Root microbiome relates to plant host evolution in maize and other Poaceae

Marie-Lara Bouffaud,1,2,3 Marie-Andrée Poirier,1,2,3 Daniel Muller1,2,3 and Yvan Moënne-Loccoz1,2,3*
1Université de Lyon, F-69622, Lyon, France.
2Université Lyon 1, Villeurbanne, France.
3CNRS, UMR5557, Ecologie Microbienne, Villeurbanne, France.

Summary
Prokaryote–eukaryote interactions are primordial, but host selection of its bacterial community remains poorly understood. Because eukaryote evolution affects numerous traits shaping the ecology of their microbiome, we can expect that many evolutionary changes in the former will have the potential to impact on the composition of the latter. Consequently, the more phylogenetically distant the eukaryotic hosts, the more distinct their associated bacterial communities should be. We tested this with plants, by comparing the bacterial communities associated with maize genotypes or other Poaceae. 16S rRNA taxonomic microarray analysis showed that the genetic distance between rhizobacterial communities correlated significantly with the phylogenetic distance (derived from chloroplastic sequences) between Poaceae genotypes. This correlation was also significant when considering specific bacterial populations from all main bacterial divisions, instead of the whole rhizobacterial community. These results indicate that eukaryotic host’s evolutionary history can be a significant factor shaping directly the assembly and composition of its associated bacterial compartment.

Introduction
The physiology and ecology of eukaryotic organisms is influenced by various types of interactions with microbial partners, ranging from parasitism to mutualism (Read et al., 2008; Provorov and Vorobyov, 2009). The mechanisms and establishment of these interactions depend on the particularities of hosts and microbial partners considered (Ammar et al., 2009; Hardoim et al., 2011). As a consequence, different eukaryotic hosts may display different types of microbial community, regardless of whether mammals and their gut microbes (Ley et al., 2008; Muegge et al., 2011), insects and their endosymbionts (Moran et al., 2008), or plants and their rhizosphere or phyllosphere microbiota (Wieland et al., 2001; Kennedy et al., 2004; Ushio et al., 2008; Redford et al., 2010; Hardoim et al., 2011; Bouffaud et al., 2012) are concerned. Indeed, microbial communities of animals (gut) and plants display striking similarities (Gopal et al., 2013).

Eukaryotic hosts from a same taxonomic group share a number of common life history traits, and many of the latter will determine selection pressures for associated microorganisms by conditioning host properties in terms of colonization surfaces, availability of carbon and energy sources, and exposure to defence mechanisms and secondary metabolites (Ammar et al., 2009; Voges et al., 2010). Such group-related life history traits have been evidenced for animals and plants, regardless of whether these groups belong to a same species or correspond to different species (Ley et al., 2008; Ammar et al., 2009; Voges et al., 2010; Hardoim et al., 2011). Indeed, eukaryotic hosts from a same taxonomic group tend to display similar microbial communities (Ley et al., 2008; Bouffaud et al., 2012). If we take into consideration the origin of these taxonomic groups, it can be expected that past evolutionary history of eukaryotic hosts will influence many of the traits likely to determine the composition of their associated microbial community.

This hypothesis has two implications. The first one is the existence of a relation between host evolution and microbiome composition. This relation was not significant in the case of sponge bacteria (e.g. Schmitt et al., 2012). It was not well established with gut bacteria, for which such a relation was indeed found when considering only great apes (Ochman et al., 2010) but not a wider range of mammals (Ley et al., 2008). However, this type of relation was evidenced more clearly with plant-associated bacteria, based on analysis of interactions between tree species and phyllosphere bacteria (Redford et al., 2010), as well as maize genetic groups and rhizobacteria (Bouffaud et al., 2012; Peiffer et al., 2013). The second implication is that the more phylogenetically distant eukaryotic hosts are, the more distinct their associated bacterial communities would be. This second implication
did not prove correct for sponge bacteria (Schmitt et al., 2012), gut bacteria associated with wild hominids (Ochman et al., 2010) or root bacteria associated with genetic groups of maize (Bouffaud et al., 2012), a crop domesticated from teosinte in Mexico and disseminated in southern and northern America and later to Europe (Doebley, 2004). Perhaps it is due to the facts that all maize lines belonged to the same species (*Zea mays*) and originated from rather recent (about 9000 years ago) domestication of teosinte. Therefore, the hypothesis that rhizobacterial community composition could be directly related to plant evolutionary history requires further appraisal based on a wider range of plant genotypes, and this was the focus of the current work.

In this study, the relation between plant evolution and rhizobacterial community composition was assessed using genotypes from maize and related Poaceae. The work was carried out using two lines from each of the two maize genetic groups displaying the most different rhizobacterial communities in Bouffaud and colleagues (2012). The experiment also included another *Z. mays* genotype (i.e. a teosinte; *Z. mays* ssp. *parviglumis*), another member of the *Panicoideae* subfamily (i.e. sorghum), and a member of another subfamily within the monocots (i.e. wheat; *Pooideae* subfamily). Chloroplastic genes, which have been widely studied to reveal plant phylogeny (Hausner et al., 2006) including for Angiosperm subfamilies (Shaw et al., 2007; Hodge et al., 2010), were used here to estimate genetic distances between these plant genotypes.

In the current work, the eight plant genotypes selected were grown in a same soil originating from a cultivated field in Europe, as described in Bouffaud and colleagues’ (2012), and rhizobacterial community composition was assessed using a 16S rRNA taxonomic microarray previously validated for rhizosphere studies (Sanguin et al., 2006; Kyselková et al., 2009) and used to compare maize lines (Bouffaud et al., 2012).

**Results**

**Relationship between maize/Poaceae evolution and rhizobacterial community**

In the current work, principal component analysis (PCA) of 16S rRNA taxonomic microarray data showed that bacterial community composition in the rhizosphere (at 21 days in cropped soil) differed from that in bulk soil (as expected), and that rhizobacterial community composition differed according to the Poaceae genotype (Fig. 1 and Supporting Information Fig. S1). Whether or not the latter differences were in proportion to plant phylogenetic distance was then investigated.

Pairwise, maximum-likelihood phylogenetic distances within the Poaceae studied ranged from $-3.24$ to $-0.95$ log based on analysis of three chloroplastic sequences, i.e. gene *rps16* and the intergenic spacers *rps16-trnK* and *atpI-atpH* (Fig. 2). The lowest distances (from $-3.24$ to $-2.94$ log) were between different maize lines and the highest distances (all at $-0.95$ log) between maize lines and wheat, which was the only genotype outside of the *Panicoideae* subfamily. These phylogenetic distances were in accordance with the taxonomic relations between the eight Poaceae studied. Therefore, our maize-based experimental set-up, which included nested comparisons.

**Fig. 1.** Principal component analysis of bulk soil and Poaceae rhizospheres at 21 days in cropped soil based on 16S rRNA microarray data. FV4 and W85 belong to maize group Northern Flint, and FV252 and Mo17 to maize group Corn Belt Dent. For each treatment (bulk soil and rhizosphere samples), mean and standard deviations are represented ($n = 5$). Statistical relations are indicated, along PC1 and PC2 axes, by letters a–e (ANOVA and Fisher’s LSD tests; $P < 0.05$).
between maize lines and (i) other maize lines from the same genetic groups of *Z. mays* ssp. *mays*, (ii) other maize lines from distinct genetic groups, (iii) another *Z. mays* from a distinct subspecies (*Z. mays* ssp. *parviglumis*, i.e. teosinte), (iv) another genus of the Panicoideae subfamily (i.e. sorghum) and (v) a member of another monocot subfamily (i.e. wheat; *Pooideae* subfamily), was appropriate to assess the relation between plant phylogenetic distance and the distance between the corresponding rhizobacterial communities.

A significant positive correlation (Pearson coefficient $r = 0.91$, $n = 17$, $P = 6 \times 10^{-7}$) was found when comparing the log value of plant phylogenetic distance (i.e. between individual maize lines or between a maize line and another Poaceae; without including the FV4/Mo17 comparison as distance was 0) with the Bray–Curtis distance (based on normalized hybridization intensity data) between the corresponding rhizobacterial communities in cropped soil at 21 days (Fig. 2). When the other comparisons possible were included, i.e. teosinte-sorghum, teosinte-wheat and sorghum-wheat, the correlation level was lower but still highly significant ($r = 0.68$, $n = 20$, $P = 0.0011$). The two correlations above between maize/Poaceae evolution and rhizobacterial community composition were maintained when replacing the Bray–Curtis distance by the Euclidean distance (derived from Fig. 1) between treatment positions along axis PC1 ($r = 0.88$, $n = 17$, $P = 3 \times 10^{-6}$ and $r = 0.70$, $n = 20$, $P = 6 \times 10^{-4}$ respectively), but not along axis 2.

**Relationship between maize/Poaceae evolution and particular rhizobacterial populations**

Among the 1033 probes of the microarray, those hybridized with at least one of the five individual plants from a given Poaceae genotype in cropped field at 21 days ranged from 99 (W85 maize) to 198 (wheat), giving a total of 268 hybridized probes. Among them, a significant positive correlation ($P < 0.05$) was found when comparing the log value of the maximum-likelihood phylogenetic distance (as above) and the normalized signal level for 90 probes, i.e. 34% of the 268 probes (Table 1).

These correlations were evidenced within each of the six main bacterial groups detected, i.e. the *Alphaproteobacteria* (for 21 of 60 probes, i.e. 35%), *Betaproteobacteria* (for 13 of 43 probes, i.e. 30%), *Gammaproteobacteria* (for 14 of 28 probes, i.e. 50%) and *Deltaproteobacteria* (for 6 of 22 probes, i.e. 27%) classes, and the *Firmicutes* (for 14 of 41 probes, i.e. 34%) and *Actinobacteria* (for 10 of 38 probes, i.e. 26%) phyla (Table 1). At lower taxonomic levels, the probes for which correlations were found targeted particularly (i) *Rhodospirillales* such as *Azospirillum*, *Glucanacetobacter*, *Rhodospirillum* and *Sphingomonadaceae* (*Alphaproteobacteria*), (ii) *Burkholderiales* such as...
Table 1. Individual 16S rRNA probes for which Poaceae rhizosphere signal was correlated with the phylogenetic distance to maize.

<table>
<thead>
<tr>
<th>Probes</th>
<th>Bacterial target</th>
<th>Correlation$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Class Alphaproteobacteria (21 of 60 probes)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Order Rhodospirillales (13 of 20 probes)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Azobr1</td>
<td>Some Azospirillum (A. brasilense and A. lipoferum) and Roseomonas fauriei</td>
<td>0.66 **</td>
</tr>
<tr>
<td>Ira1</td>
<td>Azospirillum irakense and Rhodocista sp. AR2107 and AT2107</td>
<td>0.70 **</td>
</tr>
<tr>
<td>Lipo1</td>
<td>Azospirillum lipoferum, Azospermum doebereinerae and Azospirillum largomobile</td>
<td>0.87 ****</td>
</tr>
<tr>
<td>Sphingo5B</td>
<td>Most Sphingomonadaceae</td>
<td>0.76 ***</td>
</tr>
<tr>
<td>Sphingo4</td>
<td>Most Sphingomonadaceae</td>
<td>0.84 ****</td>
</tr>
<tr>
<td>Aceto3A</td>
<td>Some Acetobacteraceae</td>
<td>0.49</td>
</tr>
<tr>
<td>Aceto3B</td>
<td>Acetobacteraceae</td>
<td>0.59 **</td>
</tr>
<tr>
<td>Gludi</td>
<td>Gluconacetobacter diazotrophicus</td>
<td>0.59 **</td>
</tr>
<tr>
<td>Inteur1</td>
<td>Gluconacetobacter intermedius and Gluconacetobacter bactereupaeae</td>
<td>0.79 ***</td>
</tr>
<tr>
<td>Acidocellia1</td>
<td>Acidocellia and Acidiphilum</td>
<td>0.56</td>
</tr>
<tr>
<td>Ac2</td>
<td>Acidocellia and Acidiphilum</td>
<td>0.53</td>
</tr>
<tr>
<td>Rhodo2</td>
<td>Rhodospirillum</td>
<td>0.74 ***</td>
</tr>
<tr>
<td><strong>Order Rhizobiales (8 of 34 probes)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phyllobact</td>
<td>Mesorhizobium/Rhizobium</td>
<td>0.52</td>
</tr>
<tr>
<td>Agro157</td>
<td>Agrobacterium (biovars 1, 2 and 3)</td>
<td>0.53</td>
</tr>
<tr>
<td>B6.437</td>
<td>Agrobacterium (G1, G3, G4, G7, some A. rubi) and most Rhizobium</td>
<td>0.85 ****</td>
</tr>
<tr>
<td>Xan</td>
<td>Xanthobacter</td>
<td>0.51</td>
</tr>
<tr>
<td>Blastochlo6</td>
<td>Blastochloris and Ochrobactrum</td>
<td>0.54</td>
</tr>
<tr>
<td>Ochro3</td>
<td>Ochrobactrum</td>
<td>0.52</td>
</tr>
<tr>
<td>Rhodopseud</td>
<td>Rhodopseudomonas</td>
<td>0.53</td>
</tr>
<tr>
<td>Barto2</td>
<td>Bartonella</td>
<td>0.68</td>
</tr>
<tr>
<td><strong>Order Betaproteobacteria (13 of 43 probes)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta2</td>
<td>Betaproteobacteria</td>
<td>0.70 **</td>
</tr>
<tr>
<td>Beta3</td>
<td>Betaproteobacteria</td>
<td>0.65 **</td>
</tr>
<tr>
<td><strong>Order Burkholderiales (10 of 35 probes)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hefri4</td>
<td>Herbaspillum huttiense, Herbaspillum putei and Herbaspillum magnetovirio</td>
<td>0.68 **</td>
</tr>
<tr>
<td>Bkoglait</td>
<td>Burkholderia glathei</td>
<td>0.84 ****</td>
</tr>
<tr>
<td>Jant1</td>
<td>Janthinobacterium</td>
<td>0.57</td>
</tr>
<tr>
<td>Nt1A</td>
<td>Most Nitrosospira</td>
<td>0.57</td>
</tr>
<tr>
<td>Varpar</td>
<td>Variorovax paradoximus</td>
<td>0.60</td>
</tr>
<tr>
<td>Alfae1</td>
<td>Acaligenes faecalis cluster</td>
<td>0.61 **</td>
</tr>
<tr>
<td>Alfae2</td>
<td>A. faecalis cluster</td>
<td>0.68</td>
</tr>
<tr>
<td>Acave4</td>
<td>Acidovorax avenae cluster</td>
<td>0.85 ****</td>
</tr>
<tr>
<td>Metpet1</td>
<td>Some Burkholderiales</td>
<td>0.54</td>
</tr>
<tr>
<td>Hypal2</td>
<td>Hydrogenophaga pallonii</td>
<td>0.86 ****</td>
</tr>
<tr>
<td><strong>Order Neisseriales (1 of 2 probes)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eikcor2</td>
<td>Eikenella corrodens and Neisseria meningitidis</td>
<td>0.56</td>
</tr>
<tr>
<td><strong>Class Gammaproteobacteria (14 of 28 probes)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Order Thiobacterales (3 of 3 probes)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Achro1</td>
<td>Achromatium</td>
<td>0.56</td>
</tr>
<tr>
<td>Frtu4</td>
<td>Francisella tularensis</td>
<td>0.77 ***</td>
</tr>
<tr>
<td>Frtu7</td>
<td>F. tularensis</td>
<td>0.88 ****</td>
</tr>
<tr>
<td><strong>Order Alteromonadales (1 of 1 probe)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alter2</td>
<td>Alteromonas</td>
<td>0.55</td>
</tr>
<tr>
<td><strong>Order Pseudomonadales (3 of 9 probes)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mor1</td>
<td>Moraxella</td>
<td>0.80 ***</td>
</tr>
<tr>
<td>Pseu1</td>
<td>Pseudomonas</td>
<td>0.50</td>
</tr>
<tr>
<td>PseubC4.6</td>
<td>Pseudomonas corrugata</td>
<td>0.75 ***</td>
</tr>
<tr>
<td><strong>Order Enterobacterales (5 of 7 probes)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kleb1</td>
<td>Klebsiella oxytoca</td>
<td>0.67 **</td>
</tr>
<tr>
<td>Pagg5</td>
<td>Pantoea agglomerans</td>
<td>0.79 ****</td>
</tr>
<tr>
<td>Pagg6</td>
<td>P. agglomerans</td>
<td>0.77 **</td>
</tr>
<tr>
<td>Ptol1</td>
<td>Pantoea toletana</td>
<td>0.55</td>
</tr>
<tr>
<td>Pho1</td>
<td>Photorhabdus</td>
<td>0.77 ***</td>
</tr>
<tr>
<td><strong>Order Xanthomonadales (1 of 4 probes)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xancam</td>
<td>Xanthomonas campestris</td>
<td>0.80 ***</td>
</tr>
<tr>
<td><strong>Order Chromatiales (1 of 2 probes)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nitmob</td>
<td>Nitrosococcus mobilis</td>
<td>0.52</td>
</tr>
</tbody>
</table>

© 2014 Society for Applied Microbiology and John Wiley & Sons Ltd, *Environmental Microbiology*, 16, 2804–2814
Table 1. cont.

<table>
<thead>
<tr>
<th>Probes</th>
<th>Bacterial target</th>
<th>Correlation*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Class Deltaproteobacteria</strong> (6 of 22 probes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Order Myxococcales (6 of 18 probes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CystbSUB1</td>
<td>Cystobacterineae</td>
<td>0.54 *</td>
</tr>
<tr>
<td>Halian2</td>
<td>Haliangium</td>
<td>0.70 **</td>
</tr>
<tr>
<td>Annmxb5</td>
<td>Anaeromxyobacter</td>
<td>0.56 *</td>
</tr>
<tr>
<td>Annmxb6</td>
<td>Anaeromxyobacter</td>
<td>0.74 ***</td>
</tr>
<tr>
<td>un43Chon10</td>
<td>Uncultured Chondromyces</td>
<td>0.64 **</td>
</tr>
<tr>
<td>Nancs7</td>
<td>Nannocystis</td>
<td>0.63 **</td>
</tr>
<tr>
<td><strong>Class Epsilonproteobacteria</strong> (3 of 5 probes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alvi2</td>
<td>Unidentified Epsilonproteobacteria affiliated to Alvinella</td>
<td>0.89 ****</td>
</tr>
<tr>
<td><strong>Order Campylobacterales</strong> (2 of 3 probes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HeliFe11</td>
<td>Helicobacter heilmanni cluster (including H. felis)</td>
<td>0.62 ****</td>
</tr>
<tr>
<td>Wolin2</td>
<td>Wolinella succinogenes</td>
<td>0.86 ****</td>
</tr>
<tr>
<td><strong>Phylum Firmicutes</strong> (14 of 41 probes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGC354A</td>
<td>Firmicutes (excepted Clostridia and some Lactobacillales)</td>
<td>0.67 **</td>
</tr>
<tr>
<td><strong>Order Bacillales</strong> (9 of 32 probes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacpich2</td>
<td>Bacillus pichinotyi</td>
<td>0.68 **</td>
</tr>
<tr>
<td>Bacpich3</td>
<td>Bacillus pichinotyi</td>
<td>0.78 ***</td>
</tr>
<tr>
<td>Bacjoe6</td>
<td>Bacillus jeotgali</td>
<td>0.68 **</td>
</tr>
<tr>
<td>Bacfor1</td>
<td>Bacillus fortis and relatives</td>
<td>0.85 ****</td>
</tr>
<tr>
<td>Bacfara2</td>
<td>B. fortis and relatives</td>
<td>0.86 ****</td>
</tr>
<tr>
<td>Bacved_3</td>
<td>Bacillus vederi</td>
<td>0.83 ***</td>
</tr>
<tr>
<td>Palv</td>
<td>Paenibacillus alvei</td>
<td>0.63 **</td>
</tr>
<tr>
<td>Paeni4.2</td>
<td>Paenibacillus popoliiae and relatives (including Paenibacillus lentimorus, Paenibacillus thyaminolyticus)</td>
<td>0.86 ****</td>
</tr>
<tr>
<td>Pbo2</td>
<td>Paenibacillus borealis</td>
<td>0.82 ****</td>
</tr>
<tr>
<td><strong>Order Selenomonadales</strong> (3 of 4 probes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diali3</td>
<td>Dialister</td>
<td>0.69 **</td>
</tr>
<tr>
<td>Diali14</td>
<td>Dialister</td>
<td>0.71 **</td>
</tr>
<tr>
<td>Mega8</td>
<td>Megasphaera</td>
<td>0.87 ****</td>
</tr>
<tr>
<td><strong>Order Chlostridiales</strong> (1 of 1 probe)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mogi12</td>
<td>Mogibacterium</td>
<td>0.90 ****</td>
</tr>
<tr>
<td><strong>Phylum Actinobacteria</strong> (10 of 38 probes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HGC236-m</td>
<td>Most Actinobacteria</td>
<td>0.52 *</td>
</tr>
<tr>
<td><strong>Order Actinomycetales</strong> (8 of 35 probes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ActiD2</td>
<td>Actinomyces</td>
<td>0.80 ***</td>
</tr>
<tr>
<td>Agrm4</td>
<td>Agromyces</td>
<td>0.68 **</td>
</tr>
<tr>
<td>Comba8</td>
<td>Corynebacterium</td>
<td>0.86 ****</td>
</tr>
<tr>
<td>Glyc3</td>
<td>Glycomyces</td>
<td>0.68 **</td>
</tr>
<tr>
<td>Kocur10</td>
<td>Kocuria</td>
<td>0.90 ****</td>
</tr>
<tr>
<td>Nono1</td>
<td>Nonomurarea</td>
<td>0.52 *</td>
</tr>
<tr>
<td>Propio1</td>
<td>Propionibacterium</td>
<td>0.90 ****</td>
</tr>
<tr>
<td>Rhodrub</td>
<td>Rhodococcus</td>
<td>0.57 *</td>
</tr>
<tr>
<td><strong>Order Coriobacteriales</strong> (1 of 2 probes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collin29</td>
<td>Collinsella</td>
<td>0.87 ****</td>
</tr>
<tr>
<td><strong>Phylum Cyanobacteria</strong> (1 of 5 probes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prochio1</td>
<td>Prochlorothrix</td>
<td>0.89 ****</td>
</tr>
<tr>
<td><strong>Phylum Tenericutes</strong> (1 of 2 probes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spiroplas4</td>
<td>Spiroplasma</td>
<td>0.66 **</td>
</tr>
<tr>
<td><strong>Phylum Bacteroidetes</strong> (3 of 8 probes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rik6</td>
<td>Rikenella</td>
<td>0.85 ****</td>
</tr>
<tr>
<td><strong>Order Cytophagales</strong> (2 of 2 probes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spiro5</td>
<td>Spirosoma</td>
<td>0.67 **</td>
</tr>
<tr>
<td>Hyme3</td>
<td>Hymenobacter</td>
<td>0.79 **</td>
</tr>
<tr>
<td><strong>Phylum Deferribacteria</strong> (1 of 2 probes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OP2.1</td>
<td>Candidate phylum OP2</td>
<td>0.78 ***</td>
</tr>
<tr>
<td><strong>Phylum Planctomycetes</strong> (3 of 8 probes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planctc9</td>
<td>Gemmata cluster</td>
<td>0.70 **</td>
</tr>
<tr>
<td>Planctc12</td>
<td>Gemmata cluster</td>
<td>0.70 **</td>
</tr>
<tr>
<td>Planctc16</td>
<td>Some Pirrellula</td>
<td>0.76 **</td>
</tr>
</tbody>
</table>

a. Pearson correlation coefficient r and probability level (*P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001) are indicated.

b. For each order, class or phylum targeted by probes, the number of probes giving a statistically significant correlation among hybridized probes is indicated in parenthesis.
Acidovorax, Alcaligenes, Burkholderia and Hydrogenophaga (Betaproteobacteria), (iii) Francisella, Moraxella, Pantoea, Photurhambus, Pseudomonas and Xanthomonas (Gammaproteobacteria), (iv) Myxococcales such as Anaeromyxobacter (Deltaproteobacteria), (v) Megaspheara, Mogibacterium, and Bacillales such as Bacillus and Peanibacillus (Firmicutes) and (vi) Collinsella, and Actinomycetales such as Actinomyces, Corynebacterium, Kocuria and Propionibacterium (Actinobacteria). Therefore, these correlations did not involve probes from a single taxonomic group but rather probes from each of the main taxonomic groups colonizing Poaceae roots. No significant negative correlation (at \(P < 0.05\)) was found for any of the 178 other probes.

**Discriminating power of probes according to plant phylogenetic distance**

Among the 90 microarray probes whose signals correlated with maize/Poaceae distances, as many as 60 probes (along with 99 other probes for which no correlation was found) did discriminate bacterial communities when comparing the eight Poaceae genotypes (and bulk soil) by ANOVA (Supporting information Table S1). Within *Z. mays* spp. *mays*, genotype separation within or across maize genetic groups involved 11 of these 60 probes (i.e. 18%), and they targeted particularly certain *Alphaproteobacteria* (including *Azospirillum*, *Sphingomonadaceae*, *Rhodopseudomonas* and *Acideacellae*), *Betaproteobacteria* (Variovorax), *Gammaproteobacteria* (*Pseudomonas* and *Deltaproteobacteria* (Anaeromyxobacter and *Nannocystis*) (Supporting information Table S1). Genotype separation between different Poaceae genera (i.e. across longer taxonomic distances) involved all 60 probes. Therefore, 60 of the 90 probes whose signals correlated with maize/Poaceae distances could also discriminate between treatments, and they did not display the same type of discriminating power, as only 11 of the 60 probes could discriminate between closely related Poaceae as well as between more distant Poaceae genotypes.

**Significance of field management and plant growth stage**

Rhizobacterial community composition may change according to soil properties (Reeve *et al.*, 2010) or plant phenoology (Baudoin *et al.*, 2009). Therefore, different soils or plant development stages were considered to assess whether the relation between hosts and rhizobacterial community composition was conserved. To this end, we carried out further comparisons using two extreme situations, i.e. the comparison between two maize lines from the same genetic group (i.e. FV4 and W85 from Northern Flint group) versus the comparison between each of these maize lines and a non-Poaceae external reference (the dicot tomato). When these three plant genotypes were grown 21 days in soil from a neighbouring field, which corresponds to the same soil type but has been, for more than 50 years, managed as meadow instead of being cultivated with maize (Supporting information Table S1), probe hybridization changes compared with the cropped soil were closer for the two maize lines FV4 and W85 (with a difference only along PCA axis 2; Supporting information Fig. S2) than for tomato versus either maize line (with a difference along PCA axis 1). In addition, probe hybridization changes taking place at rhizobacterial community level when letting the three plants grow another 21 days in cropped field soil were the same for the two maize lines (no difference along the first two PCA axes), whereas they differed between tomato and either maize line (along PCA axis 1). These results obtained with three selected plant genotypes mean that a similar relation between plant evolution and rhizobacterial community composition may also be expected when considering another soil management regime or plants at a later development stage.

**Discussion**

This work targeted the hypothesis that the evolutionary history of eukaryotic hosts will determine the composition of their microbiome. This hypothesis implies that the more phylogenetically distant the hosts, the more distinct their associated bacterial communities. Indeed, a quantitative relation (positive and linear) was evidenced here when comparing the log value of plant phylogenetic distances between various maize genotypes and related Poaceae (enabling comparisons from intra-species to subfamily levels) with the pairwise community composition distances between their associated bacterial biomes (Fig. 2). Previous investigations failed to establish this type of relation, perhaps because of sampling bias [e.g. resulting from the analysis of captive animal hosts in artificial, zoo settings (Muegge *et al.*, 2011)], negligence of host phylogeny patterns [e.g. when comparing forest trees (Ushio *et al.*, 2008)], or phylogenetic analyses restricted to comparisons of hosts at intra-species [with maize genetic groups (Bouffaud *et al.*, 2012)] or species level [with trees (Redford *et al.*, 2010)] and/or between very closely related taxa [with wild hominids (Ochman *et al.*, 2010)].

Our findings mean that many of the phenotypic particularities resulting from the evolutionary history of these Poaceae (and not only the main traits that members of a given plant group all have in common) might have the potential to influence the ability of the resulting genotypes to select bacterial taxa from a given soil. This concerns bacterial taxa (Table 1 and Supporting Information Table...
S1) of (i) contrasted relations with plants, ranging from cooperation (certain members of the Rhodospirillales and Bacillales; Vessey, 2003; Vacheron et al., 2013) and symbiosis [noticeably for certain Rhizobiales (Masson-Boivin et al., 2009)] to parasitism (e.g. Agrobacterium and Xanthomonas) and (ii) contrasted taxonomic status, as they were found in all the main phyla studied. Indeed, as many as 24% of the probes that hybridized displayed a significant correlation with plant distance (at $P < 0.01$), meaning that the relationship was relevant for a substantial fraction of the bacterial community. Till now, this information was available only in the case of a single, plant-associated bacterial genus (Bradyrhizobium), for which genetic distances between strains were related to the genetic distances between their Amphicarpaea bracteata hosts (Parker and Spoerke, 1998).

Such a relationship might involve coevolution (Kinkel et al., 2011), but probably to a restricted extent in the current experimental context. Indeed, coevolution is a process well documented for mutualists and parasites (Provorov and Vorobyov, 2009; Thrall et al., 2012), but which remains speculative at the scale of the entire rhizobacterial community (Lambers et al., 2009). In addition, although certain endophytic bacteria might be transmitted vertically (Johnston-Monje and Raizada, 2011), this is very unlikely to happen for most rhizobacteria, which are selected from soil (Bais et al., 2006). Finally, the current experiments were performed with soils selected from a single geographic area rather than using for each plant type a different soil taken in one region of the world where the corresponding plant genetic group or species formed, diversified and/or spread (Camus-Kulandaivelu et al., 2006). This is important to keep in mind because local selection of genotypes according to particular environmental conditions has had a strong influence on the resulting genetic structure of crop species (Purugganan and Fuller, 2009). The cropped soil of the study has a history of maize cultivation with commercial hybrids and not maize landraces, in a region where the crop has become important in the last 50 years only, and it has not been grown with wheat or sorghum in the last decades. For the meadow soil, there is no documented history of cropping in the records (dating back to 1955).

In this work, one teosinte representing maize’s pre-domestication ancestor was included, but the corresponding correlation data were in line with those obtained when comparisons were done between cultivated genotypes (Fig. 2). This, which is consistent with the fact that teosinte, maize and sorghum (all three from the same Panicoideae subfamily) have similar root morphology during vegetative stages (Aloni and Griffith, 1991; Singh et al., 2010), suggests that the domesticated/wild status was of negligible significance compared with plant phylogenetic distances. It would mean that for interaction with soil bacteria, domestication is only of importance as a factor triggering accelerated evolution (which materializes by longer phylogenetic distances with non-domesticated relatives). Indeed, plant domestication leads to drastic changes and acceleration in the subsequent evolution process, selecting traits facilitating grain harvest (Purugganan and Fuller, 2009).

Different plant species and, at a smaller scale, different plant genotypes within species may have specific phenotypic traits, including root properties that are likely to influence rhizobacterial community composition (Fan et al., 2001; Bais et al., 2006). These properties include bacterial root-colonization sites (de Dorlodot et al., 2007), nutrient uptake and rhizosphere depletion of minerals (Comas and Eissenstat, 2009), root mycorrhization (Becerra et al., 2007) and, in particular, root exudation patterns (Czarnota et al., 2003). Differences in rhizodeposits (carbon and energy sources) occur between genotypes of a same plant species (Czarnota et al., 2003), different plant species (Fan et al., 2001) or different plant families, paralleling phylogenetic relations (Lambert et al., 2010). Indeed, rhizobacterial community composition in the current work depended both on Poaceae plant species and plant genotype within the Z. mays species (Fig. 1 and Supporting Information Fig. S1).

The robustness of the current findings was shown when comparing very close genotypes (two maize lines) and distant ones (maize versus tomato) in the same soil, but which had been subjected to a different past management regime (permanent meadow) or was studied at a different sampling time (Supporting information Fig. S2). This approach also evidenced that soil management was a more influential factor than plant growth stage, as found for other crop species (Wieland et al., 2001). Here, cropping modified soil properties (Supporting information Table S2) known to influence bacterial community composition, such as pH (Lauber et al., 2009), organic matter (Baudoin et al., 2003) and nitrogen contents (Kennedy et al., 2004).

In conclusion, this analysis performed with maize genotypes and related Poaceae is the first study showing that the extent of diversification of eukaryotic hosts can be a significant factor determining the composition of their associated bacterial compartment. It is also an indirect clue that host–bacteria interactions are controlled by a very wide range of host traits, and thus that various types of evolutionary changes in host genotype might impact on bacterial selection processes.

**Experimental procedures**

**Plant genotypes**

Four maize (Zea mays L.) inbred lines from two genetic groups, i.e. FV4 and W85 (group Northern Flint) and FV252...
and Mo17 (Corn Belt Dent) were used (provided by INRA, St Martin d’Hýnx, France). The study also included one teosinte (Zea mays ssp. parviglumis; provided by UNAM, Cuernavaca, Mexico), one sorghum (Sorghum bicolor cv. Arprim; Semences de Provence, Fourques, France) and one wheat (Triticum aestivum L. cv. Florina; provided by AgroScope, Changins, Switzerland), as well as tomato (Solanum lycopersicum cv. Marmande; Vilmorin, La Ménitré, France) used as non-Poaceae reference.

Plant phylogenetic distances within the Poaceae were estimated using three chloroplastic sequences established for phylogeny assessment (Shaw et al., 2007; Hodge et al., 2010), i.e. gene rps16 and the intergenic spacers rps16-trnK and atpI-atpH. To obtain their sequences, seeds of the different plant lines were surface-disinfected (as described above), germinated on water agar (8 g l\(^{-1}\)) plates for 2 days in the dark at 28°C, and the plates were placed for 7 days in a growth chamber (16 h of light at 28°C and 8 h of darkness at 22°C). DNA was extracted from 200 mg of fresh leaves crushed in liquid nitrogen (as described below). PCR procedures were adapted from Shaw and colleagues (2007) and Hodge and colleagues (2010). DNA was amplified using an Eppendorf Mastercycler, in a 50 μl volume containing 1× PCR buffer with MgCl\(_2\), 0.2 mM of each dNTP (Promega, Charbonnière, France), 0.5 μM of each primer, 5% dimethyl sulfoxide, 1.75 U of Expand high-fidelity Taq polymerase (Roche Applied Science, Meylan, France), and 1 μl of template DNA (–30 ng). PCR amplification was done with one cycle of 5 min at 94°C, followed by 30 cycles of 1 min at 94°C, 1 min at 56°C, 1 min at 72°C, and a final cycle at 72°C for 5 min. PCR primers were rps16F/rps16R for rps16 (Hodge et al., 2010), rpS16x2F2/trnK (UUU) × 1 for rps16-trnK spacer, and atpI/atpH for atpI-atpH spacer (Shaw et al., 2007).

PCR products were cleaned using a MinElute PCR purification kit (Qiagen, Courtaboeuf, France) following the manufacturer instructions. DNA sequencing was performed on both strands (Beckman Coulter Genomics, Takeley, UK). The DNA sequences were aligned using BioEdit version 7.0.5.3 (Hall, 1999) (http://www.mbio.ncsu.edu/BioEdit/BioEdit.html). Phylogenetic analysis was performed by MEGA5 (Tamura et al., 2011), using the maximum likelihood method, and phylogenetic distances were obtained using the Kimura 2 parameter model.

Greenhouse experiment

Plant lines were cultivated in two soils from adjacent fields located at La Côte Saint-André near Lyon (France), one cropped with maize for more than 15 years, and the second under permanent meadow (for at least 40 years). Both are luvisols with loamy topsoil (clay 15%–16%, silt 41%–45%, sand 41%–43%), but their chemical compositions differ. In particular, the soil under meadow displays lower pH and higher organic matter content, nitrogen content and cation exchange capacity (Supporting information Table S2). The two topsoils were collected in September 2009 and sieved at 6 mm.

Wheat seeds were surface-disinfected by soaking for 3 min in 70% ethanol, 30 min in 3% calcium hypochlorite containing 0.1% sodium dodecyl sulfate, and then 20 min in a 2% sodium thiosulfate solution. After each immersion, the seeds were rinsed five times with demineralized sterile water. Maize, teosinte, sorghum and tomato lines were surface-disinfected using a 3% sodium hypochlorite solution (60 min for maize lines, 60 min for the others) and rinsed five times with demineralized sterile water. At the end of the disinfection process, seeds of maize, sorghum, wheat and tomato were left in demineralized sterile water (1 h for teosinte, 4 h for the others) to favour germination.

For each plant line, five pots (2.5 kg soil per 3 dm\(^3\) pot) sown with three seeds each (thinned to one seedling on day 10) were used (i.e. five replicates). Five non-planted pots were also included. They were placed following a randomized block design in a greenhouse with 16 h of light (25°C) and 8 h of darkness (18°C) each day. Water was adjusted manually to soil retention capacity. Sampling was carried out at 21 days for each plant line in cropped field soil, and for maize lines FV4 and W85 (group Northern Flint) and tomato in meadow field soil. Another sampling was carried at 42 days for maize lines FV4 and W85 and tomato in the cropped field soil. The root systems were dug up and shaken vigorously to separate soil loosely adhering to the roots. Each root system with closely adhering soil was frozen in liquid nitrogen immediately, then lyophilized and stored at −20°C. Both bulk soils were sampled (5 g in each of the non-planted pots) at 21 and 42 days in both soils, lyophilized and stored at −20°C.

Extraction and labelling of soil DNA

Total DNA from 0.5 g of rhizosphere or bulk soil was extracted using a protocol derived from Bürgmann and colleagues (2003). Briefly, 0.5 g of soil, 0.5 g of zirconium beads (VWR, Fontenay-sous-Bois, France), 0.5 ml of extraction buffer (5% hexadecyltrimethylammonium bromide, 1 mM of 1,4-dithio-DL-threitol, in a 0.12 M phosphate buffer [pH 8]) were processed in a bead beater (TissueLyser II Retsch; Qiagen) for 90 s at 30 m s\(^{-1}\). After 10 min centrifugation at 16000 g, supernatants were extracted twice with phenol-chloroform-isoamyl alcohol (24:24:1 v/v/v) and then with chloroform-isoamyl alcohol (24:1 v/v). Nucleic acids were precipitated overnight with potassium acetate (3 M, pH 4.8) and absolute ethanol at −20°C. After centrifugation for 30 min at 16000 g, pellets were washed with 70% ethanol and then dissolved in Rnase- and Dnase-free water.

The universal primers T7-PA and PH (derived from Bruce et al., 1992) were used to obtain 1.5 kb amplicons of the 16S rRNA gene rrs from the total community (Edwards et al., 1989), as described by Kyselková and colleagues (2009). PCR products were purified with a MinElute PCR purification kit (Qiagen), according to the manufacturer’s instructions, and DNA concentration was adjusted to 50 ng μl\(^{-1}\). Amplified DNA was converted to RNA by in vitro transcription, purified and fragmented (Sanguin et al., 2008).

Microarray analysis

The 1033 probes (previously described in Kyselková et al., 2009) were synthesized by Eurogentec (Seraing, Belgium) with a 5′ C6 – NH\(_2\) modification and spotted, as described by
Sanguin and colleagues (2008). Each probe was repeated four times per slide. Hybridization was performed in temperature-controlled Belly Dancer (Stovall Life Sciences, Greensboro, NC, USA), as described by Sanguin and colleagues (2008).

The slides were scanned at 532 nm with 10 μm resolution, using a GeneTec LS IV scanner (GenomicSolutions, Huntington, UK), and images were analyzed with GENEPIX 4.01 (Axon, Union City, CA, USA), as described by Sanguin and colleagues (2008). Data filtration was done using R 2.10.1 (R Development Core Team, 2008) (http://www.r-project.org), and hybridization of a given spot was considered positive when 80% of the spot pixels had intensity higher than the median local background pixel intensity plus twice the standard deviation of the local background, as done previously (Sanguin et al., 2008; Kyselkova et al., 2009). The intensity of the signal (median of signal minus background) was replaced by its square root, and the intensity of each spot was expressed as a fraction of the total intensity signal of the basic pattern it belongs to (Sanguin et al., 2008). A given probe was considered to be positive when at least two of the three spot replicates were hybridized.

Data analysis

Microarray data were treated by PCA based on the correlation matrix, using the ADE4 package of R (Thioulouse et al., 1997). One-way ANOVA and Fisher’s LSD tests ($P < 0.05$) were used to compare treatments, based on (i) treatment position along each of the first two PCA axes, i.e. based on PC1 coordinates as well as on PC2 coordinates, and (ii) on hybridization signals for each probe.

In addition, the effects of field management and plant growth stage were considered, for each probe in each treatment, by dividing the normalized hybridization value of each replicate for meadow field soil by the mean value for cropped field soil (all obtained at 21 days) and the normalized hybridization value of each replicate at 42 days by the mean value of 21 days (all obtained with cropped field soil) respectively. These ratios were then analyzed statistically, as described above.

Finally, similarity level between rhizobacterial communities was estimated, using (i) the Bray–Curtis distance (Bray and Curtis, 1957) from normalized hybridization intensity data and (ii) the phylogenetic distance between plants or the Euclidean distance from treatment positions along each of the first two PCA axes. Correlation analyses were performed using Pearson correlation coefficient. All analyses were done using R 2.10.1 software and (when relevant) $P < 0.05$ level.

Acknowledgements

We are grateful to J.P. Magnière (Ferme expérimentale, La Côte Saint André, France) for the choice of soil, V. Daubin (UMR CNRS 5558 LBBE) for help with plant phylogeny, J. Haurat (UMR CNRS 5557 Ecologie Microbienne and FR41’s DTAMB) for microarray spotting and D. Abrouk (UMR CNRS 5557 Ecologie Microbienne) for advice on statistics.

References


**Supporting information**

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site:

**Fig. S1.** Distribution of 16S rRNA microarray probes based on principal component analysis of bulk soil and Poaceae rhizospheres at 21 days in cropped soil. Relations between treatments are shown in Fig. 1.

**Fig. S2.** Significance of field management and plant growth stage, based on principal component analysis of bulk soil and rhizosphere samples for maize lines FV4 and W85 and for tomato, using normalized 16S rRNA microarray hybridization values obtained at 21 days in meadow soil (divided by the mean value for cropped field soil obtained also at 21 days) and at 42 days in cropped soil (divided by the mean value at 21 days in the same cropped field soil).

A. Comparison of rhizosphere and bulk soil samples based on meadow/cropped ratios at 21 days (filled squares) and 42/21-day ratios (empty squares). For each treatment, mean and standard deviation are represented (n = 5). Statistical relations are indicated along PC1 and PC2 axes by letters a-c (ANOVA and Fisher’s LSD tests; P < 0.05).

B. Contribution of microarray probes. PCA results showed that changes were of higher magnitude when comparing bacterial communities of cropped soil with meadow soil, than when comparing 42 with 21 days. Comparing field management effects, twofold differences or more in probe signal intensity were found in at least one treatment (bulk soil or rhizospheres) for 165 probes, with often higher signals with cropped soil. These probes targeted Firmicutes (19% of the 165 probes), Alphaproteobacteria (18%), Betaproteobacteria (15%), Actinobacteria (11%), Deltaproteobacteria (10%) and Gammaproteobacteria (9%). When comparing samplings, a twofold difference or more in signal intensity was found in at least one treatment (often a rhizosphere treatment) for 128 probes, with often higher signals with samples at 21 days. These probes targeted Alphaproteobacteria (22% of the 128 probes), Betaproteobacteria (20%), Actinobacteria (14%), Firmicutes (13%), Deltaproteobacteria (10%) and Gammaproteobacteria (8%).

**Table S1.** Statistical relationship between treatments for each individual probe, based on comparison of probe intensities (n = 5).

**Table S2.** Topsoil chemical composition.
