FINE-SCALE ADAPTATIONS TO ENVIRONMENTAL VARIATION AND GROWTH STRATEGIES DRIVE PHYLLOSPHERE *METHYLOBACTERIUM* DIVERSITY.

Jean-Baptiste Leducq*[1,2,3], Émilie Seyer-Lamontagne[1], Domitille Condrain-Morel[2], Geneviève Bourret[2], David Sneddon[3], James A. Foster[3], Christopher J. Marx[3], Jack M. Sullivan[3], B. Jesse Shapiro*[1,4] & Steven W. Kembel*[2]

1 - Université de Montréal

2 - Université du Québec à Montréal

3 – University of Idaho

4 – McGill University

* Corresponding authors. Email: jeanbaptiste@uidaho.edu, jesse.shapiro@mcgill.ca, kembel.steven_w@uqam.ca

Key words: *Methylobacterium* diversity, Phyllosphere community dynamics, *rpoB* barcoding, temperature adaptation, growth strategies in *Bacteria*

Abstract

Methylobacterium is a prevalent bacterial genus of the phyllosphere. Despite its ubiquity, little is known about the extent to which its diversity reflects neutral processes like migration and drift, versus environmental filtering of life history strategies and adaptations. In two temperate forests, we investigated how phylogenetic diversity within *Methylobacterium* was structured by biogeography, seasonality, and growth strategies. Using deep, culture-independent barcoded marker gene sequencing coupled with culture-based approaches, we uncovered a considerable diversity of *Methylobacterium* in the phyllosphere. We cultured different subsets of *Methylobacterium* lineages depending upon the temperature of isolation and growth (20 °C or 30 °C), suggesting long-term adaptation to temperature. To a lesser extent than temperature adaptation, *Methylobacterium* diversity was also structured across large (>100km; between forests) and small geographical scales (<1.2km within forests), among host tree species, and was dynamic over seasons. By measuring growth of 79 isolates at different temperature treatments, we observed contrasting growth performances, with strong lineage- and season-dependent variations in growth strategies. Finally, we documented a progressive replacement of lineages with a high-yield growth strategy typical of cooperative, structured communities, in favor of those characterized by rapid growth, resulting in convergence and homogenization of community structure at the end of the growing season. Together our results show how *Methylobacterium* is phylogenetically structured into lineages with distinct growth strategies, which helps explain their differential abundance across regions, host tree species, and time. This works paves the way for further investigation of adaptive strategies and traits within a ubiquitous phyllosphere genus.

Importance

Methylobacterium is a bacterial group tied to plants. Despite its ubiquity and importance to their hosts, little is known about the processes driving *Methylobacterium* community dynamics. By combining traditional culture-dependent and –independent (metabarcoding) approaches, we monitored *Methylobacterium* diversity in two temperate forests over a growing season. On the surface of tree leaves, we discovered remarkably diverse and dynamic *Methylobacterium* communities over short temporal (from June to October) and spatial scales (within 1.2 km). Because we cultured different subsets of *Methylobacterium* diversity depending on the temperature of incubation, we suspected that these dynamics partly reflected climatic adaptation.

By culturing strains in lab conditions mimicking seasonal variations, we found that diversity and environmental variations were indeed good predictors of *Methylobacterium* growth performances. Our findings suggest that *Methylobacterium* community dynamics at the surface of tree leaves results from the succession of strains with contrasted growth strategies in response to environmental variations.

Acknowledgments: FRQNT (funding), NSERC (funding), Canada Research Chairs (funding), NSF (DEB-1831838), Geneviève Lajoie, Dominique Tardif, Ariane Lafrenière, Hélène Dion-Phénix, Yves Terrat, Kenta Araya (help for sampling), Sylvain Dallaire (help for phenotyping screen), Sergey Stolyar (discussions), Gault Natural Reserve (McGill University), Station Biologique des Laurentides (UdeM), Centre d'Étude de la Forêt (CEF), QCBS and two anonymous reviewers for their helpful suggestions.

Author contributions : J.B.L., E.S.L., D.C.M., G.B., B.J.S. and S.W.K. planned fieldwork and the experiments. J.B.L. E.S.L., D.C.M. and G.B. performed experiments. J.B.L., E.S.L., S.W.K. , D.S. and J.M.S performed bioinformatic analyses. J.A.F., C.J.M. and J.M.S. provided discussion in the early stages of this study. J.B.L. B.J.S. and S.W.K. drafted the manuscript with contributions from C.J.M.

Data accessibility: raw reads for *16S rRNA* and *rpoB* barcoding on phyllosphere communities (BioProject PRJNA729807; BioSamples SAMN19164946-SAMN19165146) were deposited in NCBI under SRA Accession Numbers SRR14532212-SRR14532451. Partial nucleotide sequences from marker genes obtained by SANGER sequencing on *Methylobacterium* isolates (BioProject PRJNA730554 ; Biosamples SAMN19190155-SAMN19190401) were deposited in

NCBI under GenBank Accession Numbers MZ268514-MZ268593 (*16S rRNA* gene), MZ330152-MZ330358 (*rpoB* gene) and MZ330130-MZ330151 (*sucA*). Bioprojects, Biosamples, SRA and GenBank accession numbers are listed in **Dataset S1**. R codes and related data were deposited on Github (https://github.com/JBLED/methylo-phyllo-diversity).

Introduction

The phyllosphere, the aerial parts of plants including leaves, is a microbial habitat estimated to be as vast as twice the surface of the earth (1). Although exposed to harsh conditions including UV radiation, temperature variation, and poor nutrient availability, the phyllosphere harbors a diverse community of microorganisms, of which bacteria are the most abundant (1). A key challenge in microbial ecology and evolution is understanding the evolutionary and ecological processes that maintain diversity in habitats such as the phyllosphere. Bacteria living in the phyllosphere carry out key functions including nitrogen fixation, growth stimulation and protection against pathogens (1–3). At broad spatial and temporal scales, bacterial diversity in the phyllosphere varies as a function of geography and host plant species, potentially due to restricted migration and local adaptation to the biotic and abiotic environment (4–6), leading to patterns of cophylogenetic evolutionary association between phyllosphere bacteria and their host plants (7). Whether those eco-evolutionary processes are important at the scale of several days to several years, as microbes and their host plants migrate and adapt to changing climates, is still an open question (8). Another challenge is to link seasonal variation with plant-associated microbial community dynamics, as shifts in microbial community composition are tighly linked with host plant carbon cycling (9) and ecosystem functions including nitrogen fixation (10). More generally, we understand very little about how the ecological strategies of phyllosphere bacteria vary among lineages and in response to variation in environmental conditions throughout the growing season (9, 11).

Phenotypic traits are often phylogenetically conserved in microbes (12), and these traits influence the assembly of ecological communities through their mediation of organismal interactions with the abiotic and biotic environment (13). Recent work has shown that many microbial traits exhibit phylogenetic signal, with closely related lineages possessing more similar traits, although the phylogenetic depth at which this signal is evident differs among traits (14). Most comparative studies of microbial trait evolution have focused on broad patterns across major phyla and classes (14), although some studies have found evidence for complex patterns of biotic and abiotic niche preferences evolving within genus-level phylogenies (15, 16). Furthermore, to date the majority of studies of the diversity of plant-associated microbes have been based on the use of universal marker genes such as the bacterial *16S rRNA* gene, providing a global picture of long-term bacterial adaptation to different biomes and host plants at broad phylogenetic scales (17). However, these studies lack sufficient resolution to assess the evolutionary processes at finer spatial and temporal scales that lead to the origin of adaptations within microbial genera and species (18, 19).

The Rhizobiales genus *Methylobacterium* (*Alphaproteobacteria*, *Rhizobiales*, *Methylobacteriaceae*) is one of the most prevalent bacterial genera of the phyllosphere, present on nearly every plant (20–22). Characterized by pink colonies due to carotenoid production, *Methylobacterium* are facultative methylotrophs, able to use one-carbon compounds, such as methanol excreted by plants, as sole carbon sources (23, 24). Experimental studies have shown the important roles of *Methylobacterium* in plant physiology, including growth stimulation through hormone secretion (25–27), heavy metal sequestration (27), anti-phytopathogenic compound secretion, and nitrogen fixation in plant nodules (28), sparking increasing interest in the use of *Methylobacterium* in plant biotechnology applications (27, 29, 30). Although up to 64 *Methylobacterium* species have been described (31–39), genomic and phenotypic information was until recently limited to a small number of model species: *M. extorquens, M. populi*, *M.*

nodulans, *M. aquaticum* and *M radiotolerans*, mostly isolated from anthropogenic environments, and only rarely from plants (40–44). Aditionaly, *Methylobacterium* was mostly isolated assuming its optimal growth was in the range 25-30 °C (45), an approach that could bias strain collections toward mesophylic isolates to the exclusion of isolates from temperate forests where temperatures typically range from 10 to 20 °C during the growing season (46). Newly available genomic and metagenomic data now allow a better understanding of the distribution of *Methylobacterium* diversity across biomes (31) and suggest that they represent a stable and diverse fraction of the phyllosphere microbiota (22). However, we still understand relatively little about the drivers of the evolution and adaptation of *Methylobacterium* in natural habitats.

In this study, we assessed the diversity of *Methylobacterium* in temperate forests and asked whether *Methylobacterium* associated with tree leaves act as a single unstructured population, or if their diversity is structured by regional factors (*e.g.* a combination of isolation by distance and regional environmental variation) or by niche adaptation (*e.g.* host tree or temperature adaptation) (12). First, we assessed *Methylobacterium* diversity by combining culturing and metabarcoding approaches along with phylogenetic analysis and quantified how this diversity varied across space, time, and environment in the phyllosphere. Second, we quantified the extent of phylogenetic niche differentiation within the genus, with a focus on quantifying the evidence for adaptation to local environmental variation at different spatial, temporal and phylogenetic scales. We hypothesized that distinct phylogenetic lineages would be associated with distinct environmental niches. Third, we quantified *Methylobacterium* growth performance under fine-scale environmental variations, with a focus on temperature, to determine whether fine-scale changes in diversity over space and time might result from environmental filtering of isolates with contrasting growth strategies under local environmental conditions. We found that

Methylobacterium phyllosphere diversity consisted of deeply branching phylogenetic lineages associated with distinct growth phenotypes, isolation temperatures, and large-scale spatial effects (forest of origin), while finer-scale spatial effects, host tree species, and time of sampling were more weakly and shallowly phylogenetically structured. Over the course of a year, from spring to fall, we observed a homogenization of *Methylobacterium* community structure coinciding with the progressive replacement of isolates with high yield strategy by isolates with rapid growth. Together our results show that this ubiquitous phyllosphere genus is structured into lineages with distinct growth strategies, which helps explain their differential abundance across space and time.

Methods

Phylogenetics of plant-associated Methylobacterium diversity.

We evaluated the known *Methylobacterium* diversity and its distribution across biomes, with a special emphasis on the phyllosphere. First, we constructed a phylogeny of *Methylobacteriaceae* from the complete nucleotide sequence of *rpoB*, a highly polymorphic housekeeping gene commonly used to reconstruct robust phylogenies in bacteria, because unlikely to experience horizontal gene transfer or copy number variation (47, 48). We retrieved *rpoB* sequences from genomes publicly available in September 2020, including 153 *Methylobacteria*, 30 *Microvirga* and 2 *Enterovirga* (**Dataset S1a**), performed alignment and inferred a consensus phylogeny with MrBayes v. 3.2.7a ((49); **Supplementary Materials and Methods S1)**. For each *Methylobacterium* reference genome, we retrieved the species name and the sampling origin, when available. Additionally, we assigned each genome to a group (A, B, C) according to previously proposed subdivisions (31). We subdivided group A in nine clades (A1-A9; **Supplementary Materials and Methods S1)**.

Study sites and sample collection

The two study forests were located at the Gault Nature Reserve (Mont Saint-Hilaire, Quebec, Canada; 45.54 N 73.16 W), here referred as MSH, an old forest occupying Mount Saint-Hilaire, and the *Station Biologique des Laurentides* (Saint-Hippolyte, Quebec, Canada; 45.99 N 73.99 W), here referred to as SBL, a mosaic of natural wetlands, xeric and mesic forests (**Figure 1**, **Dataset S1b**). In August 2017, for the purpose of a pilot survey, we collected leaves from the subcanopy (3-5m) of 19 trees among dominant species in MSH (*Fagus grandifolia*, *Acer saccharum*, *Acer pensylvanicum* and *Ostrya virginiana*). In 2018, we realized a time series survey in MSH and SBL. In each forest, we marked and collected leaf samples from the subcanopy of 40 trees (representative of local tree species diversity) in 4-6 plots distributed along a 1.2 km transect (**Supplementary Materials and Methods S1)**. In MSH, the transect followed an elevation and floristic gradient dominated by tree species *F. grandifolia (FAGR)*, *A. saccharum (ACSA)*, *O. virginiana (OSVI)* and *Quercus rubra (QURU)*. In SBL, the transect followed a constant environnement dominated by *A. saccharum*, *F. grandifolia*, *A. pensylvanicum (ACPE), Abies balsamea* (*ABBA*) and *Acer rubrum* (*ACRU*). For this time series, each tree was sampled 3-4 times from June to October 2018. For each sampled plot and time point, we also sampled a negative control consisting of empty sterile bags opened and sealed on site. The leaf surface microbial community from each sample was collected with phosphate buffer and split in two equal volumes for microbial community DNA extraction and *Methylobacterium* isolation, respectively (**Supplementary Materials and Methods S1)**.

Methylobacterium isolation and development of a fine-scale single-copy molecular marker specific to Methylobacterium

For both pilot and time series surveys, we performed *Methylobacterium* isolation on MMS synthetic solid media with 0.1% methanol supplemented with yeast extract and vitamins (**Supplementary Materials and Methods S1)**. For each leaf sample, isolation was replicated at 20 °C and 30 °C to minimize biases toward mesophylic strains. Isolates from the 2017 pilot survey (*n*=80; **Dataset S1c**) were identified by PCR amplification and partial sequencing of the *16S rRNA* ribosomal gene and assigned to *Methylobacterium* clades (**Supplementary Materials and Methods S1; Dataset S1d,e)**. As an alternative to the *16S rRNA* gene, we developed a highly polymorphic marker targeting the *Methylobacteriaceae* family. We tested two candidate genes, *rpoB* (47, 48, 50, 51) and *sucA* (51–53), which, contrary to *16S rRNA*, were single-copy in *Methylobacterium* genomes and were polymorphic enough to distinguish among *Methylobacterium* groups and clades (**Supplementary Materials and Methods S1; Figure S1)**. In 20 representative *Methylobacterium* isolates from the 2017 pilot survey (**Dataset S1c,d,e**), we successfully amplified a *rpoB* hypervariable region (targeted by primers Met02-352-F and Met02-1121-R), that we choose as a specific marker for *Methylobacteriaceae* (**Supplementary Materials and Methods S1**; **Table S1**). Isolates from the 2018 timeline survey (*n*=167; **Dataset S1e,f,**) were assigned to *Methylobacterium* clades using a consensus phylogenetic tree inferred with MrBayes v. 3.2.7a (49) from nucleotide sequences of the *rpoB* marker obtained for these isolates, aligned together with *rpoB* complete nucleotide sequences available from 188 *Methylobacteriaceae* genomes (**Dataset S1a**) and partial nucleotide sequences obtained from 20 representative isolates from the pilot survey (**Supplementary Materials and Methods S1)**.

Culture-based assessment of Methylobacterium *diversity in the tree phyllosphere.*

We tested for associations between *Methylobacterium* culture-based diversity at different phylogenetic depths, with isolate characteristics as proxy for adaptive response to environmental variables through their evolution, using the *rpoB* phylogenetic tree built from timeline survey isolates as a guide. We assigned *Methylobacterium* isolates according to their phylogenetic placement. After excluding nodes supported by less than 30% of bootstraps, the tree was converted into an ultrametric tree scaled proportionally to pairwise nucleotide similarity (*PS;* **Supplementary Materials and Methods S1**). First, for each *PS* value in the tree in the range 0.926-1.000 (corresponding to *PS* range within clades), we classified isolates into discrete taxa and performed a PERMANOVA (10,000 permutations) on *Methylobacterium* community dissimilarity using the *Bray-Curtis* index (*BC*) based on taxa absolute abundance (Hellinger transformation) using the R package *vegan* (54). We tested for the relative contribution of four factors and their interactions on taxon frequency: sampling forest (*F*); temperature of isolation (*T*); sampling time (*D*) and host tree species (*H*). Second, we asked specifically which nodes within the tree were associated with *F* and *T*. For each node with at least 30% of support, and each factor, we tested for the association between embedded taxa and *F* (SBL and MSH) or *T* (20 and 30 °C) by permutation of factors between embedded nodes (100,000 permutations per node; **Supplementary Materials and Methods S1)**.

Culture-free assessment of Methylobacterium *diversity in the tree phyllosphere (barcoding)*

We evaluated the bacterial phyllosphere diversity through barcoding and sequencing of phyllosphere samples from the 2018 timeline survey. First, we evaluated the bacterial diversity targeting the *16S rRNA* gene (55) in 46 phyllosphere samples from 13 trees from both forests sampled 3-4 times throughout the 2018 growth season. We included one negative control, and one positive control consisting of mixed DNAs of *Methylobacterium* isolates typical of the phyllosphere (METH community; **Supplementary Materials and Methods S1**). Second, we evaluated the *Methylobacteriaceae* phyllosphere diversity targeting the *rpoB* marker (see above) in 184 phyllosphere samples from 53 trees representative of diversity found in MSH (n=26) and SBL (n=27), sampled 3-4 times throughout the 2018 growth season. We included four negative controls and four positive controls (METH community). Library preparation and sequencing were performed as described in **Supplementary Materials and Methods S1.** For each phyllosphere sample and controls, we estimated bacterial diversity based on Amplicon Sequence Variants (ASVs) using package *dada2* in R (56). We assessed ASV taxonomy using SILVA v.138 database for *16S rRNA* gene (57) and a *rpoB* nucleotide sequence database available for *Bacteria* (48), curated by a ML phylogenetic tree (200 permutations**; Supplementary Materials and Methods S1**). Taxonomy for *Methylobacterium* ASVs (at the clade level) was refined using blast against NCBI databases for *16S rRNA* gene (58) and using phylogenetic placement for *rpoB* (**Supplementary Materials and Methods S1**). To validate the *rpoB* barcoding accuracy in estimating *Methylobacterium* diversity, we compared *Methylobacterium* clade relative abundances estimated from *16S rRNA* and *rpoB* barcoding in a heatmap (**Supplementary Materials and Methods S1**). We also compared *Methylobacterium* diversity estimations from *rpoB* barcoding and culture-dependant approaches by matching *rpoB* partial nucleotide sequences obtained from isolates with those obtained from ASVs (**Supplementary Materials and Methods S1**). We evaluated relative contributions of sampling forest (*F*), plot within forest (*P*), host tree species (*H*), time of sampling (*T*) and their interactions on bacteria (*16S rRNA* barcoding) and *Methylobacteriaceae* (*rpoB* barcoding) community dissimilarity among phyllosphere samples (*BC* index, Hellinger transformation on ASV relative abundance), using PERMANOVA (10,000 permutations; **Supplementary Materials and Methods S1**) and principal component analysis (PCA). For *rpoB* barcoding, specifically, we reported *Methylobacterium* ASV significantly associated with the aforementioned factors (*F*, *P*, *H*, *T;* ANOVA) into the PCA (**Supplementary Materials and Methods S1)**.

Spatial and temporal dynamics of Methylobacterium *communities*

We evaluated the spatial and temporal dynamics of *Methylobacterium* communities in the timeline survey (*rpoB* barcoding) using autocorrelation analyses. In order to remove potential differences in community composition between forests, we analyzed samples from MSH and SBL separately. For each pairwise comparison between two samples from the same forest, we evaluated the effects of spatial distance (*pDist*) separating trees sampled at the same date (*spatial autocorrelation analyses*) and time (*pTime*) separating dates at which trees were sampled (*temporal autocorrelation analyses*) on *BC* dissimilarity among samples (see above). We evaluated the effects of *pDist* and *pTime* on *BC* under linear models by ANOVA (**Supplementary Materials and Methods S1**).

Ecophylogenetic structure of Methylobacterium *communities*

We quantified the ecophylogenetic structure of *Methylobacterium* communities by comparing the phylogenetic dissimilarity of co-occurring *rpoB* ASVs with the dissimilarity expected under a null model of stochastic community assembly from the pool of all ASVs, in order to quantify the evidence for different community assembly processes (59) as a function of forest, host tree species, and time of sampling. For each community of *Methylobacterium* ASVs, we calculated a measure of phylogenetic dissimilarity among co-occurring ASVs (mean nearest taxon distance (MNTD)) and compared observed MNTD to that expected under a null model of stochastic community assembly from the pool of all ASVs. We calculated the standardized effect size (SES) of MNTD (60), which expresses the difference between the observed MNTD value versus the mean and standard deviation of MNTD values obtained across 999 random draws of ASVs from the pool of observed ASVs across all samples while maintaining observed sample ASV richness (61). We evaluated the effects of forest, host tree species, and time of sampling on SES(MNTD) by ANOVA.

Monitoring of Methylobacterium *growth performance*

We evaluated the growth abilities of 79 *Methylobacterium* isolates from the timeline survey for four temperature treatments mimicking temperature variations during the growing season. Each treatment consisted of an initial pre-conditioning step (*P*) during which each isolate was incubated on solid MMS media with methanol as sole carbon source for 20 days at either 20 °C (*P20*) or 30 °C (*P30*), and a second monitoring step (*M*) during which pre-conditioned isolates were incubated on the same media and their growth monitored for 24 days at 20 °C (*P20M20* and *P20M30*) or 30 °C (*P30M20* and *P30M30*; **Figure S2**). Treatments *P20M20* and *P30M30* mimicked stable thermal environments, and treatments *P20M30* and *P30M20* mimicked variable thermal environments. For each isolate and each combination of treatments (*PXXMXX*), we realized 5 replicates, randomly spotted on 48 petri dishes according to a 6×6 grid. During the monitoring step, we took photographs of each petri dish at days 7, 13 and 24 after inoculation (**Figure S2**). Photos were converted to pixel intensities with ImageJ 1.52e and processed in R for background correction, measurement of spot intensities and correction for position-dependant competition effects (**Figure S3**; **Supplementary Materials and Methods S1).** For each isolate and temperature treatment, logistic growth curves were inferred from bacteria spot intensity variation observed over three time points during the monitoring step. From growth curves, we estimated maximum growth intensity, or yield (Y) and growth rate (r) as the inverse of lag+log time necessary to reach Y (**Figure S4** (62, 63); **Supplementary Materials and Methods S1**). We evaluated the effects of following factors on *Methylobacterium* growth abilities (Y and r) under different temperature treatments: isolate assignement to clades (C), forest of origin (F), host tree species (H), time of sampling (D), temperature of isolation (T_I; at which each isolate was isolated), temperature of incubation during pre-conditioning (T_P) and monitoring (T_M) steps, and all possible interactions between those factors (ANOVA; **Supplementary Materials and Methods S1**).

Results

Phylogenetics of plant-associated Methylobacterium diversity.

A phylogeny of 153 *Methylobacterium* isolates built from available genomic databases showed that plants (65% of strains) and especially the phyllosphere compartment (41% of strains) were the most prevalent source of *Methylobacterium* sampled to date (**Figure 2; Dataset S1a**). Phyllosphere-associated diversity was not randomly distributed in the *Methylobacterium* phylogenetic tree. Isolates from the phyllosphere represented the largest part of diversity within group A (56% of isolates) but not in groups B and C (17 and 12% of isolates, respectively). Group A was paraphyletic and most of its diversity consisted of undescribed taxa falling outside previously well-described linages. Accordingly, we subdivided *Methylobacterium* group A into 9 monophyletic clades (A1-A9).

16S rRNA *community analyses of the tree phyllosphere.*

We focused on *Methylobacterium* phyllosphere diversity variation observable at the scale of seasonal variation (within year 2018) on individual trees within two temperate forests of northeastern North America (**Figure 1a,b; Dataset S1b,g**): Mont Saint Hilaire (MSH**; Figure 1c**) and Station biologique des Laurentides (SBL**; Figure 1d**). The distribution of the phyllosphere bacterial community assessed in 46 leal samples by bacterial *16S rRNA* gene amplicon sequence variants (ASVs) was mostly explained by differences among forests (31.6% of variation explained; $p<0.001$; PERMANOVA), host tree species (15.6% of variation; $p<0.001$) and time of sampling (12.0%; $p<0.05$; **Table 1**). Although representing only 1.3% (0.0-3.2% per sample) of total *16S rRNA* sequence diversity, *Methylobacterium* was present in almost all analyzed samples (45 out of 46; **Dataset S1h**). We assigned the 15 *Methylobacterium* ASVs identified by *16S rRNA* sequencing to clades from *Methylobacterium* group A: A9 (related to *M. phyllosphaerae/M. mesophilicum/M. phyllostachyos/ M. pseudosasicola/M. organophilum;* 0.87% of total diversity, nine ASVs), A6 (related to *M. cerastii*, 0.29%; one ASV) and A1 (related to *M. gossipicola*; 0.13%, 3 ASVs; **Table S2**; **Dataset S1i**). With two rare ASVs (<0.01% of relative abundance) related to *M. komagatae*, belonging to group A (31) but unrelated to any aforementioned clade, we defined a new clade (A10). No ASVs from MSH or SBL were assigned to group B or group C.

Culture-based assessment of Methylobacterium diversity in the tree phyllosphere.

We evaluated the culturable part of *Methylobacterium* diversity from a subsample of 36 trees (18 per forest). Using *rpoB* gene partial nucleotide sequences as a marker, we identified 167 pink isolates that we assigned to *Methylobacterium* based upon their phylogenetic placement (**Dataset S1e,f; Figure 3)**. As observed for *16S rRNA* ASVs, most isolates were assigned to clades from group A typical of the phyllosphere: A9 (59.9% of isolates), A6 (24.6%), A1 (5.4%), A10 (3.6%) and A2 (related to *M. bullatum* and *M. marchantiae*; **Dataset S1d** 1.8%). Few isolates were assigned to group B (4.2% of isolates, related to *M. extorquens*) and none to group C (**Table S2**). The higher polymorphism in the *rpoB* marker revealed a considerable diversity within clades, as we identified 71 unique *rpoB* sequences, in contrast to the smaller number obtained with *16S rRNA* barcoding (15 ASVs). We determined that *Methylobacterium* diversity assessed at varying depths in the *rpoB* phylogeny was systematically explained by forest of origin (4.5±1.0% of variance explained; PERMANOVA; $p<0.001$; **Figure 3a; Dataset S1j**) and temperature of isolation (5.9±2.1% of variance explained; $p<0.001$). Temperature of isolation was the most important factor distinguishing deep phylogenetic divergences (pairwise nucleotide similarity range: 0.948-0.993), while forest of origin was slightly more important in structuring more recently diverged nodes (pairwise nucleotide similarity >0.993). Time of sampling had a slight but significant effect on diversity (2.1±0.2% of variance explained; $p<0.05$) and it was only observed for higher pairwise nucleotide similarity values (range 0.994-1.000). We did not observe any significant effects of host tree species on *Methylobacterium* isolate diversity, at any level of the phylogeny. In the phylogeny, we identified two nodes strongly associated with temperature of isolation, corresponding to clades A6 (20 °C; $p<0.001$; permutation test) and

A9+A10 (30 °C; *p*<0.001; **Figure 3b**). Other clades were evenly isolated at 20 and 30 °C and we observed no significant association between temperature of isolation and nodes embedded within clades. Nodes associated with forest of origin also roughly corresponded to certain major clades, with clades A1+A2 almost exclusively sampled in MSH (*p*<0.01). Overall, clade A9 was isolated significantly more often at SBL (*p*<0.001) but at least three of its subclades were significantly associated with either MSH or SBL (*p*<0.05).

Comparison of Methylobacterium diversity assessed by rpoB barcoding and isolation

We performed culture-independent *rpoB* amplicon sequencing from 179 leaf samples from 53 trees in both forests, allowing a monthly monitoring for most trees (**Dataset S1b,g**). We identified 283 *Methylobacteriaceae rpoB* ASVs in these samples (**Dataset S1k,l**), representing 24.6% of all sequences. Non-*Methylobacteriaceae* ASVs were mostly assigned to other Rhizobiales families (850 ASVs, 70.33% of sequence abundance) and to Caulobacterales (209 ASVs, 4.42% of sequence abundance) typical of the phyllosphere (**Supplementary Materials and Methods S1**), indicating that the *rpoB* marker can potentially be used at a broader taxonomic scale (**Figure S5a**). Within *Methylobacteriaceae*, ASVs were mostly classified as *Methylobacterium* (200 ASVs, 23.05% of sequence relative abundance), and *Enterovirga* (78 ASVs, 1.56%; **Dataset S1k**). We assigned most of *Methylobacterium* ASVs to previously cultured clades A9 (45.2% of *Methylobacterium* sequence abundance), A6 (24.3%), A1 (6.1%) and A10 (1.0%; **Dataset S1k; Table S2; Figure S5b**). Estimates of *Methylobacterium* diversity based on *rpoB* sequences from culture-independent sequencing were generally concordant with estimates based on *16S rRNA* barcoding (**Figure S5c; Table S2**) and estimates from cultured isolates (**Figure S5d; Table S2**). The major exception was group B, representing 19.1% of

Methylobacterium sequence abundance (*rpoB* barcoding) but not detected by *16S rRNA* barcoding, and representing 4.2% of isolates (**Table S2**). Clade A4 (related to *M. gnaphalii* and *M. brachytecii*) represented 1.7% of *Methylobacterium* sequence abundance (*rpoB* barcoding) but was not detected by *16S rRNA* barcoding, nor was it isolated. Other clades could be detected by *rpoB* barcoding with low sequence abundance (<0.3%) but not by *16S rRNA* barcoding, and were unevenly isolated (<1.8% of isolates).

Fine-scale temporal and spatial distribution of Methylobacterium *diversity assessed by* rpoB *barcoding*

The community composition of the 200 *Methylobacterium* ASVs was mostly explained by spatial variation at both large (distance between forests: 100 km) and local scales (distance between plots within forest: 150-1,200 m), as well as sampling date during the growing season (1-5 months; proportion of variation explained: 32.4%, 8.0% and 4.8%, respectively; $p<0.001$; PERMANOVA; **Table 1**). We observed slight but significant effects of host tree species, and of the interaction between host tree species and plots within forests, on *Methylobacterium* community composition (explaining 7.1% and 4.3% of variation in community composition; $p<0.001$ and $p<0.01$, respectively; PERMANOVA; **Table 1**). A large proportion of *Methylobacterium* ASVs (83 out of 200) were significantly associated with one or either forest (ANOVA; **Figure 4a; Dataset S1m**), regardless their clade membership. The only exception was clade A1, which was almost exclusively observed (and isolated; see **Figure 3b**) in the MSH forest. We found 25 ASVs whose relative abundance significantly increased throughout the growing season (ANOVA; $p<0.05$), mostly belonging to clades A1 (n=11). Four ASVs increased significantly in frequency over time in both forests, and mostly belonged to group B (n=3), (**Dataset S1m**). We found no clear association between ASV or clade with host tree species, nor plots within forests (data not shown). *Methylobacterium* diversity was heterogeneously distributed at local spatial scale, as we observed a significant increase of community dissimilarity (Bray-Curtis index; *BC*) with geographical distance separating two samples within MSH (spatial autocorrelation analysis; ANOVA; $p<0.001$) but not SBL ($p>0.05$, **Table 2, Figure 4b**). We also observed a significant increase of community dissimilarity over time separating two sampling dates in both forests (temporal autocorrelation analysis; $p<0.001$; **Table 2**), indicating that community composition changed during the growing season. This effect was more marked in MSH than in SBL (**Figure 4c**). The overall community *BC* dissimilarity consistently decreased from June to October in both MSH (from 0.624 to 0.297) and SBL (from 0.687 to 0.522; **Table 2, Figure 4d**), indicating that the observed change of diversity over time resulted from a progressive homogeinization of *Methylobacterium* community between the beginning and the end of the growing season at the scale of a forest, although without affecting its heterogeneous spatial distributions in MSH (**Table 2, Figure 4e**). *Methylobacterium* communities were strongly phylogenetically clustered (**Figure 4f**), with all communities containing ASVs that were much more closely related than expected by chance (mean SES(MNTD) (± standard deviation) = -4.8 ± 0.9, all SES(MNTD) *p*-values <0.05 compared with null model of random community assembly). While all communities were strongly phylogenetically clustered, SES(MNTD) differed among host tree species (ANOVA; F=6.4, $p<0.001$) and forests (ANOVA; F=10.9, P<0.001), and decreased during the growing season (ANOVA; F=95.2, $p<0.001$).

Effect of short scale temperature variation in combination with other environmental and genetic factors on Methylobacterium growth performances

We measured growth of 79 *Methylobacterium* isolates (sampled in 2018 in both forests; MSH: n=32, SBL: n=47) in conditions mimicking temperature variations during the growing season (**Figures S2-S4** ; **Dataset S1n**). Clade membership explained a large part of variation in growth rate (r) and yield (Y; 7.6 and 30.6% of variation explained, respectively; ANOVA; $p<0.001$; **Figures 5a,b, Table 3; Dataset S1o**). Group B isolates ($Y = 12.2 \pm 5.0$) have higher yield than group A ($Y = 5.4 \pm 3.5$). Isolates from clades A1, A2 and B had the highest growth rate (r range: 0.101±0.032 – 0.121±0.031). Other clades (A6, A9 and A10) had on average slower growth (r range: 0.082±0.021 – 0.088±0.024). Time of sampling, host tree species and forest also explained significant variation in growth rate (5.4%, $p<0.001$; 2.2%, $p<0.01$ and 1.5%, $p<0.05$, respectively; ANOVA) and limited or no significant variation in yield (1.3%; $p<0.001$; 1.3%; $p<0.01$; 0.2%; $p>0.05$, respectively; **Table 3**). Among the aforementioned factors only the interaction between time of sampling and clade membership explained significant variation in growth rate (2.9%; $p<0.001$), while all possible pairwise interactions between these factors explained significant variation in yield (range 1.4 – 5.9%; $p<0.01$; **Table 3**). In both SBL and MSH, growth rate increased consistently from June (r = 0.075±0.018 and 0.085±0.033, respectively) to September/October (r = 0.097±0.031 and 0.103±0.027, respectively; **Figure 5c**). The temperature of isolation (at which each isolate was originally isolated) had very limited effect on growth rate (1.0%; $p<0.01$) and yield (0.6%; $p<0.05$). These effects were independent of temperatures during pre-conditioning and monitoring steps (no significant interaction in the ANOVA). Temperature of incubation had significant effects on growth performance. Temperature during the monitoring step explained respectively 2.0% and 15.8% of variation in yield and growth rate ($p<0.01$ and $p<0.001$, respectively; ANOVA; **Table 3**), regardless of clade membership, time of sampling, and other environmental factors (no significant interaction in the ANOVA). Isolates incubated at 20 °C had on average higher yield (Y=6.9±5.4) but slower growth

(r=0.077±0.022) than isolates incubated at 30 °C (Y=4.9±3.6; r=0.100±0.030; **Figure 5d**). Temperature during the pre-conditioning step had no effect on growth rate (p>0.05; ANOVA), and limited effect on yield (1.4%; p<0.05; ANOVA; **Table 3**).

Discussion

Methylobacterium is ubiquitous on leaves in the temperate forests of Québec and its diversity in this habitat is quite similar to what has been described in the phyllosphere throughout the world, with three main clades A9 (*M. brachiatum*, *M. pseudosasicola*), A6 (related to *M. cerastii*) and A1 (related to *M. gossipicola*) dominating diversity. Our barcoding approach based on a clade-specific *rpoB* marker revealed previously undocumented diversity within these clades, as well as within several other clades that were not detected by a classical *16S rRNA* marker: B (related to *M. extorquens*), A2 (related to *M. bullatum* and *M. marchantiae*), A4 (related to *M. gnaphalii* and *M. brachytecii*) and A10 (related to *M. komagatae*). This diversity, like that of the overall phyllosphere community, was mostly determined by differences between forests, with barcoding approaches suggesting combined effects of restricted migration, local adaptation to host tree species, and climatic conditions at large geographical scales (>100km). With higher molecular resolution, we observed that *Methylobacterium* diversity was spatially structured even at the scale of a forest (within 1.2 km), and also showed a clear pattern of temporal dynamics and succession over the course of a growing season. This result indicates that, although representing a stable proportion of the plant leaf microbiota between years (22) *Methylobacterium* diversity is highly dynamic within the course of a season. A finer analysis of *Methylobacterium* diversity suggested that clade identity partly explained *Methylobacterium* geographical distribution at large scales (between forests) but not at finer scales (plots), nor was it an indicator of adaptation to a particular host tree species, nor a determinant of temporal dynamics. These results are consistent with previous observations that geographic origin is a stronger driver of phyllosphere *Methylobacterium* diversity than host identity (22). The distribution of *Methylobacterium* diversity at small temporal and geographical scales likely resulted from more contemporaneous community assembly events selecting for phenotypic traits that evolved among deeply diverging lineages of *Methylobacterium*, as has been observed in other bacterial (16) and plant clades (64). We found further evidence for deterministic community assembly as *Methylobacterium* communities were strongly phylogenetically clustered compared to the expectation under a stochastic model of community assembly, indicating that the leaf habitat acts as an ecological filter selecting for a non-random subset of *Methylobacterium* diversity.

We explored mechanisms explaining the temporal dynamics of *Methylobacterium* diversity at the scale of a growing season. Because we observed contrasting *Methylobacterium* culturable diversity between 20 and 30 °C, we suspected that adaptation to temperature variation during the growing season could explain part of these temporal dynamics. By monitoring *Methylobacterium* isolate growth under different temperature treatments, we confirmed that temperature affected isolate growth performances but interestingly, independantly from the temperature at which isolates were obtained. The fact that most tested isolates also grew slower but more efficiently at 20 °C than at 30 °C (**Figure 5d**), regardless of their phylogenetic and environmental characteristics, is in line with a temperature-dependent trade-off between growth rate and yield described in many bacteria (reviewed in (63)). High yield strategies are typical of cooperative bacterial populations, while fast growth-strategies are typical of competitive populations (63). These observations also stress the importance of considering incubation temperature when interpreting results from previous culture-based assessments of *Methylobacterium* diversity.

We provide two lines of evidence that factors other than direct adaptation to temperature drive *Methylobacterium* responses to temperature variation, by affecting their growth strategy in different competitive conditions rather than by affecting their metabolism directly. First, clade identity was one of the main predictors of overall isolate performance, with some clades (A1, A2, B) possessing a rapid growth strategy under all temperature conditions, while others (clades A6, A9, A10) had systematically slower growth. These clade-specific growth strategies could explain why certain *Methylobacterium* isolates are less competitive and less frequently isolated at higher temperatures. Still, we cannot rule out that clade-specific growth strategy also reflect experimental conditions. Second, we observed strong associations between isolate growth performance and time of sampling, regardless of clade membership, suggesting that growth strategies also respond to seasonal variations in environmental conditions, and to the level of establishment and competition in the phyllosphere community (63). These associations are unlikely to be driven by the direct effects of temperature on metabolic rates because isolation temperature had little effect on growth strategies, in contrast to clade identity and time of sampling which had more significant effects. Together, these observations could explain why isolates from clades A1 and B with fast-growth strategies consistently increase in frequency during this period due to changes in selection for different ecological strategies, leading to the homogeneization of the community.

Taken together, our temporal survey of diversity dynamics and screening for growth performance suggest the following timeline of the dynamics of the *Methylobacterium* phyllosphere community. At the very beginning of the growing season, a pool of bacteria with mixed ecological strategies and genotypes colonizes newly emerging leaves. Due to the stochasticity of this colonization, we initially observe strong dissimilarity among phyllosphere communities, regardless of their spatial position. During the summer, conditions allow the progressive establishment of a diverse *Methylobacterium* community with a high yield strategies (63), dominated by increasingly closely phylogenetically related strains. At the end of the growing season, with migration, environmental conditions shifting and leaves senescing, isolates with a fast-growth strategy are able to grow rapidly, dominating the phyllosphere community and leading to its further homogeneization before leaves fully senesce. This scenario provides an explanation for the observation of community convergence and increasing homogeneity of phyllosphere communities throughout the growing season (65, 66).

Our study illustrates that *Methylobacterium* is a complex group of divergent lineages with different ecological strategies and distributions, reflecting long-term adaptation to contrasting local environments. Based upon a similar observation, some authors recently proposed to reclassify *Methylobacterium* group B within a new genus (*Methylorubrum*) that they argue is ecologically and evolutionarily distinct from other *Methylobacterium* clades (31). Although clade B was well supported as a distinct clade in our analyses, our results suggest that it is in fact embedded within clade A, which would render the genus *Methylobacterium* paraphyletic if clade B is defined as a distinct genus (**Figure S1**). Furthermore, group B was not particularly ecologically distinct in comparison with other major clades (**Figure 2**). Our results emphasize the fact that thorough genomic investigations are needed to clarify the taxomonic status of *Methylobacterium*. Beyond any taxonomic considerations, neither clade identity assessed by individual genetic markers nor the tremendous ecological diversity among *Methylobacterium* clades can predict all of the spatial and temporal variation in *Methylobacterium* diversity in nature. In order to define the niches of *Methylobacterium* clades and to understand the metabolic mechanisms underlying their contrasting life strategies, future characterization of their functions and genome structure will be required using phylogenomic approaches.

In conclusion, we find that *Methylobacterium* adaptive responses to local environmental variation in the phyllosphere are driven by both long-term inherited ecological strategies that differ among major clades within the genus, as well by seasonal changes affecting habitat characteristics and community structure in the phyllosphere habitat. Overall, our study combining culture-free and culture-based approaches provides novel insights into the factors driving fine-scale adaptation of microbes to their habitats. In the case of *Methylobacterium,* our approach revealed the particular importance of considering organismal life-history strategies to help understand the fine-scale diversity and dynamics of this ecologically important taxon.

References

1. Vorholt JA. 2012. Microbial life in the phyllosphere. 12. Nature Reviews Microbiology 10:828–840.

2. Lindow SE, Brandl MT. 2003. Microbiology of the Phyllosphere. Appl Environ Microbiol 69:1875–1883.

3. Fürnkranz M, Wanek W, Richter A, Abell G, Rasche F, Sessitsch A. 2008. Nitrogen fixation by phyllosphere bacteria associated with higher plants and their colonizing epiphytes of a tropical lowland rainforest of Costa Rica. 5. The ISME Journal 2:561–570.

4. Laforest-Lapointe I, Messier C, Kembel SW. 2016. Host species identity, site and time drive temperate tree phyllosphere bacterial community structure. Microbiome 4:27.

5. Laforest-Lapointe I, Messier C, Kembel SW. 2016. Tree phyllosphere bacterial communities: exploring the magnitude of intra- and inter-individual variation among host species. PeerJ 4:e2367.

6. Noble AS, Noe S, Clearwater MJ, Lee CK. 2020. A core phyllosphere microbiome exists across distant populations of a tree species indigenous to New Zealand. PLOS ONE 15:e0237079.

7. Kembel SW, O'Connor TK, Arnold HK, Hubbell SP, Wright SJ, Green JL. 2014. Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. PNAS 111:13715–13720.

8. Shade A, Gregory Caporaso J, Handelsman J, Knight R, Fierer N. 2013. A meta-analysis of changes in bacterial and archaeal communities with time. 8. The ISME Journal 7:1493–1506.

9. Malik AA, Martiny JBH, Brodie EL, Martiny AC, Treseder KK, Allison SD. 2020. Defining trait-based microbial strategies with consequences for soil carbon cycling under climate change. 1. The ISME Journal 14:1–9.

10. Moyes AB, Kueppers LM, Pett-Ridge J, Carper DL, Vandehey N, O'Neil J, Frank AC. 2016. Evidence for foliar endophytic nitrogen fixation in a widely distributed subalpine conifer. New Phytologist 210:657–668.

11. Lajoie G, Kembel SW. 2019. Making the Most of Trait-Based Approaches for Microbial Ecology. Trends in Microbiology 27:814–823.

12. Nemergut DR, Schmidt SK, Fukami T, O'Neill SP, Bilinski TM, Stanish LF, Knelman JE, Darcy JL, Lynch RC, Wickey P, Ferrenberg S. 2013. Patterns and Processes of Microbial Community Assembly. Microbiol Mol Biol Rev 77:342–356.

13. Cavender-Bares J, Kozak KH, Fine PVA, Kembel SW. 2009. The merging of community ecology and phylogenetic biology. Ecology Letters 12:693–715.

14. Martiny JBH, Jones SE, Lennon JT, Martiny AC. 2015. Microbiomes in light of traits: A phylogenetic perspective. Science 350.

15. Tromas N, Taranu ZE, Castelli M, Pimentel JSM, Pereira DA, Marcoz R, Shapiro BJ, Giani A. 2020. The evolution of realized niches within freshwater Synechococcus. Environmental Microbiology 22:1238–1250.

16. Tromas N, Taranu ZE, Martin BD, Willis A, Fortin N, Greer CW, Shapiro BJ. 2018. Niche Separation Increases With Genetic Distance Among Bloom-Forming Cyanobacteria. Front Microbiol 9.

17. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, Prill RJ, Tripathi A, Gibbons SM, Ackermann G, Navas-Molina JA, Janssen S, Kopylova E, Vázquez-Baeza Y, González A, Morton JT, Mirarab S, Zech Xu Z, Jiang L, Haroon MF, Kanbar J, Zhu Q, Jin Song S, Kosciolek T, Bokulich NA, Lefler J, Brislawn CJ, Humphrey G, Owens SM, Hampton-Marcell J, Berg-Lyons D, McKenzie V, Fierer N, Fuhrman JA, Clauset A, Stevens RL, Shade A,

Pollard KS, Goodwin KD, Jansson JK, Gilbert JA, Knight R. 2017. A communal catalogue reveals Earth’s multiscale microbial diversity. 7681. Nature 551:457–463.

18. Poretsky R, Rodriguez-R LM, Luo C, Tsementzi D, Konstantinidis KT. 2014. Strengths and Limitations of 16S rRNA Gene Amplicon Sequencing in Revealing Temporal Microbial Community Dynamics. PLOS ONE 9:e93827.

19. Ranjan R, Rani A, Metwally A, McGee HS, Perkins DL. 2016. Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. Biochemical and Biophysical Research Communications 469:967–977.

20. Corpe WA, Rheem S. 1989. Ecology of the methylotrophic bacteria on living leaf surfaces. FEMS Microbiol Ecol 5:243–249.

21. Keppler F, Hamilton JTG, Braß M, Röckmann T. 2006. Methane emissions from terrestrial plants under aerobic conditions. 7073. Nature 439:187–191.

22. Knief C, Ramette A, Frances L, Alonso-Blanco C, Vorholt JA. 2010. Site and plant species are important determinants of the Methylobacterium community composition in the plant phyllosphere. ISME J 4:719–728.

23. Clarke PH. 1983. The biochemistry of Methylotrophs by C. Anthony Academic Press; London, New York, 1982 xvi + 432 pages. £24.00, $49.50. FEBS Letters 160:303–303.

24. Anthony C. 1991. Assimilation of carbon by methylotrophs. Biotechnology 18:79–109.

25. Ivanova EG, Doronina NV, Trotsenko YuA. 2001. Aerobic Methylobacteria Are Capable of Synthesizing Auxins. Microbiology 70:392–397.

26. Madhaiyan M, Poonguzhali S, Lee HS, Hari K, Sundaram SP, Sa TM. 2005. Pink-pigmented facultative methylotrophic bacteria accelerate germination, growth and yield of sugarcane clone Co86032 (Saccharum officinarum L.). Biol Fertil Soils 41:350–358.

27. Madhaiyan M, Poonguzhali S, Sa T. 2007. Metal tolerating methylotrophic bacteria

reduces nickel and cadmium toxicity and promotes plant growth of tomato (Lycopersicon esculentum L.). Chemosphere 69:220–228.

28. Dourado MN, Aparecida Camargo Neves A, Santos DS, Araújo WL. 2015. Biotechnological and Agronomic Potential of Endophytic Pink-Pigmented Methylotrophic Methylobacterium spp. Biomed Res Int 2015.

29. Ryu JH (Chungbuk NU, Madhaiyan M (Chungbuk NU, Poonguzhali S (Chungbuk NU, Yim WJ (Chungbuk NU, Indiragandhi P (Chungbuk NU, Kim KA (Chungbuk NU, Anandham R (Chungbuk NU, Yun JC (National I of AS and T, Kim KH (The U of S, Sa TM (Chungbuk NU. 2006. Plant Growth Substances Produced by Methylobacterium spp. and Their Effect on Tomato (Lycopersicon esculentum L.) and Red Pepper (Capsicum annuum L.) Growth. Journal of Microbiology and Biotechnology.

30. Lee HS, Madhaiyan M, Kim CW, Choi SJ, Chung KY, Sa TM. 2006. Physiological enhancement of early growth of rice seedlings (Oryza sativa L.) by production of phytohormone of N2-fixing methylotrophic isolates. Biol Fertil Soils 42:402–408.

31. Green PN, Ardley JK. 2018. Review of the genus Methylobacterium and closely related organisms: a proposal that some Methylobacterium species be reclassified into a new genus, Methylorubrum gen. nov. International Journal of Systematic and Evolutionary Microbiology 68:2727–2748.

32. Chen W-M, Cai C-Y, Li Z-H, Young C-C, Sheu S-Y. 2019. Methylobacterium oryzihabitans sp. nov., isolated from water sampled from a rice paddy field. International Journal of Systematic and Evolutionary Microbiology, 69:3843–3850.

33. Feng G-D, Chen W, Zhang X-J, Zhang J, Wang S-N, Zhu H. 2020. Methylobacterium nonmethylotrophicum sp. nov., isolated from tungsten mine tailing. International Journal of Systematic and Evolutionary Microbiology, 70:2867–2872.

34. Jia L juan, Zhang K shuai, Tang K, Meng J yu, Zheng C, Feng F ying. 2020. Methylobacterium crusticola sp. nov., isolated from biological soil crusts. International Journal of Systematic and Evolutionary Microbiology, 70:2089–2095.

35. Kim J, Chhetri G, Kim I, Lee B, Jang W, Kim MK, Seo T. 2020. Methylobacterium terricola sp. nov., a gamma radiation-resistant bacterium isolated from gamma ray-irradiated soil. International Journal of Systematic and Evolutionary Microbiology, 70:2449–2456.

36. Kim J, Chhetri G, Kim I, Kim MK, Seo T. 2020. Methylobacterium durans sp. nov., a radiation-resistant bacterium isolated from gamma ray-irradiated soil. Antonie van Leeuwenhoek 113:211–220.

37. Ten LN, Li W, Elderiny NS, Kim MK, Lee S-Y, Rooney AP, Jung H-Y. 2020. Methylobacterium segetis sp. nov., a novel member of the family Methylobacteriaceae isolated from soil on Jeju Island. Arch Microbiol 202:747–754.

38. Jiang L, An D, Wang X, Zhang K, Li G, Lang L, Wang L, Jiang C, Jiang Y. 2020. Methylobacterium planium sp. nov., isolated from a lichen sample. Arch Microbiol 202:1709–1715.

39. Pascual JA, Ros M, Martínez J, Carmona F, Bernabé A, Torres R, Lucena T, Aznar R, Arahal DR, Fernández F. 2020. Methylobacterium symbioticum sp. nov., a new species isolated from spores of Glomus iranicum var. tenuihypharum. Curr Microbiol 77:2031–2041.

40. Marx CJ, Bringel F, Chistoserdova L, Moulin L, Farhan Ul Haque M, Fleischman DE, Gruffaz C, Jourand P, Knief C, Lee M-C, Muller EEL, Nadalig T, Peyraud R, Roselli S, Russ L, Goodwin LA, Ivanova N, Kyrpides N, Lajus A, Land ML, Médigue C, Mikhailova N, Nolan M, Woyke T, Stolyar S, Vorholt JA, Vuilleumier S. 2012. Complete Genome Sequences of Six Strains of the Genus Methylobacterium. J Bacteriol 194:4746–4748.

41. Tani A, Ogura Y, Hayashi T, Kimbara K. 2015. Complete Genome Sequence of

Methylobacterium aquaticum Strain 22A, Isolated from Racomitrium japonicum Moss. Genome Announc 3.

42. Minami T, Ohtsubo Y, Anda M, Nagata Y, Tsuda M, Mitsui H, Sugawara M, Minamisawa K. 2016. Complete Genome Sequence of Methylobacterium sp. Strain AMS5, an Isolate from a Soybean Stem. Genome Announc 4.

43. Morohoshi T, Ikeda T. 2016. Complete Genome Sequence of Methylobacterium populi P-1M, Isolated from Pink-Pigmented Household Biofilm. Genome Announc 4.

44. Belkhelfa S, Labadie K, Cruaud C, Aury J-M, Roche D, Bouzon M, Salanoubat M, Döring V. 2018. Complete Genome Sequence of the Facultative Methylotroph Methylobacterium extorquens TK 0001 Isolated from Soil in Poland. Genome Announc 6.

45. Green PN. 2015. Methylobacterium, p. 1–8. *In* Bergey's Manual of Systematics of Archaea and Bacteria. American Cancer Society.

46. Charron G, Leducq JB, Bertin C, Dube AK, Landry CR. 2014. Exploring the northern limit of the distribution of Saccharomyces cerevisiae and Saccharomyces paradoxus in North America. FEMS Yeast Res 14:281–8.

47. Vos M, Quince C, Pijl AS, Hollander M de, Kowalchuk GA. 2012. A Comparison of rpoB and 16S rRNA as Markers in Pyrosequencing Studies of Bacterial Diversity. PLOS ONE 7:e30600.

48. Ogier J-C, Pagès S, Galan M, Barret M, Gaudriault S. 2019. rpoB, a promising marker for analyzing the diversity of bacterial communities by amplicon sequencing. BMC Microbiol 19:171.

49. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. Systematic Biology 61:539–542.

50. Küpfer M, Kuhnert P, Korczak BM, Peduzzi R, Demarta A. 2006. Genetic relationships of Aeromonas strains inferred from 16S rRNA, gyrB and rpoB gene sequences. International Journal of Systematic and Evolutionary Microbiology, 56:2743–2751.

51. Ee R, Madhaiyan M, Ji L, Lim Y-L, Nor NM, Tee K-K, Chen J-W, Yin W-F. 2016. Chania multitudinisentens gen. nov., sp. nov., an N-acyl-homoserine-lactone-producing bacterium in the family Enterobacteriaceae isolated from landfill site soil. International Journal of Systematic and Evolutionary Microbiology, 66:2297–2304.

52. Frank RAW, Price AJ, Northrop FD, Perham RN, Luisi BF. 2007. Crystal Structure of the E1 Component of the Escherichia coli 2-Oxoglutarate Dehydrogenase Multienzyme Complex. Journal of Molecular Biology 368:639–651.

53. Bell RL, González-Escalona N, Stones R, Brown EW. 2011. Phylogenetic evaluation of the ‘Typhimurium’ complex of Salmonella strains using a seven-gene multi-locus sequence analysis. Infection, Genetics and Evolution 11:83–91.

54. Dixon P. 2003. VEGAN, a package of R functions for community ecology. Journal of Vegetation Science 14:927–930.

55. Redford AJ, Bowers RM, Knight R, Linhart Y, Fierer N. 2010. The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. Environmental Microbiology 12:2885–2893.

56. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: High resolution sample inference from Illumina amplicon data. Nat Methods 13:581–583.

57. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596.

58. Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, Madden TL, Matten WT, McGinnis SD, Merezhuk Y, Raytselis Y, Sayers EW, Tao T, Ye J, Zaretskaya I. 2013. BLAST: a more efficient report with usability improvements. Nucleic Acids Res 41:W29-33.

59. Stegen JC, Lin X, Fredrickson JK, Chen X, Kennedy DW, Murray CJ, Rockhold ML, Konopka A. 2013. Quantifying community assembly processes and identifying features that impose them. ISME J 7:2069–2079.

60. Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO. 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26:1463–1464.

61. Kembel SW. 2009. Disentangling niche and neutral influences on community assembly: assessing the performance of community phylogenetic structure tests. Ecology Letters 12:949–960.

62. Zwietering MH, Jongenburger I, Rombouts FM, Riet K van ’t. 1990. Modeling of the Bacterial Growth Curve. Appl Environ Microbiol 56:1875–1881.

63. Lipson DA. 2015. The complex relationship between microbial growth rate and yield and its implications for ecosystem processes. Front Microbiol 6.

64. Cavender-Bares J, Ackerly DD, Baum DA, Bazzaz FA. 2004. Phylogenetic Overdispersion in Floridian Oak Communities. The American Naturalist 163:823–843.

65. Copeland JK, Yuan L, Layeghifard M, Wang PW, Guttman DS. 2015. Seasonal Community Succession of the Phyllosphere Microbiome. MPMI 28:274–285.

66. Maignien L, DeForce EA, Chafee ME, Eren AM, Simmons SL. Ecological Succession and Stochastic Variation in the Assembly of Arabidopsis thaliana Phyllosphere Communities. mBio 5:e00682-13.

Figure and table legends

Figure 1 - Sampling design. a) Locations of the two sampled forests MSH (green) and SBL (orange) in the province of Québec (Canada). b) Time line survey in each forest in 2018 (2-4 time points available per tree). c-d). Detailed map of each forest and each plot within forests (squares; 6 to 10 trees were sampled per plot; see **Dataset S1b**). For each plot, trees are indicated by points colored according to their taxonomy (color code on bottom left): ABBA (*Abies balsamea*), ACRU (*Acer rubrum*). ACSA (*Acer saccharum*), OSVI (*Ostrya virginiana*), QURU (*Quercus rubra*), FAGR (*Fagus grandifolia*), ASPE (*Acer Pennsylvanicum*). Shades of grey indicate elevation (50 m elevation scale)

Figure 2 - ***Methylobacterium* phylogeny and ecology.** Most of *Methylobacterium* diversity is found in association with plants, especially in the phyllosphere. Phylogenetic consensus tree (nodal posterior probabilities indicated next to the branches) from *rpoB* complete nucleotide sequences available for 153 *Methylobacterium* genomes and rooted on 32 *Methylobacteriaceae* outgroups (*Microvirga*, *Enterovirga*; no shown; see **Dataset S1a**). For each genome, species name, the anthropogenic origin (black squares) and/or environmental origin (color code on top right) are indicated. Groups A, B, C adapted from Green *et* Ardley (31).

Figure 3 - Tests for phylogenetic association of traits with culture-based estimation of *Methylobacterium* diversity. a) Part of variance (PERMANOVA; x-axis) in *Methylobacterium* isolated diversity explained by forest of origin, host tree species, sampling date, temperature of isolation and their interactions (see Venn diagram on top left for color code) in function of pairwise nucleotide similarity (PS; y-axis; see **Dataset S1j**) in a phylogenetic tree (partial *rpoB*

nucleotide sequences of 187 isolates and 188 *Methylobacteriaceae* reference sequences). b) Permutation test for node association with forest of origin and temperature of isolation (color code on top) mapped on the *rpoB* phylogeny (scaled on PS values). Frames in the tree indicate nodes significantly associated with at least one factor (ANOVA; Bonferroni correction; $p<0.001$: “***”; $p<0.01$:”**”; $p<0.05$:”*”). For each isolate (names in bold), colored boxes at the tip of the tree indicate forest of origin and temperature of isolation.

Figure 4 - Short-scale spatial and temporal dynamics of *Methylobacterium* communities assessed by *rpoB* barcoding. a) A principal component analysis (PCA) on *Methylobacterium* ASVs relative abundance shows that 179 phyllosphere samples cluster according to forest of origin (MSH: open triangles, SBL: full triangles) and date of sampling (detail showed only for MSH). The significant association of 83 and 25 ASVs with forest of origin and/or sampling date, respectively is shown (points colored according to clade assignation; legend on top right). b) Spatial and c) temporal autocorrelation analyses conducted in each forest separately. Points represent Bray-Curtis (*BC*) dissimilarity in function of pairwise geographic (*pDist*; b) or pairwise time (*pTime*; c) distance separating two communities. For each forest and variable, the predicted linear regression is indicated (full line: $p<0.001$; dotted line: $p>0.05$; ANOVA). d) *BC* in function of sampling time for each forest. e) Detail of spatial autocorrelation analyzes in MSH, conducted for each sampling time point separately. f) Standardized effect size of mean nearest taxon phylogenetic distance (SES(MNTD)) between forests and across sampling dates. Negative values of SES(MNTD) indicate communities contain ASVs that are phylogenetically clustered compared to a null model of stochastic community assembly.

Figure 5 - Analysis of 79 *Methylobacterium* isolate growth performances under 4 different temperature treatments. a) Average growth curves (growth intensity in function of time) for each clade (line: mean value; frame: 1/3 of standard deviation; point: average maximal growth). b) Growth rate (*r*) in function of yiel (*Y*). Each point represents the average *r*/*Y* values for an isolate and a temperature treatment (79 isolates x 4 treatments), colored according to clade membership. Ellipsoides are centered on average values per clade and represent 30% of confidence interval (standard deviation). c) *r* (log scale) in function of time at which samples strains were isolated from were collected, colored according to the forest of origin. Points: real data; bars: average *r* value per forest (n=2) and time (n=4) category. d) *r* in function of Y, corrected for clade assignement (residuals of the *r*~Clade and *Y*~Clade linear regressions). Each point represents the average *r*/*Y* residual values for an isolate and a temperature, colored according to the monitoring temperature (legend on top right).

Table 1 - PERMANOVA analysis of variance in *Bacteria* and *Methylobacterium* community diversity. Part of variance in dissimilarity (R^2; Bray-Curtis index) among samples associated with four factors and their possible interactions (*F*: forest of origin; *D*: date of sampling; *H*: host tree species; *P*: plot within forest) and their significance are shown (10,000 permutations on ASV relative abundance, Hellinger transformation; "***": p<0.00l; "**": p<0.01; "*": p<0.05.). For *16S rRNA*, *P* was omitted to conserve degrees of freedom.

	***Bacteria* (*16S rRNA*)**		***Methylobacterium (rpoB)***	
Samples	**46**		**179**	
Factor	**R^2**	**Pr(>F)**	**R^2**	**Pr(>F)**
Forest of origin (F)	0.316***	<0.000	0.324***	<0.001
Host tree specie (H)	0.156***	<0.001	0.071***	<0.001
Time of sampling (D)	0.120*	0.016	0.048***	<0.001
Plot within forests (P)	-	-	0.080***	<0.001
F:H	0.020	0.080	0.004	0.110
H:D	0.239	0.217	0.074**	0.028
H:P	-	-	0.043**	0.007
D:P	-	-	0.058	0.455
H:D:P	-	-	0.085	0.052
Residuals	0.150	-	0.213	-

Table 2 - Summary of statistics from autocorrelation analyzes on 179 phyllosphere *Methylobacterium* samples assessed by *rpoB* barcoding (200 ASVs). *Spatial autocorrelation general models* : pairwise dissimilarity between two communities (Bray-Curtis index; *BC*) as a function of pairwise spatial distance separating two sampled trees (*pDist*) and date of sampling (*Date*) and their interaction (*pDist:Date*). *Spatial autocorrelation models per date*: *BC* as a function of pairwise spatial distance (*pDist*). *Temporal autocorrelation: general models*: *BC* as a function of pairwise spatial time separating two sampled trees (*pTime*). For each model, the average and standard deviation of the intercept (mean *BC* value) are indicated. For each factor (*pDist*, *Date*, *pDist:Date* and *pTime*), the average and standard deviation of estimates (slope) are indicated. Significance of estimates was assessed by ANOVA ("***": p<0.001; "**": p<0.01; "*": p<0.05).

Categories (*n*)		**Intercept (sd)**		**Estimates*10^{-3} (sd)**	
*Spatial autocorrelation general models: lm(BC~pDist*D)*					
Site (within dates)		*BC*	*pDist*	*D*	*pDist:Date*
MSH		0.5965 (0.0107)	**-0.0041 (0.0192)*** **	**-2.7648 (0.1313)*** **	**0.0007 (0.0002)** **
SBL		0.6493 (0.0097)	0.0157 (0.0145)	**-1.5575 (0.1646)*** **	0.0000 (0.0002)
Spatial autocorrelation models per date: lm(BC~pDist)					
Site	**Date**	*BC*	*pDist*		
MSH	**27 Jun.**	0.6237 (0.0340)	-0.0425 (0.0725)		
	6 Aug.	0.4919 (0.0112)	**0.0503 (0.0192)** **		
	7 Sept.	0.3746 (0.0059)	**0.0313 (0.0099)** **		
	18 Oct.	0.2966 (0.0045)	**0.0795 (0.0073)*** **		
SBL	**20 Jun.**	0.6868 (0.0146)	0.0082 (0.0216)		
	16 Jul.	0.5819 (0.0113)	0.0215 (0.0174)		
	16 Aug.	0.5415 (0.0105)	0.0114 (0.0150)		
	20 Sept.	0.5222 (0.0089)	0.0145 (0.0130)		
Temporal autocorrelation general models (BC~pTime)					
Site		*BC*	*pTime*		
MSH		0.4086 (0.0032)	**1.0786 (0.0607)*** **		
SBL		0.5789 (0.0030)	**0.3012 (0.0617)*** **		

Table 3 - Variance in yield (*Y*) and growth rate (*r*) measured in 79 *Methylobacterium* isolates grown under 4 temperature treatments. *Y* and *r* values were transformed in log to meet normal distribution. For each factor following factors: clade (*C*), forest of origin (*F*), host tree species (*H*), time of sampling (*D*), temperature of incubation during pre-conditioning (T_P) and monitoring (T_M) steps, temperature of isolation (T_I) and their interactions, significance of *Y* and *r* responses are shown ("***": p<0.00l; "**": p<0.01; "*": p<0.05; see **Dataset S1o** for details).

	Rate	Yield
Forest (*F*)	0.015**	0.002
Host tree species (*H*)	0.022***	0.013***
Date of sampling (*D*)	0.054***	0.013***
Pre-conditioning temperature (T_P)	0.001	0.014***
Monitoring temperature (T_M)	0.158***	0.020***
Temperature of isolation (T_I)	0.010**	0.006*
Clade (*C*)	0.076***	0.306***
F:D	0.005	0.035***
H:D	0.003	0.019***
H:C	0.005	0.059***
D:C	0.029**	0.028***
F:C	0.011	0.014**
F:TI	0.003	0.021***
H:TI	0.000	0.021***
C:TI	0.024***	0.007
F:H:D	0.012**	0.023***
F:H:C	0.010**	0.002
F:D:C	0.001	0.008**
H:D:C	0.000	0.013***
H:D:TI	0.016***	0.019***
other interactions (sum)	0.177	0.079
Residuals	0.371	0.279

Supplemental material legends

Supplementary Materials and Methods S1 – Detailled materials and methods and supplementary references.

Figure S1 - ML phylogenetic trees from *sucA* (a) and *rpoB* (b) concatenated hypervariable (HV) regions and consensus clade tree (c).

Figure S2 - Experimental design of *Methylobacterium* monitoring for growth performance under four temperature treatments.

Figure S3 - Example of image analysis of *Methylobacterium* monitoring for growth performance under four temperature treatments.

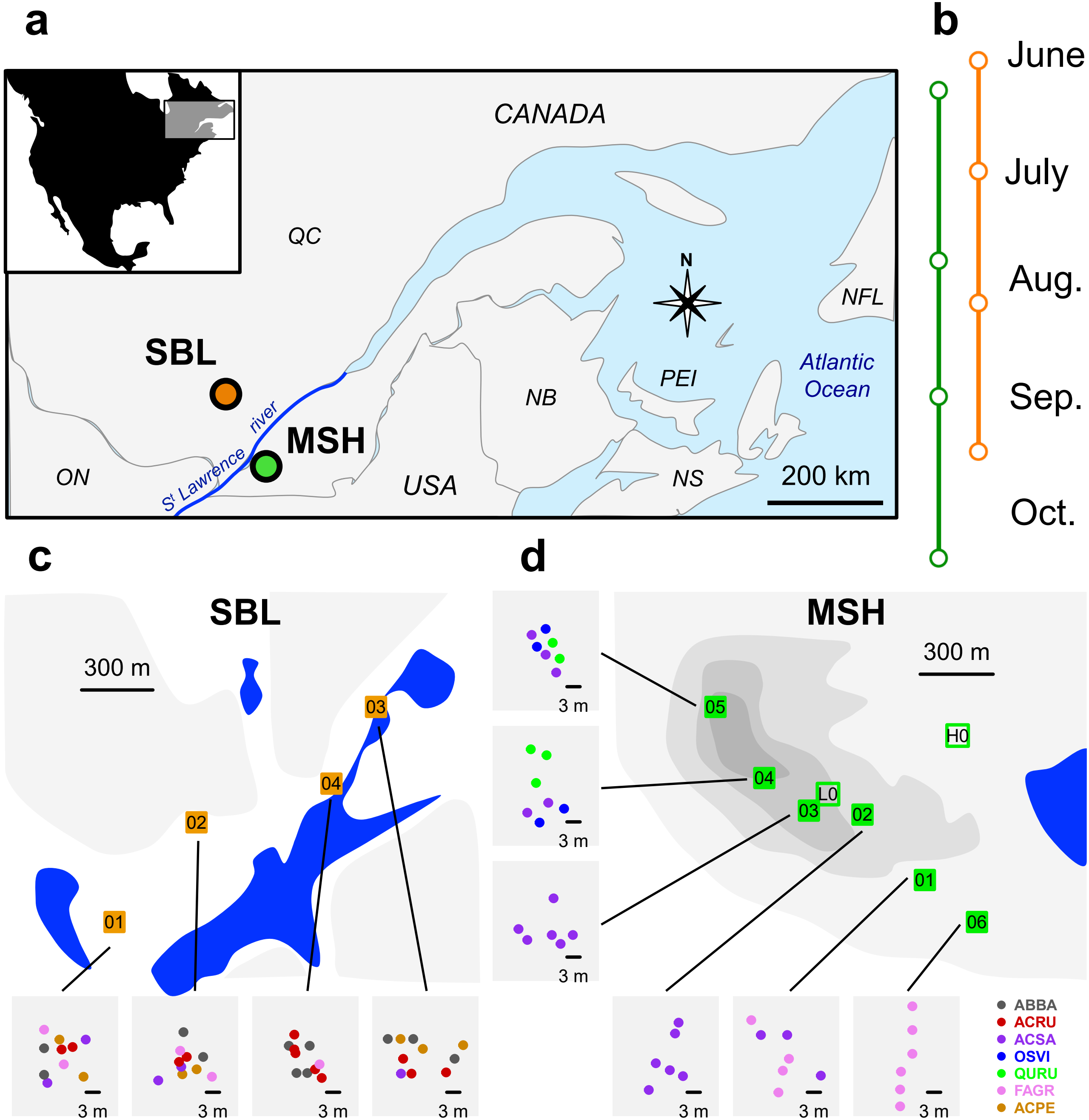
Figure S4 - Prediction of log normal growth curve, growth rate and yield for 79 isolates incubated under four temperature treatments.

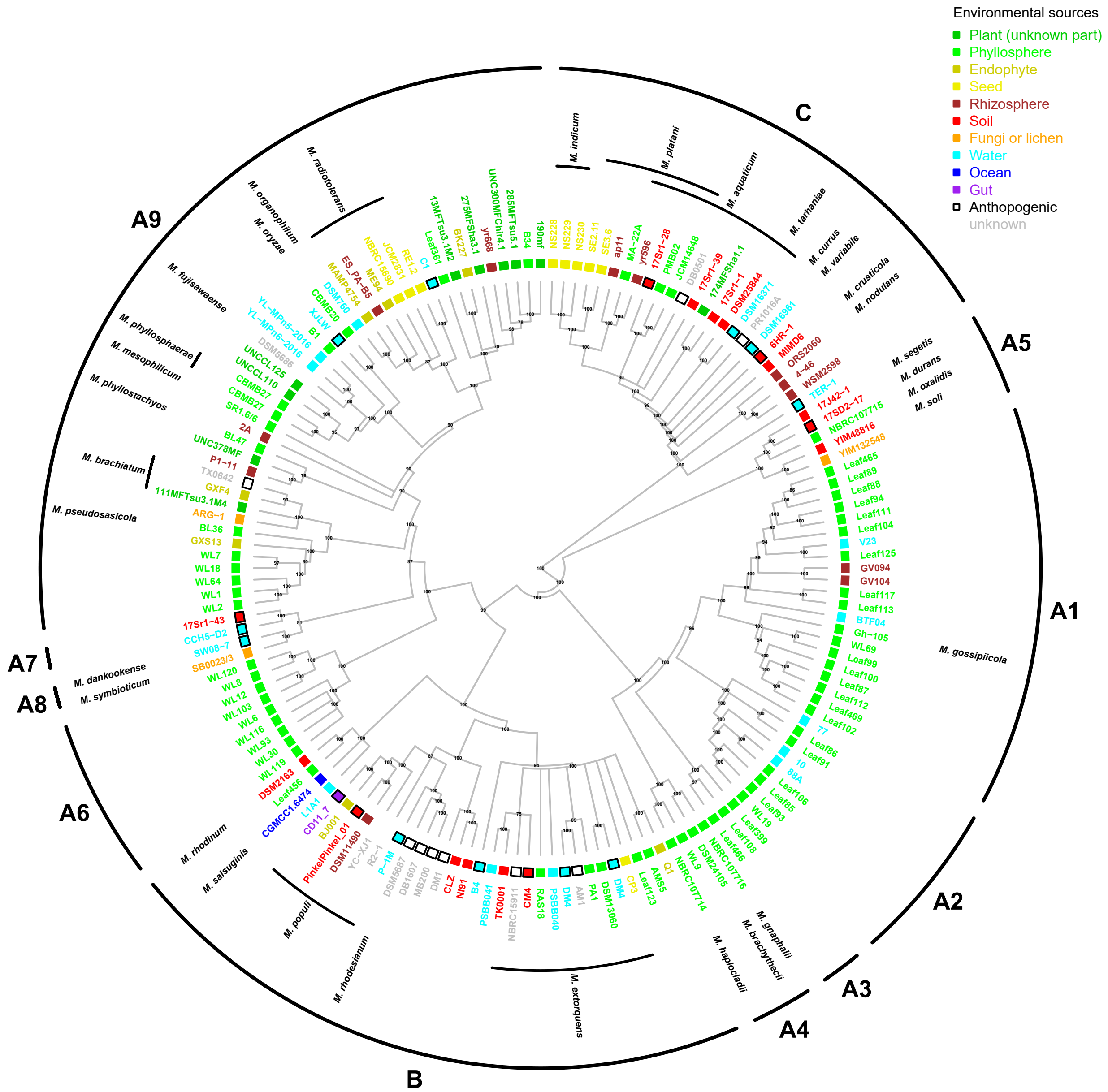
Figure S5 – *Alphaproteobacteria* (a) and *Methylobacterium* (b) diversity assessed by *rpoB* barcoding, comparison of *Methylobacterium* diversity assessement from *rpoB* barcoding and *16s rRNA* barcoding (c) and comparison of *Methylobacterium* diversity assessment from *rpoB* barcoding and isolation (d).

Table S1 - Primers used to amplify hyper variable regions in genes *sucA* and *rpoB* and sequence amplification success in 20 *Methylobacterium* isolates from a pilot survey in MSH in august 2017.

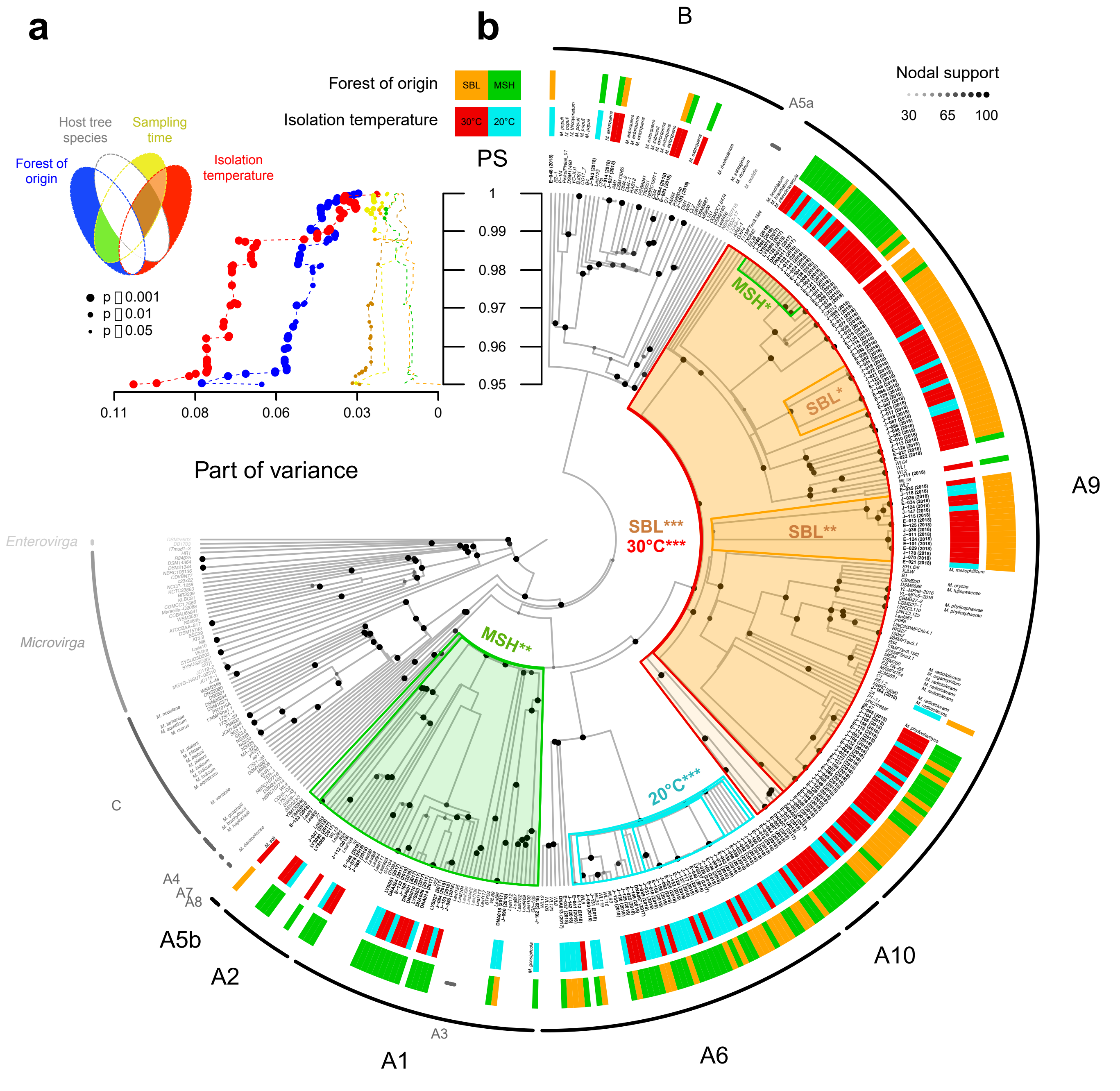
Table S2 - *Methylobacterium* diversity assessed by culture-dependant (isolates; *rpoB* sanger sequencing) and culture-free approaches (*16S rRNA* and *rpoB* barcoding) and comparison between different methods.

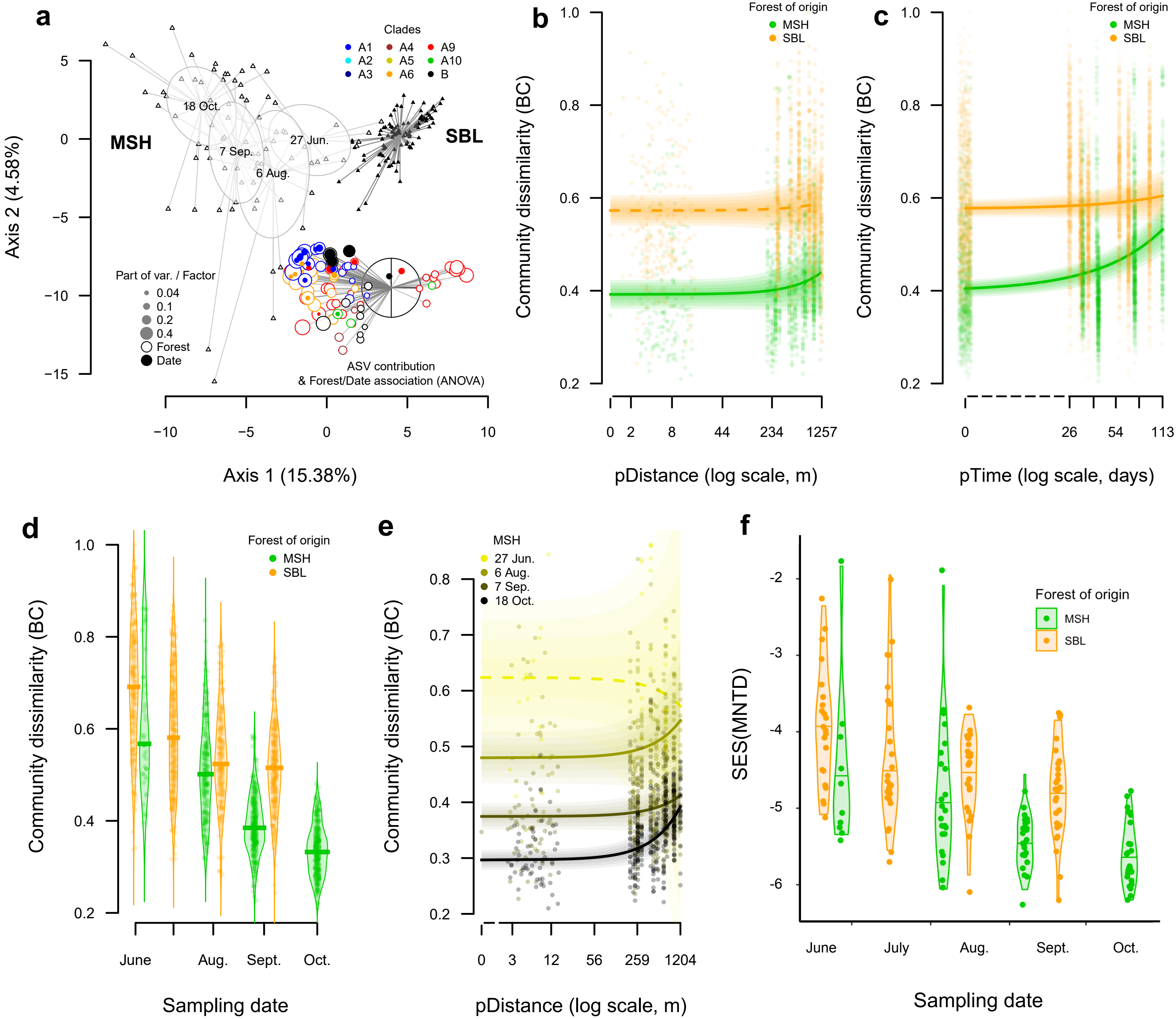
Dataset S1 – a) List of reference *Methylobacteriaceae* genomes used in this study ; b) List of phyllosphere samples and their deposited accession numbers ; c) List of 80 methylotrophic isolates from MSH (pilot survey in august 2017) ; d) Clade assignment of 76 *Methylobacterium* isolates from the 2017 pilot survey based on BLAST ; e) List of isolates and nucleotide sequence obtained from pilot and timeline surveys, and they deposited accession numbers ; f) List of 167 *Methylobacterium* isolates from timeline survey (2018) ; g) List of *16s rRNA* and *rpoB* barcoding libraries and they deposited accession numbers ; h) List of *16s rRNA* ASV nucleotide sequences, taxonomy and their absolute abundance in 46 phyllosphere samples ; i) Summary of *16s rRNA* ASV taxonomic assignation ; j) Tests for phylogenetic association of traits with culture-based estimation of *Methylobacterium* diversity for different phylogenetic depth ; k) List of *rpoB* ASV nucleotide sequences, taxonomy and their absolute abundance in 184 phyllosphere samples ; l) Summary of *rpoB* ASV taxonomic assignation ; m) Detail of ANOVA analysis for each *Methobacterium* ASV relative abundance ; n) Average rate and yield values for 79 *Methylobacterium* isolates monitored under four temperature treatments ; o) Detailled ANOVA results for yield and growth rate.


a
b
CANADA
QC
SBL
MSH
St Lawrence river
ON
USA
NB
PEI
NS
NFL
Atlantic Ocean
N
200 km
June
July
Aug.
Sep.
Oct.
c
d
SBL
300 m
01
02
03
04
3 m
MSH
300 m
05
04
03
02
01
06
H0
L0
ABBA
ACRU
ACSA
OSVI
QURU
FAGR
ACPE

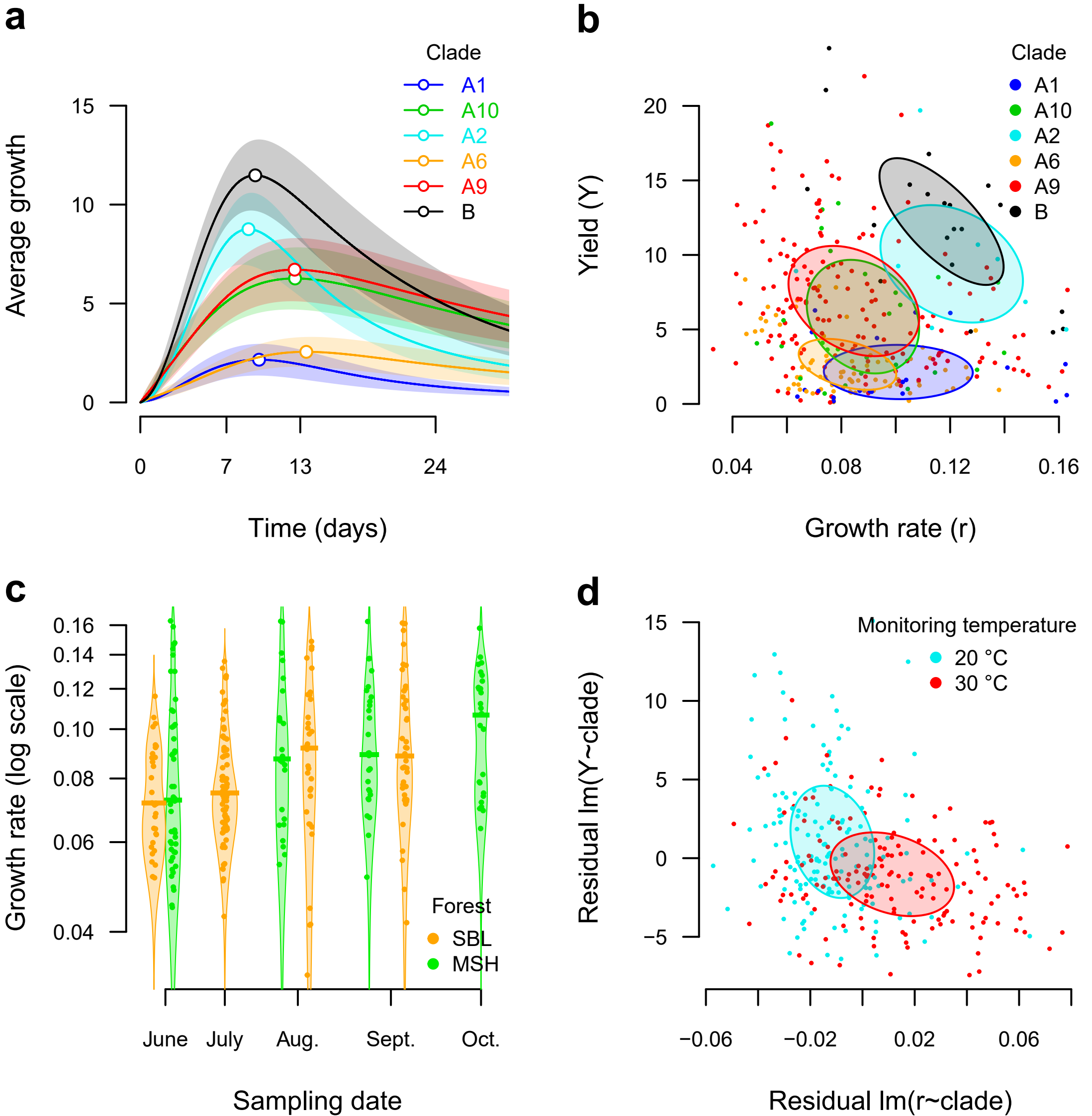


Environmental sources
Plant (unknown part)
Phyllosphere
Endophyte
Seed
Rhizosphere
Soil
Fungi or lichen
Water
Ocean
Gut
Anthopogenic
unknown
A1
A2
A3
A4
A5
A6
A7
A8
A9
B
C
M. indicum
M. platani
M. aquaticum
M. tarhaniae
M. currus
M. variabile
M. crusticola
M. nodulans
M. segetis
M. durans
M. oxalidis
M. soli
M. gossipiicola
M. gnaphalii
M. brachythecii
M. haplocladii
M. extorquens
M. rhodesianum
M. populi
M. salsuginis
M. rhodinum
M. dankookense
M. symbioticum
M. pseudosasicola
M. brachiatum
M. phyllostachyos
M. mesophilicum
M. phyllosphaerae
M. fujisawaense
M. oryzae
M. organophilum
M. radiotolerans
190mf
B34
285MFTsu5.1
UNC300MFChir4.1
yr668
275MFSha3.1
BK227
13MFTsu3.1M2
Leaf361
C1
RE1.2
JCM2831
NBRC15690
ME94
ES_PA-B5
MAMP4754
DSM760
CBMB20
XJLW
B1
YL-MPn5-2016
YL-MPn6-2016
DSM5686
UNCCL125
UNCCL110
CBMB27
CBMB27
SR1.6/6
2A
BL47
UNC378MF
P1-11
TX0642
GXF4
111MFTsu3.1M4
ARG-1
BL36
GXS13
WL7
WL18
WL64
WL1
WL2
17Sr1-43
CCH5-D2
SW08-7
SB0023/3
WL120
WL8
WL12
WL103
WL6
WL116
WL93
WL30
WL119
DSM2163
Leaf456
CGMCC1.6474
L1A1
CD11_7
BJ001
PinkelPinkel_01
DSM11490
YC-XJ1
R2-1
P-1M
DSM5687
DB1607
MB200
DM1
CLZ
NI91
B4
PSBB041
TK0001
NBRC15911
CM4
RAS18
PSBB040
DM4
AM1
PA1
DSM13060
DM4
CP3
Leaf123
Q1
AMS5
NBRC107714
WL9
DSM24105
NBRC107716
Leaf466
Leaf108
Leaf399
WL19
Leaf93
Leaf85
Leaf106
88A
10
Leaf91
Leaf86
77
Leaf102
Leaf469
Leaf112
Leaf87
Leaf100
Leaf99
WL69
Gh-105
BTF04
Leaf113
Leaf117
GV104
GV094
Leaf125
V23
Leaf104
Leaf111
Leaf94
Leaf88
Leaf89
Leaf465
YIM132548
YIM48816
NBRC107715
17SD2-17
17J42-1
TER-1
WSM2598
4-46
ORS2060
MIMD6
6HR-1
DSM16961
PR1016A
DSM16371
DSM25844
17Sr1-1
174MFSha1.1
17Sr1-39
DB0501
JCM14648
PMB02
17Sr1-28
yr596
MA-22A
ap11
SE3.6
SE2.11
NS230
NS229
NS228


a
b
Host tree species
Sampling time
Forest of origin
Isolation temperature
p ≤ 0.001
p ≤ 0.01
p ≤ 0.05
PS
1
0.99
0.98
0.97
0.96
0.95
0.11
0.08
0.06
0.03
0
Part of variance
Forest of origin
SBL
MSH
Isolation temperature
30°C
20°C
Nodal support
30
65
100
B
A5a
A9
A10
A6
A3
A1
A2
A5b
A4
A7
A8
C
Enterovirga
Microvirga
SBL***
30°C***
SBL**
SBL*
MSH*
MSH**
20°C***



a
Clades
A1
A2
A3
A4
A5
A6
A9
A10
B
MSH
SBL
18 Oct.
7 Sep.
6 Aug.
27 Jun.
Part of var. / Factor
0.04
0.1
0.2
0.4
Forest
Date
ASV contribution
& Forest/Date association (ANOVA)
Axis 1 (15.38%)
Axis 2 (4.58%)
b
Forest of origin
MSH
SBL
Community dissimilarity (BC)
pDistance (log scale, m)
0 2 8 44 234 1257
c
Forest of origin
MSH
SBL
Community dissimilarity (BC)
pTime (log scale, days)
0 26 54 113
d
Forest of origin
MSH
SBL
Community dissimilarity (BC)
June
Aug.
Sept.
Oct.
Sampling date
e
MSH
27 Jun.
6 Aug.
7 Sep.
18 Oct.
Community dissimilarity (BC)
pDistance (log scale, m)
0 3 12 56 259 1204
f
Forest of origin
MSH
SBL
SES(MNTD)
June
July
Aug.
Sept.
Oct.
Sampling date


a
Clade
A1
A10
A2
A6
A9
B
Average growth
15
10
5
0
0
7
13
24
Time (days)
b
Clade
A1
A10
A2
A6
A9
B
Yield (Y)
20
15
10
5
0
0.04
0.08
0.12
0.16
Growth rate (r)
c
Growth rate (log scale)
0.16
0.14
0.12
0.10
0.08
0.06
0.04
Forest
SBL
MSH
June
July
Aug.
Sept.
Oct.
Sampling date
d
Monitoring temperature
20 °C
30 °C
Residual lm(Y~clade)
15
10
5
0
−5
−0.06
−0.02
0.02
0.06
Residual lm(r~clade)

SUPPLEMENTARY MATERIALS AND METHODS

FINE-SCALE ADAPTATIONS TO ENVIRONMENTAL VARIATION AND GROWTH STRATEGIES DRIVE PHYLLOSPHERE *METHYLOBACTERIUM* DIVERSITY.

Jean-Baptiste Leducq[1,2,3], Émilie Seyer-Lamontagne[1], Domitille Condrain-Morel[2], Geneviève Bourret[2], David Sneddon[3], James A. Foster[3], Christopher J. Marx[3], Jack M. Sullivan[3], B. Jesse Shapiro[1,4] & Steven W. Kembel[2]

1 - Université de Montréal
2 - Université du Québec à Montréal
3 – University of Idaho
4 – McGill University

Phylogenetics of plant-associated *Methylobacterium* diversity

We aimed to assess the known diversity of *Methylobacterium* and its distribution across biomes and especially the phyllosphere. First, we constructed a phylogeny of *Methylobacteriaceae* from the complete sequence of *rpoB*, a highly polymorphic housekeeping gene commonly used to reconstruct robust phylogenies in bacteria, because unlikely to experience horizontal gene transfer or copy number variation (1, 2). We retrieved this gene from all complete and draft *Methylobacteriaceae* genomes publicly available in September 2020, including 153 *Methylobacteria*, 30 *Microvirga* and 2 *Enterovirga* (**Dataset S1a**), using blast of the *rpoB* complete sequence from the *M. extorquens* strain TK001 against NCBI databases *refseq_genomes* and *refseq_rna* (3) available for *Methylobacteriaceae* (Uncultured/environmental samples excluded). Nucleotide sequences were converted in amino-acid in MEGA7 (4, 5), aligned according to the protein sequence, an converted back in nucleotides. The *Methylobacteriaceae* phylogenetic tree was inferred from the *rpoB* nucleotide sequence alignment (4064 bp). Nodal support values (Bayesian posterior probabilities) were estimated using MrBayes v. 3.2.7a (6). Bayesian analyses consisted of paired independent runs, each using four Metropolis coupled chains that consisted of 5 million generations, after which standard deviation of split frequencies had stabilized to less than 0.03. For each of the paired runs, trees were sampled every 1000 generations and the first 1 million generations were treated as the burn-in and discarded. The remaining trees from the two runs (*n*=1802) were combined to determine split frequencies (nodal posterior probabilities). The consensus tree and nodal support values were determined in PAUP v. 4. To improve the presentation of the tree, branch lengths were computed in R, using the Grafen method (7) (function *compute.brlen* in package *ape* ; **Figure 2**). For each *Methylobacterium* reference strain, we retrieved the species name and the sampling origin, when available. Additionally, we assigned each strain to a group (A, B, C) according to previously proposed subdivisions (8). Because group A consisted in several paraphyletic groups branching deeply in the *rpoB* phylogeny, we subdivided *Methylobacterium* group A into 9 clades (A1-A9), using a ~92% pairwise similarity (PS) cut-off on the *rpoB* complete sequence. PS was calculated in MEGA7 from nucleotide sequences as PS = *1-pdistance*.

Study sites and phyllosphere sampling

Study forests

The two study forests were located in Gault Nature Reserve (Mont Saint-Hilaire, Quebec, Canada ; 45.54 N 73.16 W), here referred as MSH, an old forest occupying the hill of Mount Saint-Hilaire, and *Station Biologique des Laurentides* (Saint-Hippolyte, Quebec, Canada ; 45.99 N 73.99 W), here referred as SBL, a mosaic of natural wetlands, xeric and mesic forests (**Figure 1; Dataset S1b**).

In august 2017, we realized a pilot survey in MSH. We choose two plots (MSH-L0 and MSH-H0) with similar tree species composition and different elevations (175 and 220 m, respectively). Forests were dominated by merican beech (*Fagus grandifolia*), sugar maple (*Acer saccharum*), striped maple (*Acer pensylvanicum*), birch (*Betula alleghaniensis*) and northern red oak (*Quercus rubra*). We collected leaves from 18 randomly chosen trees among dominant species (9 trees per forest), for which we were able to sample the lower part of the canopy (3-5m), hence excluding *B. alleghaniensis* and *Q. rubra*. Additionally, we sampled one American hophornbeam (*Ostrya virginiana*) in MSH-H0. For each tree (n=19), sampling was replicated tree times in different parts of the subcanopy (3-5m), whenever possible.

In 2018, we realized a timeline survey in MSH and SBL. In MSH, we choose 6 plots distributed along a 1.2 km ecological and altitudinal transect (MSH1-6). Lower plots MSH1,6 (170-190 m) were dominated by *F. grandifolia* and *A. saccharum*. Medium plots MSH2,3 (225-270 m) were dominated by *B. alleghaniensis*, *Q. rubra* and *A. saccharum*. Higher plots MSH4,5 (290-315 m) were dominated by *A. saccharum*, *Q. rubra* and *O. virginiana*. In SBL, we choose 4 plots (SBL1-4) distributed along a 1.2 km transect in a transition zone dominated by *A. saccharum*, *F. grandifolia*, *A. pensylvanicum,* balsam fir (*Abies balsamea*), red maple (*Acer rubrum*), and paper birch (*Betula papyrifera*). During this aurvey, we tagged 80 trees (40 per forest, 6-10 per plots) from the dominant species for which we were able to sample the subcanopy, hence excluding *Betula ssp* from any plot and *Q. rubra* from MSH2,3. Each tree was sampled 3-4 times from June to October 2018. For each time point and each tree, sampling was replicated twice in different parts of the subcanopy, whenever possible.

Sample collection

Sampling consisted of collecting about 2-10g of leaves into Fisherbrand® sterile bags with a pole pruner while wearing nitrile gloves. For each plot, we realized a negative controls consisting in empty sterile bags opened and sealed on site. Samples were sealed and conserved up to 48h at 4°C until further processing. Samples were randomized before processing. Each bag was unsealed under a sterile hood and 50ml of sterile phosphate buffer was added (KH_2PO_4 100mM solution poured into K_2HPO_4 100mM solution until pH 7.3). Bags were sealed again and vigorously agitated for 5 minutes. For each bag, about 45 ml of phosphate buffer containing the microbial community was transferred in sterile 50 ml Falcon® tubes and centrifuged for 30 minutes at 3,900 rpm at 4°C. Supernatant was removed in order to left the pellet in less than 1 ml of phosphate buffer. The pellet was resuspended and split in two equal volumes in two sterile 1.5 ml tubes. The first tube was directly stored at -80 °C for future DNA extraction, metagenomics and community analysis. The second tube was completed with 500 µl of 50% glycerol (minimum final glycerol concentration: 25%) for future isolations and stored at -80 °C.

Samples were randomized before DNA extraction. DNA extraction was performed using a DNeasy PowerSoil Pro Kit (Qiagen) with the following modifications: each sample (~500 µl) was thawed on ice and centrifuged in order to resuspend the pellet containing the microbial community in a maximum 200 µl volume before proceeding to extraction. DNA was eluted in 50µl of solution C6 and stored at -20 °C. DNA was quantified using Qubit™ dsDNA HS Assay Kit and was detectable in all samples but negative controls with an average of 1.3 ng/µl.

Bacterial community analysis of the timeline survey by *16S rRNA* barcoding

Samples

We evaluated the bacterial phyllosphere diversity through Illumina sequencing of the *16S rRNA* ribosomal gene using primers 799F-1115R targeting the V5-V6 region and excluding chloroplastic DNA (9). *16S rRNA* gene amplicon sequencing was performed on 46 phyllosphere samples from 13 trees from both forests sampled 3-4 times throughout the 2018 growth season (**Dataset S1b**). We included one negative control (see section Sample collection) and one positive control (METH community) consisting in mixed genomic DNAs from 18 *Methylobacterium* isolates representative of diversity in SBL and MSH (**Dataset S1c,f**), one

Escherichia coli strain and one *Sphingomonas sp.* isolate from MSH (isolate DNA022; **Dataset S1c**).

Library preparation and sequencing

PCRs contained 1µL of sample DNA, 5 µL of Phusion Hot Start II Buffer (thermoscientific Fisher®), 0.5 µL of dNTP mix 10 µM, 0.5 µL of each primer at 10µM, 0.75µL of DMSO and 0.25 µL of Phusion Hot Start II Buffer DNA Polymerase (thermoscientific Fisher®) for a final volume of 25µL. PCRs were carried out in a thermocycler MasterCycler ProS Eppendorf© with the following steps: 30″ at 98 °C; then, 35 cycles of 15″ at 98 °C, 30″ at 64 °C and 30″ at 72 °C; and a final extension of 10′ at 72 °C. PCR products were loaded on 1% agarose gel ran at 120V for 45′ to control for amplification and fragment size with a 100 bp ladder (Invitrogen™). For each sample and the ladder, 1 µL of PCR product was mixed with 5 µL of dye buffer EZ-VISION® TREE (VWR™).

Samples were randomized before PCR amplification (random attribution of barcodes). PCR amplification (~400bp) was obtained for all samples and positive control but not negative controls. The *16S rRNA* amplicon library was prepared according to QIAseq FX DNA Library Kit (QIAGEN) protocole. The library was controlled for DNA concentration (Qbit™), quality (qPCR, NEBNext® Library Quant Kit for Illumina) and size distribution (Bioanalyzer DNA High Sensitivity, Agilent). Library concentration was adjusted according to the qPCR value at 6pM and sequenced with 1% phiX on Miseq (Illumina) with Miseq reagent kit v3 600-cycles (Paired-end 300 pb).

Amplicon Sequence Variant (ASV) definition with dada2

We obtained an average of 32,159 (range 12,162-55,181) paired reads per phyllosphere sample, 33,528 for the METH community (positive control) and 1,545 for the negative control. We processed *16S rRNA* reads in order to obtain an ASV (Amplicon Sequence Variants) abundance table per sample, using package *dada2* in R (10) with following modifications. According to sequence quality profiles, 3' ends of forward and reverse reads from each sequence were trimmed 50 and 100bp, respectively (option truncLen in function *filterAndTrim*). Additionally, 5' ends of both reads were trimmed 20bp in order to remove the primer part (option trimLeft in function *filterAndTrim*). After trimming, reads with expected error higher than 2 (option maxEE in function *filterAndTrim*) were discarded. Sample inference was performed using pseudo-pooling of samples (option pool=pseudo in function *dada*). Forward and reverse reads were merged together (function *mergePairs*) and chimeras were removed (function *removeBimeraDenovo*). After quality filtering (27% of sequences discarded), merging (13% discarded) and chimera removing (14% discarded), we conserved an average of 14,871 (range 8,096-21,529) sequences per phyllosphere sample, 16,367 for the METH community and 1,077 for the negative control, for a total of 24,733 unique variants (ASVs).

Contaminant ASVs filtering and rarefaction

We assigned taxonomy of each ASV identified in negative control and METH community using blast against NCBI nucleotide collection (nr/nt), uncultured/environmental samples excluded. In negative control (22 ASVs), 99.6% of diversity (proportion of sequences) corresponded to taxa typically associated with human oral and skin microbiomes (*Saccharibacteria*, *Leptotrichia, Corynebacterium*, *Actinomyces*, *Cutibacterium*, *Pseudoglutamicibacter*, *Prevotella* and *Kocuria*). In METH community (47 ASVs), 61.6% of diversity (proportion of sequences) corresponded to genomic DNA from bacterial orders mixed in the community: *Sphingomonadales* (1 ASV; 5.4% of diversity), *Enterobacterales* (*Escherichia coli*; 4 ASVs; 18.5% of diversity) and *Rhizobiales* (*Methylobacterium ssp.*; 18 ASVs; 36.4% of diversity). Although the remaining diversity was dominated by *Thermotogales* (*Fervidobacterium sp.*, 36.4% of sequences, 3 ASVs), these ASVs were absent from the negative control and phyllosphere samples, suggesting that these contaminants only affected the METH community. The remaining diversity corresponded to ASVs abundant in phyllosphere samples and thus likely corresponded to cross contamination among samples before the library preparation step. Each of these contaminant ASVs (*Thermotogales* excluded) had less than 0.3% of relative abundance in the METH community. We thus used this value as a conservative threshold to remove likely contaminant ASVs. In clear, any ASV that had a maximum relative abundance calculated across samples (negative control excluded) lower than 0.3% was discarded. We performed rarefaction curves on each sample to determine a conservative number of sequences to conserve per sample. Accordingly, the randomly picked 5,000 sequences per sample, hence excluding the negative control. After rarefaction, we obtained 732 ASVs, 725 of which were present in phyllosphere samples (42-292 per sample) and 24 in the METH community.

Taxonomic assignation of ASVs using SILVA

We aimed to assign taxonomy of each of the 732 identified ASVs (725 in phyllosphere samples), with emphasize on *Methylobacterium*, hence limiting taxonomic assignation at the genus level within *Methylobacteriaceae*, at the family level within *Rhizobiales*, at the order level within *Alphaproteobacteria*, at the class level within *Proteobacteria*, and at the phylum level within *Bacteria*. We used SILVA v.138 (11) as a database with *assignTaxonomy* function in R package *dada2* (**Dataset S1h**).

Among phyllosphere samples, 100% of diversity (proportion of sequences) was assigned to *Bacteria* (725 ASVs; **Dataset S1i**). The phyllosphere bacterial community was dominated in both forests by *Actinobacteria* (36.3% of sequences, 232ASVs), *Bacteroidota* (22.1%, 98 ASVs), *Deinococcota* (5.7%, 28 ASVs) and *Proteobacteria* (29.8%, 249 ASVs). *Proteobacteria* consisted in *Gammaproteobacteria* (9.1%, 87 ASVs) and *Alphaproteobacteria* (20.7%, 162 ASVs). In *Alphaproteobacteria*, more than 75% of diversity was found within the order *Rhizobiales* (13.1%, 72 ASVs), dominated by the family *Beijerinckiaceae* (13.0%; 67 ASVs). *Methylobacteriaceae* did not exist as a separate family in the SILVA database and *Methylobacterium* (annotated *Methylobacterium-Methylorubrum* (8)) was embedded within *Beijerinckiaceae* (15 ASVs, 1.3%). *Beijerinckiaceae* contained two "true" *Beijerinckiaceae* genera: *Methylocella* (4 ASVs, 0.5%), *Methylorosula* (3 ASVs, 1.5%), two incorrectly classified genera: *Psychroglaciecola* (*Methylobacteriaceae*; (12)), *Roseiarcus* (*Roseiarcaceae*; (13)) and one unknown genera annotated as *1174-901-12* but representing 9.5% of total diversity (38 ASVs). We blasted sequences from the two most abundant ASVs annotated "*1174-901-12* " by SILVA against NCBI nucleotide collections (nr/nt, uncultured/environmental samples excluded) and obtained 100% identity with *Lichenibacterium minor* and *Lichenibacterium ramalinae,* respectively. These species were recently isolated from lichens colonizing birch trunks in boreal forests and represent the unique members of the newly described *Rhizobiales* family *Lichenibacteriaceae* (14). Finally, 5 ASVs (0.2% of diversity) were not assigned at the genus level.

Methylobacterium was present in almost all analyzed samples (45 out of 46), representing 1.3% [0.0-3.2%] of total sequence abundance. Using blast against NCBI databases (3) *refseq_genomes* and *refseq_rna* available for *Methylobacteriaceae* (Uncultured/environmental samples excluded), we determined that the 15 *Methylobacterium* ASVs identified by *16S rRNA* sequencing mostly belonged to clades typical of the phyllosphere: A9 (*M. phyllosphaerae/M. mesophilicum/M. phyllostachyos/ M. pseudosasicola/M. organophilum;* 0.87% of sequences, nine ASVs), A6 (0.29%; one ASV) and A1 (*M. gossipicola*; 0.13%, 3 ASVs; **Dataset S1i**). No ASV was assigned to group B or group C. We defined a new clade (A10) with two rare ASVs assigned to *M. komagatae* (<0.01%) but unrelated to any aforementioned clade.

16S rRNA community analyzes

We evaluated factors shaping phyllosphere microbial diversity estimated from *16S rRNA* barcoding, namely: forest of origin, host tree species and time of sampling. Relative abundances were normalized by Hellinger transformation to account for rare taxa, using function *decostand* in R package *vegan* (15). We performed a PERMANOVA to evaluate the contribution of each factor to microbial diversity (Hellinger transformation) using function *adonis* from R package *vegan* (**Table 1**; 10,000 permutations). As observed for PCA, forest of origin explained most of the variation (31.6%; $p<0.001$), followed by host tree species (15.6%; $p<0.001$) and sampling time (11.9%; $p<0.05$).

Methylobacterium isolation from a pilot survey in MSH in august 2017

Isolation

We performed *Methylobacterium* isolation from samples collected during the pilot survey in 2017 in MSH (57 samples from 19 trees; plots H0 and L0; 9-10 trees per plots; **Dataset S1b**). For each sample, we spread 10µL of phyllosphere microbial community glycerol stock (see section Sample collection) in a petri dish containing MMS synthetic solid media with 0.1% methanol sterilized by filtration (0.22µm filters), as sole carbon source to select for *Methylobacterium*, and 50mg/L of Cycloheximide to reduce fungal contamination (16). Petri dishes were incubated two weeks at 20 °C and 30 °C. Both temperatures were tested for all samples in order to minimize biases toward mesophylic isolates. From each petri dish, 0 to 3 pink colonies were isolated to maximize representativeness in term of color and size, spread separately on new MMS synthetic solid media (0.1% methanol, 50mg/L Cycloheximide) enriched with Sigma® RPMI1640 vitamins solution and 0.05g/L yeast extract to boost cell growth (17) and incubated 2-4 weeks at 30 °C or 20 °C according to the temperature of isolation. After incubation, colonies from petri dishes with contamination or with at least two different types of colonies (based on color) were spread on new MMS media whenever possible. Clean petri dishes were swabbed with 2mL of sterile distilled water. Colonies were collected in approximately 1mL of liquid and centrifuged in 1.5 mL Eppendorf tube, 10 minutes at 4°C (21130 rcf). Pictures of pellets were taken and tubes with evidence of contamination (two different colors in the pellet) were discarded. Clean pellets were resuspended; 450 µL were directly stored at -80 °C for future DNA extraction; 50µL were lysed by 10 minutes of incubation at 98°C, centrifuged 10 minutes at 4°C (21130 rcf) and stored at -20 °C for future isolate identification based upon amplification by PCR of marker genes, and 500µL were mixed with 500µL of glycerol 50% for culture stock (final glycerol concentration: 25%, storage at -80 °C).

16S rRNA V4 region amplification

We identified isolates from the pilot survey by PCR amplification and sequencing of the V4 region from *16S rRNA* ribosomal gene using primers 515F GTGCCAGCMGCCGCGGTAA (18) and 786R GGACTACHVGGGTWTCTAAT (19), universal for bacteria. For isolates DNA001-DNA024, identification was performed from genomic DNA extracted using DNeasy Blood & Tissue Kit (QIAGEN). For isolates LYS001-LYS096, identification was directly performed from cell lysate (see above). PCRs contained 1µL of cell lysates of genomic DNA, 5 µL of Phusion Hot Start II Buffer (thermoscientific Fisher®), 0.5 µL of dNTP mix 10 µM, 1 µL of each primer at 3µM, 0.75µL of DMSO and 0.25 µL of Phusion Hot Start II Buffer DNA Polymerase (thermoscientific Fisher®) for a final volume of 25µL. PCRs were carried out in a thermocycler MasterCycler ProS Eppendorf© with the following steps: 3′ at 98 °C; then, 35 cycles of 45″ at 98 °C, 1′ at 50 °C and 1′30″ at 72 °C; and a final extension of 10′ at 72 °C. PCR products were then sequenced by Sanger sequencing.

Methylobacterium isolate identification

We obtained *16S rRNA* nucleotide sequences from the V4 region for 80 pink colonies isolated for 18 out of 19 sampled trees (**Dataset S1c**). Sequences were manually curated according to the original chromatogram and classified using blast of the V4-*16S rRNA* sequence against NCBI databases (3) *refseq_genomes* and *refseq_rna* available for *Methylobacteriaceae* (Uncultured/environmental samples excluded). Four isolates did not closely match any *Methylobacteriaceae* reference and actually corresponded to *Stenotrophomonas sp*. (n=1), *Deinococcus sp*. (n=1) and *Sphingomonas sp.* (n=2). We assigned the 76 remaining isolates to previously identified *Methylobacterium* clades. We identified 8 unique V4-*16S rRNA* sequence variants. One variant present in 11 isolates was assigned to clade A9 (at least 99% of similarity with *M. phyllosphaerae/M. mesophilicum/M. phyllostachyos/ M. pseudosasicola/M. organophilum*); two variants in 19 isolates to clade A6 (*M. sp.*; 100% of similarity) and two variants in 12 isolates to clade A10 (*M. komagatae*; 100% of similarity). The three remaining variants identified in 34 isolates had 100% similarity with either clade A1 (*M. gossipicola*), A2 (*M. sp.*) or A3 (*M. sp.*). Because of the close V4-*16S rRNA* similarity (>99%) among some representative sequences of different clades, we affined the assignation of 24 representative isolates using sequencing and phylogenies of partial nucleotide sequences of two candidate marker genes: *sucA* and *rpoB* (Details in section Development of a *Methylobacterium*-specific molecular marker). We confirmed previous assignations to clades A6, A9 and A10 and distinguished isolates from clades A1 (n=30) and A2 (n=4; **Dataset S1c**).

Development of a *Methylobacterium*-specific molecular marker

Development of a fine-scale single-copy molecular marker specific to *Methylobacterium*

As an alternative to the *16S rRNA* gene, we developed a highly polymorphic marker targeting specifically – but not exclusively – isolates from the *Methylobacteriaceae* family. We choose two candidate genes, *rpoB* and *sucA*. Gene *sucA* is part of the 2-oxoglutarate dehydrogenase (OGDH) complex which catalyzes the decarboxylation of 2-oxoglutarate (20). It has been used as a marker gene to reconstruct phylogenies in *Salmonella* (21) and *Enterobacteriaceae* (22). Gene *rpoB* encodes the beta subunit of RNA polymerase and is widely used to reconstruct phylogeny in bacteria (1, 2, 22, 23). We retrieved complete *rpoB* and *sucA* nucleotide sequences available for

153 *Methylobacterium* genomes and 32 *Methylobacteriaceae* outgroups (*Microvirga*, *Enterovirga*; see **Dataset S1a**) and confirmed that both genes were single-copy in all genomes, contrary to *16S rRNA*. We performed alignment based on the amino-acid sequence in MEGA7 with default parameters (5). Based on the alignment, we identified five hypervariable (HV) regions (three in *sucA*, two in *rpoB*) flanked by well-conserved regions across *Methylobacteriaceae*, for which we designed specific primers (**Table S1**). We tested each primer pair targeting a HV region on 20 representative *Methylobacterium* isolates from the pilot survey. PCRs contained 1µL of cell lysates of genomic DNA, 5 µL of Phusion Hot Start II Buffer (thermoscientific Fisher®), 0.5 µL of dNTP mix 10 µM, 1 µL of each primer at 3µM, 0.75µL of DMSO and 0.25 µL of Phusion Hot Start II Buffer DNA Polymerase (thermoscientific Fisher®) for a final volume of 25µL. PCRs were carried out in a thermocycler MasterCycler ProS Eppendorf© with the following steps: 3′ at 98 °C; then, 35 cycles of 45″ at 98 °C, 30″ at 60 °C and 1′30″ at 72 °C; and a final extension of 10′ at 72 °C. PCR products were then sequenced by Sanger sequencing.

We successfully amplified and obtained clean sequenced for both *rpoB* HV regions in the 20 tested isolates (**Table S1; Dataset S1c,d**). We choose the first HV region targeted by primers Met02-352-F (AAGGACATCAAGGAGCAGGA) and Met02-1121-R (ACSCGGTAKATGTCGAACAG) as specific marker for *Methylobacteriaceae* for the rest of this study.

A consensus *Methylobacterium* phylogeny by coupling *sucA* and *rpoB* phylogenies

In order to validate *Methylobacterium* clade definition, we performed ML phylogenetic trees (100 permutations, complete deletion) for *rpoB* and *sucA* partial nucleotide sequences, separately. For *sucA*, we concatenated the three HV regions (1,663 bp) available for 189 reference genomes and the 14 tested isolates for which we obtained sequences for all of the three HV regions (**Figure S1a**). For *rpoB*, we concatenated the two HV regions (1,244 bp) available for 163 reference genomes and the 20 tested isolates (**Figure S1b**). We observed a strong congruence between both phylogeny topologies (summarized in **Figure S1c**). All defined clades were monophyletic in the *sucA* and *rpoB* phylogenies and supported by more than 80% of bootstraps in both for most clades but A1 (49 and 79%, respectively), B (99 and 68%, respectively) and A5 (89 and 45%, respectively). The consensus clade tree shows that group C is the more basal group of *Methylobacterium*. Among the remaining clades, we distinguished three groups of sister clades, among and within which phylogenetic relationships remain mostly unsolved: A1/A2/A3, A4/B and A5/(A6/((A7/A8)/(A9/A10))).

Culture-based assessment of *Methylobacterium* diversity of the timeline survey

Isolate isolation and identification

We monitored temporal and host-associated trends in *Methylobacterium* diversity in the phyllosphere by performing isolation on 28 trees sampled in 2018 in MSH and SBL (Timeline survey; **Dataset S1f**). Isolation was performed at 20 and 30 °C on MMS media with methanol as sole carbon source, as described in section Isolation**.** In a first batch, isolation was performed on all samples collected from 8 trees (4 per forest; 2 *Acer saccharum* and 2 *Fagus grandifolia*) between June and October 2018 (3-4 time replicates per tree). From this survey, we obtained 98 pink isolates (36 in MSH, 62 in SBL), for which we were able to amplify the *rpoB* marker using primers Met02-352-F and Met02-1121-R and get readable nucleotide sequences **(**details in section Development of a *Methylobacterium*-specific molecular marker). We obtained the highest average isolation success per sample in date 1 for MSH (27 June, n=5.0) and in date 2 for SBL (16 July, n=5.8). We thus selected these dates for a second isolation batch focusing on host-associated diversity. In each forest, we selected 10 trees sampled at the aforementioned dates and representative of diversity found in each forest. We repeated isolation and identification as described above and obtained 69 isolates (37 in MSH, 32 in SBL). Combining both batchs, we obtained 167 *Methylobacterium* isolates, 56.3% of which came from SBL, 43.7% from MSH; and 32.9% of which were isolated at 20 °C, 67.1% at 30 °C.

Isolate assignation to *Methylobacterium* clades

Among the 167 *Methylobacterium* isolates from the timeline survey, we identified 71 unique sequence variants, which is almost a 10-fold increase in comparison with V4-*16S rRNA* (see **section** *Methylobacterium* isolation from a pilot survey in MSH in august 2017). We assigned the 167 isolates to *Methylobacterium* clades (**Dataset S1f; Table S2)** using a phylogenetic tree (**Figure 3b**) inferred from an alignment combining partial nucleotide sequence for the *rpoB* marker sequenced for all isolates (first *rpoB* HV region), *rpoB* complete nucleotide sequences available for 185 *Methylobacteriaceae* complete genomes and partial nucleotide sequence for 2 *rpoB* HV regions we previously obtained for 20 representative *Methylobacterium* isolates from the pilot survey (1,244bp; detail in section Development of a *Methylobacterium*-specific molecular marker). Alignment was manually curated in MEGA7 (5), using the complete amino acid sequence as a guide, and sites not present in >70% of the sequences were removed using phyutility (v2.6). Nodal branch supports in the phylogenetic tree were estimated using MrBayes v. 3.2.7a as described in section Phylogenetics of plant-associated *Methylobacterium* diversity with the following modifications: standard deviation of split frequencies had stabilized to less than 0.05. Nodes with less than 30% of posterior probability were collapsed. the 167 isolates were assigned to clades within which their were embedded, using reference genomes and 20 isolates from the 2017 pilot survey as references. For these references, most clades were monophyletic, but clade A1 that formed a monophyletic group with A3, and clade A5 that splitted in two monophyletic groups (A5a, A5b). Isolates were assigned to A1 (n=9; 5.4%), A2 (n=3; 1.8%), A6 (n=41; 24.6%), A9 (n=100; 59.9%), A10 (n=6; 3.6%), A5b (n=1; 0.6%) and B (n=7; 4.2%).

Visual scaling of the *rpoB* phylogenetic tree according to pairwise nucleotide similarity.

We aimed to assess *Methylobacterium* isolate diversity a different depths within the *rpoB* phylogenies, with more emphases on the tips of the tree. We normalized the phylogenetic tree (**Figure 3b)** so it was scaled proportionally to nucleotide pairwise similarity (*PS*;). In other terms, we aimed to find a visual consensus between tree topology and evolutionary rate (here assumed to be proportional with PS). First, we imported the phylogenetic tree in Newick format in R using function *read.tree* in package *ape* (24). We converted the *read.file* object in a matrix filed with node names, with isolates in rows and nodes in columns. Nodes were ordered from the tip to the root of the tree, and then stacked to the root, so that the last column corresponded to the root of the tree, the before last column contained the closest embedded node(s), and so on. Each column of the matrix was labeled with a level number, *L*=1 corresponding to tips of the tree (sequences). For each node, we also calculated the median PS value (*PS* = 1-*pdistance*) among embedded sequences, using a *pdistance* matrix calculated for all possible pairwise sequences in MEGA7 (complete deletion). *PS* values were transformed in inverse log scale (*PSc*= -log(1.005-*PS*)) to optimize resolution at the tips of the tree. Hence, each node was associated with a *L* and a *PSc* value. We used Pearson's correlation coefficient (r^2) between log(*L*) and *PSc* as an estimator of tree scaling (r^2 =-0.6071 for the initial tree). Then we iteratively moved nodes among levels in the matrix while respecting level hierarchy, until r^2 stabilized close to -1 (r^2=-0.9948; 3,000 iterations), meaning that *L* was roughly proportional to *PS*. We used the resulting matrix as guide tree for graphical representations in next sections.

PERMANOVA test at different depths in the *rpoB* phylogenetic tree

We aimed to test for association between *Methylobacterium* isolate diversity assessed at different depths within the *rpoB* phylogeny with sampling and isolation characteristics as proxy for *Methylobacterium* adaptive response to environmental variables through their evolution (**Figure 3a**). For each *PS* value in the phylogenetic tree in the range 0.950-1.000 (roughly corresponding to *PS* range within clades), we classified isolates into discrete taxa and performed a PERMANOVA (10,000 permutations) on *Methylobacterium* community dissimilarity using the *Bray-Curtis* index (*BC*) based on taxa absolute abundance (Hellinger transformation) using R package *vegan* (15). We tested for the relative contribution of four factors and their interactions on taxon frequency: forest of origin (*F*); temperature of isolation (*T*); sampling time (*D*) and host tree species (*H*). For each test, we performed permutations among isolates within batches of isolation to limit batch contribution in the explained variance (**Dataset S1j**).

Permutation test for node association

We asked specifically which nodes within the *Methylobacterium* phylogenetic tree were associated with the two major factors contributing to overall diversity, namely forest of origin and temperature of isolation (**Figure 3a**). For every level in the *rpoB* tree, we independently tested for association between taxa abundance (see above) and forest of origin (SBL and MSH) or temperature of isolation (20 and 30 °C) by permutation (100,000 permutations per level and per factor; **Figure 3b**). For each category of association (node-factor), we calculated p-values according to the following formula: $p=(b+1)/(m+1)$ where b was the number of expected values higher than the observed value and m, the number of permutations (25) and applied Bonferroni correction on p-values. To test for association with forest of origin, permutations were performed among isolates obtained at the same temperature and from the same batch of isolation. To test for association with temperature of isolation, permutations were performed among isolates from the same forest and from the same batch of isolation.

Methylobacterium community analysis of the timeline survey by *rpoB* barcoding

Samples, library preparation and sequencing

We evaluated the *Methylobacteriaceae* phyllosphere diversity through Illumina sequencing of the Methylobacteriaceae-specific marker specific using primers Met02-352-F and Met02-1121-R targeting the first hypervariable region of gene *rpoB* **(**details in section Development of a *Methylobacterium*-specific molecular marker). Amplicon sequencing was performed on 184 phyllosphere samples from 53 trees representative of diversity found in MSH (n=26) and SBL (n=27), and 48 of which allowed a monthly monitoring of diversity (3-4 samples per tree; **Dataset S1b,g**).

Library preparation and sequencing were performed as described in section Bacterial community analysis, with following modification. PCR amplification, library preparation and sequencing were proceeded in four different libraries, each containing a random subset of 46 phyllosphere samples, one negative control (see section Sample collection) and one positive controls (METH community; see section Bacterial community analysis). PCRs contained 1µL of sample DNA, 5 µL of Phusion Hot Start II Buffer (thermoscientific Fisher®), 0.5 µL of dNTP mix 10 µM, 0.5 µL of each primer at 10µM, 1.5µL of DMSO and 0.25 µL of Phusion Hot Start II Buffer DNA Polymerase (thermoscientific Fisher®) for a final volume of 20µL. PCRs were carried out in a thermocycler MasterCycler ProS Eppendorf© with the following steps: 30″ at 98 °C; then, 35 cycles of 15″ at 98 °C, 30″ at 60 °C and 60″ at 72 °C; and a final extension of 10′ at 72 °C. PCR amplification (~800bp) was obtained for all samples and positive controls but not negative controls.

Amplicon Sequence Variant (ASV) definition with dada2

We obtained an average of 30,423 (range 7,510–70,039) paired reads per phyllosphere sample (*n*=184), 29,297 (range 13,368–37,491) for METH communities (*n*=4) and 2,226 (range 467–3,490) for negative controls (*n*=4). We processed *rpoB* reads in order to obtain an ASV

(Amplicon Sequence Variants) abundance table per sample, using package *dada2* in R (10) with following modifications. Read trimming, learning error and concatenating steps were processed separately for each sequencing run (*n*=4). According to sequence quality profiles, 3' ends of forward and reverse reads from each sequence were trimmed 50 and 100bp, respectively (option truncLen in function *filterAndTrim*). Additionally, 5' ends of both reads were trimmed 20bp in order to remove the primer part (option trimLeft in function *filterAndTrim*). After trimming, reads with expected error higher than 2 (option maxEE in function *filterAndTrim*) were discarded. Sample inference was performed using pseudo-pooling of samples (option pool=pseudo in function *dada*). Because forward and reverse reads together (410 bp) did no cover the whole amplicon size (750 pb), they were concatenated together (function *mergePairs,* option *justConcatenate=T*), adding a "*nnnnnnnnnn*" string between forward and reverse sequences. Sequencing runs were combined (function *mergeSequenceTables*) and chimeras were removed (function *removeBimeraDenovo*). After quality filtering (29% of sequences discarded), concatenation (2% discarded) and chimera removing (27 % discarded), we conserved an average of 12,994 (range 3,935–25,807) sequences per phyllosphere sample, 13,915 (range 6,129–20,430) per METH community and 38 (range 7–100) per negative control, for a total of 44,518 unique variants (ASVs).

ASVs filtering and rarefaction

Before processing to rarefaction, we checked diversity within negative and positive controls (METH communities), using blast of the most abundant ASVs against NCBI RefSeq Genome Database (refseq_genomes, limited to *Alphaproteobacteria*) and Nucleotide collection (nr/nt), uncultured/environmental samples excluded. We recovered very few sequences in negative controls (7-100 sequences per replicate, 51 ASVs), most of ASVs present in only one replicate, suggesting very limited and scattered contamination. The most abundant ASV (45 sequences in a single replicate) corresponded to *Rhodococcus sp.*, a typical contaminant of DNA extraction kits (26). Other ASVs (1-11 sequences per sample) were mostly assigned to *Rhizobiales* families typical of the phyllosphere (*Beijerinckiaceae*, *Lichenibacteriaceae*), suggesting very limited cross contamination between samples from a same sequencing run.

In METH communities (6,129-20,430 sequences per replicate, 243 ASVs), we found very good congruence in ASV absolute abundance across the four replicates (Pearson's correlation coefficient: 98.7-99.4%). Nucleotide sequences of the 19 most abundant ASVs (97.7% of total diversity) were exactly identical to *rpoB* partial sequences of 18 *Methylobacterium ssp.* (0.9-20.9% per ASV) and one *Sphingomonas sp.* (0.5%) isolates from which DNA was mixed to built the METH community. We thus considered them as "true ASVs". No ASV was assigned to *E. coli*, nor *Fervidobacterium sp.* although also present in the METH community and detectable through *16S rRNA* barcoding amplicon sequencing (see section Bacterial community analysis), confirming that the *rpoB* marker specificity is limited at least to *Alphaproteobacteria*. We considered that the 224 remaining ASVs (2.3% of diversity) could be either sequencing or PCR errors, contaminants or chimeric ASVs (i.e. resulting from concatenation of forward and reverse reads from different origins). We estimated the potential origin of these "false" ASVs by comparing their nucleotide sequences with those from the 19 trues ASVs, as well as homologous sequences from nine reference genomes that we identified through quick blast search of sequences from the most abundant false ASVs, and belonging to *Caulobacterales* (*n*=2) and *Rhizobiales* families *Beijerinckiaceae* (*n*=2), *Bradyrhizobiaceae* (*n*=2), *Lichenibacteriaceae* (*n*=2) and *Methylocystaceae* (*n*=1). For each putatively false ASV, we calculated nucleotide pairwise similarity (PS) between its nucleotide sequence and all references sequences. In order to identify chimeric ASVs, we did it separately for forward and reverse reads. We identified 12 relatively abundant chimeric ASVs (1.1% of total diversity) merely corresponding to scattered combinations of forward and reverse reads from true *Methylobacterium* ASVs. Other false ASVs corresponded to either *Methylobacterium* errors variants of true ASVs and/or contaminant (38 ASVs, 0.5% of total diversity), likely contaminant from phyllosphere samples (*Beijerinckiaceae*, *Bradyrhizobiaceae*, *Lichenibacteriaceae* and *Caulobacterales*; 77 ASVs, 0.5% of total diversity), chimeric combinations of ASVs among the aforementioned taxa (41 ASV, 0.1% of total diversity) and ASVs unrelated to any of the reference sequences (PS<90%; 56 ASVs, 0.2% of total diversity). Taken separately, false ASVs never exceeded 0.25% of relative abundance in one of the four replicated METH communities, while true ASVs always have at least 0.62% of relative abundance.

Accordingly, we filtered out ASVs that did not have at least 0.5% of relative abundance in a least one phyllosphere sample. We performed ASV filtering and rarefaction as described in section Bacterial community analysis, with a random sampling of 3,000 sequences in each phyllosphere sample and METH community (negative controls excluded; **Dataset S1k**). After rarefaction, we conserved 1,400 ASVs, with an average of 203 (range 48-355) per phyllosphere sample and 29 (range 25-35) per METH community. In METH communities, chimeric ASVs were correctly filtered out and remaining "false" ASVs (25 ASVs; 0.52% of total diversity) mostly corresponded to contamination from the most abundant taxa found across phyllosphere samples (see next section).

ASV taxonomic assignation

We assigned taxonomy of the 1,400 identified ASVs. We used a *rpoB* complete nucleotide sequence database available for *Bacteria* (44,673 reference sequences), previously developed by Ogier et al. (2) and last updated in December 2017. We formatted the database in R according to SILVA v.138 format (11) and processed to taxonomic assignation of the 100 most abundant ASVs with *assignTaxonomy* function in R package *dada2* (10). We considered assignation supported by at least 50% of bootstrap (*minBoot* =50). The 100 most abundant ASVs were assigned at the Class level to *Alphaproteobacteria*, and at the Order level to *Rhizobiales*, confirming specificity of the *rpoB* marker for this order. However, only 31 ASVs were assigned at the family level, to *Methylobacteriaceae* (28 ASVs) and *Bradyrhizobiaceae* (3 ASVs), respectively. We retrieved nucleotide sequences of the 13 most abundant ASVs that were not assigned at the family level and performed blast against NCBI databases (3) *refseq_genomes* and *refseq_rna* available for *Alphaproteobacteria* (Uncultured/environmental samples excluded). Nine ASVs matched *rpoB* sequences from at least one of five Rhizobiales genomes recently added in the NCBI databases (Query Cov > 99%; Per. Ident >98%): two from *Lichenibacterium* (*Lichenibacteriaceae*), a newly described genus (14), and three from *Group RH* a new, yet unnamed, *Beijerinckiaceae* genus (27). Accordingly, we reduced the *rpoB* database to *Alphaproteobacteria* to decrease computation time, and included the five aforementioned sequences. Additionally, we removed all sequences annotated as *Methylobacteriaceae* and replaced them by complete *rpoB* sequences from *Methylobacteriaceae* genomes available on September 2020 (**Figure 2**; **Dataset S1a**) combined with *rpoB* partial nucleotide sequences available from 20 isolates from the pilot survey (**Dataset S1c**). We annotated *Methylobacterium* reference sequences at the species level according to clades. The final *rpoB* database for *Alphaproteobacteria* contained 3062 reference sequences that we used to assign taxonomy of the 1,400 ASVs.

Most of ASVs (1,132) and diversity (94.91% of sequences) were assigned to *Rhizobiales* and 198 to *Caulobacterales* (4.32% of diversity). Within *Rhizobiales*, diversity was mostly assigned to *Beijerinckiaceae* (131 ASVs, 22.53% of diversity), *Lichenibacteriaceae* (262 ASVs, 25.68% of diversity) and *Methylobacteriaceae* (231 ASVs, 23.78% of diversity). A large proportion of *Rhizobiales* diversity (360 ASVs, 21.20% of diversity) was not assigned at the family level. To validate ASV taxonomy, we performed a phylogeny of 1,344 ASVs that could be exactly aligned upon they amino-acid sequence (*Rhodobacterales*, *Rickettsiales*, *Rhodospirillales*, some *Sphingomonadales* and unassigned *Alphaproteobacteria* ASVs excluded; **Figure S5a**). The ML tree (200 replicates) shows good support of main taxonomic groups identified with the *rpoB* database. Most of unassigned *Rhizobiales* diversity was found at the tip of long branches within monophyletic groups with few of no reference sequences: *Lichenibacteriaceae*, *Beijerinckiaceae* (*Group RH*) and *Methylobacteriaceae* (*Enterovirga*) clades, and in a monophyletic group sister of *Lichenibacteriaceae* and *Beijerinckiaceae*, and emcompassing some ASVs assigned to *Methylocystaceae* (*Methylocystaceae*-like group). We corrected ASV taxonomy according to their phylogeny (**Dataset S1k**).

After phylogenetic correction of taxonomy, ASV diversity was mostly found within *Caulobacterales* (209 ASVs, 4.42% of diversity) and within *Rhizobiales* (1,133 ASVs, 94,96% of diversity) in families *Methylobacteriaceae* (283 ASVs, 24.65% of diversity), *Beijerinckiaceae* (*GroupRH*; 165 ASVs; 24.58% of diversity), *Lichenibacteriaceae* (307 ASVs; 31.75% of diversity) and *Methylocystaceae-like* (171 ASVs, 11.04% of diversity). In *Methylobacteriaceae*, ASVs were mostly classified in *Methylobacterium* (200 ASVs, 23.05% of diversity), and *Enterovirga* (78 ASVs, 1.56% of diversity; **Dataset S1l**).

Methylobacterium ASV assignation to clades and of diversity between barcoding and isolation

We assigned the 200 *Methylobacterium* ASVs to clades using using a ML phylogenetic tree based on *rpoB* partial nucleotide sequences from the 283 Methylobacteriaceae ASV (including *Microvirga*: n=5 and *Enterovirga*: n=78), partial sequences including both *rpoB* variable regions available from 20 isolates from the pilot survey (**Dataset S1c,d**) and complete *rpoB* sequences from 185 references Methylobacteriaceae genomes (**Dataset S1a**). ASV Sequences were manually aligned to sequences from reference genomes based upon the amino-acid sequences. Missing positions in ASVs nucleotides sequences were replaced by "Ns". Both alignment and phylogeny were performed in MEGA7 (5). Only *Methylobacterium* ASVs embedded within nodes supported by at least 30% of bootstraps (200 permutations, pairwise deletion) and containing references assigned to a single clade, were assigned (**Figure S5b**).

We compared *Methylobacterium* diversity estimations from culture-dependant (*16s rRNA* and *rpoB* barcoding) and –independent methods (isolation). First, for *16s rRNA* barcoding (15 *Methylobacterium* ASVs) and *rpoB* barcoding comparison (200 *Methylobacterium* ASVs), we used data availalble for both methods from 41 phyllosphere samples from SBL (n = 27) and MSH (n=14) and the METH community (n=1; four replicates combined for *rpoB* barcoding). For each sample and each method, we calculated *Methylobacterium* ASVs sequence relative abundances (after excluding non-*Methylobacterium* ASVs) and combined these relative abundances for each clade (A1, A6, A9, A10 for *16S rRNA*; A1, A2, A3, A4, A6, A9, A10, B and unassigned ASVs for *rpoB*; number of ASVs per clade and method summarized in **Table S2**). We displayed results in a heatmap with samples in rows and clade relative abundance estimated by either *16S rRNA* or *rpoB* barcoding in columns using Euclidean distance to calculate similarity among relative abudances and samples **(Figure S5c**). Samples mostly clustered according to their origin (SBL or MSH). Relative abundances clustered according to clade (A1, A6, A9 and A10) rather than to method (*rpoB* or *16S rRNA* barcoding). We observed some inconsistency in clade A9 relative abundance when comparing both methods. This could be due to the high relative abundance of clade B detected by *rpoB* (19.1% of sequences) but not by *16S rRNA* barcoding (no sequence). By removing clades that were not detected by *16S rRNA* barcoding (B, A2, A3, A4, A5, unknown) in the calculation of clade relative abundance assessed by *rpoB* barcoding, we observed almost perfect correlation in relative abundance of clades A1, A6, A9 and A10 between both methods (data not showed). Second, for the culture-dependant (*rpoB* barcoding; 200 *Methylobacterium* ASVs from 184 phyllosphere samples) and –independent methods comparison (isolation; *rpoB* sequences from 167 isolates), we calculated pairwise nucleotide similarity (PS) among nucleotide sequences from isolates and ASVs by keeping only comparable regions (Ns removed; 333 bp left). We identified 123 isolates (out of 167: 73.7%) that had 100% identity (PS=1) with at least one ASV, for a total of 53 ASVs (out of 200), representing 71.2% of *Methylobacterium* diversity (sequence relative abundance) estimated from *rpoB* barcoding. Using a more relaxed threshold assuming up to 2 sequencing errors of mutations (PS>0.994), we found a match between 155 isolates (92.8%) and 124 ASVs representing 85.9% of diversity (**Figure S5d**). In both cases, we found a good congruence between relative abundances of sequence variants estimated from both methods (**Table S2**), indicating that our survey based on isolation was a good estimator of phyllosphere *Methylobacterium* diversity. The only exceptions were clade B, for which only 7 isolates (out of 167) were isolated in comparison with high relative abundance estimated from *rpoB* barcoding (19.1%), and clade A4 that represented 1.4% of *Methylobacterium* diversity based on barcoding but was not isolated. All analyses were conducted in R (28).

PERMANOVA analysis of *Methylobacterium* community dissimilarity

We performed PERMANOVA analyzes of 184 phyllosphere samples to evaluate relative contributions of forest of origin (*F*), plot within forest (*P*), host tree species (*H*), time of sampling (*T*) and their interactions, in *Methylobacterium* community dissimilarity (Bray distance). PERMANOVAs were conducted in R using function *adonis* from package *vegan* (15). *Methylobacterium* ASV absolute abundances (200 ASVs) were corrected by Hellinger transformation (*decostand* function) to account for rare ASVs and to correct for heterogeneity in *Methylobacterium* abundance between samples. For each analysis, 10,000 permutations were conducted between samples within randomized sequencing runs (*n*=4; *strata* option in *adonis* function) to control for variations due to sequencing errors. We tested following models: (i) a general model including all samples (*n*=184) and factors (*S**T*H*P*); (ii) two forest-specifics models conducted separately on MSH (*n*=85) and SBL (*n*=99) samples, hence excluding *S* from models (*T**H*P*; **Table 1**).

Methylobacterium ASVs association with environmental factors

We tested for the association of each *Methylobacterium* ASV with forest of origin (*F*), plot within forest (*P*), host tree species (*H*) and time of sampling (*T*) and their interactions (**Dataset S1m**). For each ASV independantly, we evaluated by ANOVA the contribution of these factors on the ASV relative abundance (*fx*; Hellinger transformation) under a linear model: lm($fx \sim S*T*H*P$). Then for each factor and interaction separately, we retrieve all *p*-values associated to the contribution (part of variance) to ASV relative abundance, and applied Bonferroni correction on *p*-values. In a principal component analysis (PCA) based on *Bray-Curtis* dissimilarity among 184 *Methylobacterium* communities (Hellinger transformation on 200 *Methylobacterium* ASV relative abundances), we reported ASVs contributions to the PCA only for ASVs significantly associated with one or either forest in the ANOVA (displayed in **Figure 4a**).

Spatial and temporal autocorrelation analysis on *Methylobacterium* ASVs

We quantified spatial and temporal dynamics of *Methylobacterium* community using autocorrelation analysis based on Bray-Curtis pairwise dissimilarity (*BC*) between 184 phyllosphere samples. In order to remove large-scale spatial variation due to strong difference in community composition between forests (**Table 1**), we analyzed MSH and SBL separately. For each possible pair of phyllosphere samples, we calculated *BC* on ASV relative abundance (200 ASV, Hellinger transformation), *pDist* as the spatial distance separating trees where the two samples came from (in meters) and *pTime* as the time separating dates when the two communities were sampled (in days). We evaluated the effects of *pDist* and *pTime* on *BC* under three different linear models by using ANOVA (**Table 2**). (i) *Spatial autocorrelation general models*: in MSH and SBL (two models), we evaluated the effect of *pDist* on *BC*. In order to take into account variations in community composition among dates, only pairwise comparisons among samples from a same date (*D*) were considered, and *D* as well as the *pDist:D* interaction were included in the model (**Figure 4b**). (ii) *Spatial autocorrelation models per date*. In each forest (n=2) and sampling date taken separately (n=4), we evaluated the effect of *pDist* on *BC* (eight models; **Figure 4e** only for MSH). (iii) *Temporal autocorrelation: general models*: in MSH and SBL (two models), we evaluated the effect of *pTime* on *BC*, regardless spatial scales (**Figure 4c**). For each model, we reported the average and standard deviation (*sd*) of the intercept, corresponding to the average and *sd BC* values among all the considered pairwise comparisons (BC distributions from model (ii) displayed in **Figure 4d**). For each factor (*pDist*, *Date*, *pDist:Date* and *pTime*), we also reported the average and standard deviation of estimates (slope), which significance was assessed by ANOVA on the linear model (**Table 2**).

Methylobacterium growth performances under four temperature treatments

We tested for the adaptive response of *Methylobacterium* isolates from the phyllosphere, to temperature variations during tree growing season. For 80 *Methylobacterium* isolated in 2018 in forests MSH (n=32) and SBL (n=47), we evaluated growth abilities under four temperature treatments, as a proxy for adaptation to temperature variations. Each treatment consisted in a first pre-conditioned (*P*) step during which each isolate was incubated for 20 days to either 20 (*P20*) or 30 °C (*P30*), and a second monitoring step (*M*) during which each pre-conditioned isolate was incubated and their growth monitored for 24 days at 20 °C (*P20M20* and *P20M30*) or 30 °C (*P30M20* and *P30M30;* **Figure S2**). We expected that treatments P20M20 and P30M30 mimicked stable thermal environments and that treatments P20M30 and P30M20 mimicked variable thermal environments.

Growing conditions

Pre-conditioning step (*P*; **Figure S2a,b**): For each isolate (*n*=80), two negative controls (*n*=2) and each temperature treatment (*n*=2), 10µL of cellular culture (thawed from -80 °C glycerol stocks) or sterile water for controls, were spread on 20mL of solid MMS media containing 0.1% methanol, 50mg/L of yeast extract and 50mg/L of vitamin mix (Sigma® RPMI1640). After 20 days of incubation at 20 °C (*P20*) or 30 °C (*P30*), petri dishes with no evidence of contamination were swabbed with 2mL of sterile distilled water. Colonies were collected in approximately 1mL of liquid and centrifuged in 1.5 mL Eppendorf tube, 10 minutes at 4°C (21130 rcf). Tubes with evidence of contamination (two different colors in the pellet) were discarded. Clean pellets were suspended by pipetting and cell concentrations were adjusted with sterile water to the same optic density OD_{630}=0.2, equivalent to about 1.6×10^{8} cells/mL. No dilution was applied to negative controls.

Monitoring step (*M*; **Figure S2c,d**): For each pre-conditioned culture *P20* (n=82) and *P30* (n=82), 10µL (approximately 1.6×10^6 cells) were spotted on new petri dishes containing the same MMS media as used in the pre-conditioning step. Spots were distributed on petri dishes according to a 6×6 square grid (36 spots per plate), each dish containing 17 isolates and one negative control, for each of which one culture came from *P20* treatment and one from *P30*. *P20* and *P30* replicates from the same isolate were spotted next to each other in order to facilitate *P* treatment comparisons. Isolate positions were randomized elsewhere. Each petri dish was duplicated, one copy for incubation at 20 °C (*M20* treatment) and one for incubation at 30 °C (*M30*).

For each isolate and each combination of treatments (*PXXMXX*), we realized 5 replicates, randomly distributed in two series (*PXXM20* and *PXXM30*) of 24 petri dishes (**Figure S2e**). Within each *M* treatment, each petri dish had a different set and display of isolates, and replicates from the same isolate were in different petri dishes and in different positions. During the monitoring step, we took pictures of each petri dish with a Nexus LG device, at days 7, 13 and 24 after inoculation.

Image analysis

Pictures from each petri dish (n=24) and each time point (n=3) were first analyzed with ImageJ 1.52e software. Each original picture (**Figure S3a**) was converted in grey scale. Areas outside of the agar, as well as every visible particle other than bacteria spots within the agar area, were manually cropped using the elliptic tool. The picture was duplicated (**Figure S3b**). The first copy was used to measure raw bacteria spot intensities (*BW*). In the second copy, used for background correction (*BACK*), bacteria spots were cropped using the elliptic tool, by keeping as much background area as possible. Both copies having exactly the same dimensions (about 1,800x1,800 pixels per petri dish) were converted in matrix of pixel intensities using the */transform/image_to_results* tool and normalized in R in a 500x500-pixel matrix by averaging intensities. Cropped areas and pixels with intensities out of the range 50-200 were considered as missing values in *BACK* and *BW*. In order to reconstruct background intensities, missing values in the *BACK* file (including cropped positions of bacteria spots) were predicted from values with known intensity (**Figure S3c**). First, in order to tighten mesh in cropped areas, 15,000 pixels with missing intensities were randomly sampled and their intensity predicted from average known intensities in a 30x30 pixels windows. Second, intensities for remaining pixels with missing values were predicted in row (X-axis) and column (Y-axis), separately according to known and predicted values in 50-pixel windows, using function *runmean* (package *CaTools*). Third, for each pixel, predicted values in X and Y axis were averaged. Reconstructed intensities of the *BACK* files were then subtracted to the *BW* file in order to remove background (**Figure S3d**). For each spot, the growth area was determined as followed (**Figure S3e,f**). X and Y coordinates of the approximate central position of the bacteria spot was manually retrieved from the corrected *BW* file (36 per picture) loaded in ImageJ. In R, a 30x30-pixel window centered on these coordinates was defined. For each possible circle inscribed within this area, pixel intensity distribution outside and within the circle were compared using a t-test (minimum circle area: 200 pixels). The circle with the largest t value was considered as the approximate profile of the bacteria spot, for which the average growth intensity (within the circle) and the average background intensity (outside the circle, within the 30x30-pixel window) were reported.

For of each spot, the average intensity was corrected by subtracting average local background intensity (**Figure S3f**). After correcting for local background, spot intensities from negative controls were almost indistinguishable from background (I = 0.10±0.29).

We observed in average slightly but consistent higher spot intensity at the border (I= 5.8±5.5) than at the center of petri dishes (I = 3.5±5.6), regardless of petri dish, isolate identity, time point, temperature treatment or other factors, suggesting that replicates located close to the border took advantage of less competition for nutrients (**Figure S3g**). Although positions of replicates from the same isolate were randomized, some could just by chance be systematically located close to the center or the border of a dish. We thus corrected for this border effect by predicting I values in function of their position in the petri dish with the polynomial regression: $I \sim X^2Y^2+X^2Y+XY^2+X+Y$, were X and Y are the average coordinates of the spot on the petri dish (**Figure S3h**). Residuals from the polynomial regression were used as corrected I values (**Figure S3i**).

Growth profile analysis

On average, spot intensity increased between days 7 (I = 4.66±4.11) and 13 (I = 5.32±4.91), followed by a average decrease in intensity at day 24 (I = 4.22±4.14), illustrating that after reaching a maximal intensity (or yield), isolates eventually underwent starvation because of nutrient depletion in their immediate environment. For each isolate, each replicate and each treatment, we estimated yield (*Y*) and the growth rate (*r*; (29)). Because our survey was limited to three time points (7, 13 and 24 days), we estimated those values from predicted growth curves in the range 0-36 days, assuming that intensity was null at day 0 (I_0=0) and that growth curves followed a log normal distribution (**Figure S4a**; (30)). For each spot, we estimated the predicted growth curve, in the range 0-36 days after inoculation, from the best log normal model fitting assumed (I_0) and known intensity values (I_7, I_{13} and I_{24}). We defined *Y* as the maximal intensity predicted from the log normal growth curve and *r* as the inverse of *log+lag* times necessary to reach Y (**Figure S4b**). We observe a good congruence between *Y* values predicted from log normal curves and maximum intensities observed from the three time points (**Figure S4c**), as well as a good congruence between predicted lag+log time and time at which the maximum intensity was actually observed (7, 13 or 24 days; **Figure S4d**). For a majority of replicates (87%), the predicted *Y* was reached before the 24th day of incubation (**Figure S4e**). From this point, we discarded the remaining replicates and considered 79 out of 80 isolates for which average *Y* and *r* could be predicted from at least one replicate (**Dataset S1n**).

In order to assess factors affecting *Methylobacterium* growth abilities under different temperature treatments, we constructed linear models (*lm* function in R) predicting *Y* and *r* (log transformations to meet normal distribution) in function of clade assignement of isolates (*C*), forest of origin (*F*), host tree species (*H*), time of sampling (*D*), temperature of isolation (T_I; at which each isolate was isolated), temperature of incubation during pre-conditioning (T_P) and monitoring (T_M) steps, and all possible interactions between factors (log$(Y$ or $r) \sim H*S*D*C*T_P*T_M*T_I$) and evaluated the contribution of these factors in log(*r*) and log(*Y*) using an ANOVA (**Table 3; Dataset S1o**).

REFERENCES

1. Vos M, Quince C, Pijl AS, Hollander M de, Kowalchuk GA. 2012. A Comparison of rpoB and 16S rRNA as Markers in Pyrosequencing Studies of Bacterial Diversity. PLOS ONE 7:e30600.
2. Ogier J-C, Pagès S, Galan M, Barret M, Gaudriault S. 2019. rpoB, a promising marker for analyzing the diversity of bacterial communities by amplicon sequencing. BMC Microbiol 19:171.
3. Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, Madden TL, Matten WT, McGinnis SD, Merezhuk Y, Raytselis Y, Sayers EW, Tao T, Ye J, Zaretskaya I. 2013. BLAST: a more efficient report with usability improvements. Nucleic Acids Res 41:W29-33.
4. Tamura K, Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–526.
5. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 33:1870–1874.
6. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. Systematic Biology 61:539–542.
7. Grafen A. 1989. The phylogenetic regression. Philos Trans R Soc Lond B Biol Sci 326:119–157.
8. Green PN, Ardley JK. 2018. Review of the genus Methylobacterium and closely related organisms: a proposal that some Methylobacterium species be reclassified into a new genus, Methylorubrum gen. nov. International Journal of Systematic and Evolutionary Microbiology 68:2727–2748.
9. Redford AJ, Bowers RM, Knight R, Linhart Y, Fierer N. 2010. The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. Environmental Microbiology 12:2885–2893.
10. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: High resolution sample inference from Illumina amplicon data. Nat Methods 13:581–583.
11. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-

based tools. Nucleic Acids Res 41:D590–D596.

12. Qu Z, Jiang F, Chang X, Qiu X, Ren L, Fang C, Peng F. 2014. Psychroglaciecola arctica gen. nov., sp. nov., isolated from Arctic glacial foreland soil. International Journal of Systematic and Evolutionary Microbiology, 64:1817–1824.

13. Kulichevskaya IS, Danilova OV, Tereshina VM, Kevbrin VV, Dedysh SN. 2014. Descriptions of Roseiarcus fermentans gen. nov., sp. nov., a bacteriochlorophyll a-containing fermentative bacterium related phylogenetically to alphaproteobacterial methanotrophs, and of the family Roseiarcaceae fam. nov. Int J Syst Evol Microbiol 64:2558–2565.

14. Pankratov TA, Grouzdev DS, Patutina EO, Kolganova TV, Suzina NE, Berestovskaya JJ. 2020. Lichenibacterium ramalinae gen. nov, sp. nov., Lichenibacterium minor sp. nov., the first endophytic, beta-carotene producing bacterial representatives from lichen thalli and the proposal of the new family Lichenibacteriaceae within the order Rhizobiales. Antonie van Leeuwenhoek 113:477–489.

15. Dixon P. 2003. VEGAN, a package of R functions for community ecology. Journal of Vegetation Science 14:927–930.

16. Green PN. 2006. Methylobacterium, p. 257–265. *In* Dworkin, M, Falkow, S, Rosenberg, E, Schleifer, K-H, Stackebrandt, E (eds.), The Prokaryotes: Volume 5: Proteobacteria: Alpha and Beta Subclasses. Springer, New York, NY.

17. Tani A, Sahin N, Matsuyama Y, Enomoto T, Nishimura N, Yokota A, Kimbara K. 2012. High-Throughput Identification and Screening of Novel Methylobacterium Species Using Whole-Cell MALDI-TOF/MS Analysis. PLoS ONE 7:e40784.

18. Turner S, Pryer KM, Miao VP, Palmer JD. 1999. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. The Journal of Eukaryotic Microbiology 46:327–338.

19. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences of the United States of America 108 Suppl 1:4516–4522.

20. Frank RAW, Price AJ, Northrop FD, Perham RN, Luisi BF. 2007. Crystal Structure of the E1 Component of the Escherichia coli 2-Oxoglutarate Dehydrogenase Multienzyme Complex. Journal of Molecular Biology 368:639–651.

21. Bell RL, González-Escalona N, Stones R, Brown EW. 2011. Phylogenetic evaluation of the 'Typhimurium' complex of Salmonella strains using a seven-gene multi-locus sequence analysis. Infection, Genetics and Evolution 11:83–91.

22. Ee R, Madhaiyan M, Ji L, Lim Y-L, Nor NM, Tee K-K, Chen J-W, Yin W-F. 2016. Chania multitudinisentens gen. nov., sp. nov., an N-acyl-homoserine-lactone-producing bacterium in the family Enterobacteriaceae isolated from landfill site soil. International Journal of Systematic and Evolutionary Microbiology, 66:2297–2304.

23. Küpfer M, Kuhnert P, Korczak BM, Peduzzi R, Demarta A. 2006. Genetic relationships of Aeromonas strains inferred from 16S rRNA, gyrB and rpoB gene sequences. International Journal of Systematic and Evolutionary Microbiology, 56:2743–2751.

24. Paradis E, Schliep K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526–528.

25. Phipson B, Smyth GK. 2010. Permutation p-values should never be zero: calculating exact p-values when permutations are randomly drawn. Statistical Applications in Genetics and Molecular Biology 9.

26. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses | BMC Biology | Full Text.

27. Wegner C-E, Gorniak L, Riedel S, Westermann M, Küsel K. 2019. Lanthanide-Dependent Methylotrophs of the Family *Beijerinckiaceae*: Physiological and Genomic Insights. Appl Environ Microbiol 86:e01830-19, /aem/86/1/AEM.01830-19.atom.

28. R-Developement-Core-Team. 2011. R: A language and environment for statistical computing. ISBN 3-900051-07-0. R Foundation for Statistical Computing. Vienna, Austria, 2013. url: http://www. R-project. org.

29. Lipson DA. 2015. The complex relationship between microbial growth rate and yield and its implications for ecosystem processes. Front Microbiol 6.

30. Zwietering MH, Jongenburger I, Rombouts FM, Riet K van 't. 1990. Modeling of the Bacterial Growth Curve. Appl Environ Microbiol 56:1875–1881.

Figure S1 - ML phylogenetic trees from *sucA* (a) and *rpoB* (b) concatenated hypervariable (HV) regions. Trees were drawn from sequences obtained for 20 representative isolates from 2017 pilot survey (black circles) and reference genomes. Trees with the highest log likelihood are shown. Bootstraps: only values for node supported by at least 50% of replicated trees are displayed. Phylogenetic tree was rooted on *Microvirga* and *Enterovirga* outgroups (Compressed). a) The *sucA* ML tree was inferred from 3 aligned concatenated HV regions (1,663 bp) available for 189 reference genomes and 14 tested isolates. b) The *rpoB* ML tree was inferred from 2 aligned and concatenated HV regions (1,244 bp) available for 163 reference genomes and the 20 tested isolates. c) Consensus clade tree from *sucA* and *rpoB* ML phylogenies. Only tree topology among clades supported by both phylogenies is shown, regardless bootstrap support. For each consensus node, the minimum (most conservative) bootstrap support found between phylogenies is shown (grey scale, legend on top).

Figure S2 - Experimental design of *Methylobacterium* monitoring for growth performance under four temperature treatments. a) 79 isolates (two showed in this example: pink and cyan) and two negative controls (not showed) were tested for ability to grow under different temperature treatments. b) Pre-conditioning step (*P*): For each isolate and negative controls and each temperature treatment (20 and 30 °C), 10µL of cellular culture from stock were spread on solid Methanol-MMS media. c) After 20 days of incubation at 20 °C (*P20*) or 30 °C (*P30*), petri dishes were swabbed and collected cell concentrations adjusted to OD_{630}=0.2. d) Monitoring step (*M*): Each pre-conditioned culture *P20* (*n*=81) and *P30* (*n*=81), was spotted on new Methanol-MMS media (five replicates per culture, per *P* treatment). Each petri dish was duplicated, one copy for incubation at 20 °C (*M20* treatment) and one for incubation at 30 °C (*M30*). Pictures of petri dishes were took 7, 13 and 24 after inoculation. e) Three examples of spot organization on petri dishes (24 per *M* treatment, 17 isolates + one negative control per dish). Open circle represent negative controls.

Figure S3 - Example of image analysis of *Methylobacterium* monitoring for growth performance under four temperature treatments. a) Original picture. b) The original picture was converted in grey scale in ImajeJ. Areas outside of the agar, as well as every visible particle other than bacteria spot within the agar area, were manually cropped (black). The picture was duplicated. A copy was used for background correction (*BACK*; bacteria spots cropped). Another copy was used to measure raw bacteria spot intensities (*BW*). c) Reconstruction of background intensities. d) Correction of raw intensities by subtracting background values. e) Definition of growth area: detail of a spot (top) and average pixel intensities in function of the radius of concentric circles drawn from the center of the colony. Growth area is defined as the circle with maximal *T* value in *t*-test comparison between intensities outside and within the area (here in red). f) Comparison of intensity distributions outside (red; local background) and within (blue; spot) growing area. Dotted lines indicate average intensity values. g) Border effect: spot intensities after correction for local background (*I*) are shown for 48 petri dishes and 3 time points in function of their average position of each petri. Because of less competition for nutrients, *I* values (proportional to point size) are in average higher close to the border of the petri dish (spot positioned according to the original picture). h) Expected *I* values (scale on top) in function of *X*/*Y* position on the petri dish predicted from a polynomial regression ($I \sim X^2Y^2+X^2Y+XY^2+X+Y$). i) Corrected *I* values (residuals from the polynomial regression).

Figure S4 - Prediction of log normal growth curve, growth rate and yield for 79 isolates incubated under four temperature treatments. a) Log normal best prediction for four different observed cases (legend on bottom). Curves were predicted in the range 0-36 days form values observed at T_7, T_{13} and T_{24}, assuming null intensity at T_0. Models assuming that intensity remained null until T_i were tested in the range 0-7 days. b) determination of yield (*Y*, maximum intensity) and growth rate ($r = 1/log{+}lag$) from predicted growth curve. c) comparison of predicted yield and maximum observed intensity. d) comparison of predicted log+lag values with time (T_7, T_{13} or T_{24}) at which maximum intensity was observed. e) All predicted growth curves showed separately for each temperature treatment. Replicates for which maximum intensity was not reach at day 36 according to the model were discarded ($log{+}lag \geq 36$).

Figure S5 – *Alphaproteobacteria* and *Methylobacterium* diversity assessed by *rpoB* barcoding. a) Unrooted ML phylogenetic tree based on *rpoB* partial nucleotide sequences from 1,344 ASVs (410 bp). Only nodes supported by at least 50% of bootstraps (200 permutations) are shown. ASVs are labeled according to their taxonomic assignation based on *rpoB* nucleotide sequence database (legend on bottom left). b) Unrooted ML phylogenetic tree based on *rpoB* partial nucleotide sequences from 283 *Methylobacteriaceae* ASV (points) and 232 references isolates and genomes (unlabeled tips). Only nodes supported by at least 50% of bootstraps (200 permutations) are shown. Full circle indicate 200 *Methylobacterium* ASVs assigned to clades (colored) when clustering with identified reference sequences with at least 50% of support. Unassigned ASVs are indicated in grey. c) Heatmap showing the comparison of *Methylobacterium* diversity assessement from *rpoB* barcoding and *16s rRNA* barcoding. For 41 phyllosphere samples and the METH community (in rows; colors indicating sample origin; legend on bottom left), the relative sequence abundances (color scale on top left) of ASVs assigned to the same clade were combined (in columns). For *rpoB* barcoding, all clades were detected (Unk. indicates unassigned ASVs). For *16s rRNA* barcoding only clades A1, A6, A9 and A10 were detected. d) Comparison of *Methylobacterium* diversity assessment from *rpoB* barcoding (Y-axis) and isolation (X-axis). Number of *Methylobacterium* isolates in function of ASV relative abundance assuming 98.5% of nucleotide identity between *rpoB* sequences obtained by SANGER sequencing in isolates and *rpoB* sequences of ASVs (maximum of 6 nucleotide mismatches). Points are colored according to assignement to clades (legend on top right). The proportion of unmatched diversity (no match between ASV and isolate sequences) is displayed in pie charts for ASVs (top left) and isolates (bottom right).

Figure S1 - ML phylogenetic trees from *sucA* (a) and *rpoB* (b) concatenated hypervariable (HV) regions. Trees were drawn from sequences obtained for 20 representative isolates from 2017 pilot survey (black circles) and reference genomes. Trees with the highest log likelihood are shown. Bootstraps: only values for node supported by at least 50% of replicated trees are displayed. Phylogenetic tree was rooted on *Microvirga* and *Enterovirga* outgroups (Compressed). a) The *sucA* ML tree was inferred from 3 aligned concatenated HV regions (1,663 bp) available for 189 reference genomes and 14 tested isolates. b) The *rpoB* ML tree was inferred from 2 aligned and concatenated HV regions (1,244 bp) available for 163 reference genomes and the 20 tested isolates. c) Consensus clade tree from *sucA* and *rpoB* ML phylogenies. Only tree topology among clades supported by both phylogenies is shown, regardless bootstrap support. For each consensus node, the minimum (most conservative) bootstrap support found between phylogenies is shown (grey scale, legend on top).

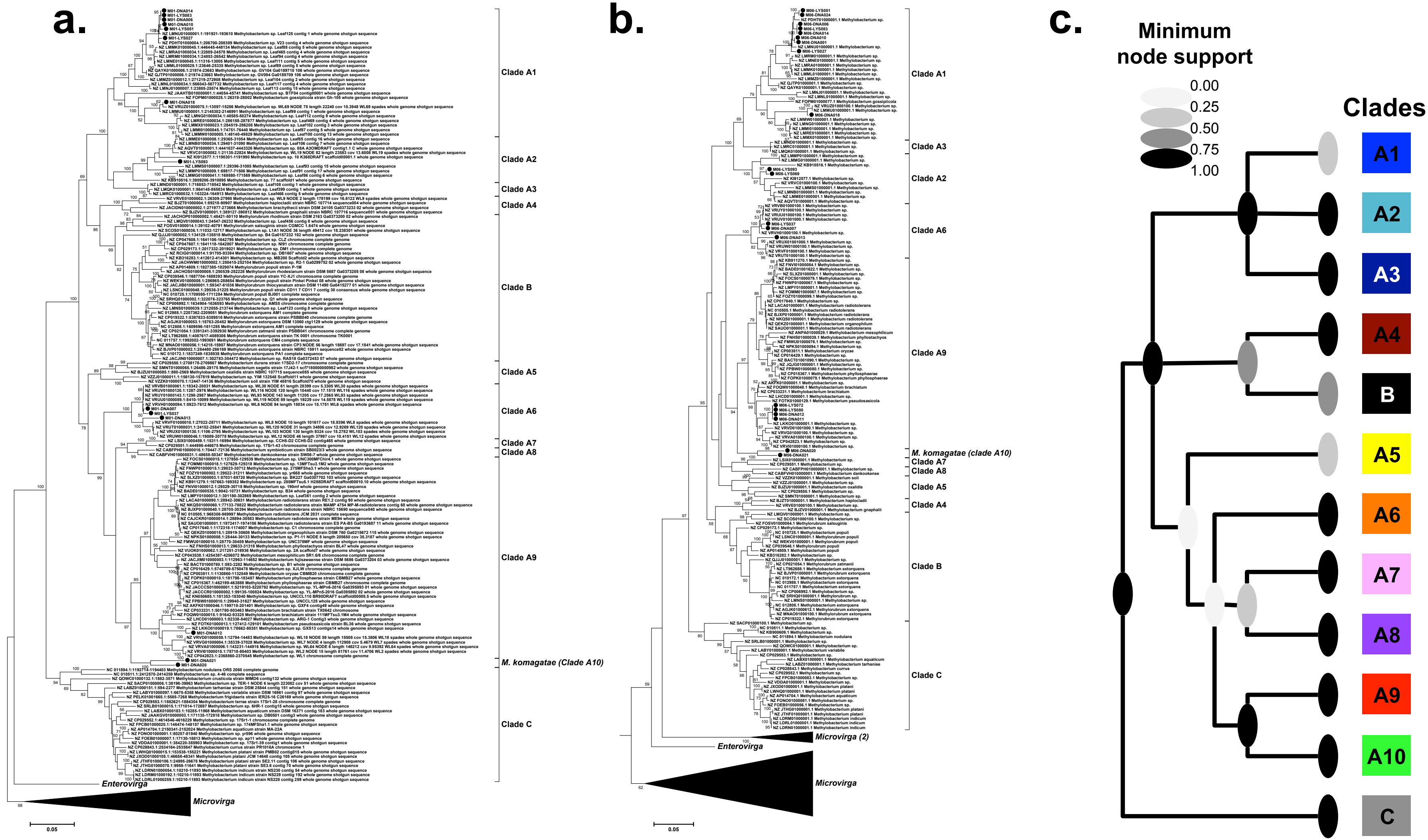
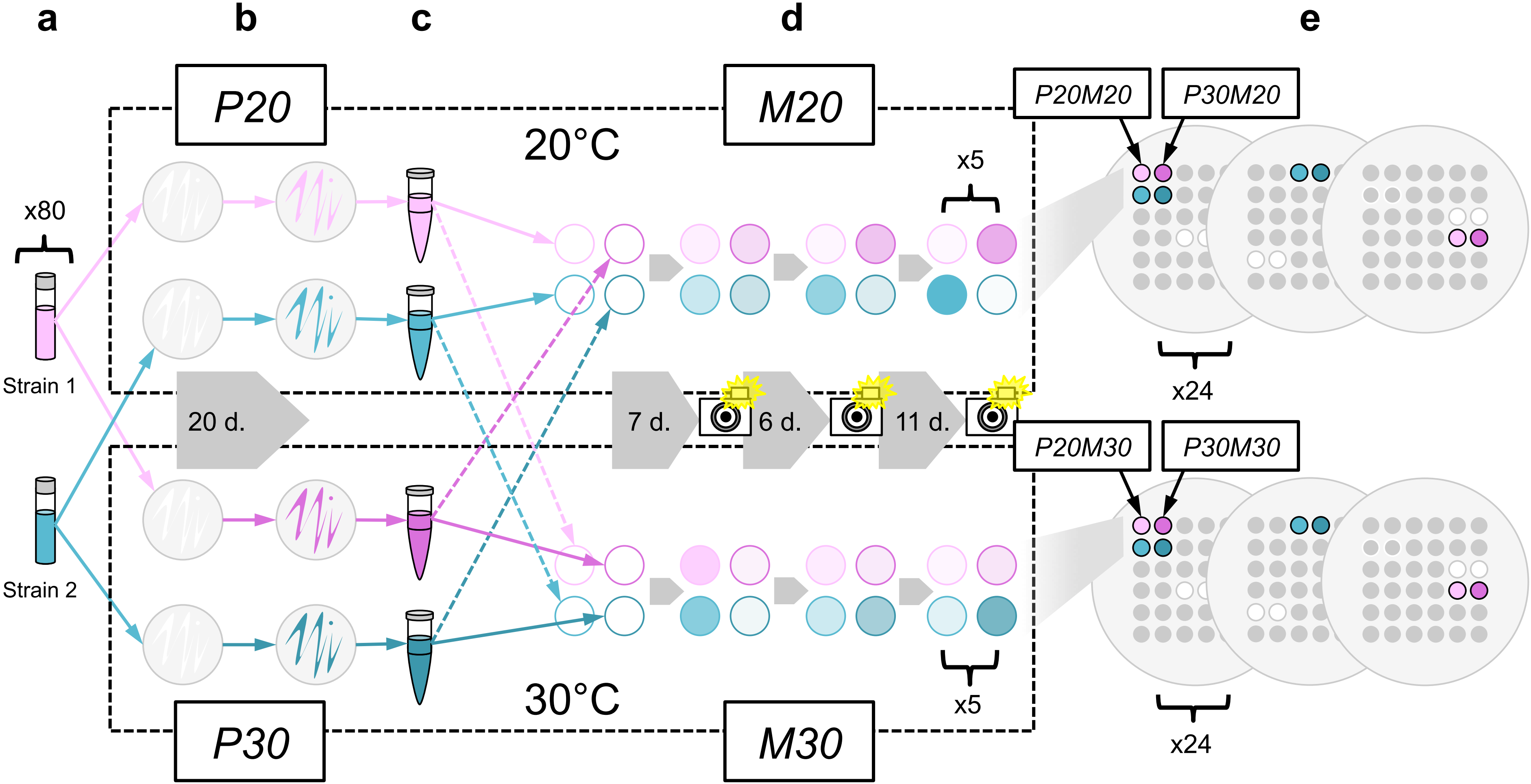

a.
Clade A1
Clade A2
Clade A3
Clade A4
Clade B
Clade A5
Clade A6
Clade A7
Clade A8
Clade A9
M. komagatae (Clade A10)
Clade C
Enterovirga
Microvirga
0.05
b.
Clade A1
Clade A3
Clade A2
Clade A6
Clade A9
M. komagatae (clade A10)
Clade A7
Clade A8
Clade A5
Clade A4
Clade B
Clade C
Microvirga (2)
Enterovirga
Microvirga
0.05
c.
Minimum node support
0.00
0.25
0.50
0.75
1.00
Clades
A1
A2
A3
A4
B
A5
A6
A7
A8
A9
A10
C

Figure S2 - Experimental design of *Methylobacterium* monitoring for growth performance under four temperature treatments. a) 79 isolates (two showed in this example: pink and cyan) and two negative controls (not showed) were tested for ability to grow under different temperature treatments. b) Pre-conditioning step (*P*): For each isolate and negative controls and each temperature treatment (20 and 30 °C), 10µL of cellular culture from stock were spread on solid Methanol-MMS media. c) After 20 days of incubation at 20 °C (*P20*) or 30 °C (*P30*), petri dishes were swabbed and collected cell concentrations adjusted to OD_{630}=0.2. d) Monitoring step (*M*): Each pre-conditioned culture *P20* (*n*=81) and *P30* (*n*=81), was spotted on new Methanol-MMS media (five replicates per culture, per *P* treatment). Each petri dish was duplicated, one copy for incubation at 20 °C (*M20* treatment) and one for incubation at 30 °C (*M30*). Pictures of petri dishes were took 7, 13 and 24 after inoculation. e) Three examples of spot organization on petri dishes (24 per *M* treatment, 17 isolates + one negative control per dish). Open circle represent negative controls.



a
b
c
d
e
P20
M20
P20M20
P30M20
20°C
x80
x5
Strain 1
x24
20 d.
7 d.
6 d.
11 d.
P20M30
P30M30
Strain 2
x5
30°C
x24
P30
M30

Figure S3 - Example of image analysis of *Methylobacterium* monitoring for growth performance under four temperature treatments. a) Original picture. b) The original picture was converted in grey scale in ImajeJ. Areas outside of the agar, as well as every visible particle other than bacteria spot within the agar area, were manually cropped (black). The picture was duplicated. A copy was used for background correction (*BACK*; bacteria spots cropped). Another copy was used to measure raw bacteria spot intensities (*BW*). c) Reconstruction of background intensities. d) Correction of raw intensities by subtracting background values. e) Definition of growth area: detail of a spot (top) and average pixel intensities in function of the radius of concentric circles drawn from the center of the colony. Growth area is defined as the circle with maximal *T* value in *t*-test comparison between intensities outside and within the area (here in red). f) Comparison of intensity distributions outside (red; local background) and within (blue; spot) growing area. Dotted lines indicate average intensity values. g) Border effect: spot intensities after correction for local background (*I*) are shown for 48 petri dishes and 3 time points in function of their average position of each petri. Because of less competition for nutrients, *I* values (proportional to point size) are in average higher close to the border of the petri dish (spot positioned according to the original picture). h) Expected *I* values (scale on top) in function of *X*/*Y* position on the petri dish predicted from a polynomial regression ($I \sim X^2Y^2+X^2Y+XY^2+X+Y$). i) Corrected *I* values (residuals from the polynomial regression).

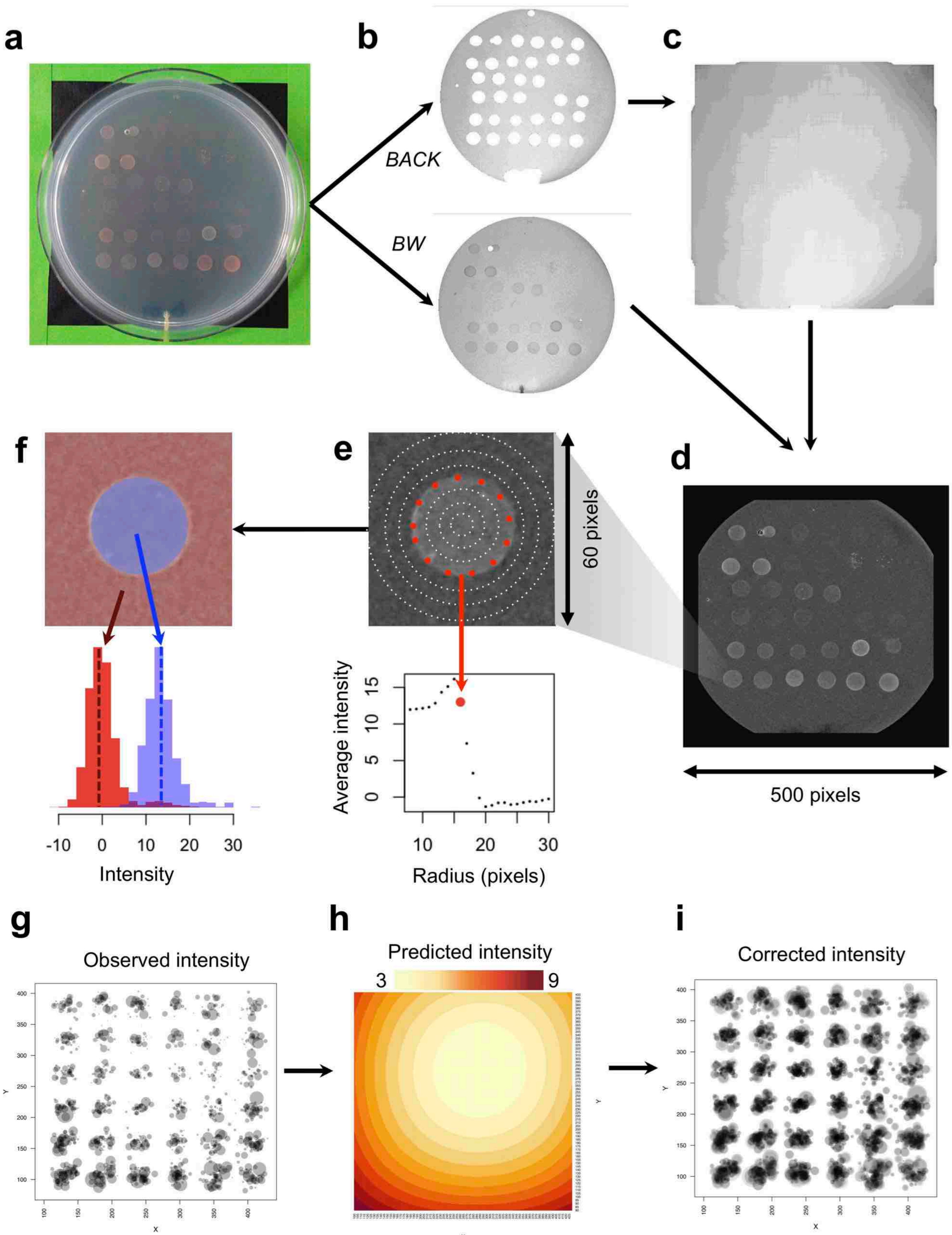
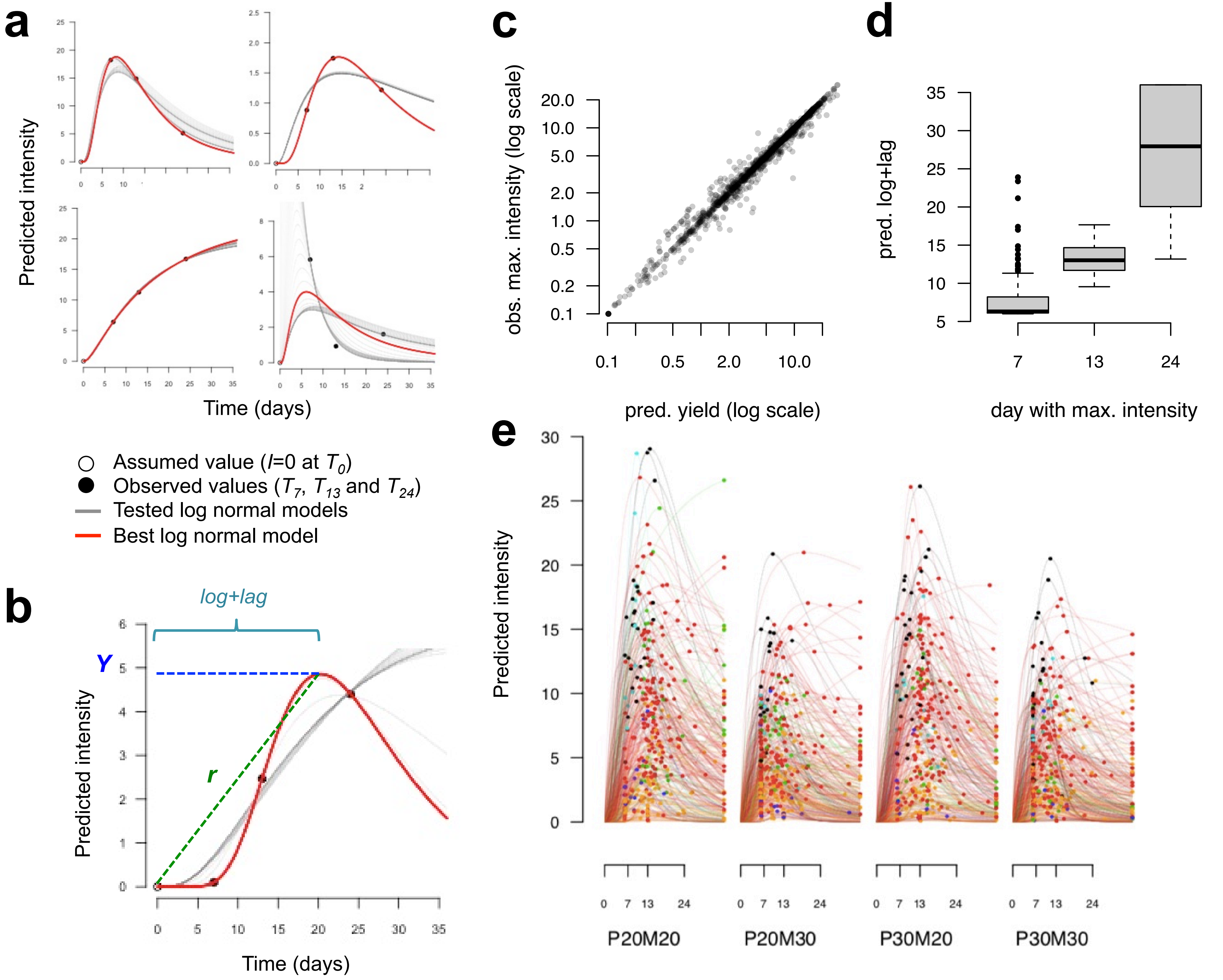
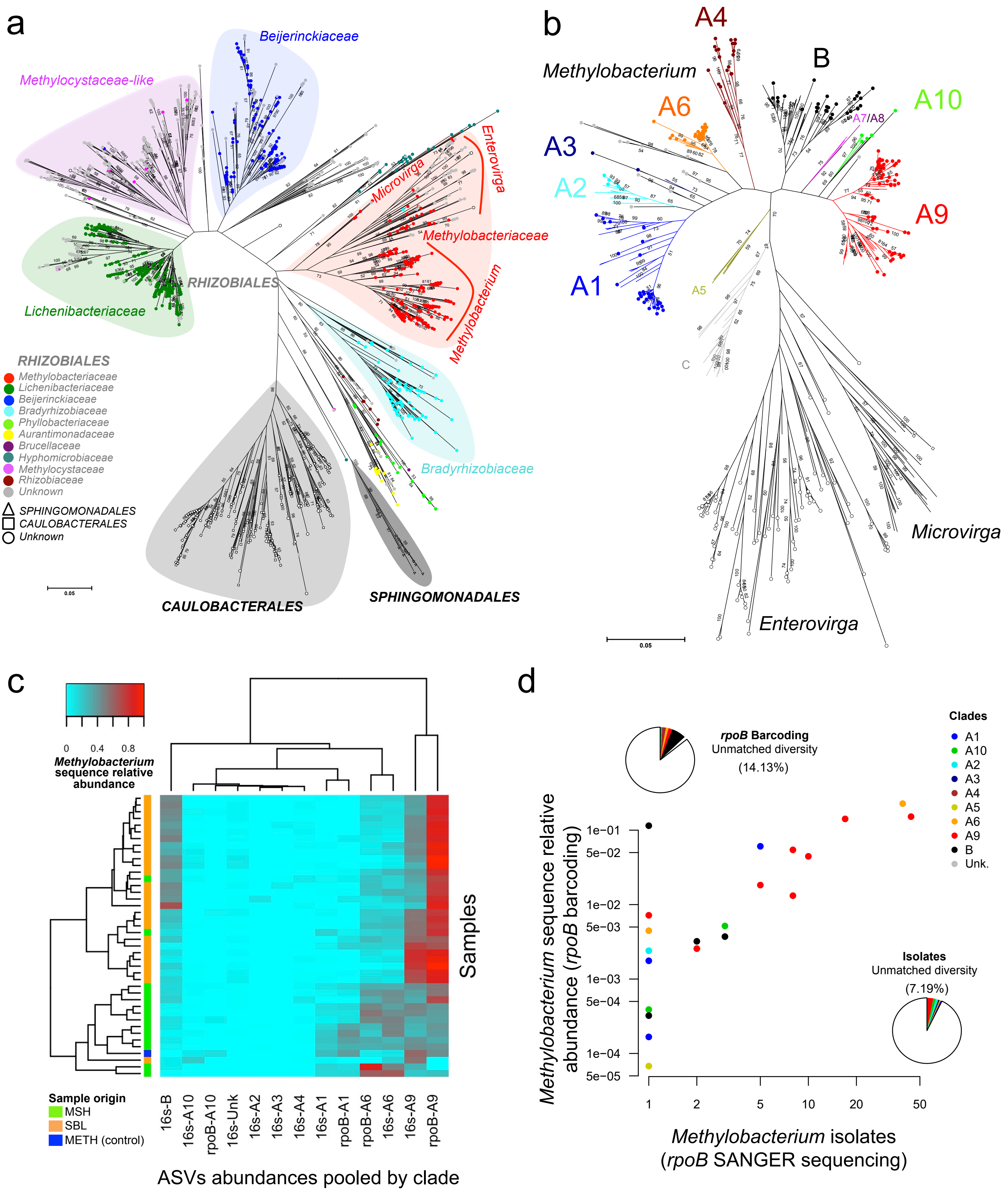

a
b
c
BACK
BW
d
e
f
60 pixels
500 pixels
Average intensity
Radius (pixels)
Intensity
g
h
i
Observed intensity
Predicted intensity
Corrected intensity
3
9

Figure S4 - Prediction of log normal growth curve, growth rate and yield for 79 isolates incubated under four temperature treatments. a) Log normal best prediction for four different observed cases (legend on bottom). Curves were predicted in the range 0-36 days form values observed at T_7, T_{13} and T_{24}, assuming null intensity at T_0. Models assuming that intensity remained null until T_i were tested in the range 0-7 days. b) determination of yield (*Y*, maximum intensity) and growth rate ($r = 1/log{+}lag$) from predicted growth curve. c) comparison of predicted yield and maximum observed intensity. d) comparison of predicted log+lag values with time (T_7, T_{13} or T_{24}) at which maximum intensity was observed. e) All predicted growth curves showed separately for each temperature treatment. Replicates for which maximum intensity was not reach at day 36 according to the model were discarded ($log{+}lag \geq 36$).



a
Predicted intensity
Time (days)
Assumed value (I=0 at T0)
Observed values (T7, T13 and T24)
Tested log normal models
Best log normal model
b
log+lag
Y
r
Predicted intensity
Time (days)
c
obs. max. intensity (log scale)
pred. yield (log scale)
d
pred. log+lag
day with max. intensity
e
Predicted intensity
P20M20
P20M30
P30M20
P30M30

Figure S5 – *Alphaproteobacteria* and *Methylobacterium* diversity assessed by *rpoB* barcoding. a) Unrooted ML phylogenetic tree based on *rpoB* partial nucleotide sequences from 1,344 ASVs (410 bp). Only nodes supported by at least 50% of bootstraps (200 permutations) are shown. ASVs are labeled according to their taxonomic assignation based on *rpoB* nucleotide sequence database (legend on bottom left). b) Unrooted ML phylogenetic tree based on *rpoB* partial nucleotide sequences from 283 *Methylobacteriaceae* ASV (points) and 232 references isolates and genomes (unlabeled tips). Only nodes supported by at least 50% of bootstraps (200 permutations) are shown. Full circle indicate 200 *Methylobacterium* ASVs assigned to clades (colored) when clustering with identified reference sequences with at least 50% of support. Unassigned ASVs are indicated in grey. c) Heatmap showing the comparison of *Methylobacterium* diversity assessement from *rpoB* barcoding and *16s rRNA* barcoding. For 41 phyllosphere samples and the METH community (in rows; colors indicating sample origin; legend on bottom left), the relative sequence abundances (color scale on top left) of ASVs assigned to the same clade were combined (in columns). For *rpoB* barcoding, all clades were detected (Unk. indicates unassigned ASVs). For *16s rRNA* barcoding only clades A1, A6, A9 and A10 were detected. d) Comparison of *Methylobacterium* diversity assessment from *rpoB* barcoding (Y-axis) and isolation (X-axis). Number of *Methylobacterium* isolates in function of ASV relative abundance assuming 98.5% of nucleotide identity between *rpoB* sequences obtained by SANGER sequencing in isolates and *rpoB* sequences of ASVs (maximum of 6 nucleotide mismatches). Points are colored according to assignement to clades (legend on top right). The proportion of unmatched diversity (no match between ASV and isolate sequences) is displayed in pie charts for ASVs (top left) and isolates (bottom right).


a
Methylocystaceae-like
Beijerinckiaceae
Lichenibacteriaceae
RHIZOBIALES
Microvirga
Enterovirga
Methylobacteriaceae
Methylobacterium
Bradyrhizobiaceae
CAULOBACTERALES
SPHINGOMONADALES
RHIZOBIALES
Methylobacteriaceae
Lichenibacteriaceae
Beijerinckiaceae
Bradyrhizobiaceae
Phyllobacteriaceae
Aurantimonadaceae
Brucellaceae
Hyphomicrobiaceae
Methylocystaceae
Rhizobiaceae
Unknown
SPHINGOMONADALES
CAULOBACTERALES
Unknown
0.05
b
Methylobacterium
A4
B
A6
A10
A7/A8
A3
A2
A9
A1
A5
C
Microvirga
Enterovirga
0.05
c
0 0.4 0.8
Methylobacterium sequence relative abundance
Samples
Sample origin
MSH
SBL
METH (control)
16s-B
16s-A10
rpoB-A10
16s-Unk
16s-A2
16s-A3
16s-A4
16s-A1
rpoB-A1
rpoB-A6
16s-A6
16s-A9
rpoB-A9
ASVs abundances pooled by clade
d
rpoB Barcoding
Unmatched diversity
(14.13%)
Clades
A1
A10
A2
A3
A4
A5
A6
A9
B
Unk.
Isolates
Unmatched diversity
(7.19%)
Methylobacterium sequence relative abundance (rpoB barcoding)
1e-01
5e-02
1e-02
5e-03
1e-03
5e-04
1e-04
5e-05
1 2 5 10 20 50
Methylobacterium isolates
(rpoB SANGER sequencing)