Bacterial phylogeny predicts volatile organic compound composition and olfactory response of an aphid parasitoid

Tim Goelen[1], Islam S. Sobhy[1,2], Christophe Vanderaa[3], Felix Wäckers[4], Hans Rediers[1], Tom Wenseleers[3], Hans Jacquemyn[5] and Bart Lievens[1]

[1]Laboratory for Process Microbial Ecology and Bioinspirational Management (PME&BIM), Dept of Microbial and Molecular Systems, KU Leuven, BE-3001 Leuven, Belgium

[2]Dept of Plant Protection, Faculty of Agriculture, Suez Canal Univ., Ismailia, Egypt

[3]Laboratory of Socio-Ecology & Social Evolution, Biology Dept, KU Leuven, Leuven, Belgium

[4]Biobest, Westerlo, Belgium, and: Lancaster Environment Centre, Lancaster Univ., Lancaster, UK

[5]Laboratory of Plant Conservation and Population Biology, Biology Dept, KU Leuven, Leuven, Belgium.

Corresponding author: Bart Lievens, Laboratory for Process Microbial Ecology and Bioinspirational Management (PME&BIM), Dept of Microbial and Molecular Systems, KU Leuven, BE-3001 Leuven, Belgium. E-mail: bart.lievens@kuleuven.be

Decision date: 27-May-2020

ABSTRACT

There is increasing evidence that microorganisms emit a wide range of volatile compounds (mVOCs, microbial volatile organic compounds) that act as insect semiochemicals, and therefore play an important role in insect behaviour. Although it is generally believed that phylogenetically closely related microbes tend to have similar phenotypic characteristics and therefore may elicit similar responses in insects, currently little is known about whether the evolutionary history and phylogenetic relationships among microorganisms have an impact on insect-microbe interactions. In this study, we tested the hypothesis that phylogenetic relationships among 40 *Bacillus* strains isolated from diverse environmental sources predicted mVOC composition and the olfactory response of the generalist aphid parasitoid *Aphidius colemani*. Results revealed that phylogenetically closely related *Bacillus* strains emitted similar blends of mVOCs and elicited a comparable olfactory response of *A. colemani* in Y-tube olfactometer bioassays, varying between attraction and repellence. Analysis of the chemical composition of the mVOC blends showed that all *Bacillus* strains produced a highly similar set of volatiles, but often in different concentrations and ratios. Benzaldehyde was produced in relatively high concentrations by strains that repel *A. colemani*, while attractive mVOC blends contained relatively higher amounts of acetoin, 2,3-butanediol, 2,3-butanedione, eucalyptol and isoamylamine. Overall, these results indicate that bacterial phylogeny had a strong impact on mVOC compositions and as a result on the olfactory responses of insects.

Keywords: *Aphidius colemani*, *Bacillus*, phylogenetic signal, *rpoB*, semiochemical, VOCs

Introduction

Microorganisms release a plethora of volatiles (further referred to as mVOCs, microbial volatile organic compounds), many of which play an important role in intra- and inter-kingdom interactions (Schulz-Bohm et al. 2017, Tilocca et al. 2020). For example, it has recently been shown that mVOCs act as insect semiochemicals that affect insect behaviour (Leroy et al. 2011b, Davis et al. 2013). In some cases, mVOCs strongly attract insects by signalling the presence of suitable resources such as appropriate food sources or oviposition sites (Leroy et al. 2011a, Becher et al. 2012, Rering et al. 2018, Sobhy et al. 2018, 2019). This chemical communication between insects and microorganisms is believed to drive a mutualistic relationship, in which not only the insects profit from the microorganisms, but also the microorganisms benefit from the insects by being dispersed to new niches where they can continue to develop or complete their life cycle (Um et al. 2013, Christiaens et al. 2014). Nevertheless, some mVOCs have been found to repel insects (Burkepile et al. 2006, Stensmyr et al. 2012, Goelen et al. 2020).

In most types of ecological interactions (e.g. plant-pollinator interactions, seed-disperser interactions), both ecological correlates and evolutionary history can have a strong impact on the structure of the interactions (Rezende et al. 2007, Gómez et al. 2010). For microorganisms, however, phylogeny has been argued to be a less accurate indicator of an organism's ecology owing to evolutionary processes more common to microorganisms such as gene loss, horizontal gene transfer and convergent evolution (Morrisey et al. 2016). Nonetheless, recent studies have indicated that in some cases the phylogeny of microorganisms reflected phenotypic characteristics and key functional traits (Martiny et al. 2013, 2015, Goberna and Verdú 2016), and therefore can be used to predict how microbes interact with their living and non-living environment (Nuismer and Harmon 2015).

Despite an increased understanding of the role of mVOCs as insect semiochemicals, little is still known about whether microbial phylogeny predicts mVOC composition and the associated response of insects. Recent research suggests that production of insect-attracting metabolites is a basic and general trait in yeasts, and not limited to a specific type of yeast metabolism. For example, odours released by nine phylogenetically and ecologically distinct yeast species were all found to be attractive to the vinegar fly *Drosophila melanogaster* (Becher et al. 2018). In line with this study, Madden et al. (2018) proposed a 'dispersal–encounter hypothesis' whereby yeasts produce insect-attracting mVOCs to get dispersed between ephemeral, spatially disparate sugar resources, and the insects, in turn, obtain the benefits of an honest signal from the yeasts for the sugar resources. Similar mutualistic interactions may occur between insects and bacteria, although compelling evidence is still scant (Leroy et al. 2011b).

Recent Y-tube olfactometer bioassays have shown that olfactory responses of aphid parasitoids (*Aphidius colemani* Viereck, Hymenoptera: Braconidae) to volatiles from bacteria occurring in the habitat of the parasitoids differed strongly between bacterial strains, ranging from significant attraction to repellence (Goelen et al. 2020). However, to what extent phylogenetic relationships between bacteria mediate mVOC composition and the olfactory response of parasitoids is less clear. In this study, we tested the hypothesis that bacterial phylogeny predicts microbial volatile composition and the olfactory response of insects. Therefore, we used a large set of *Bacillus* strains isolated from a variety of habitats and the generalist aphid parasitoid *Aphidius colemani*. The genus *Bacillus* represents a heterogeneous group of ubiquitous Gram-positive bacteria that occupy diverse ecological systems and have been isolated from various habitats, including soil, water, dust, air, insects and plant-associated habitats such as roots, leaves, honeydew and nectar (Logan and De Vos 2009). Members of the genus

Bacillus produce a wide array of volatiles (Kai et al. 2009), some of which can promote plant growth without physical contact (Ping and Boland 2004) or have antimicrobial activity (Gao et al. 2017). Furthermore, it has been shown that mVOCs produced by *Bacillus* strains may affect insect behaviour. For example, Rockett (1987) and Poonam et al. (2002) showed that *Bacillus* volatiles stimulate oviposition in gravid *Culex* (Diptera: Culicidae) females. Additionally, the melon fruit fly *Bactrocera cucurbitae* Coquillett (Diptera: Tephritidae) and the parasitoid *A. colemani* have been shown to be attracted to mVOC blends emitted by *Bacillus* spp. (Mishra and Sharma 2018, Goelen et al. 2020). Specifically, we first analysed the chemical composition of the mVOC blends produced by 40 phylogenetically diverse *Bacillus* strains using gas chromatography-mass spectrometry (GC-MS). Subsequently, we evaluated the olfactory response of *A. colemani* to these blends in a Y-tube olfactometer bioassay. Finally, we tested whether phylogenetic relationships between *Bacillus* strains could accurately predict mVOC composition and insect response using both univariate and multivariate tests for phylogenetic signal and trait correlations.

Materials and Methods

Study organisms

A collection of 40 *Bacillus* strains was used in this study, representing (at least) 20 different species (based on *rpoB* sequence similarity with identified strains), including ten type strains (Table 1). About half of the strains were isolated from sources associated with the habitat of *A. colemani*, including aphids, honeydew and nectar, as well as from *Aphidius* wasps and one of its hyperparasitoids (*Dendrocerus aphidum*). The other half represented species that may occur in the parasitoid's habitat (Logan and De Vos 2009, Samuni-Blank et al. 2014, Miliute et al. 2015), but originated from other habitats like wooden barrels, soil and soil conditioners, and sludge and sediment

samples (Table 1). All isolates were stored in tryptic soy broth (TSB; Oxoid, Hampshire, UK) containing 25% glycerol at -80°C until use. Olfactory responses were investigated using female adults of *A. colemani* that had been obtained in the form of parasitized aphid mummies from Biobest (Westerlo, Belgium) (Aphidius-system®). Mummies were placed inside a nylon insect cage (20 cm × 20 cm × 20 cm, BugDorm, MegaView Science Co., Ltd., Taichung, Taiwan) and kept under controlled conditions (22°C, 70% RH, 16L:8D-h) until parasitoid emergence. All behavioural experiments were performed with <24-h-old, food- and water-starved females.

Production of mVOCs

For production of mVOCs the procedure by Goelen et al. (2020) was followed. Briefly, bacterial stock cultures were plated on tryptic soy agar (TSA; Oxoid, Hampshire, UK) and incubated at 25°C for 24h, followed by a re-streak on TSA and incubation at 25°C for another 24h. Subsequently, single colonies were inoculated in 10 mL TSB and incubated overnight at 25°C in a rotary shaker at 120 rpm. Next, cells were washed and diluted to a cell suspension with an optical density (OD 600 nm) of 1, of which 1.5 mL was inoculated in a 250 mL Erlenmeyer flask containing 150 mL GYP25 medium. The medium was prepared by filter-sterilizing (pore size 0.22 µm; Rapid-Flow™, Thermo Scientific, Waltham, USA) a solution of 5% w/v glucose, 0.5% w/v peptone (Bacto™ Peptone, BD Biosciences, San Jose, USA) and 0.25% w/v yeast extract (Sigma-Aldrich, Saint Louis, USA). Erlenmeyer flasks were covered with sterile silicone plugs and incubated at 25°C in an orbital shaker at 120 rpm. Cultivation for each strain was carried out in triplicate, and non-inoculated, blank medium was included as a control. After an incubation period of 48h, the media were centrifuged at 10,000 g for 15 min, and filter-sterilized to obtain cell-free culture supernatants. The samples were then stored in small aliquots in sterile, amber glass vials at -20°C until further use.

Chemical analysis of mVOCs

For all samples, mVOC composition of the cell-free media was determined by headspace solid phase micro extraction gas chromatography followed by mass spectrometry detection (HS-SPME-GC-MS). Volatile compounds were separated by a Thermo Trace 1300 GC system (Thermo Fisher Scientific, Watham, USA) equipped with a MXT-5 column (30 m length × 0.18 mm inner diameter × 0.18 µm film thickness; Restek, Bellefonte, USA), and detected by a ISQ mass spectrometer (Thermo Fisher Scientific, Waltham, USA). An amount of 1.75 g NaCl was added to 5 mL of the samples and was kept at 60°C under constant agitation in a TriPlus RSH SMPE auto sampler (Thermo Fisher Scientific, Watham, USA). Volatile collection and separation conditions were as described previously (Goelen et al. 2020), and identification and quantification of the compounds was performed as in Reher et al. (2019). Briefly, chromatograms were analysed using AMDIS v2.71 (Stein 1999) to deconvolute overlapping peaks. The NIST MS Search v2.0g software in combination with the NIST2011, FFNSC and Adams libraries were used to manually identify the empirical spectra, taking into account the expected retention time. The compound elution profiles were extracted and integrated using a file with all our target compounds containing the expected retention times and spectrum profiles. The extraction was performed for every peak in every chromatogram over a time restricted window using weighted non-negative least square analysis (Lawson and Hanson 1995). Finally, the peak areas were computed from the extracted profiles and summarized in a single table.

Olfactometer bioassays

Insect olfactory response was evaluated using the procedure described previously (Sobhy et al. 2018, Goelen et al. 2020). A glass Y-tube olfactometer (base: 20 cm;

arms:12 cm with a 60° angle at the Y-junction; inner diameter: 1.5 cm), connected to an air pump producing a unidirectional airflow of 400 mL/min from the arms to the base, was placed on a table that was homogeneously illuminated by four 24W T5 TL-fluorescent tubes (16 x 549 mm, 1350 Lumen, 5500K; True-Light®, Naturalite Benelux, Ansen, The Netherlands) at a height of 45 cm. To avoid any visual distraction, a white curtain was placed around the olfactometer. Additionally, to improve insect responsiveness, the olfactometer was mounted at a 20° incline, by which the insects were stimulated to move towards the bifurcation.

For each bacterial strain, 150 µL of the cell-free cultivation medium was loaded on a filter paper (37 mm; Macherey-Nagel, Düren, Germany) and subsequently put in one of the olfactometer odour chambers. The second chamber received another filter paper on which 150 µL blank medium was loaded as a control. The bioassay was carried out by releasing twelve groups of five adult females (n = 60) at the base of the olfactometer and assessing their response after 10 min. Individuals that had walked at least 1 cm into one of the arms at the time of evaluation were considered to have made a choice for the odour source introduced by that olfactometer arm. If the released females did not make a choice within a 10 min period, they were considered as 'non-responders' and were discarded from statistical analysis. For every release, new parasitoid females were used. The filter papers inside the odour chambers were renewed after every two releases. Further, the Y-tube olfactometer arms were flipped 180° after each six releases to minimize any spatial effect on parasitoid choice. At the same time, the Y-tube was also renewed by a cleaned Y-tube. At the end of the assay, all olfactometer parts were thoroughly rinsed with tap water, distilled water, acetone and finally pentane, after which the parts were kept overnight at 150°C. All bioassays were conducted at 21 ± 2°C, 60 ± 5% RH and performed between 09:00 and 16:00 h. As the mVOC

composition of the three biological replicates was highly similar, olfactory response was determined for one of the three biological replicates.

Data analysis

Chemical analysis

To visualize the differences in the mVOC composition, a non-metric multidimensional scaling (NMDS) was performed on the strain*average volatile peak area matrix by using a Bray-Curtis distance matrix (Vegan package) in R 3.3.2 (R Core Team 2019). To test for significant differences in chemical composition of mVOC profiles between bacterial clades (see further, six major clades were identified) and between groups of strains eliciting different olfactory responses in *A. colemani* (i.e. attractive, neutral or repellent), a two-way permutational multivariate analysis of variance (perMANOVA) was carried out on the strain*volatile peak area matrix. Volatile peak area was used as dependent variable, and bacterial clade and groups of strains evoking a similar olfactory response in *A. colemani* were used as independent variables. Statistical significance was estimated using 1000 permutations. This analysis was performed using the adonis function (Vegan package) in R 3.3.2 (R Core Team 2019).

Olfactometer bioassays

For each bacterial strain, parasitoid olfactory response was analysed using a Generalized Linear Mixed Model (GLMM) based on a binomial distribution (choice is binary: for either control side or treatment side) with a logit link function (logistic regression) using *Bacillus* strain as fixed factor (performed in R with the 'glmer' function from the lme4 package). Each release of one cohort of five individuals served as a replicate. To adjust for overdispersion and to prevent pseudoreplication, the release of each cohort (n = 12) was included in the model as a random factor. The number of parasitoids choosing for

the control or treatment side in each cohort was entered as response variable. To examine the preference of the investigated parasitoids for mVOCs produced by each of the tested *Bacillus* strains, we tested the null hypothesis (H_0) that the parasitoids show no preference for any olfactometer arm (i.e. 50:50 response) by testing H_0: logit = 0 which equals a 50:50 distribution. In addition, an analysis of variance Type III Wald chi-square test was performed on the GLMM to determine if there was an overall difference between the olfactory responses of all tested *Bacillus* strains. Results were presented by calculating the Preference Index (PI) by dividing the difference between the number of parasitoids choosing the bacterial odours and the parasitoids choosing the control by the total number of responding insects.

Testing for phylogenetic signal

Both multivariate and univariate analyses were performed to evaluate the presence of a phylogenetic signal in mVOC composition and insect behaviour. First, phylogenetic principal component analysis (pPCA) was used to evaluate the presence of phylogenetic signal in the complete, multivariate volatile composition dataset. pPCA is a method to summarise and visualise the phylogenetic resemblance of a multivariate trait dataset into two principal components (PCs) showing global or local phylogenetic structures (Jombart et al. 2010). The first PC denotes the global structure and reveals the mVOCs that are more similar in related strains than in more distant strains. The local structure is depicted in the second PC which reveals the mVOCs that create dissimilarities among closely related strains (Jombart et al. 2010).

In order to test for statistical significance of the phylogenetic dependence of the PCs, Pagel's λ (Pagel 1999), Abouheif's C_{mean} (Abouheif 1999) and Moran's I (Pavoine et al. 2008) were calculated. All three indices are suited to correctly identify phylogenetic signals, even at moderate Brownian motion of traits and low sample sizes

(Münkemüller et al. 2012). A Mantel test was used to test for correlations between the phylogenetic distance matrices and the distance matrices derived from the PCs from the pPCA. The PC distance matrix was calculated with Euclidean distances using the 'dist' function in R 3.3.2 (R Core Team 2019) while phylogenetic distances were measured by selecting the complement of Abouheif proximity. Additionally, presence of phylogenetic signal was tested for each mVOC individually by calculating the same three indices and performing separate Mantel tests. To correct for multiple testing, *P*-values were adjusted using the false discovery rate (FDR) correction method. Finally, the same analyses were done to evaluate any potential phylogenetic signal in the olfactory response of *A. colemani*, using the behavioural data expressed in PI-values as dependent variable.

In all calculations, a maximum-likelihood (ML) tree based on concatenated sequences of partial 16S rRNA gene (1252-1262 bp) and partial *rpoB* sequences (1102-1105 bp) was used. PCR amplification, sequencing and sequence alignment were performed as described previously (Bosmans et al. 2015), using the primers mentioned in Table A1 (Supplementary material Appendix). The tree was constructed by implementing the Kimura-2 model (gamma distributed with invariant sites (G+I)) with 1000 bootstraps. The branch length information of the ML tree was included in all analyses of phylogenetic signal and depiction of phylogenetic trees. All analyses were performed with functions in the packages 'adephylo' (Jombart et al. 2010), 'ape' (Paradis et al. 2004) and 'phytools' (Revell 2012) in R 3.3.2 (R Core Team 2019) using the volatile composition data in the form of mean centred, log transformed peak areas. *P* values < 0.05 were regarded as statistically significant. Finally, a phylogenetic heatmap was constructed using the ML tree and the volatile composition data in the form of mean centred, log transformed peak areas of compounds that showed phylogenetic

signal in all three indices and the Mantel test, using the 'phylo.heatmap' function (Phytools package) in R 3.3.2 (R Core Team 2019).

Results

mVOC composition

In total, 159 compounds were detected in the headspace of the tested cell-free media, including aldehydes, alkanes, amines, aromatics, esters, ketones, organic acids, pyrazines and terpenoids (Fig. A1, Supplementary material Appendix). On average, the investigated *Bacillus* strains produced highest amounts of acetoin, benzaldehyde and 3-methyl-butanoic acid. While most compounds were produced by all strains tested, allyl acetate, ammonium acetate and 4-methyl-pentan-2-one were only produced by a small subset of the investigated strains (Fig. A1, Supplementary material Appendix). NMDS ordination of the mVOC composition (Fig. 1) showed a clear clustering pattern, which corresponded well to the phylogenetic position of the strains. Phylogenetic analyses grouped the tested bacteria into six well-supported clades (bootstrap values between 93 and 100) (Fig. 2), and mVOC composition differed significantly between bacterial clades (perMANOVA: pseudo-F5 = 10.7, P < 0.001).

Olfactory response

Olfactory response of *A. colemani* to the volatile emissions of the tested *Bacillus* strains varied significantly between strains ($\chi^2_{(39)}$ =116.75, P < 0.001; Fig. 2; Table A2, Supplementary material Appendix). Female parasitoids significantly preferred the mVOC blends of ten *Bacillus* strains over the blank medium. These included strains that were not only isolated from sources from the parasitoid's habitat (floral nectar, honeydew, aphids, *Aphidius* parasitoids and the hyperparasitoid *Dendrocerus aphidum*),

but also from sources that were not associated with the parasitoid, i.e. whisky barrels (Table 1). By contrast, for five strains, parasitoids significantly preferred the blank medium (Fig. 2; Table A2, Supplementary material Appendix). Again, this was not related to the habitat from which the strains were collected (Table 1). Strain ST12.14/237 (having almost 99% *rpoB* sequence identity with *B. velezensis* and *B. amyloliquefaciens*) yielded the highest PI-value (PI = 0.42), while *B. flexus* DSM 1316 elicited a strong significantly repellent response in *A. colemani* (PI = -0.55). The mVOC blends of the remaining 25 strains had no statistically significant effect on the olfactory response of *A. colemani*, and were further regarded as "neutral" responses. However, in general, strains belonging to clades A, B, C and D yielded positive PI-values (exception: ST12.14/138, PI = -0.05), whereas strains from clade E and clade F elicited a negative PI-value, except for *B. muralis* LMG 20238 (PI = 0.25) (Fig. 2). The perMANOVA revealed that there was a statistically significant difference in the mVOC chemical composition correlating with the response of *A. colemani* (perMANOVA: pseudo-F_2 = 10.3, P < 0.001). Although all tested strains produced a highly similar set of volatiles, benzaldehyde was produced consistently in relatively higher concentrations by strains that repelled *A. colemani*, while attractive mVOC blends contained relatively higher amounts of acetoin, 2,3-butanediol, 2,3-butanedione, eucalyptol and isoamylamine (Fig. A1, Supplementary material Appendix). Furthermore, the interaction between the independent variables olfactory response and bacterial clade was also significant (perMANOVA: pseudo-F_2 = 4.06, P = 0.003), indicating that the response of *A. colemani* to *Bacillus* mVOCs depends on the clade to which the strain belongs.

Phylogenetic signal analysis

From the pPCA of the 159 mVOCs detected (Fig. 3), the two PCs with the most positive and negative eigenvalues were retained and kept in the analysis, i.e. PC1 and PC39, showing the global and local phylogenetic structure, respectively (Fig. 3A). Together, these PCs explained 43.4% of the total variance in mVOC composition. The first PC (PC1) explained 26.8% of the variance, while the second PC (PC39) explained 16.6% of the variance (Fig. 3B). The global structure (PC1), revealing the mVOCs that are more similar in related than in distant strains, was determined by a large number of compounds. The most important compounds driving the global structure were acetoin, 2,3-butanedione, isoamylamine, nonan-2-ol, tetrametyl-pyrazine and 2,3-butanediol (highest positive loadings), and methyl-methacrylate, 2-ethyl-hexanoic acid, 2-methyl-propanoic acid, cyclohexanone and phenylacetaldehyde (highest negative loadings) (Fig. 3B). The most important compounds associated with the local structure (PC39), corresponding to compounds that create dissimilarities between related strains, were cis-2-tert-butyl-cyclohexanol acetate (highest positive loading) and acetone, 2,6-di-tert-butyl-P-benzoquinone, n-decanal, trimethylamine and 3-hydroxy-3-methyl-2-butanone (highest negative loadings) (Fig. 3B). Strains from clade A showed the highest scores of PC1, meaning that they produced the highest amounts of the compounds with highest positive loadings such as acetoin and 2,3-butanedione (Fig. 3C). Especially, strains related to *B. amyloliquefaciens / B. velezensis* produced the highest amounts of these compounds. Conversely, strains from clade E and F were found to have the highest negative scores for PC1. This clustering of positive and negative scores of PC1 into the different clades suggests a clear phylogenetic pattern in the mVOC emissions of *Bacillus*. Indeed, the univariate analysis of phylogenetic signal on PC1 was significant for all three indices as well as for the Mantel test (Table 2). This was further confirmed by the multivariate Mantel test on the complete mVOC dataset, which showed a significant correlation between phylogenetic distances and relative amounts of mVOCs

produced by the *Bacillus* strains (Mantel $Z = 6.29\times10^{10}$, $P < 0.001$) (Table 2). Further, there was also a phylogenetic signal in the olfactory response of *A. colemani* to the volatile emissions of *Bacillus* strains (Table 2). When focusing on the individual compounds, it becomes clear that the phylogenetic position of the *Bacillus* strains reflects the production of certain compounds. More specifically, significant phylogenetic signal was found for 30.2%, 27.7%, 26.4% and 26.4% of the compounds detected for Pagels's λ, Moran's I, Abouheif's C_{mean}, and the Mantel Test, respectively (Table 2; Table A3, Supplementary material Appendix). In total, 15.7% of the mVOCs detected (25 compounds) were statistically significant according to all three indices and the Mantel Test (Fig. 4).

Discussion

Phylogenetic conservatism of mVOC profiles

In this study, we have shown that phylogenetically related *Bacillus* strains tended to emit similar blends of mVOCs and elicited a comparable olfactory response in *A. colemani*. The observed olfactory response was largely in agreement with results of a previous study where a smaller set of *Bacillus* strains was tested against the same parasitoid (Goelen et al. 2020), indicating the robustness of our bioassay and results. Furthermore, we identified the key mVOCs that significantly explained the phylogenetic signal. In total, 15.7% of the detected mVOCs showed a significant phylogenetic signal for all statistical tests performed (i.e. Pagel's λ, Moran's I, Abouheif's C_{mean} and Mantel test). Among these, acetoin, 2,3-butanediol and 2,3-butanedione were among the most important drivers of the global structure in the pPCA. These compounds represent typical fermentation products, and are well known to be produced by *Bacillus* species (Dettwiler et al. 1993, Asari et al. 2016). These compounds were particularly produced by a phylogenetically related subset of *Bacillus*

strains, i.e. strains from clade A and B, whereas they were absent or only produced in low amounts by strains belonging to clade C to clade F. However, it is important to note that the compounds detected in this study were only tentatively identified through GC-MS analysis and that the precise composition of volatile blends emitted by microorganisms may depend on the nutrient medium used to cultivate them (Gonzalez et al. 2019). GC peak enhancement with co-injection of authentic standards should be performed to achieve absolute identification of the compounds and analyses ideally should be repeated for different media to investigate how culturing conditions affect the composition of mVOC blends and associated response of parasitoids. Under the prevailing culturing conditions, our results further showed that several mVOCs significantly contributed to the local structure in the pPCA which cannot be explained by phylogeny. Instead, such patterns could arise from evolutionary events, such as gene loss, horizontal gene transfer and interaction with the environment, leading to differences between closely related strains (Goberna and Verdú 2016).

Insect response

Phylogenetic relationships of the investigated *Bacillus* strains predicted the olfactory response of *A. colemani*. While strains belonging to clade A, B, C and D produced attractive mVOC blends, strains belonging to clade E and F produced blends that were repellent to *A. colemani*. Closer inspection of the chemical composition of these mVOC blends revealed that all *Bacillus* strains produced a highly similar set of volatiles, but often in significantly different concentrations and ratios. This suggests that mVOCs elicit a different response in the insects depending on the presence of background mVOCs and the mVOC concentrations (Bruce et al. 2005, Mumm and Hilker 2005). Benzaldehyde was produced in relatively high concentrations by strains that repel *A. colemani*, and in lower amounts by strains that elicit attractive behaviour, and is

therefore an important candidate to explain the differences observed for both groups of bacteria. Previous research has shown that benzaldehyde can induce both electrophysiological and behavioural responses in several insects (James 2005), including *Aphidius* and other parasitoids and aphid natural enemies (Han and Chen 2002, Simpson et al. 2011). A preliminary coupled gas chromatography-electroantennography (GC-EAG) analysis of the mVOC blends produced by strains ST18.16/133 and ST18.16/150 revealed that benzaldehyde, amongst a number of others compounds, elicited an electrophysiological response in *A. colemani*. Further testing of the olfactory response of *A. colemani* against different concentrations of synthetic benzaldehyde confirmed that high concentrations yield a negative response, while concentrations lower or equal to 100 ng evoked parasitoid attraction (T. Goelen, unpublished results). A similar, albeit less pronounced, phylogenetic pattern was observed for phenylacetaldehyde, which has also been shown to affect insect behaviour (Dötterl et al. 2006, Bruce and Pickett 2011).

Attraction does not necessarily have to be caused by the most abundant compounds as has been shown for plant volatiles, where often minor constituents of volatile blends affect insect behaviour (McCormick et al. 2014). It has been shown that volatile concentration is a very important factor in affecting the behavioural response of arthropods, since the same compounds can evoke a different response depending on the concentration (Gadino et al. 2012). Therefore, concentrations of mVOCs produced by repellent strains can be too high and cause repellence or even mask otherwise attractive compounds (Aartsma et al. 2017). Moreover, certain compounds can induce or inhibit the response to other compounds (Turner and Ray 2009). Conversely, attractive mVOC profiles, in particular the profiles emitted by strains belonging to clade A and B, contained relatively higher amounts of acetoin, 2,3-butanediol, 2,3-butanedione, eucalyptol and isoamylamine when compared to repellent mVOC profiles. All these

volatile compounds have been shown to affect insect behaviour and even elicit insect attraction (Bengtsson et al. 2009, Turner et al. 2011, Kuhns et al. 2014, Rebora et al. 2017). Nevertheless, although our data suggest that these compounds may have contributed to the observed insect behaviour, additional experiments using GC-EAG and pure compounds are required to confirm the effects of these mVOCs on parasitoid olfactory response.

Parasitoids responding to ubiquitous bacteria

Under the prevailing culturing conditions, the olfactory response of *A. colemani* was strongly correlated to the phylogeny of the tested *Bacillus* strains and not to their source of origin, suggesting that *A. colemani* also responds to bacteria which may never encounter the parasitoid naturally. Strains from the clades that were attractive to *A. colemani* represented isolates that were or were not associated with the parasitoid's environment. Likewise, strains from the repellent clades originated from sources that were associated or not associated with the habitat of *A. colemani*. The question then rises why ubiquitous *Bacillus* species produce volatiles that affect parasitoids and why *A. colemani* responds to them. Most mVOCs are considered as side-products of primary and secondary metabolism, and are formed mainly by oxidation of glucose from various intermediates (Korpi et al. 2009). The fact that they can act as insect attractants could be a merely coincidental interaction as *Bacillus* species are known to produce volatile products that are typically associated with plants (e.g. eucalyptol, geraniol, limonene and phenylacetaldehyde) and insects (e.g. acetoin, 2,3-butanediol, nonan-2-ol) (Knudsen et al. 1993, Löfsteft et al. 2008, Bruce and Pickett 2011), and thereby may accidentally mimic host plant, food or insect host associated cues. This overlap between plant and microbial volatiles has recently also been observed in yeasts, which emit insect attracting volatiles typically associated with flowers (Ljunggren et al. 2019).

However, there could also be a deeper ecological association between bacteria and insects, as is, for example, the case for yeasts. Collective evidence suggests that yeast volatiles mediate mutualistic interactions between yeasts and insects, in which the insects exploit the mVOCs as semiochemicals to detect suitable oviposition sites and food sources and even use them to discriminate between sources which best support their growth and survival (Scheidler et al. 2015, Becher et al. 2018, Madden et al. 2018, Rering et al. 2018). In turn, the yeasts may benefit from the insects as vectors to disperse to new habitats (Christiaens et al. 2014) or to survive unfavourable conditions (Pozo et al. 2018). However, further research is needed to investigate these scenarios for bacteria. Although microbial emissions may signal a number of advantages for insects, responding to ubiquitous microbes like *Bacillus* spp. may also pose potential caveats for parasitoids. Optimal foraging assumes that insects only respond to signals from which they benefit. Future experiments should therefore also be performed to investigate to what extent responding to semiochemicals from widespread microorganisms may pose potential trade-offs in parasitoid foraging success.

Conclusion

In conclusion, we have shown that the phylogeny of *Bacillus* species predicted both mVOC composition and the olfactory response of an aphid parasitoid, *A. colemani*. A specific subset of mVOCs was the main driver for the phylogenetic signal in *Bacillus*, which are possible candidates facilitating olfactory response in *A. colemani*, as these volatiles describe the difference between attractive and repellent clades.

Declarations

Acknowledgements

We would like to thank Sergio Álvarez-Pérez for his assistance in the statistical analyses, and Maarten Blockhuys and Murat Bakırdöven who aided in the course of the experiments.

Funding

This work was supported by the Flemish Research Foundation (FWO) [project 1S15116316N].

Author contributions

The first, second, fourth, fifth, seventh and eighth author conceived the ideas and designed methodology. The first and third author collected the data. The sixth author contributed to equipment and reagents, and contributed to the mVOC analysis. The fourth author provided insects required for experiments. The first, third, sixth, seventh and eighth author analysed the data. The first, seventh and eighth author led the writing of the manuscript. All authors contributed critically to the drafts of this manuscript and gave final approval for publication.

Competing interests

The authors have declared that no competing interests exist.

References

Aartsma, Y. et al. 2017. Herbivore-induced plant volatiles and tritrophic interactions across spatial scales. - New Phytol. 216: 1054–1063.

Abouheif, E. 1999. A method to test the assumption of phylogenetic independence in comparative data. - Evol. Ecol. Res. 1: 895–909.

Asari, S. et al. 2016. Multiple effects of *Bacillus amyloliquefaciens* volatile compounds : plant growth promotion and growth inhibition of phytopathogens. - FEMS Microbiol. Ecol. 92: fiw070.

Becher, P. G. et al. 2012. Yeast, not fruit volatiles mediate *Drosophila melanogaster* attraction, oviposition and development. - Funct. Ecol. 26: 822–828.

Becher, P. G. et al. 2018. Chemical signaling and insect attraction is a conserved trait in yeasts. - Ecol. Evol. 8: 2962–2974.

Bengtsson, J. M. et al. 2009. Field attractants for *Pachnoda interrupta* selected by means of GC-EAD and single sensillum screening. - J. Chem. Ecol. 35: 1063–1076.

Bosmans, L. et al. 2015. Assessment of the genetic and phenotypic diversity among rhizogenic Agrobacterium biovar 1 strains infecting solanaceous and cucurbit crops. - FEMS Microbiol. Ecol. 91: 1–16.

Bruce, T. J. A. and Pickett, J. A. 2011. Perception of plant volatile blends by herbivorous insects – Finding the right mix. - Phytochemistry 72: 1605–1611.

Bruce, T. J. A. et al. 2005. Insect host location : a volatile situation. - Trends Plant Sci. 10: 269–274.

Burkepile, D. E. et al. 2006. Chemically mediated competition between microbes and animals: Microbes as consumers in food webs. - Ecology 87: 2821–2831.

Christiaens, J. F. et al. 2014. The fungal aroma gene *ATF1* promotes dispersal of yeast cells through insect vectors. - Cell Rep. 9: 425–432.

McCormick, A. et al. 2014. Little peaks with big effects: Establishing the role of minor plant volatiles in plant-insect interactions. - Plant, Cell Environ. 37: 1836–1844.

Davis, T. S. et al. 2013. Microbial volatile emissions as insect semiochemicals. - J. Chem. Ecol. 39: 840–859.

Dettwiler, B. et al. 1993. A simulation model for the continuous production of acetoin and butanediol using *Bacillus subtilis* with integrated pervaporation separation. – Biotechnol. Bioeng. 41: 791–800.

Dötterl, S. et al. 2006. Nursery pollination by a moth in *Silene latifolia*: The role of odours in eliciting antennal and behavioural responses. - New Phytol. 169: 707–718.

Gadino, A. N. et al. 2012. Olfactory response of *Typhlodromus pyri* (Acari : Phytoseiidae) to synthetic methyl salicylate in laboratory bioassays. - J. Appl. Entomol. 136: 476–480.

Gao, Z. et al. 2017. Identification of endophytic *Bacillus velezensis* ZSY-1 strain and antifungal activity of its volatile compounds against *Alternaria solani* and *Botrytis cinerea*. - Biol. Control 105: 27–39.

Goberna, M. and Verdú, M. 2016. Predicting microbial traits with phylogenies. - ISME J. 10: 959–967.

Goelen, T. et al. 2020. Volatiles of bacteria associated with parasitoid habitats elicit distinct olfactory responses in an aphid parasitoid and its hyperparasitoid. - Funct. Ecol. 34: 507-520.

Gómez, J. M. et al. 2010. Ecological interactions are evolutionary conserved across the entire tree of life. - Nature 465: 918–921.

Gonzalez, M. et al. 2019. Yeast smell like what they eat: Analysis of volatile organic compounds of *Malassezia furfur* in growth media supplemented with different lipids. - Molecules 24: 419.

Han, B. and Chen, Z. 2002. Behavioral and electrophysiological responses of natural enemies to synomones from tea shoots and kairomones from tea aphids, *Toxoptera aurantii*. - J. Chem. Ecol. 28: 2203–2219.

James, D. G. 2005. Further field evaluation of synthetic herbivore-induced plant volatiles as attractants for beneficial insects. - J. Chem. Ecol. 31: 481–495.

Jombart, T. et al. 2010. adephylo: New tools for investigating the phylogenetic signal in biological traits. - Bioinformatics 26: 1907–1909.

Kai, M. et al. 2009. Bacterial volatiles and their action potential. - Appl. Microbiol. Biotechnol. 81: 1001–1012.

Knudsen, J. T. et al. 1993. Floral scents-a checklist of volatile compounds by head-space techniques. - Phytochemistry 33: 253–280.

Korpi, A. et al. 2009. Microbial volatile organic compounds. - Crit. Rev. Toxicol. 39: 139–193.

Kuhns, E. H. et al. 2014. Eucalyptol is an attractant of the redbay ambrosia beetle, *Xyleborus glabratus*. - J. Chem. Ecol. 40: 355–362.

Lawson, C. L. and Hanson, R. J. 1995. Solving least squares problems. - Siam.

Leroy, P. D. et al. 2011a. Microorganisms from aphid honeydew attract and enhance the efficacy of natural enemies. - Nat. Commun. 2: 348.

Leroy, P. D. et al. 2011b. The semiochemically mediated interactions between bacteria and insects. – Chemoecology 21: 113-122.

Ljunggren, J. et al. 2019. Yeast volatomes differentially effect larval feeding in an insect herbivore. - Appl. Environ. Microbiol. 85: e01761-19.

Löfsteft, C. et al. 2008. Identification of a sex pheromone produced by sternal glands in females of the caddisfly *Molanna angustata* Curtis. - J. Chem. Ecol. 34: 220–228.

Logan, N. A. and De Vos, P. 2009. Genus *Bacillus*. - In: De Vos, P. et al. (eds.), Bergey's Manual of Systematic Bacteriology. vol. 3. Springer, pp. 21–128.

Madden, A. A. et al. 2018. The ecology of insect– yeast relationships and its relevance to human industry. – Proc. R. Soc. B 285: 20172733.

Martiny, A. C. et al. 2013. Phylogenetic conservatism of functional traits in microorganisms. - ISME J. 7: 830–838.

Martiny, J. B. H. et al. 2015. Microbiomes in light of traits: A phylogenetic perspective. - Science 350: aac9323.

Miliute, I. et al. 2015. Bacterial endophytes in agricultural crops and their role in stress tolerance: a review. - Zemdirbyste-Agriculture102: 465–478.

Mishra, M. and Sharma, K. 2018. Characterization of culturable gut bacterial isolates from wild population of melon fruit fly (*Bactrocera cucurbitae*) and assessing their attractancy potential for sustainable pest management. - Phytoparasitica 46: 583–594.

Morrissey, E. M. et al. 2016. Phylogenetic organization of bacterial activity. - The ISME Journal 10: 2336–2340.

Mumm, R. and Hilker, M. 2005. The significance of background odour for an egg parasitoid to detect plants with host eggs. - Chem. Senses 30: 337–343.

Münkemüller, T. et al. 2012. How to measure and test phylogenetic signal. - Methods Ecol. Evol. 3: 743–756.

Nuismer, S. L. and Harmon, L. J. 2015. Predicting rates of interspecific interaction from phylogenetic trees. – Ecol. Lett. 18: 17–27.

Pagel, M. 1999. Inferring the historical patterns of biological evolution. - Nature 401: 877–884.

Paradis, E. et al. 2004. APE: Analyses of phylogenetics and evolution in R language. - Bioinformatics 20: 289–290.

Pavoine, S. et al. 2008. Testing for phylogenetic signal in phenotypic traits: New matrices of phylogenetic proximities. - Theor. Popul. Biol. 73: 79–91.

Ping, L. and Boland, W. 2004. Signals from the underground: bacterial volatiles promote growth in *Arabidopsis*. - Trends Plant Sci. 9: 263–266.

Poonam, S. et al. 2002. Oviposition attractancy of bacterial culture filtrates response of *Culex quinquefasciatus*. - Mem. Inst. Oswaldo Cruz 97: 359–362.

Pozo, M. I. et al. 2018. Surviving in the absence of flowers: do nectar yeasts rely on overwintering bumblebee queens to complete their annual life cycle? FEMS Microb. Ecol. 94: fiy196.

R Core Team 2019. A language and environment for statistical computing. - In: Cowles, R. S. et al. (eds.), R Foundation for statistical computing. <https://www.R-project.org/>.

Rebora, M. et al. 2017. Antennal responses to volatile organic compounds in a stonefly. - J. Insect Physiol. 98: 231–237.

Reher, T. et al. 2019. Evaluation of hop (*Humulus lupulus*) as a repellent for the management of *Drosophila suzukii*. - Crop Prot. 124: 104839.

Rering, C. C. et al. 2018. Nectar-inhabiting microorganisms influence nectar volatile composition and attractiveness to a generalist pollinator. – New Phytol. 220: 655-658.

Revell, L. J. 2012. phytools: An R package for phylogenetic comparative biology (and other things). - Methods Ecol. Evol. 3: 217–223.

Rezende, E. L. et al. 2007. Effects of phenotypic complementarity and phylogeny on the nested structure of mutualistic networks. - Oikos 116: 1919–1929.

Rockett, C. L. 1987. Bacteria as ovipositional attractants for *Culex pipiens* (Diptera:Culicidae). - Gt. Lakes Entomol. 20: 151–155.

Samuni-Blank, M. et al. 2014. The role of abiotic environmental conditions and herbivory in shaping bacterial community composition in floral nectar. – PLoS ONE 9: e99107.

Scheidler, N. H. et al. 2015. Volatile codes: Correlation of olfactory signals and reception in *Drosophila*-yeast chemical communication. – Sci. Rep. 5: 14059.

Schulz-Bohm, K. et al. 2017. Microbial volatiles: Small molecules with an important role in intra- and inter-kingdom interactions. - Front. Microbiol. 8: 1–10.

Simpson, M. et al. 2011. Insect attraction to synthetic herbivore-induced plant volatile-treated field crops. - Agric. For. Entomol. 13: 45–57.

Sobhy, I. S. et al. 2018. Sweet scents : nectar specialist yeasts enhance nectar attraction of a generalist aphid parasitoid without affecting survival. - Front. Plant Sci. 9: 1–13.

Sobhy, I. S. et al. 2019. Associative learning and memory retention of nectar yeast volatiles in a generalist parasitoid. - Anim. Behav. 153: 137–146.

Stein, S. E. 1999. An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data. - J. Am. Soc. Mass Spectrom. 10: 770–781.

Stensmyr, M. C. et al. 2012. A conserved dedicated olfactory circuit for detecting harmful microbes in *Drosophila*. - Cell 151: 1345–1357.

Tilocca, B. et al. 2020. Scent of a killer: microbial volatilome and its role in the biological control of plant pathogens. – Front. Microbiol. doi.org/10.3389/fmicb.2020.00041.

Turner, S. L. and Ray, A. 2009. Modification of CO_2 avoidance behaviour in *Drosophila* by inhibitory odorants. - Nature 461: 277–281.

Turner, S. L. et al. 2011. Ultra-prolonged activation of CO_2-sensing neurons disorients mosquitoes. - Nature 474: 87–91.

Um, S. et al. 2013. The fungus-growing termite *Macrotermes natalensis* harbors bacillaene-producing *Bacillus* sp. that inhibit potentially antagonistic fungi. - Sci. Rep. 3: 3250.

Figure Legends

Figure 1. Non-metric multidimensional scaling (NMDS) ordination plot based on Bray-Curtis dissimilarities of the mVOC composition of the 40 *Bacillus* strains investigated in this study (n = 3) (stress value = 0.159). Symbol shapes indicate the clade to which each of the strains belongs (see also Fig. 2): circle = clade A, square = clade B, hexagon = clade C, diamond = clade D, upward facing triangle = clade E, and downward facing triangle = clade F. Symbol colours indicate the effect of the mVOCs on the olfactory response of *Aphidius colemani*, i.e. green = significantly attractive, grey = neutral, and red = significantly repellent. Symbol sizes are proportional to the absolute values of the Preference Index (PI) as determined in the olfactometer bioassay. The black cross refers to the blank medium. mVOC composition differed significantly between the bacterial clades defined (perMANOVA: pseudo-F_5 = 10.7, P < 0.001).

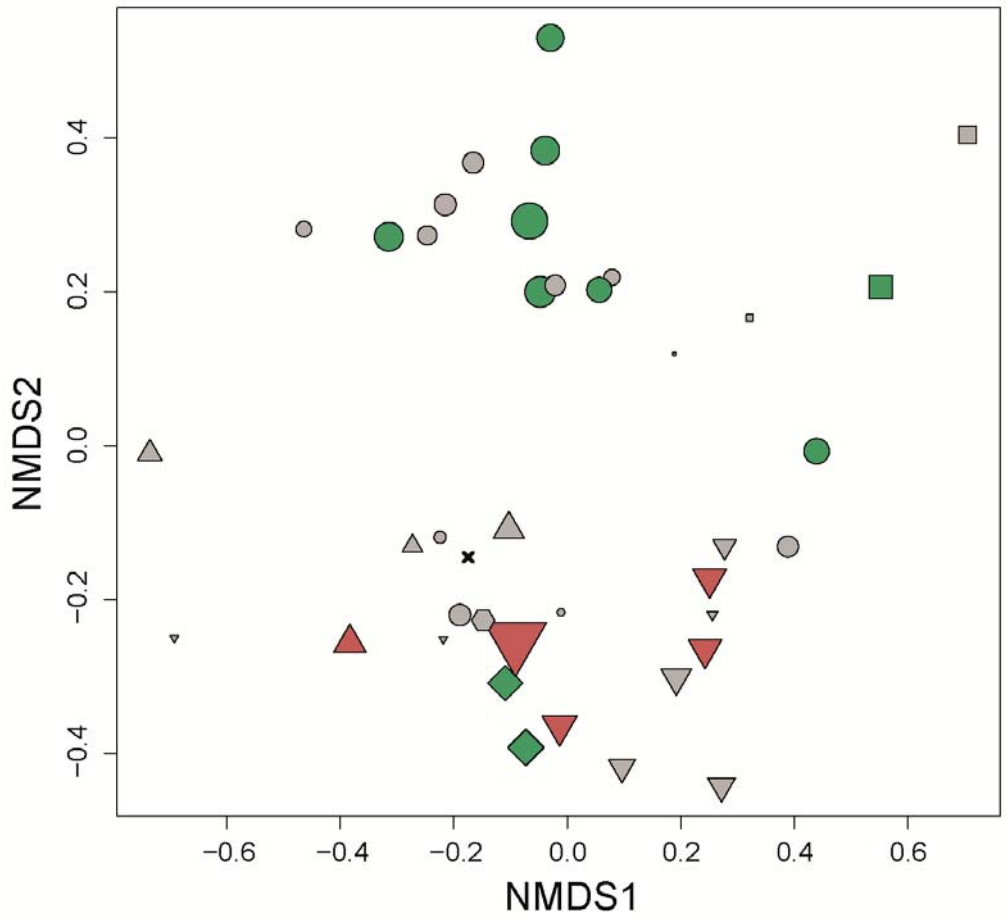

Figure 2. Maximum likelihood phylogenetic tree based on a concatemer of 16S rRNA gene and *rpoB* sequences of all 40 *Bacillus* strains investigated in this study (black). Additionally, the most closely related type strains (light grey) were added as a reference. Based on a sequence identity cut-off of 85%, six major clades can be distinguished. Bars depict the Preference Index (PI) for *Aphidius colemani* females when having the choice between the bacterial mVOCs and the blank medium. Bar colours indicate the effect of the mVOCs on the olfactory response of *A. colemani*, i.e. green = significantly attractive, grey = neutral, and red = significantly repellent.

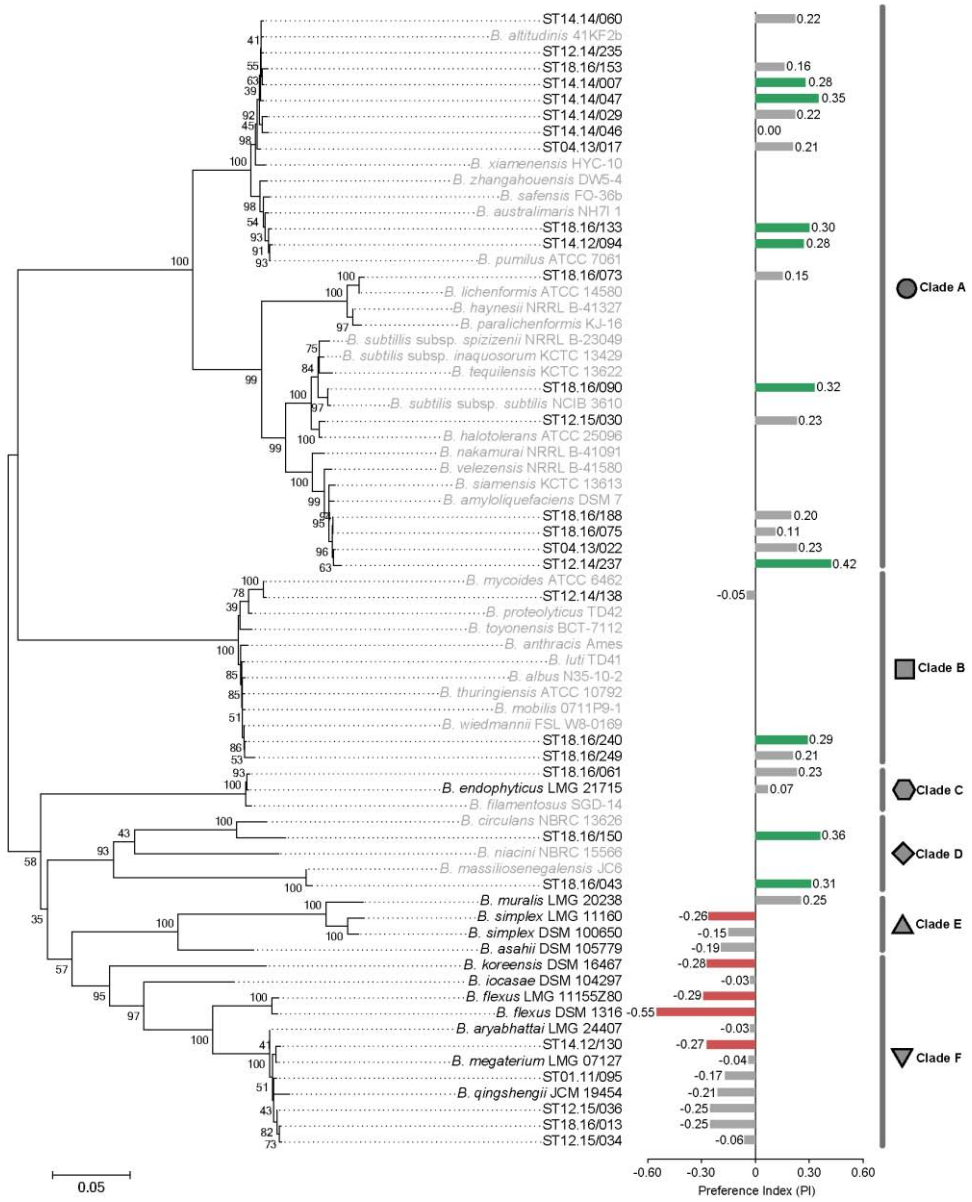

Figure 3. Results of phylogenetic principal component analysis (pPCA). A) Bar chart of eigenvalues of all PCs with PC1 and PC39 in blue and orange, respectively. B) Biplot showing the global (PC1; blue axis) and local structure (PC39; orange axis) in the pPCA. PC1 denotes the global structure and reveals the mVOCs that are more similar in related than in distant strains. The local structure is depicted in the second PC, which reveals the mVOCs that create dissimilarities among closely related strains. Only mVOCs with highest loadings (absolute value >0.1) on the PCs are shown. Symbol colours, shapes and sizes are as explained in Fig. 1. Vector numbers refer to the different mVOCs: (1) propanal, (2) trimethylamine, (3) acetone, (4) 2-methyl-2-propanol, (5) 1-propanol, (6) 2,3-epoxy-2,3-dimethylbutane, (7) butanal, (8) 2,3-butanedione, (9) ammonium acetate, (10) fluoro-benzene, (11) acetoin, (12) isoamylamine, (13) methyl-methacrylate, (14) 3-hydroxy-3-methyl-2-butanone, (15) 2-methyl-2-pentanol, (16) 2-methyl-propanoic acid, (17) 2,3-butanediol, (18) 2-hexanone, (19) ethyl-butanoate, (20) pinacol, (21) 1,3-dimethyl-benzene, (22) cyclohexanone, (23) 2,5-hexanedione, (24) benzaldehyde, (25) α-methylstyrene, (26) 2,3,5-trimethyl-pyrazine, (27) eucalyptol, (28) limonene, (29) 2-ethyl-1-hexanol, (30) Indane, (31) phenylacetaldehyde, (32) 2-isopropyl-5-methyl-pyrazine, (33) para-cymene, (34) acetophenone, (35) 2-methyl-6-methylene-2-octanol, (36) 4-methylbenzaldehyde, (37) tetramethyl-pyrazine, (38) 3,7-dimethyl-octan-3-ol, (39) nonan-2-ol, (40) 2-ethyl-hexanoic acid, (41) 2,3-dihydro-4-methyl-1H-indene, (42) menthol, (43) verbenone, (44) n-decanal, (45) 3,5-dimethyl-benzaldehyde, (46) indole, (47) nonanoic acid, (48) cis-2-tert-butyl-cyclohexanol acetate, (49) 4-tert-butylcyclohexyl acetate, (50) isobutyl 3-hydroxy-2,2,4-trimethylpentanoate, (51) (1-hydroxy-2,4,4-trimethylpentan-3-yl) 2-methylpropanoate, (52) 3-hydroxy-2,4,4-trimethylpentyl 2-methylpropanoate, (53) 2-ethyl-3-hydroxyhexyl 2-methylpropanoate, (54) 2,6-di-tert-butyl-P-benzoquinone, (55) 4-(2,6,6-trimethyl-2-cyclohexen-1-yl)-3-penten-2-one, (56) 2,2,4-trimethyl-1,3-

pentanediol diisobutyrate, (57) 1,1-(1,2-dimethyl-1,2-ethenediyl)bis-benzene, (58) 3,5-di-tert-butyl-4-hydroxybenzaldehyde, (59) 1,1-(1,1,2,2-tetramethyl-1,2-ethanediyl)bis-benzene, (60) butyl isobutyl phthalate, and (61) methyl-dehydroabietate. C) Maximum likelihood phylogenetic tree based on a concatemer of 16S rRNA gene and *rpoB* sequences of all 40 *Bacillus* strains investigated in this study and results of the pPCA on the mVOCs produced by the strains. The phylogenetic tree was divided into six major clades based on a 85% sequence identity cut-off. Positive and negative scores on PC1 (global structure) and PC39 (local structure) are indicated by blue and orange circles, respectively. Circle size is proportional to the absolute score values.

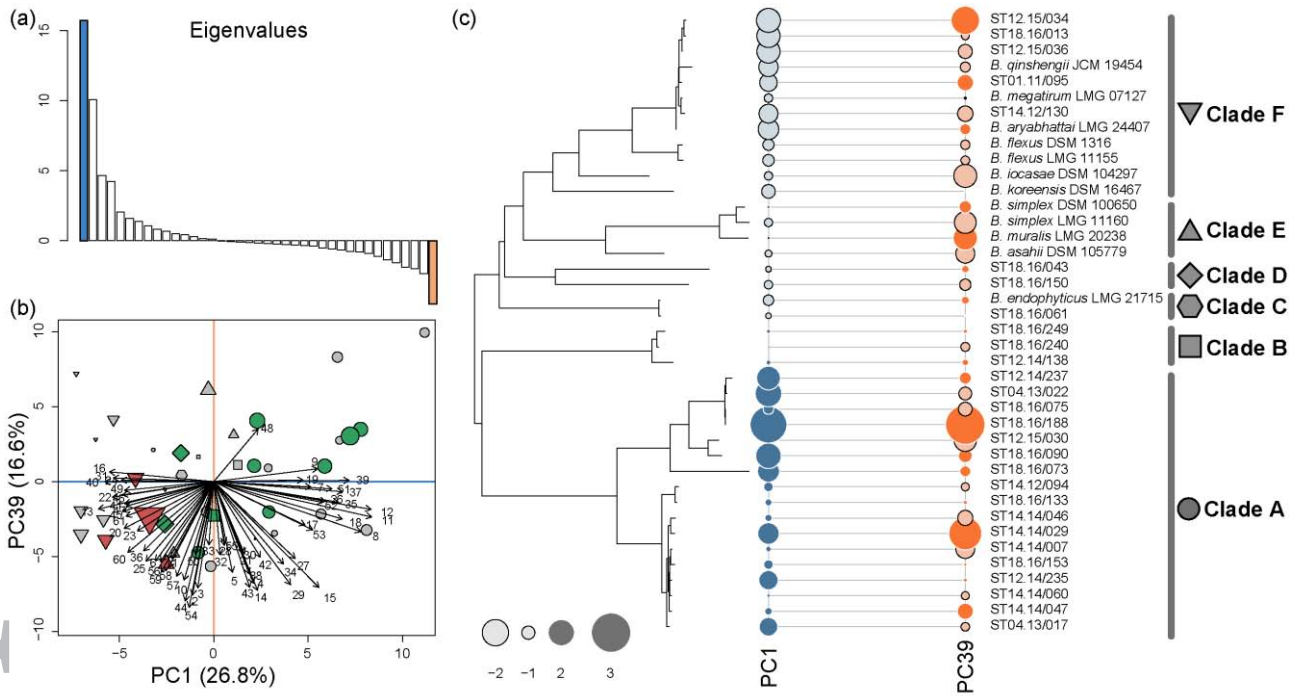

Figure 4. Phylogenetic heatmap using the maximum likelihood phylogenetic tree based on a concatemer of 16S rRNA gene and *rpoB* sequences of all 40 *Bacillus* strains used in this study. The phylogenetic tree was divided into six major clades based on a 85% sequence identity cut-off. The heatmap depicts 25 mVOCs that show a significant phylogenetic signal for all three indices calculated (Pagels's λ, Moran's I and Abouheif's C_{mean}) and the Mantel Test. Data are presented in the form of mean centred, log transformed average peak areas of compounds (n = 3). mVOC compositions were clustered by using Manhattan distances and a Ward.D2 clustering algorithm. [a]TMP = trimethylpentan; MP = methylpropanoate.

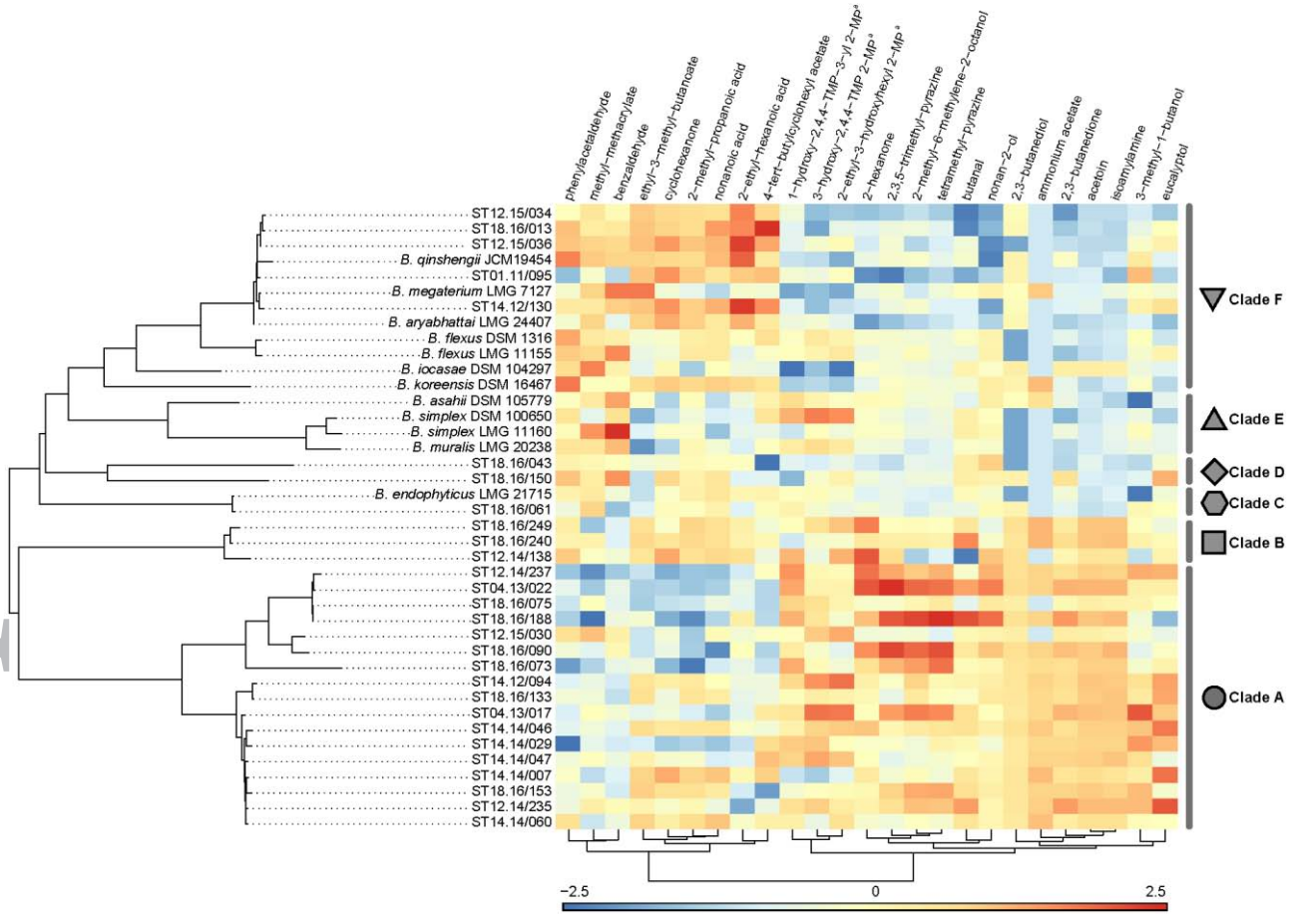

Table Legends

Table 1. Bacterial isolates used in this study.

Isolate identifier[a]	16S rRNA gene Accession N°[b]	*rpoB* Accession N°[c]	Clade[d]	Phylogenetic affiliation[e]				Source of isolation
				Phylum	Family	Closest match in GenBank to identified species (Accession N°)	Identity (%)	
T14.14/060	MN220674	MN232848	A	Firmicutes	Bacillaceae	*Bacillus altitudinis* (CP024204.1)	99.82	Whisky barrel
ST12.14/235	MN220665	MN232841	A	Firmicutes	Bacillaceae	*Bacillus altitudinis* (CP024204.1)	99.64	Floral nectar *Centaurea cyanus*
ST14.14/007	MN220[illegible]	MN232844	A	Firmicutes	Bacillaceae	*Bacillus altitudinis* (CP024204.1)	99.91	Whisky barrel
ST14.14/047	MN220673	MN232847	A	Firmicutes	Bacillaceae	*Bacillus altitudinis* (CP024204.1)	99.46	Whisky barrel
ST14.14/029	MN220671	MN232846	A	Firmicutes	Bacillaceae	*Bacillus altitudinis* (CP024204.1)	99.73	Whisky barrel
ST14.14/046	MN220672	MN232845	A	Firmicutes	Bacillaceae	*Bacillus altitudinis* (CP024204.1)	99.00	Whisky barrel
ST04.14/017	MN220662	MN232843	A	Firmicutes	Bacillaceae	*Bacillus altitudinis* (CP024204.1)	99.28	Soil conditioner
LMG 24407[#]	EF114313	GCA_900101665.1	F	Firmicutes	Bacillaceae	*Bacillus aryabhattai* (GCA_900101665.1)	100.00	Cryogenic tube
DSM 105779[#]	[illegible]	GCA_003570725.1	E	Firmicutes	Bacillaceae	*Bacillus asahii* (GCA_003570725.1)	100.00	Soil
ST18.16/150	MN220680	MN232831	D	Firmicutes	Bacillaceae	*Bacillus circulans* (CP026033.1)	92.20	*Macrosiphum euphorbiae*
LMG 21715[#]	AF295[illegible]	GCA_900115845	C	Firmicutes	Bacillaceae	*Bacillus endophyticus* (GCA_900115845)	100.00	Cotton plants
ST18.16/061	MN220676	MN232836	C	Firmicutes	Bacillaceae	*Bacillus filamentosus* (CP026635.1)	99.28	*Myzus persicae* var. *nicotianae*
DSM 1316	MN220[illegible]3	MN232837	F	Firmicutes	Bacillaceae	*Bacillus flexus* (CP040367.1)	99.10	Unknown
LMG 11155[#]	BCV[illegible]4	GCA_001591565.1	F	Firmicutes	Bacillaceae	*Bacillus flexus* (GCA_003184905.1)	100.00	Unknown
ST12.15/030	MN220666	MN232850	A	Firmicutes	Bacillaceae	*Bacillus halotolerans* (CP029364.1)	99.91	Dried figs
DSM 104297[#]	KX[illegible]2210	GCA_003184905.1	F	Firmicutes	Bacillaceae	*Bacillus iocasae* (GCA_003184905.1)	100.00	Deep-sea sediment
DSM 16467[#]	LILC01000014	GCA_001274935.1	F	Firmicutes	Bacillaceae	*Bacillus koreensis* (GCA_001274935.1)	100.00	Rhizosphere
ST18.16/073	MN220[illegible]	MN232851	A	Firmicutes	Bacillaceae	*Bacillus licheniformis* (CP034569.1)	100.00	*Myzus persicae* var. *nicotianae*
ST12.15/036	MN220667	MN232829	F	Firmicutes	Bacillaceae	*Bacillus megaterium* (CP001982.1)	99.55	Dried figs
ST18.16/013	MN220[illegible]	MN232827	F	Firmicutes	Bacillaceae	*Bacillus megaterium* (CP001982.1)	99.64	*Myzus persicae* var. *nicotianae*
ST12.15/034	MN220686	MN232828	F	Firmicutes	Bacillaceae	*Bacillus megaterium* (CP001982.1)	99.64	Dried figs
ST01.11/095	MN220661	MN232839	F	Firmicutes	Bacillaceae	*Bacillus megaterium* (CP028084.1)	99.82	Activated sludge
ST14.12/130	MN220669	MN232832	F	Firmicutes	Bacillaceae	*Bacillus megaterium* (CP001982.1)	99.19	Floral nectar *Epipactis palustris*

LMG 7127[#]	JJMH01000057	GCA_000832985.1	F	Firmicutes	Bacillaceae	*Bacillus megaterium* (GCA_000832985.1)	100.00	Unknown
LMG 20238[#]	LMBV01000055	GCA_001439925.1	E	Firmicutes	Bacillaceae	*Bacillus muralis* (GCA_001439925.1)	100.00	Mural
ST12.14/138	MN[illegible]	MN232833	B	Firmicutes	Bacillaceae	*Bacillus mycoides* (CP009692.1)	99.55	Floral nectar *Symphytum officinale*
ST18.16/133	MN220679	MN232849	A	Firmicutes	Bacillaceae	*Bacillus pumilus* (CP029464.1)	100.00	*Aphidius ervi*
ST14.12/094	MN2[illegible]0668	MN232842	A	Firmicutes	Bacillaceae	*Bacillus pumilus* (LT906438.1)	99.91	Floral nectar *Epipactis palustris*
JCM 19454[#]*	JX29[illegible]3295	MN232838*	F	Firmicutes	Bacillaceae	*Bacillus qingshengii* (MN232838)	100.00	Weathered tuff surface
DSM 100650	MN220659	MN232853	E	Firmicutes	Bacillaceae	*Bacillus simplex* (CP017704.1)	96.56	Cleanroom facility
LMG 11160[#]	BCVO01[illegible]	GCA_002243645.1	E	Firmicutes	Bacillaceae	*Bacillus simplex* (GCA_002243645.1)	100.00	Unknown
ST18.16/153	[illegible]220681	MN232857	A	Firmicutes	Bacillaceae	*Bacillus* sp. (*B. altitudinis* [CP009108.1], *B. pumilus* [CP007436.1])	99.91	*Macrosiphum euphorbiae*
ST18.16/188	MN220683	MN232855	A	Firmicutes	Bacillaceae	*Bacillus* sp. (*B. velezensis* [CP029296.1], *B. amyloliquefaciens* [CP006845.1])	100.00	*Macrosiphum euphorbiae*
ST18.16/075	MN220677	MN232852	A	Firmicutes	Bacillaceae	*Bacillus* sp. (*B. velezensis* [CP029296.1], *B. amyloliquefaciens* [CP006845.1])	100.00	*Myzus persicae* var. *nicotianae*
ST04.13/022	MN220663	MN232854	A	Firmicutes	Bacillaceae	*Bacillus* sp. (*B. velezensis* [CP034176.1], *B. amyloliquefaciens* [CP007242.1])	99.91	Soil conditioner
ST12.14/237	MN220682	MN232840	A	Firmicutes	Bacillaceae	*Bacillus* sp. (*B. velezensis* [CP034176.1], *B. amyloliquefaciens* [CP007242.1])	98.91	Floral nectar *Centaurea cyanus*
ST18.16/043	MN220686	MN232830	D	Firmicutes	Bacillaceae	*Bacillus* sp. X1 (CP008855.1)	80.24	*Dendrocerus aphidum*
ST18.16/090	MN22[illegible]	MN232856	A	Firmicutes	Bacillaceae	*Bacillus subtilis* (CP035230.1)	100.00	*Aphidius ervi*
ST18.16/240	MN220684	MN232834	B	Firmicutes	Bacillaceae	*Bacillus wiedmannii* (CP024684.1)	99.73	*Myzus persicae* var. *nicotianae*/ *Capsicum annuum* honeydew
ST18.16/249	MN220685	MN232835	B	Firmicutes	Bacillaceae	*Bacillus wiedmannii* (CP024684.1)	99.00	*Myzus persicae* var. *nicotianae*/ *Capsicum annuum* honeydew

[a] #Type strains acquired from available culture collections (LMG, DSMZ and JCM). 16S rRNA gene sequences were obtained from the EZBiocloud database (referred to by EZBiocloud 16S rRNA gene accession numbers); *rpoB* sequences were extracted from full genome sequences obtained from the EZBiocloud database (referred to by EZBiocloud full genome accession numbers). *No full genome or *rpoB* sequence was available for this strain.
[b]1252-1262 bp.
[c]1102-1105 bp.
[d]Clades defined using a sequence identity cut-off of 85% based on a concatemer of 16S rRNA gene and *rpoB* sequences (see Fig. 2).
[e]Based on a B[illegible] search of the *rpoB* sequences against GenBank (July 2019). Only closest matches to identified species are reported.

Table 2. Overview of results of the phylogenetic signal tests on the behavioural response of *Aphidius colemani* and the mVOC profiles.

Variable	Index values	Phylogenetic signal index			
		Pagel's λ	**Moran's I**	**Abouheif's C_{mean}**	**Mantel test**
Behavioural response (40)		**0.775*****	**0.952*****	**0.603*****	**97.4*****
Chemical composition mVOCs (159)	Range	4.44×10^{-5} - 1.00	-0.21 - 0.85	-0.19 - 0.86	5.54 - 29.0
	Mean	0.27	0.17	0.14	20.73
	Significant compounds	**48 (30.2%)**	**44 (27.7%)**	**42 (26.4%)**	**42 (26.4%)**
	PC1	**0.61*****	**0.72*****	**0.72*****	**137*****
	PC39	0.00	-0.28	-0.30	91.2
	Mantel test		***Z* = 6.29×10^{10}**	***P* < 0.001**	

[a]Pagel's λ, Moran's I, and Abouheif's C_{mean} were calculated and a Mantel test was performed using the preference index data (behavioural response of *A. colemani*) and mean centered, log transformed peak area data of every mVOC produced by the *Bacillus* strains (chemical composition mVOCs), in combination with a phylogenetic tree based on a concatemer of 16S rRNA gene and *rpoB* sequences. The same tests were performed on the eigenvectors (PC1 and PC39) resulting from the pPCA. Finally, a Mantel test was used to analyse the complete dataset of all mVOC produced by the *Bacillus* strains. Values in bold indicate a significant phylogenetic signal (***, $P < 0.001$).