:10.1016/j.nbt.2013.05.004

1076 Cummins, J., Gahan, C.G.M., n.d. Signature tagged mutagenesis in the functional genetic analysis of
1077 gastrointestinal pathogens. Gut Microbes 3, 93–103. doi:10.4161/gmic.19578

1078 Currie, C.R., Scott, J.A., Summerbell, R.C., Malloch, D., 1999. Fungus-growing ants use antibiotic-
1079 producing bacteria to control garden parasites 398, 701–705. doi:10.1038/19519

1080 Daffonchio, D., Borin, S., Frova, G., Manachini, P.L., Sorlini, C., 1998. PCR fingerprinting of whole
1081 genomes: the spacers between the 16S and 23S rRNA genes and of intergenic tRNA gene regions
1082 reveal a different intraspecific genomic variability of Bacillus cereus and Bacillus licheniformis
1083 [corrected]. Int. J. Syst. Bacteriol. 48 Pt 1, 107–16.

1084 Damiani, C., Ricci, I., Crotti, E., Rossi, P., Rizzi, A., Scuppa, P., Esposito, F., Bandi, C., Daffonchio,
1085 D., Favia, G., 2008. Paternal transmission of symbiotic bacteria in malaria vectors. Curr. Biol.
1086 18, R1087–8. doi:10.1016/j.cub.2008.10.040

1087 De Lipthay, J.R., Enzinger, C., Johnsen, K., Aamand, J., Sørensen, S.J., 2004. Impact of DNA
1088 extraction method on bacterial community composition measured by denaturing gradient gel
1089 electrophoresis. Soil Biol. Biochem. 36, 1607–1614. doi:10.1016/j.soilbio.2004.03.011

1090 Di Bella, J.M., Bao, Y., Gloor, G.B., Burton, J.P., Reid, G., 2013. High throughput sequencing
1091 methods and analysis for microbiome research. J. Microbiol. Methods 95, 401–14.
1092 doi:10.1016/j.mimet.2013.08.011

1093 Dillon, R.J., Dillon, V.M., 2004. The gut bacteria of insects: nonpathogenic interactions. Annu. Rev.
1094 Entomol. 49, 71–92. doi:10.1146/annurev.ento.49.061802.123416

1095 Doudoumis, V., Tsiamis, G., Wamwiri, F., Brelsfoard, C., Alam, U., Aksoy, E., Dalaperas, S., Abd-
1096 Alla, A., Ouma, J., Takac, P., Aksoy, S., Bourtzis, K., 2012. Detection and characterization of
1097 Wolbachia infections in laboratory and natural populations of different species of tsetse flies
1098 (genus Glossina). BMC Microbiol. 12, S3. doi:10.1186/1471-2180-12-S1-S3

1099 Douglas, A.E., 2011. Is the regulation of insulin signaling multi-organismal? Sci. Signal. 4, pe46.
1100 doi:10.1126/scisignal.2002669

1101 Dunning Hotopp, J.C., Clark, M.E., Oliveira, D.C.S.G., Foster, J.M., Fischer, P., Muñoz Torres,
1102 M.C., Giebel, J.D., Kumar, N., Ishmael, N., Wang, S., Ingram, J., Nene, R. V, Shepard, J.,
1103 Tomkins, J., Richards, S., Spiro, D.J., Ghedin, E., Slatko, B.E., Tettelin, H., Werren, J.H., 2007.
1104 Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. Science
1105 317, 1753–1756. doi:10.1126/science.1142490

1106 Durvasula, R., Gumbs, A., Panackal, A., Kruglov, O., Aksoy, S., Merrifield, R.B., Richards, F.F.,
1107 Beard, C.B., 1997. Prevention of insect-borne disease : An approach using transgenic 94, 3274–
1108 3278.

1109 Earl, S.C., Rogers, M.T., Keen, J., Bland, D.M., Houppert, A.S., Miller, C., Temple, I., Anderson,
1110 D.M., Marketon, M.M., 2015. Resistance to Innate Immunity Contributes to Colonization of the
1111 Insect Gut by Yersinia pestis. PLoS One 10, e0133318. doi:10.1371/journal.pone.0133318

1112 Engel, P., Martinson, V.G., Moran, N. a, 2012. Functional diversity within the simple gut microbiota
1113 of the honey bee. Proc. Natl. Acad. Sci. U. S. A. 109, 11002–7. doi:10.1073/pnas.1202970109

1114 Engel, P., Moran, N., 2013. The gut microbiota of insects - diversity in structure and function. FEMS
1115 Microbiol. Rev. 37, 699–735. doi:10.1111/1574-6976.12025

1116 Evans, J.D., Lopez, D.L., 2004. Bacterial probiotics induce an immune response in the honey bee
1117 (Hymenoptera: Apidae). J. Econ. Entomol. 97, 752–6.

1118 Favia, G., Ricci, I., Damiani, C., Raddadi, N., Crotti, E., Marzorati, M., Rizzi, A., Urso, R., Brusetti,
1119 L., Borin, S., Mora, D., Scuppa, P., Pasqualini, L., Clementi, E., Genchi, M., Corona, S., Negri,
1120 I., Grandi, G., Alma, A., Kramer, L., Esposito, F., Bandi, C., Sacchi, L., Daffonchio, D., 2007.
1121 Bacteria of the genus Asaia stably associate with Anopheles stephensi, an Asian malarial
1122 mosquito vector. Proc. Natl. Acad. Sci. U. S. A. 104, 9047–51. doi:10.1073/pnas.0610451104

1123 Fedhila, S., Daou, N., Lereclus, D., Nielsen-LeRoux, C., 2006. Identification of Bacillus cereus
1124 internalin and other candidate virulence genes specifically induced during oral infection in
1125 insects. Mol. Microbiol. 62, 339–55. doi:10.1111/j.1365-2958.2006.05362.x

1126 Feinstein, L.M., Sul, W.J., Blackwood, C.B., 2009. Assessment of bias associated with incomplete
1127 extraction of microbial DNA from soil. Appl. Environ. Microbiol. 75, 5428–33.
1128 doi:10.1128/AEM.00120-09

1129 Frydman, H.M., Li, J.M., Robson, D.N., Wieschaus, E., 2006. Somatic stem cell niche tropism in
1130 Wolbachia. Nature 441, 509–512. doi:10.1038/nature04756

1131 Fukatsu, T., Hosokawa, T., 2002. Capsule-transmitted gut symbiotic bacterium of the Japanese
1132 common plataspid stinkbug, Megacopta punctatissima. Appl. Environ. Microbiol. 68, 389–96.
1133 doi:10.1128/AEM.68.1.389-396.2002

1134 Fünfhaus, A., Poppinga, L., Genersch, E., 2013. Identification and characterization of two novel
1135 toxins expressed by the lethal honey bee pathogen Paenibacillus larvae, the causative agent of
1136 American foulbrood. Environ. Microbiol. doi:10.1111/1462-2920.12229

1137 Gaby, J.C., Buckley, D.H., 2012. A comprehensive evaluation of PCR primers to amplify the nifH
1138 gene of nitrogenase. PLoS One 7, e42149. doi:10.1371/journal.pone.0042149

1139 Gal, M., Preston, G.M., Massey, R.C., Spiers, A.J., Rainey, P.B., 2003. Genes encoding a cellulosic
1140 polymer contribute toward the ecological success of Pseudomonas fluorescens SBW25 on plant
1141 surfaces. Mol. Ecol. 12, 3109–3121. doi:10.1046/j.1365-294X.2003.01953.x

1142 Garcia-Armisen, T., Vercammen, K., Rimaux, T., Vrancken, G., Vuyst, L. De, Cornelis, P., 2012.
1143 Identification of a five-oxidoreductase-gene cluster from Acetobacter pasteurianus conferring
1144 ethanol-dependent acidification in Escherichia coli. Microbiologyopen 1, 25–32.
1145 doi:10.1002/mbo3.4

1146 Garcia-Gonzalez, E., Poppinga, L., Fünfhaus, A., Hertlein, G., Hedtke, K., Jakubowska, A.,
1147 Genersch, E., 2014. Paenibacillus larvae chitin-degrading protein PlCBP49 is a key virulence
1148 factor in American Foulbrood of honey bees. PLoS Pathog. 10, e1004284.
1149 doi:10.1371/journal.ppat.1004284

1150 Gavriel, S., Jurkevitch, E., Gazit, Y., Yuval, B., 2011. Bacterially enriched diet improves sexual
1151 performance of sterile male Mediterranean fruit flies. J. Appl. Entomol. 135, 564–573.
1152 doi:10.1111/j.1439-0418.2010.01605.x

1153 Gonella, E., Crotti, E., Rizzi, A., Mandrioli, M., Favia, G., Daffonchio, D., Alma, A., 2012.
1154 Horizontal transmission of the symbiotic bacterium Asaia sp. in the leafhopper Scaphoideus
1155 titanus Ball (Hemiptera: Cicadellidae). BMC Microbiol. 12 Suppl 1, S4. doi:10.1186/1471-
1156 2180-12-S1-S4

1157 Gonella, E., Negri, I., Marzorati, M., Mandrioli, M., Sacchi, L., Pajoro, M., Crotti, E., Rizzi, A.,
1158 Clementi, E., Tedeschi, R., Bandi, C., Alma, A., Daffonchio, D., 2011. Bacterial endosymbiont
1159 localization in Hyalesthes obsoletus, the insect vector of Bois noir in Vitis vinifera. Appl.
1160 Environ. Microbiol. 77, 1423–35. doi:10.1128/AEM.02121-10

1161 Gupta, P., Samant, K., Sahu, A., 2012. Isolation of cellulose-degrading bacteria and determination of
1162 their cellulolytic potential. Int. J. Microbiol. 2012. doi:10.1155/2012/578925

1163 Hamdi, C., Balloi, A., Essanaa, J., Crotti, E., Gonella, E., Raddadi, N., Ricci, I., Boudabous, A.,
1164 Borin, S., Manino, A., Bandi, C., Alma, A., Daffonchio, D., Cherif, A., 2011. Gut microbiome
1165 dysbiosis and honeybee health. J. Appl. Entomol. 135, 524–533. doi:10.1111/j.1439-
1166 0418.2010.01609.x

1167 Hammer, T.J., McMillan, W.O., Fierer, N., 2014. Metamorphosis of a butterfly-associated bacterial
1168 community. PLoS One 9, e86995. doi:10.1371/journal.pone.0086995

1169 Hardy, R.W., Holsten, R.D., Jackson, E.K., Burns, R.C., 1968. The acetylene-ethylene assay for n(2)
1170 fixation: laboratory and field evaluation. Plant Physiol. 43, 1185–207.

Hayashi, T., Hosokawa, T., Meng, X.-Y., Koga, R., Fukatsu, T., 2015. Female-Specific Specialization of a Posterior End Region of the Midgut Symbiotic Organ in Plautia splendens and Allied Stinkbugs. Appl. Environ. Microbiol. 81, 2603–2611. doi:10.1128/AEM.04057-14

Hazen, T.C., Dubinsky, E. a, DeSantis, T.Z., Andersen, G.L., Piceno, Y.M., Singh, N., Jansson, J.K., Probst, A., Borglin, S.E., Fortney, J.L., Stringfellow, W.T., Bill, M., Conrad, M.E., Tom, L.M., Chavarria, K.L., Alusi, T.R., Lamendella, R., Joyner, D.C., Spier, C., Baelum, J., Auer, M., Zemla, M.L., Chakraborty, R., Sonnenthal, E.L., D'haeseleer, P., Holman, H.-Y.N., Osman, S., Lu, Z., Van Nostrand, J.D., Deng, Y., Zhou, J., Mason, O.U., 2010. Deep-sea oil plume enriches indigenous oil-degrading bacteria. Science 330, 204–208. doi:10.1126/science.1195979

Hooke, R., 1667. Micrographia: or some physiological descriptions of minute bodies made by magnifying glasses: with observations and inquiries thereupon., Eebo Editi. ed. United States.

Hosokawa, T., Kikuchi, Y., Nikoh, N., Shimada, M., Fukatsu, T., 2006. Strict host-symbiont cospeciation and reductive genome evolution in insect gut bacteria. PLoS Biol. 4, e337. doi:10.1371/journal.pbio.0040337

Hurwitz, I., Fieck, A., Read, A., Hillesland, H., Klein, N., Kang, A., Durvasula, R., 2011a. Paratransgenic control of vector borne diseases. Int. J. Biol. Sci. 7, 1334–44.

Hurwitz, I., Hillesland, H., Fieck, A., Das, P., Durvasula, R., 2011b. The paratransgenic sand fly: a platform for control of Leishmania transmission. Parasit. Vectors 4, 82. doi:10.1186/1756-3305-4-82

Huys, G., Vanhoutte, T., Joossens, M., Mahious, A.S., De Brandt, E., Vermeire, S., Swings, J., 2008. Coamplification of eukaryotic DNA with 16S rRNA gene-based PCR primers: possible consequences for population fingerprinting of complex microbial communities. Curr. Microbiol. 56, 553–7. doi:10.1007/s00284-008-9122-z

Im, S.H., Galko, M.J., 2012. Pokes, sunburn, and hot sauce: Drosophila as an emerging model for the biology of nociception. Dev. Dyn. 241, 16–26. doi:10.1002/dvdy.22737

Ishmael, N., Dunning Hotopp, J.C., Ioannidis, P., Biber, S., Sakamoto, J., Siozios, S., Nene, V., Werren, J., Bourtzis, K., Bordenstein, S.R., Tettelin, H., 2009. Extensive genomic diversity of closely related Wolbachia strains. Microbiology 155, 2211–22. doi:10.1099/mic.0.027581-0

Jurkevitch, E., 2011. Riding the Trojan horse: Combating pest insects with their own symbionts. Microb. Biotechnol. 4, 620–627. doi:10.1111/j.1751-7915.2011.00249.x

Kaltenpoth, M., Göttler, W., Herzner, G., Strohm, E., 2005. Symbiotic bacteria protect wasp larvae from fungal infestation. Curr. Biol. 15, 475–9. doi:10.1016/j.cub.2004.12.084

Kane, M.D., Breznak, J.A., 1991. Effect of host diet on production of organic acids and methane by cockroach gut bacteria. Appl. Environ. Microbiol. 57, 2628–34.

Kellogg, C. a., Piceno, Y.M., Tom, L.M., DeSantis, T.Z., Zawada, D.G., Andersen, G.L., 2012. PhyloChip™ microarray comparison of sampling methods used for coral microbial ecology. J. Microbiol. Methods 88, 103–109. doi:10.1016/j.mimet.2011.10.019

1208 Kikuchi, Y., Fukatsu, T., 2013. Live imaging of symbiosis: spatiotemporal infection dynamics of a
1209 GFP-labelled Burkholderia symbiont in the bean bug Riptortus pedestris. Mol. Ecol.
1210 doi:10.1111/mec.12479

1211 Kikuchi, Y., Hayatsu, M., Hosokawa, T., Nagayama, A., Tago, K., Fukatsu, T., 2012. Symbiont-
1212 mediated insecticide resistance. Proc. Natl. Acad. Sci. U. S. A. 109, 8618–22.
1213 doi:10.1073/pnas.1200231109

1214 Kikuchi, Y., Hosokawa, T., Fukatsu, T., 2011. Specific developmental window for establishment of
1215 an insect-microbe gut symbiosis. Appl. Environ. Microbiol. 77, 4075–81.
1216 doi:10.1128/AEM.00358-11

1217 Kim, J.K., Han, S.H., Kim, C.-H., Jo, Y.H., Futahashi, R., Kikuchi, Y., Fukatsu, T., Lee, B.L., 2014.
1218 Molting-associated suppression of symbiont population and up-regulation of antimicrobial
1219 activity in the midgut symbiotic organ of the Riptortus-Burkholderia symbiosis. Dev. Comp.
1220 Immunol. 43, 10–4. doi:10.1016/j.dci.2013.10.010

1221 Kim, O.-S., Cho, Y.-J., Lee, K., Yoon, S.-H., Kim, M., Na, H., Park, S.-C., Jeon, Y.S., Lee, J.-H., Yi,
1222 H., Won, S., Chun, J., 2012. Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence
1223 database with phylotypes that represent uncultured species. Int. J. Syst. Evol. Microbiol. 62,
1224 716–21. doi:10.1099/ijs.0.038075-0

1225 Köhler, T., Dietrich, C., Scheffrahn, R.H., Brune, A., 2012. High-resolution analysis of gut
1226 environment and bacterial microbiota reveals functional compartmentation of the gut in wood-
1227 feeding higher termites (Nasutitermes spp.). Appl. Environ. Microbiol. 78, 4691–701.
1228 doi:10.1128/AEM.00683-12

1229 Kounatidis, I., Crotti, E., Sapountzis, P., Sacchi, L., Rizzi, A., Chouaia, B., Bandi, C., Alma, A.,
1230 Daffonchio, D., Mavragani-Tsipidou, P., Bourtzis, K., 2009. Acetobacter tropicalis is a major
1231 symbiont of the olive fruit fly (Bactrocera oleae). Appl. Environ. Microbiol. 75, 3281–8.
1232 doi:10.1128/AEM.02933-08

1233 Lambertsen, L., Sternberg, C., Molin, S., 2004. Mini-Tn7 transposons for site-specific tagging of
1234 bacteria with fluorescent proteins. Environ. Microbiol. 6, 726–32. doi:10.1111/j.1462-
1235 2920.2004.00605.x

1236 Langille, M.G.I., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., Clemente,
1237 J.C., Burkepile, D.E., Vega Thurber, R.L., Knight, R., Beiko, R.G., Huttenhower, C., 2013.
1238 Predictive functional profiling of microbial communities using 16S rRNA marker gene
1239 sequences. Nat. Biotechnol. 31, 814–21. doi:10.1038/nbt.2676

1240 Lee, F.J., Rusch, D.B., Stewart, F.J., Mattila, H.R., Newton, I.L.G., 2014. Saccharide breakdown and
1241 fermentation by the honey bee gut microbiome. Environ. Microbiol. doi:10.1111/1462-
1242 2920.12526

1243 Lee, K.-A., Kim, S.-H., Kim, E.-K., Ha, E.-M., You, H., Kim, B., Kim, M.-J., Kwon, Y., Ryu, J.-H.,
1244 Lee, W.-J., 2013. Bacterial-Derived Uracil as a Modulator of Mucosal Immunity and Gut-
1245 Microbe Homeostasis in Drosophila. Cell 153, 797–811. doi:10.1016/j.cell.2013.04.009

1246 Lemke, T., Stingl, U., Egert, M., Friedrich, M.W., Brune, A., 2003. Physicochemical conditions and
1247 microbial activities in the highly alkaline gut of the humus-feeding larva of Pachnoda ephippiata

1248 (Coleoptera: Scarabaeidae). Appl. Environ. Microbiol. 69, 6650–8.
1249 doi:10.1128/AEM.69.11.6650-6658.2003

1250 Lisdiyanti, P., Kawasaki, H., Seki, T., Yamada, Y., Uchimura, T., Komagata, K., 2001. Identification
1251 of Acetobacter strains isolated from Indonesian sources, and proposals of Acetobacter syzygii
1252 sp. nov., Acetobacter cibinongensis sp. nov., and Acetobacter orientalis sp. nov. J. Gen. Appl.
1253 Microbiol. 47, 119–131.

1254 Liu, N., Zhang, L., Zhou, H., Zhang, M., Yan, X., Wang, Q., Long, Y., Xie, L., Wang, S., Huang, Y.,
1255 Zhou, Z., 2013. Metagenomic insights into metabolic capacities of the gut microbiota in a
1256 fungus-cultivating termite (Odontotermes yunnanensis). PLoS One 8, e69184.
1257 doi:10.1371/journal.pone.0069184

1258 Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., Lai, T.,
1259 Steppi, S., Jobb, G., Förster, W., Brettske, I., Gerber, S., Ginhart, A.W., Gross, O., Grumann, S.,
1260 Hermann, S., Jost, R., König, A., Liss, T., Lüssmann, R., May, M., Nonhoff, B., Reichel, B.,
1261 Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A.,
1262 Schleifer, K.-H., 2004. ARB: a software environment for sequence data. Nucleic Acids Res. 32,
1263 1363–71. doi:10.1093/nar/gkh293

1264 Mapelli, F., Marasco, R., Rolli, E., Barbato, M., Cherif, H., Guesmi, A., Ouzari, I., Daffonchio, D.,
1265 Borin, S., 2013. Potential for plant growth promotion of rhizobacteria associated with Salicornia
1266 growing in Tunisian hypersaline soils. Biomed Res. Int. 2013. doi:10.1155/2013/248078

1267 Marasco, R., Rolli, E., Ettoumi, B., Vigani, G., Mapelli, F., Borin, S., Abou-Hadid, A.F., El-Behairy,
1268 U. a., Sorlini, C., Cherif, A., Zocchi, G., Daffonchio, D., 2012. A Drought Resistance-Promoting
1269 Microbiome Is Selected by Root System under Desert Farming. PLoS One 7.
1270 doi:10.1371/journal.pone.0048479

1271 Mariconti, M., Epis, S., Gaibani, P., Dalla Valle, C., Sassera, D., Tomao, P., Fabbi, M., Castelli, F.,
1272 Marone, P., Sambri, V., Bazzocchi, C., Bandi, C., 2012. Humans parasitized by the hard tick
1273 Ixodes ricinus are seropositive to Midichloria mitochondrii: is Midichloria a novel pathogen, or
1274 just a marker of tick bite? Pathog. Glob. Health 106, 391–6.
1275 doi:10.1179/2047773212Y.0000000050

1276 Martinson, V., Moy, J., Moran, N., 2012. Establishment of characteristic gut bacteria during
1277 development of the honeybee worker. Appl. Environ. Microbiol. 78, 2830–2840.
1278 doi:10.1128/AEM.07810-11

1279 Marzorati, M., Alma, A., Sacchi, L., Pajoro, M., Palermo, S., Brusetti, L., Raddadi, N., Balloi, A.,
1280 Tedeschi, R., Clementi, E., Corona, S., Quaglino, F., Bianco, P.A., Beninati, T., Bandi, C.,
1281 Daffonchio, D., 2006. A novel Bacteroidetes symbiont is localized in Scaphoideus titanus, the
1282 insect vector of Flavescence Dorèe in Vitis vinifera. Appl. Environ. Microbiol. 72, 1467–1475.
1283 doi:10.1128/AEM.72.2.1467-1475.2006

1284 Matsuura, Y., Kikuchi, Y., Miura, T., Fukatsu, T., 2015. *Ultrabithorax* is essential for bacteriocyte
1285 development. Proc. Natl. Acad. Sci. 201503371. doi:10.1073/pnas.1503371112

1286 Matthew, C.Z., Darby, A.C., Young, S. a., Hume, L.H., Welburn, S.C., 2005. The rapid isolation and
1287 growth dynamics of the tsetse symbiont Sodalis glossinidius. FEMS Microbiol. Lett. 248, 69–
1288 74. doi:10.1016/j.femsle.2005.05.024

1289 Merrell, D.S., Camilli, A., 2000. Detection and analysis of gene expression during infection by in
1290 vivo expression technology. Phil. Trans. R. Soc. Lond. B 355, 587–599.
1291 doi:10.1098/rstb.2000.0600

1292 Mitraka, E., Stathopoulos, S., Siden-Kiamos, I., Christophides, G.K., Louis, C., 2013. Asaia
1293 accelerates larval development of Anopheles gambiae. Pathog. Glob. Health 107, 305–11.
1294 doi:10.1179/2047773213Y.0000000106

1295 Mølbak, L., Molin, S., Kroer, N., 2007. Root growth and exudate production define the frequency of
1296 horizontal plasmid transfer in the Rhizosphere. FEMS Microbiol. Ecol. 59, 167–76.
1297 doi:10.1111/j.1574-6941.2006.00229.x

1298 Montagna, M., Chouaia, B., Mazza, G., Prosdocimi, E.M., Crotti, E., Mereghetti, V., Vacchini, V.,
1299 Giorgi, A., De Biase, A., Longo, S., Cervo, R., Lozzia, G.C., Alma, A., Bandi, C., Daffonchio,
1300 D., 2015. Effects of the diet on the microbiota of the red palm weevil (coleoptera:
1301 dryophthoridae). PLoS One 10, e0117439. doi:10.1371/journal.pone.0117439

1302 Morales-Jiménez, J., Vera-Ponce de León, A., García-Domínguez, A., Martínez-Romero, E., Zúñiga,
1303 G., Hernández-Rodríguez, C., 2013. Nitrogen-fixing and uricolytic bacteria associated with the
1304 gut of Dendroctonus rhizophagus and Dendroctonus valens (Curculionidae: Scolytinae).
1305 Microb. Ecol. 66, 200–10. doi:10.1007/s00248-013-0206-3

1306 Moran, N., McCutcheon, J.P., Nakabachi, A., 2008. Genomics and evolution of heritable bacterial
1307 symbionts. Annu. Rev. Genet. 42, 165–90. doi:10.1146/annurev.genet.41.110306.130119

1308 Mostafa, H.E., Heller, K.J., Geis, A., 2002. Cloning of Escherichia coli lacZ and lacY Genes and
1309 Their Expression in Gluconobacter oxydans and Acetobacter liquefaciens 68, 2619–2623.
1310 doi:10.1128/AEM.68.5.2619

1311 Nikodinovic, J., Barrow, K.D., Chuck, J.-A., 2003. High yield preparation of genomic DNA from
1312 Streptomyces. Biotechniques 35, 932–4, 936.

1313 Nikolaki, S., Tsiamis, G., 2013. Microbial diversity in the era of omic technologies. Biomed Res. Int.
1314 2013. doi:10.1155/2013/958719

1315 Ohkuma, M., 2003. Termite symbiotic systems: efficient bio-recycling of lignocellulose. Appl.
1316 Microbiol. Biotechnol. 61, 1–9. doi:10.1007/s00253-002-1189-z

1317 Park, D.-S., Oh, H.-W., Jeong, W.-J., Kim, H., Park, H.-Y., Bae, K.S., 2007. A culture-based study
1318 of the bacterial communities within the guts of nine longicorn beetle species and their exo-
1319 enzyme producing properties for degrading xylan and pectin. J. Microbiol. 45, 394–401.

1320 Perotti, M.A., Allen, J.M., Reed, D.L., Braig, H.R., 2007. Host-symbiont interactions of the primary
1321 endosymbiont of human head and body lice. FASEB J. 21, 1058–66. doi:10.1096/fj.06-6808com

1322 Pobigaylo, N., Wetter, D., Szymczak, S., Schiller, U., Kurtz, S., Meyer, F., Nattkemper, T.W.,
1323 Becker, A., 2006. Construction of a large signature-tagged mini-Tn5 transposon library and its
1324 application to mutagenesis of Sinorhizobium meliloti. Appl. Environ. Microbiol. 72, 4329–37.
1325 doi:10.1128/AEM.03072-05

1326 Prado, S.S., Rubinoff, D., Almeida, R.P.P., 2006. Vertical Transmission of a Pentatomid Caeca-
1327 Associated Symbiont. Ann. Entomol. Soc. Am. 99, 577–585. doi:10.1603/0013-
1328 8746(2006)99[577:VTOAPC]2.0.CO;2

1329 Prosdocimi, E.M., Novati, S., Bruno, R., Bandi, C., Mulatto, P., Giannico, R., Casiraghi, M., Ferri,
1330 E., 2013. Errors in ribosomal sequence datasets generated using PCR-coupled "panbacterial"
1331 pyrosequencing, and the establishment of an improved approach. Mol. Cell. Probes 27, 65–7.
1332 doi:10.1016/j.mcp.2012.07.003

1333 Pruesse, E., Peplies, J., Glöckner, F.O., 2012. SINA: accurate high-throughput multiple sequence
1334 alignment of ribosomal RNA genes. Bioinformatics 28, 1823–9.
1335 doi:10.1093/bioinformatics/bts252

1336 Raddadi, N., Gonella, E., Camerota, C., Pizzinat, A., Tedeschi, R., Crotti, E., Mandrioli, M., Bianco,
1337 P.A., Daffonchio, D., Alma, A., 2011. "Candidatus Liberibacter europaeus" sp. nov. that is
1338 associated with and transmitted by the psyllid Cacopsylla pyri apparently behaves as an
1339 endophyte rather than a pathogen. Environ. Microbiol. 13, 414–26. doi:10.1111/j.1462-
1340 2920.2010.02347.x

1341 Rizzi, A., Crotti, E., Borruso, L., Jucker, C., Lupi, D., Colombo, M., Daffonchio, D., 2013.
1342 Characterization of the bacterial community associated with larvae and adults of anoplophora
1343 chinensis collected in Italy by culture and culture-independent methods. Biomed Res. Int. 2013.
1344 doi:10.1155/2013/420287

1345 Russell, J.A., Funaro, C.F., Giraldo, Y.M., Goldman-Huertas, B., Suh, D., Kronauer, D.J.C., Moreau,
1346 C.S., Pierce, N.E., 2012. A veritable menagerie of heritable bacteria from ants, butterflies, and
1347 beyond: broad molecular surveys and a systematic review. PLoS One 7, e51027.
1348 doi:10.1371/journal.pone.0051027

1349 Ryu, J.-H., Kim, S.-H., Lee, H.-Y., Bai, J.Y., Nam, Y.-D., Bae, J.-W., Lee, D.G., Shin, S.C., Ha, E.-
1350 M., Lee, W.-J., 2008. Innate immune homeostasis by the homeobox gene caudal and
1351 commensal-gut mutualism in Drosophila. Science 319, 777–82. doi:10.1126/science.1149357

1352 Sagaram, U.S., Deangelis, K.M., Trivedi, P., Andersen, G.L., Lu, S.E., Wang, N., 2009. Bacterial
1353 diversity analysis of huanglongbing pathogen-infected citrus, using phyloChip arrays and 16S
1354 rRNA gene clone library sequencing. Appl. Environ. Microbiol. 75, 1566–1574.
1355 doi:10.1128/AEM.02404-08

1356 Sandionigi, A., Vicario, S., Prosdocimi, E.M., Galimberti, A., Ferri, E., Bruno, A., Balech, B.,
1357 Mezzasalma, V., Casiraghi, M., 2014. Toward a better understanding of Apis mellifera and
1358 Varroa destructor microbiomes: introducing "PhyloH" as a novel phylogenetic diversity analysis
1359 tool. Mol. Ecol. Resour. 1–14. doi:10.1111/1755-0998.12341

1360 Santaella, C., Schue, M., Berge, O., Heulin, T., Achouak, W., 2008. The exopolysaccharide of
1361 Rhizobium sp. YAS34 is not necessary for biofilm formation on Arabidopsis thaliana and
1362 Brassica napus roots but contributes to root colonization. Environ. Microbiol. 10, 2150–63.
1363 doi:10.1111/j.1462-2920.2008.01650.x

1364 Santo Domingo, J.W., Kaufman, M.G., Klug, M.J., Holben, W.E., Harris, D., Tiedje, J.M., 1998.
1365 Influence of diet on the structure and function of the bacterial hindgut community of crickets.
1366 Mol. Ecol. 7, 761–767. doi:10.1046/j.1365-294x.1998.00390.x

1367 Sassera, D., Beninati, T., Bandi, C., Bouman, E. a P., Sacchi, L., Fabbi, M., Lo, N., 2006. "Candidatus
1368 Midichloria mitochondrii", an endosymbiont of the tick Ixodes ricinus with a unique
1369 intramitochondrial lifestyle. Int. J. Syst. Evol. Microbiol. 56, 2535–40. doi:10.1099/ijs.0.64386-
1370 0

1371 Schramm, A., Fuchs, B.M., Nielsen, J.L., Tonolla, M., Stahl, D. a., 2002. Fluorescence in situ
1372 hybridization of 16S rRNA gene clones (Clone-FISH) for probe validation and screening of
1373 clone libraries. Environ. Microbiol. 4, 713–720. doi:10.1046/j.1462-2920.2002.00364.x

1374 Scully, E.D., Geib, S.M., Hoover, K., Tien, M., Tringe, S.G., Barry, K.W., Glavina del Rio, T.,
1375 Chovatia, M., Herr, J.R., Carlson, J.E., 2013a. Metagenomic profiling reveals lignocellulose
1376 degrading system in a microbial community associated with a wood-feeding beetle. PLoS One
1377 8, e73827. doi:10.1371/journal.pone.0073827

1378 Scully, E.D., Hoover, K., Carlson, J.E., Tien, M., Geib, S.M., 2013b. Midgut transcriptome profiling
1379 of Anoplophora glabripennis, a lignocellulose degrading cerambycid beetle. BMC Genomics 14,
1380 850. doi:10.1186/1471-2164-14-850

1381 Sharon, G., Segal, D., Ringo, J.M., Hefetz, A., Zilber-Rosenberg, I., Rosenberg, E., 2010. Commensal
1382 bacteria play a role in mating preference of Drosophila melanogaster. Proc. Natl. Acad. Sci. U.
1383 S. A. 107, 20051–6. doi:10.1073/pnas.1009906107

1384 Shin, S.C., Kim, S.-H., You, H., Kim, B., Kim, A.C., Lee, K.-A., Yoon, J.-H., Ryu, J.-H., Lee, W.-
1385 J., 2011. Drosophila microbiome modulates host developmental and metabolic homeostasis via
1386 insulin signaling. Science 334, 670–674. doi:10.1126/science.1212782

1387 Stoll, S., Feldhaar, H., Fraunholz, M.J., Gross, R., 2010. Bacteriocyte dynamics during development
1388 of a holometabolous insect, the carpenter ant Camponotus floridanus. BMC Microbiol. 10, 308.
1389 doi:10.1186/1471-2180-10-308

1390 Storelli, G., Defaye, A., Erkosar, B., Hols, P., Royet, J., Leulier, F., 2011. Lactobacillus plantarum
1391 promotes drosophila systemic growth by modulating hormonal signals through TOR-dependent
1392 nutrient sensing. Cell Metab. 14, 403–414. doi:10.1016/j.cmet.2011.07.012

1393 Stouthamer, R., Breeuwer, J.A., Hurst, G.D., 1999. Wolbachia pipientis: microbial manipulator of
1394 arthropod reproduction. Annu. Rev. Microbiol. 53, 71–102. doi:10.1146/annurev.micro.53.1.71

1395 Sudakaran, S., Salem, H., Kost, C., Kaltenpoth, M., 2012. Geographical and ecological stability of
1396 the symbiotic mid-gut microbiota in European firebugs, Pyrrhocoris apterus (Hemiptera,
1397 Pyrrhocoridae). Mol. Ecol. 21, 6134–51. doi:10.1111/mec.12027

1398 Sugimura, M., Watanabe, H., Lo, N., Saito, H., 2003. Purification, characterization, cDNA cloning
1399 and nucleotide sequencing of a cellulase from the yellow-spotted longicorn beetle, Psacothea
1400 hilaris. Eur. J. Biochem. 270, 3455–3460. doi:10.1046/j.1432-1033.2003.03735.x

1401 Toenshoff, E.R., Szabó, G., Gruber, D., Horn, M., 2014. The pine bark Adelgid, Pineus strobi,
1402 contains two novel bacteriocyte-associated gammaproteobacterial symbionts. Appl. Environ.
1403 Microbiol. 80, 878–85. doi:10.1128/AEM.03310-13

1404 Toju, H., Fukatsu, T., 2011. Diversity and infection prevalence of endosymbionts in natural
1405 populations of the chestnut weevil: relevance of local climate and host plants. Mol. Ecol. 20,
1406 853–68. doi:10.1111/j.1365-294X.2010.04980.x

1407 Tsiamis, G., Tzagkaraki, G., Chamalaki, A., Xypteras, N., Andersen, G., Vayenas, D., Bourtzis, K.,
1408 2012. Olive-mill wastewater bacterial communities display a cultivar specific profile. Curr.
1409 Microbiol. 64, 197–203. doi:10.1007/s00284-011-0049-4

1410 Turnbaugh, P.J., Ley, R.E., Hamady, M., Fraser-Liggett, C.M., Knight, R., Gordon, J.I., 2007. The
1411 human microbiome project. Nature 449, 804–10. doi:10.1038/nature06244

1412 Veneti, Z., Clark, M.E., Karr, T.L., Savakis, C., Bourtzis, K., 2004. Heads or tails: Host-parasite
1413 interactions in the Drosophila-Wolbachia system. Appl. Environ. Microbiol. 70, 5366–5372.
1414 doi:10.1128/AEM.70.9.5366-5372.2004

1415 Veneti, Z., Clark, M.E., Zabalou, S., Karr, T.L., Savakis, C., Bourtzis, K., 2003. Cytoplasmic
1416 incompatibility and sperm cyst infection in different Drosophila-Wolbachia associations.
1417 Genetics 164, 545–552.

1418 Vestheim, H., Jarman, S.N., 2008. Blocking primers to enhance PCR amplification of rare sequences
1419 in mixed samples - a case study on prey DNA in Antarctic krill stomachs. Front. Zool. 5, 12.
1420 doi:10.1186/1742-9994-5-12

1421 Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naive Bayesian classifier for rapid
1422 assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ. Microbiol. 73,
1423 5261–7. doi:10.1128/AEM.00062-07

1424 Wang, Y., Qian, P.-Y., 2009. Conservative fragments in bacterial 16S rRNA genes and primer design
1425 for 16S ribosomal DNA amplicons in metagenomic studies. PLoS One 4, e7401.
1426 doi:10.1371/journal.pone.0007401

1427 Westermann, A.J., Gorski, S.A., Vogel, J., 2012. Dual RNA-seq of pathogen and host. Nat. Rev.
1428 Microbiol. 10, 618–630. doi:10.1038/nrmicro2852

1429 Wilkinson, T.., 1998. The elimination of intracellular microorganisms from insects: an analysis of
1430 antibiotic-treatment in the pea aphid (Acyrthosiphon pisum). Comp. Biochem. Physiol. Part A
1431 Mol. Integr. Physiol. 119, 871–881. doi:10.1016/S1095-6433(98)00013-0

1432 Wilkinson, T.L., Ishikawa, H., 2000. Injection of essential amino acids substitutes for bacterial supply
1433 in aposymbiotic pea aphids (Acyrthosiphon pisum). Entomol. Exp. Appl. 94, 85–91.
1434 doi:10.1046/j.1570-7458.2000.00607.x

1435 Xie, W., Meng, Q., Wu, Q., Wang, S., Yang, X., Yang, N., Li, R., Jiao, X., Pan, H., Liu, B., Su, Q.,
1436 Xu, B., Hu, S., Zhou, X., Zhang, Y., 2012. Pyrosequencing the Bemisia tabaci transcriptome
1437 reveals a highly diverse bacterial community and a robust system for insecticide resistance.
1438 PLoS One 7, e35181. doi:10.1371/journal.pone.0035181

1439 Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.-H., Whitman, W.B.,
1440 Euzéby, J., Amann, R., Rosselló-Móra, R., 2014. Uniting the classification of cultured and
1441 uncultured bacteria and archaea using 16S rRNA gene sequences. Nat. Rev. Microbiol. 12, 635–
1442 645. doi:10.1038/nrmicro3330

1443 Zabalou, S., Riegler, M., Theodorakopoulou, M., Stauffer, C., Savakis, C., Bourtzis, K., 2004.
1444 Wolbachia-induced cytoplasmic incompatibility as a means for insect pest population control.
1445 Proc. Natl. Acad. Sci. U. S. A. 101, 15042–15045. doi:10.1073/pnas.0403853101

1446 Zouache, K., Raharimalala, F.N., Raquin, V., Tran-Van, V., Raveloson, L.H.R., Ravelonandro, P.,
1447 Mavingui, P., 2011. Bacterial diversity of field-caught mosquitoes, Aedes albopictus and Aedes
1448 aegypti, from different geographic regions of Madagascar. FEMS Microbiol. Ecol. 75, 377–89.
1449 doi:10.1111/j.1574-6941.2010.01012.x

1450

1451 **Table 1.** Examples of screening tests to assess the symbiotic potential of the isolate collection.

Test	References	Description
Endoglucanase	(Compant et al., 2005; Gupta et al., 2012)	Congo-red staining
Xylanase	(Park et al., 2007)	Congo-red staining
Pectin degradation	(Engel et al., 2012; Park et al., 2007)	CTAB staining
Protease secretion	(Garcia-Armisen et al., 2012)	Degragation halo in milk agar plates
EPS production	(Santaella et al., 2008)	Sucrose-rich medium
Urease	(Arioli et al., 2007)	Colorimetric assay
Uricase	(Morales-Jiménez et al., 2013)	Degradation halo in uric acid agar plates
Nitrogen fixation	(Gaby and Buckley, 2012)	PCR assay
Ammonia production	(Cappuccino and Sherman, 1992)	Colorimetric assay

1452

1453

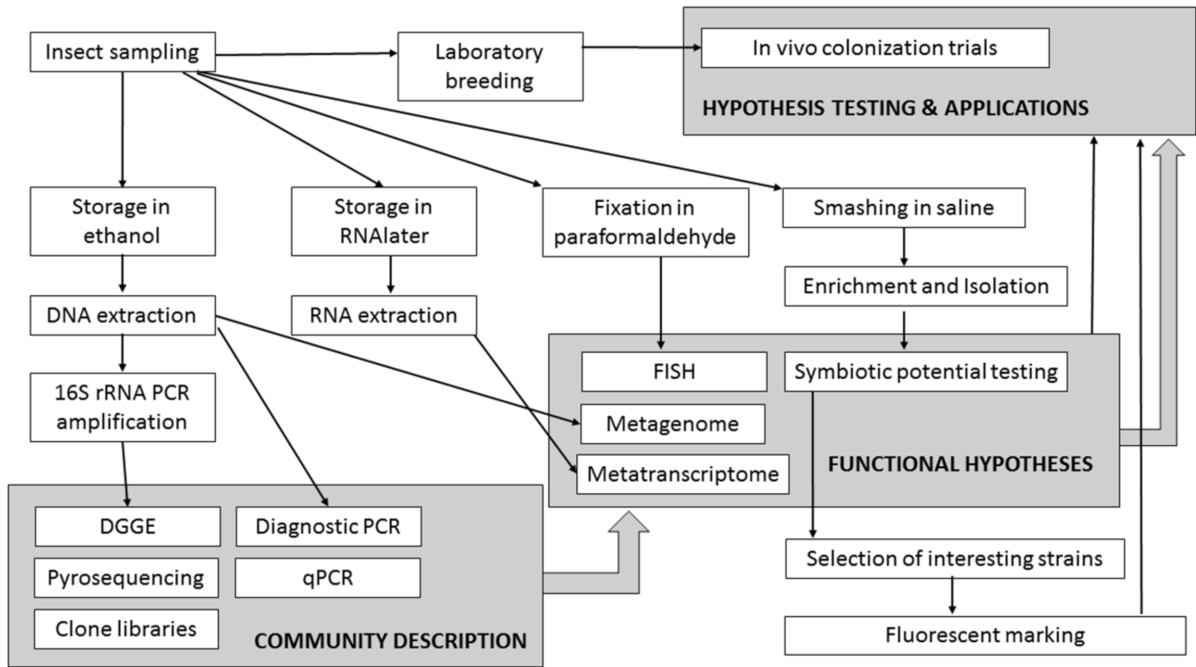
1454 **Table 2.** Examples of marking systems used to label bacteria in insect-related studies.

Fluorophore	**Transformation method**	**Strains/plasmids required**	**Target bacteria**	**Ref.**
Chromosomal GFP (mini tn7); Insertion site attTn7	Conjugation and transposition	Delivery strain: *E. coli* XL1-Blue/miniTn7 (AKN67); Helper strain: *E. coli* HB101/pRK600 (AKN 67); Helper strain: *E. coli* SM10 ëpir/pUX-BF13 (AKN68)	*Pseudomonas*, *Asaia*	(Gonella et al., 2012; Lambertsen et al., 2004)
Chromosomal dsREd (mini tn5); Insertion site casual	Conjugation and transposition	Delivery strain: *E.coli* Mv1190ëpir/TTN151 (LAM#5); Helper strain: *E. coli* HB101/pRK600 (#331)	*Pseudomonas*, *Asaia*	(Crotti et al., 2009; Mølbak et al., 2007)
Episomal GFP	Electroporation	*E. coli* pHM2-Gfp	Gram negatives	(Favia et al., 2007)

1455

1456 **Figure 1.** Schematic representation of the workflow for the characterization of the microbial

1457 community associated to insects.



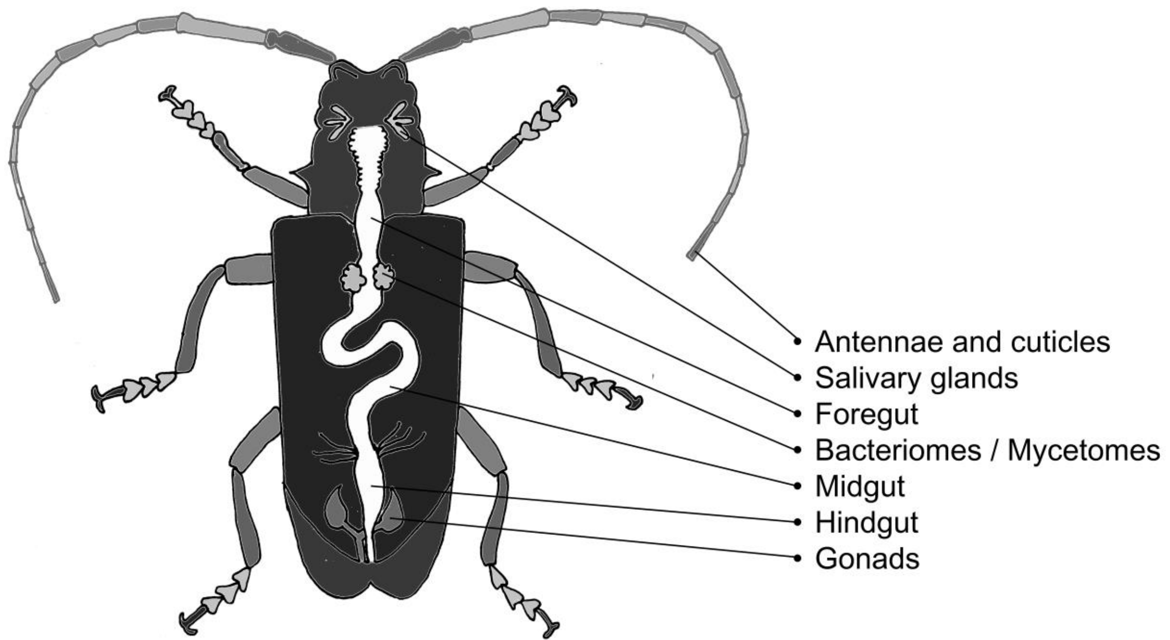
1458

1459

1460

1461

1462 **Figure 2.** Insect organs that could be the target of microbiological analysis.



1463