MPMI Vol. 28, No. 10, 2015, pp. 1102–1116. http://dx.doi.org/10.1094/MPMI-02-15-0023-R

*e*-Xtra*

# Comparative Genomics Within the *Bacillus* Genus Reveal the Singularities of Two Robust *Bacillus amyloliquefaciens* Biocontrol Strains

**M. C. Magno-Pérez-Bryan,[1] P. M. Martínez-García,[2,3] J. Hierrezuelo,[1] P. Rodríguez-Palenzuela,[2] E. Arrebola,[1] C. Ramos,[3] A. de Vicente,[1] A. Pérez-García,[1] and D. Romero[1]**

[1]Instituto de Hortofruticultura Subtropical y Mediterránea "La Mayora" (IHSM-UMA-CSIC), Departamento de Microbiología, Facultad de Ciencias, Universidad de Málaga, Bulevar Louis Pasteur 31 (Campus Universitario de Teatinos), 29071 Málaga, Spain; [2]Centro de Biotecnología y Genómica de Plantas UPM-INIA (CBGP), Parque Científico y Tecnológico de la Universidad Politécnica de Madrid, Campus de Montegancedo, 28223 Pozuelo de Alarcón, Madrid, Spain; [3]Instituto de Hortofruticultura Subtropical y Mediterránea "La Mayora" (IHSM-UMA-CSIC), Área de Genética, Facultad de Ciencias, Universidad de Málaga

Submitted 3 February 2015. Accepted 27 May 2015.

***Bacillus amyloliquefaciens* CECT 8237 and CECT 8238, formerly known as *Bacillus subtilis* UMAF6639 and UMAF6614, respectively, contribute to plant health by facing microbial pathogens or inducing the plant's defense mechanisms. We sequenced their genomes and developed a set of ad hoc scripts that allowed us to search for the features implicated in their beneficial interaction with plants. We define a core set of genes that should ideally be found in any beneficial *Bacillus* strain, including the production of secondary metabolites, volatile compounds, metabolic plasticity, cell-to-cell communication systems, and biofilm formation. We experimentally prove that some of these genetic elements are active, such as i) the production of known secondary metabolites or ii) acetoin and 2-3-butanediol, compounds that stimulate plant growth and host defense responses. A comparison with other *Bacillus* genomes permits us to find differences in the cell-to-cell communication system and biofilm formation and to hypothesize variations in their persistence and resistance ability in diverse environmental conditions. In addition, the major protection provided by CECT 8237 and CECT 8238, which is different from other *Bacillus* strains against bacterial and fungal melon diseases, permits us to propose a correlation with their singular genetic background and determine the need to search for additional blind biocontrol-related features.**

The use of beneficial bacteria to combat pests or plant diseases has rapidly grown over the last decades. Among the different microbial species examined, members of the *Bacillus* genus are potential candidates for use as biological control agents (Pérez-García et al. 2011). The bacterial attributes contributing to this preferential development of biocontrol products based on *Bacillus* species are the production of a variety of secondary metabolites, the efficient colonization of habitats, and the formation of spores, which ensure the persistence and the ease of production of reliable formulates (Cawoy et al. 2015). However, the selection of a promising biocontrol agent must be preceded by the investigation of the arsenal of bacterial features that contribute to its bioactivity to i) ensure the best results in terms of plant protection and ii) promote its use as a trusted tool for disease control, either alone or included in integrated pest management programs (Xu and Jeger 2013).

Next-generation sequencing systems offer a unique opportunity to investigate which bacterial genetic determinants support the development of these biological control activities (Liu et al. 2012). The acquisition of genome sequences of myriad bacterial species has generated a vast amount of information that, in combination with the appropriate bioinformatics tools, has increased our knowledge of the genetic potential that may contribute to bacterial fitness, adaptation to the environment, and interaction with plants (Cai et al. 2014; Earl et al. 2007). In addition, by comparison with related pathogenic bacteria, it is possible to disregard the presence of any genetic factor related to pathogenicity to either plants or humans that could compromise their commercialization as beneficial bacteria (Moreno Switt et al. 2012).

In previous research conducted in our laboratory, we proved that two *Bacillus* strains, UMAF6614 and UMAF6639, were good candidates as biocontrol agents against fungal and bacterial diseases of cucurbits. In the melon phyllosphere, the direct antagonism toward fungal and bacterial pathogens, as mediated by the production of iturin and fengycin lipopeptides, appears to be the main mechanism (Romero et al. 2007b; Zeriouh et al. 2011). In the melon rhizoplane, both strains are able to indirectly contribute to plant health by promoting plant growth and induced systemic resistance (ISR) (García-Gutiérrez et al. 2012, 2013). Either way, this multifaceted mode of action appears to be related to the production of the lipopeptide surfactin, which acts as i) a self-trigger of biofilm formation, which ensures persistence of these biocontrol agents and the efficient antagonism toward the pathogens (Zeriouh et al. 2014), and ii) an interkingdom communication molecule that activates the immune system of plants (ISR) (García-Gutiérrez et al. 2013; Ongena et al. 2007).

In addition to lipopeptides, we presumed that additional features of these two *Bacillus* strains must be implicated in their

Genome sequences were deposited in the National Center for Biotechnology Information (NCBI) GenBank database under accession numbers CP006058 (CECT 8237) and CP006960 (CECT 8238).

Corresponding author: D. Romero; E-mail: diego_romero@uma.es; Telephone: +34952134274; Fax: +34952136645

*The *e*-Xtra logo stands for "electronic extra" and indicates that seven supplementary figures and four supplementary tables are published online.

remarkable biocontrol activity and adaptation to the environment. In this study, we report the complete genome sequences of both biocontrol agents. We have performed bioinformatics analysis to i) identify and specifically diagnose these strains as *Bacillus amyloliquefaciens*, ii) define genomic regions poorly conserved within the *Bacillus* genus, and iii) determine new genes potentially implicated in the efficient biocontrol activity of CECT 8237 (formerly UMAF6639) and CECT 8238 (formerly UMAF6614).

## RESULTS AND DISCUSSION

### The strains CECT 8237 and CECT 8238 are identified as *Bacillus amyloliquefaciens*.

Strains CECT 8237 and CECT 8238 (formerly UMAF6639 and UMAF6614, respectively) have outstanding biocontrol activity against a variety of microbial diseases of cucurbits (García-Gutiérrez et al. 2012; Romero et al. 2007b; Zeriouh et al. 2011). Therefore, we sequenced their genomes and made use of customized scripts to search their whole arsenal of biocontrol-related features. In previous studies, these strains were classified as *Bacillus subtilis*, based on sequence homology of the 16S rRNA and metabolic profiles (Romero et al. 2004). However, the recurrent microbe misidentification associated with these typical analyses led us to use the most reliable and accurate multilocus phylogeny analysis (Liu et al. 2013). In this study, the CECT 8237 and CECT 8238 strains clustered with the *B. amyloliquefaciens* group (Fig. 1A), which was an unexpected finding and contrary to their initial identification as *B. subtilis*. However, this was not unprecedented, because *B. amyloliquefaciens* and *B. subtilis* are phenotypically and genetically related species (Cawoy et al. 2015).

### Plant-associated *B. amyloliquefaciens* strains gain skills in competition and environmental adaptability.

A previous genomic study proposed that *B. amyloliquefaciens* strains could be sorted into two distinct clades, i) the clade of the plant-associated strains and ii) the clade that is associated not with plants but with industrial applicability (Borriss et al. 2011). In our analysis, which included more genome sequences than in the previous study, we found this clear separation of the two clades. The CECT 8237 and CECT 8238 strains appeared to be more closely related to the strains of the plant-associated clade (Fig. 1A and B). However, it was noteworthy that both strains evolved in different branches than did strain FZB42, a paradigm of the group of agriculturally relevant *B. amyloliquefaciens* strains (Borriss et al. 2011). This separation would tentatively reflect the divergent evolution of these strains imposed by the habitat they actually occupy, and thus, we should ideally find genomic differences between plant-associated and industrially relevant strains. In general, the genome sizes of plant-associated strains of bacteria were larger than industrial strains, but this variation in size did not correlate with an increase in the number of putative open reading frames or predicted proteins (Table 1). Thus, the differences between these clades should be noted in the functionalities of their predicted proteins. To test this idea, we first defined the set of genes contained in all members of each clade (in other words, their pan-genomes) and used this information to search specifically for the genes exclusive to each clade. All plant-associated and industrially relevant *B. amyloliquefaciens* strains included in this analysis are listed in Table 1. In the first step of this analysis, we found that 5,453 coding sequences (CDS) formed the pan-genome of clade I (plant-associated strains) and that 4,753 CDS corresponded to

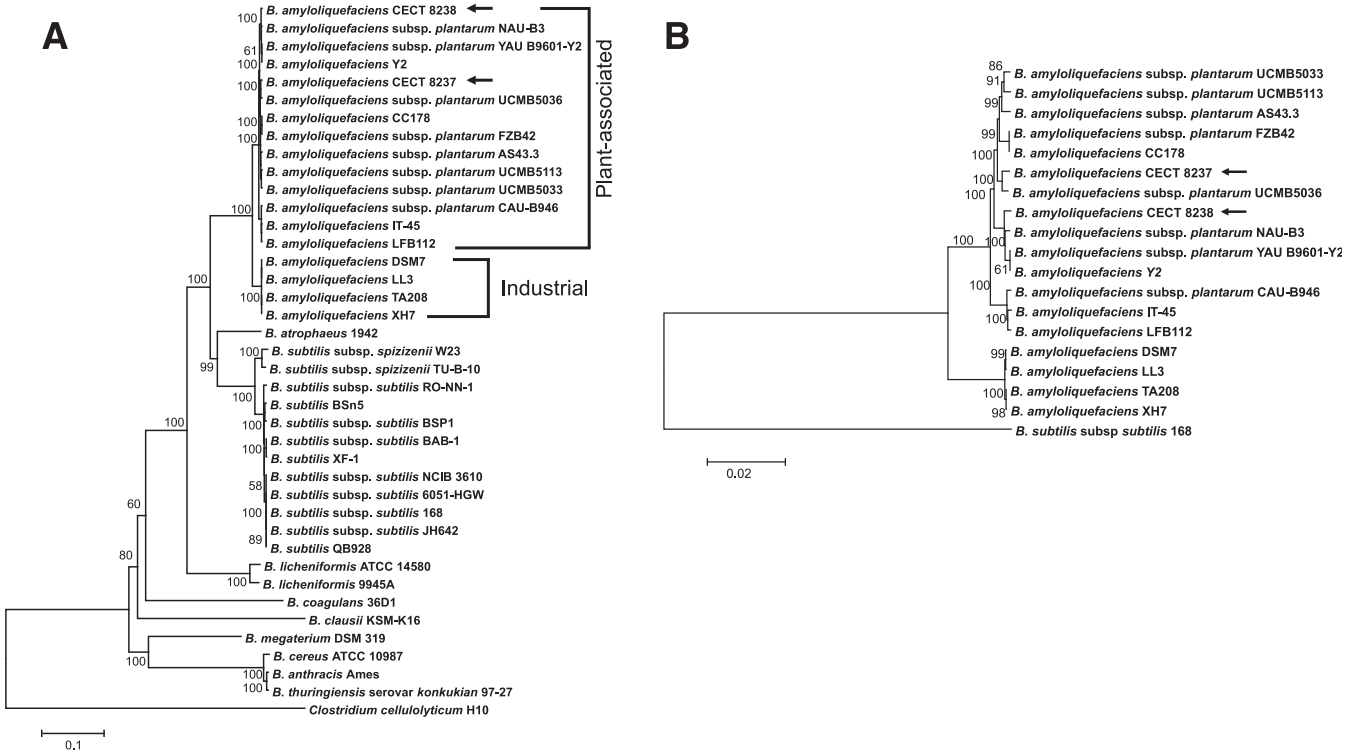


**Fig. 1.** The isolates CECT 8237 and CECT 8238 cluster with the group of plant-associated *Bacillus amyloliquefaciens* subsp. *plantarum*. **A,** Phylogenetic analysis of the isolates CECT 8237 and CECT 8238, and the closely related species *B. amyloliquefaciens*, *B. subtilis*, *B. atrophaeus*, and *B. licheniformis* and other representatives of the *Bacillus* genus, such as *B. cereus*, *B. thuringiensis*, *B. coagulans*, *B. megaterium*, and *B. anthracis*. *Clostridium cellulolyticum* H10 was used as an outgroup. **B,** Phylogenetic analysis of the isolates CECT 8237 and CECT 8238 and the plant-associated *B. amyloliquefaciens* strains. The *B. subtilis* subsp. *subtilis* 168 was used as an outgroup. In both analyses, 11 concatenated genes (*nusA*, *rpoA*, *dnaA*, *rpoB*, *gyrA*, *gyrB*, *rpoC*, *spoVG*, *sigW*, *sigH*, and *sigB*) were handled to build a neighbor-joining tree, using MEGA 5 bootstrap values (10,000 repetitions), which are shown on the branches. The topology was identical for the trees produced by the minimum evolution and maximum parsimony methods. The sequences from all of the strains were extracted from published genome sequences. Arrows indicate the location of the CECT 8237 and CECT 8238 strains in the tree.

clade II (industrial strains). It was remarkable that 1,365 CDS appeared exclusively in any strain of clade I, though only 612 CDS were present in at least one industrial strain but absent in all members of the other clade. From these CDS, we selected those (16 CDS) found in all plant-associated strains (Supplementary Table S2) and built a phylogenetic tree (Supplementary Fig. S1), which provided a similar distribution of strains as previously seen (Fig. 1); the FZB42 strains in separate branches from CECT 8237 or CECT 8238. Then, these exclusive CDS were classified into different clusters of orthologous groups (COG) categories. In order to analyze whether CDS only present in one set of strains were enriched in any COG categories with respect to the other set, an enrichment test of the computed COG categories was performed. After the application of hypergeometric tests, we found that these exclusive genes were not significantly enriched in genes belonging to any COG class.

However, we focused on those genes that showed homology to any CDS in the database and permitted their classification into different COG categories, 157 and 28 genes in plant-associated and industrial strains, respectively (Supplementary Table S3). The relative proportion of COG categories represented in one set of strains against the other was plotted in sector graphs (Fig. 2). We found that the unique CDS of clade I (plant-associated strains) primarily sorted into the following categories: metabolism of amino acids (E), carbohydrates (G),

**Table 1.** A comparison of the genomic features of *Bacillus amyloliquefaciens* CECT 8237 and CECT 8238 with other plant-associated or industrial-relevant *Bacillus amyloliquefaciens* strains

| | Features | | | | | | |
|---|---|---|---|---|---|---|---|
| **Strains** | **Genome size (bp)** | **Plasmid size (bp)** | **G+C content (%)** | **Number of genes** | **Protein-coding sequences** | **tRNAs** | **rRNAs** |
| Plant associated | | | | | | | |
| CECT 8237 | 4,034,636 | – | 46,34 | 4,059 | 3,918 | 82 | 27 |
| CECT 8238 | 4,005,145 | – | 46,49 | 4,049 | 3,894 | 82 | 27 |
| AS43.3 | 3,961,368 | – | 46,59 | 3,979 | 3,861 | 89 | 29 |
| CAU-B946 | 4,019,861 | – | 46,51 | 3,948 | 3,823 | 95 | 30 |
| CC178 | 3,916,828 | – | 46,5 | 4,074 | 3,950 | 86 | 27 |
| FZB42 | 3,918,589 | – | 46,48 | 3,813 | 3,693 | 89 | 30 |
| IT-45 | 3,928,857 | 8,009 | 46,62/40,47[z] | 3,976/9[z] | 3,851/9[z] | 95 | 30 |
| LFB112 | 3,942,754 | – | 46,7 | 4,023 | 3,859 | 94 | 30 |
| NAU-B3 | 4,196,170 | 8,438 | 46/40,3[z] | 4,123/5[z] | 4,001/5[z] | 92 | 30 |
| UCMB5033 | 4,071,167 | – | 46,2 | 4,095 | 3,877 | 86 | 30 |
| UCMB5036 | 3,910,324 | – | 46,6 | 3,842 | 3,636 | 89 | 29 |
| UCMB5113 | 3,889,532 | – | 46,7 | 3,854 | 3,656 | 89 | 29 |
| YAU B9601-Y2 | 4,242,774 | – | 45,85 | 4,110 | 3,981 | 91 | 30 |
| Y2 | 4,238,624 | – | 45,85 | 4,348 | 4,234 | 85 | 29 |
| Industrial | | | | | | | |
| DSM7 | 3,980,199 | – | 46,1 | 4,043 | 3,892 | 94 | 30 |
| LL3 | 3,995,227 | 6,758 | 45,7/42[z] | 4,346/9[z] | 4,219/9[z] | 72 | 22 |
| TA208 | 3,937,511 | – | 45,8 | 4,177 | 4,089 | 70 | 18 |
| XH7 | 3,939,203 | – | 45,8 | 4,286 | 4,190 | 75 | 21 |

[z] Data from the chromosome sequence (first column) and plasmid sequence (second column).

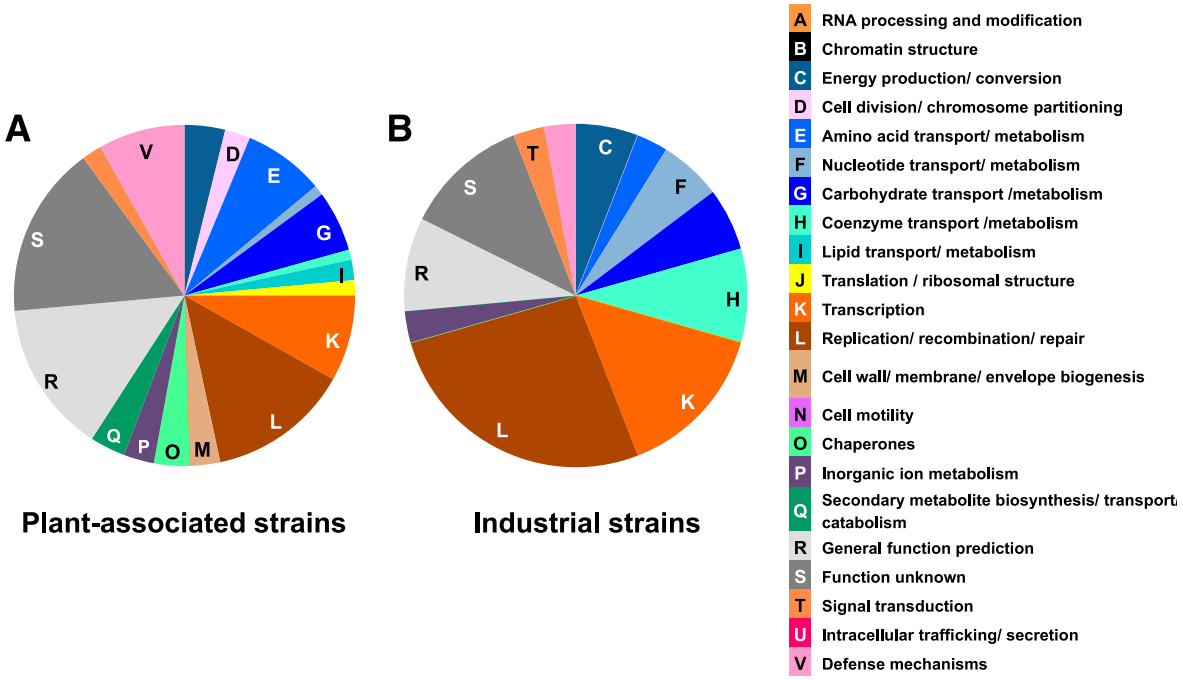


**Fig. 2.** The classification of genes into clusters of orthologous groups (COG) reveals the differences between plant-associated and industrially relevant *Bacillus amyloliquefaciens* strains. Circular charts symbolize the percentages of genes classified into the different COG specifically detected in **A,** *B. amyloliquefaciens* strains associated with plants or **B,** industrial-relevant strains. A color code is specified in the figure legend.

or lipids (I), synthesis of secondary metabolites (Q), general function prediction (R), or defense mechanisms (V) (Fig. 2A). The enhancement of their genomes in genes coding these functionalities might reflect the better adaptation of these strains to, for example, the changeable environment of plant surfaces (Carvalhais et al. 2013). One might assume that the presence of endoglucanase, *N*-acetylglucosamine-6-phosphate deacetylase, or maltose phosphorylase enzymes would allow these strains to utilize the diversity of alternative carbon sources available in the exudates of plant surfaces. It is also ecologically advantageous to have the genes involved in the mobilization and bioavailability of scarce micronutrients, such as sulfur and iron, not only as a self-benefit but also for the good of plant health (Miethke et al. 2006). The overrepresentation of genes involved in the synthesis of secondary metabolites would reflect the versatility of these strains to compete with a wide range of microbes that they might encounter in the plant environment. Contrary to this finding, the unique CDS of the clade II (industrial strains) appeared dispersed among the different categories, although some were overrepresented, such as metabolism of coenzymes (H) and replication, recombination, and repair (L) (Fig. 2B). In view of these observations, it could be said that the *B. amyloliquefaciens* strains associated with plants possess a specific core of genetic features, which would predict their better performance in the context of plant health than those strains not associated with plants. (CECT 8238, NAU-B3, YAU B9601-Y2, and Y2 strains clustered together as a result of the alignment corresponding to genes distinctive of plant-associated strains).

### *B. amyloliquefaciens* CECT 8237 and CECT 8238 have acquired gene clusters that support their biocontrol skills.

Once we defined the core genes distinctive to strains associated with plants, we investigated whether the CECT 8237 or CECT 8238 strains could bear unique and differentiable features to the other members of the same clade. To answer this question, we compared their genomes with those of other plant-associated *B. amyloliquefaciens* strains (Table 1) that were available in the database, using the MAUVE genome alignment software (Fig. 3, outer circle). Specifically, the level of conservation along the chromosomal regions is represented by different tones of blue, darker, meaning highly conserved, and softer, which indicates poorly conserved sequences. Despite the level of conservation of their genome sequences, we could identify a number of unique "windows" (white areas), which, by definition, were absent in the other *B. amyloliquefaciens* strains and, thus, a potential reservoir of distinctive bacterial features. Separately, we performed a bioinformatics analysis, based on the methodology shown in Supplementary Figure S7, to compare the studied strains with 76 *Bacillus* complete genome sequences, including those of plant-associated and industrial *B. amyloliquefaciens* strains (Supplementary Table S1, strains for comparative analysis). As a result, this analysis permitted us to screen for atypical regions (AR) that were nonconserved along the *Bacillus* genus (Supplementary Fig. S2) and that were composed of consecutive genes hypothetically related to the same physiological function (Fig. 3, green boxes). Using a minimum of seven consecutive genes as the screening window, we found three different scenarios.

First, we detected putative gene clusters in both strains that coincided with highly conserved regions (Fig. 3, outer line, dark blue boxes) and, thus, shared with most of the plant-associated *B. amyloliquefaciens* strains. Unsurprisingly, and as previously noted, the analysis of these gene clusters revealed that they were involved in the biosynthesis of known secondary metabolites, such as fengycin, difficidin, macrolactin, or bacillaene. Second, certain atypical regions overlapped with partially conserved regions (Fig. 3, outer line, light blue boxes),

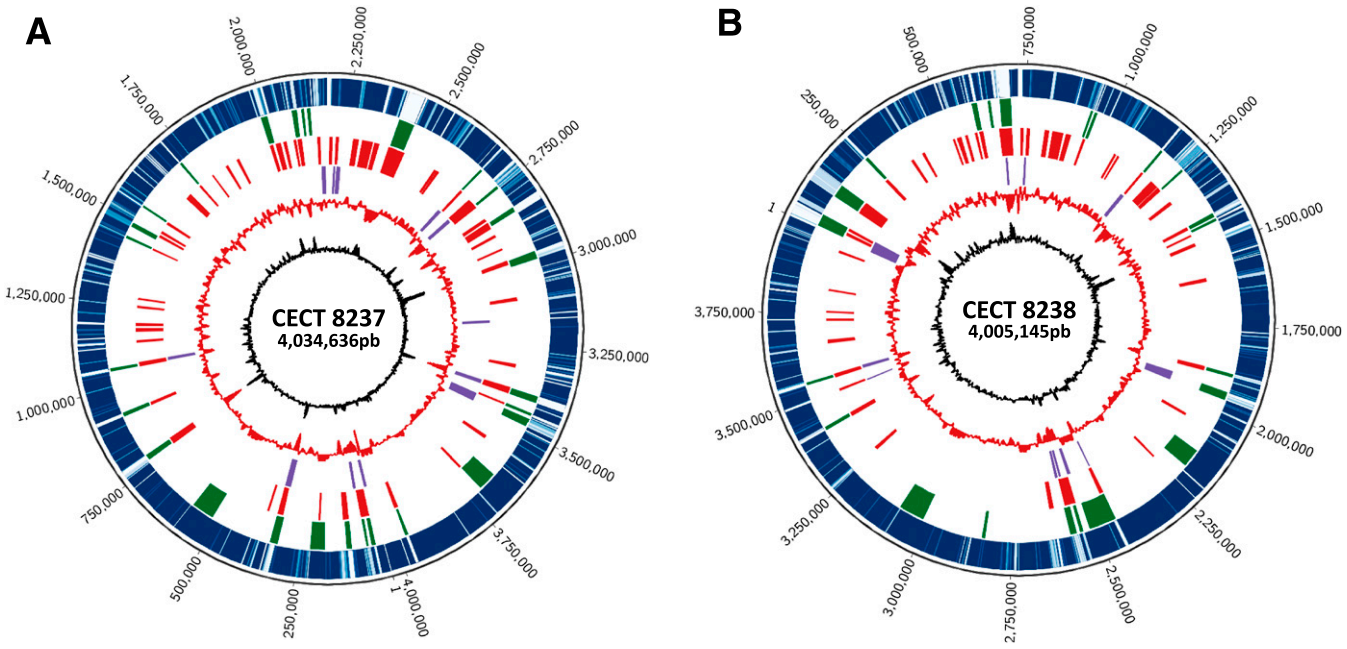


**Fig. 3.** *Bacillus amyloliquefaciens* CECT 8237 and CECT 8238 possess unique genomic regions in comparison with other plant-associated *B. amyloliquefaciens* strains. **A,** The genomes of CECT 8237 or **B,** CECT 8238 were compared with those of other *Bacillus* species available in the database and the results were organized in circles. The outer scale designates the genome's coordinates in base pairs (bp). The first circle shows the comparative genome analysis (MAUVE genome alignment software) with several *B. amyloliquefaciens* potential biological control agents. Regions highly conserved are represented in dark blue; regions slightly conserved are represented in light blue; and specific regions detected only in CECT 8237 or CECT 8238 are represented in white. The second circle shows the nonconserved (green) regions of CECT 8237 and CECT 8238 genomes among the *Bacillus* genus. The third circle (red) shows the regions of DNA acquired horizontally using Alien Hunter software. The fourth circle (purple) shows the predicted prophages (Prophage finder tool). The fifth circle (red) shows the percentage of G+C in relation to the average G+C content in a 2,000-bp window. The inner circle (black) shows the trinucleotide composition in relation to the average trinucleotide composition in a 2,000-bp window.

indicating that some genes of these AR are shared with any of the *B. amyloliquefaciens* plant-associated strains. For example, some regions harbor genes related to fatty-acid metabolism, cell envelope biogenesis, vitamin biosynthesis proteins, and phage-related proteins. In the atypical regions of CECT 8238, we found genes related to peptide synthetase and thioesterase as well as genes dedicated to sporulation, ABC transporter proteins, or *N*-acetyltransferases. Third, and even more interesting, we found 13 gene clusters coincident with unique genomic windows in CECT 8237 (Fig. 3A, outer line, white boxes), indicating that they were absent in other *B. amyloliquefaciens* strains and poorly or not conserved in the *Bacillus* genus. A total of 28 atypical regions were detected in the *B. amyloliquefaciens* CECT 8237 genome, and 20 of them might be considered genomic islands, according to the Alien Hunter analysis, local deviations in the trinucleotide usage patterns, and variation in the percent GC content (Fig. 3A). These data are suggestive of genetic acquisition after successive events of horizontal transfer. According to this idea, four of these genomic islands also overlapped with putative prophages, as predicted by the Prophage finder algorithm, and two other atypical regions contained genes related to phage proteins or mobile element domains as resolvases and transposases. The size of the genomic islands ranged from 4 to 24 kb, and most of the genes encoded hypothetical proteins with unknown functionality. In addition, we identified other genes that are tentatively associated with the interaction of bacteria with the host plants, such as the *BAMY6639_15165* gene encoding a putative adhesion protein with a predicted sorting signal for its covalent anchorage to the bacterial cell envelop, which might hypothetically mediate cell-to-cell interactions or adherence to the host surfaces (AR22). In the category of essential metabolism, we found xanthine dehydrogenases (AR15), which participate in purine metabolism and lipases (AR12), *N*-acetyltransferases (AR2, AR5), and malonyl-CoA transacylase (AR18), which is involved in fatty-acid metabolism. Finally, it was interesting to detect genes that were hypothetically involved in the transport of bacitracin (AR3), a branched cyclic dodecylpeptide antibiotic, which is nonribosomally produced by *B. licheniformis* and *B. subtilis* strains (Dintner et al. 2014) but is not homologous to the biosynthetic machinery. This finding led us to hypothesize about the putative mechanism by which this strain might possess immunity to bacitracin and related metabolites.

A similar number of atypical regions (26) were detected in the *B. amyloliquefaciens* CECT 8238 genome, and 15 of these regions appeared to be genomic islands (from 4 to 33 kb) (Fig. 3B). Seven of these genomic islands were identified as putative prophages, and an additional region contained genes related to phage proteins or mobile element domains such as peptidases, terminases, and integrases. As observed in the other strain, most of the genes contained in these genomic islands are hypothetical proteins, but it was also possible to identify genes coding proteins with known functionalities. For the genes overlapping unique genomic windows (Fig. 3B, outer line, white boxes), we found interesting features, i.e., the enzymes beta-xylosidase and sucrose-6-phosphate hydrolase, which enable this bacterium to utilize sucrose as an alternative carbon source to glucose, and the xanthine dehydrogenases and *N*-acetyltransferases, which (as previously indicated for CECT 8237) are involved in fatty-acid metabolism. We also detected the biosynthetic genes of macrolactin, difficidin, or bacillaene, metabolites related to direct antagonism of pathogenic microbes. It is worth noting the presence of the ICEBs1 excisionase, which has been previously reported to be part of the type IV secretion system in *B. subtilis* 168, a hypothetical mechanism for horizontal transfer of bacterial goods. Further analysis of this region led us to identify all of the putative genes implied in this process (*ydcQ*, *yddE*, *yddD*, *yddG*, *yddB*, *yddC*, and *yddH*) (Alvarez-Martinez and Christie 2009).

To confirm the specificity of some of these atypical regions in our strains, we designed pairs of primers on the *BAMY6639_17480* gene in CECT 8237 and on the *BAMY6614_00315* gene in CECT 8238 (details below). In this analysis, we included a collection of strains and isolates from our own laboratory collection. The primers were highly specific and provided a product from the DNA of the corresponding strain CECT 8237 or CECT 8238 but not in the rest of the *Bacillus* strains included in this analysis (Fig. 4A and B). As a control of the diagnostic polymerase chain reaction (PCR), we also tested a pair of primers that amplified a fragment of the gene *bmyB*, implicated in the synthesis of bacillomycin. As expected from the available genome information, a product was obtained in some of the *B. amyloliquefaciens* and *B. subtilis* strains but not in any of the *B. cereus* group or *B. subtilis*-type

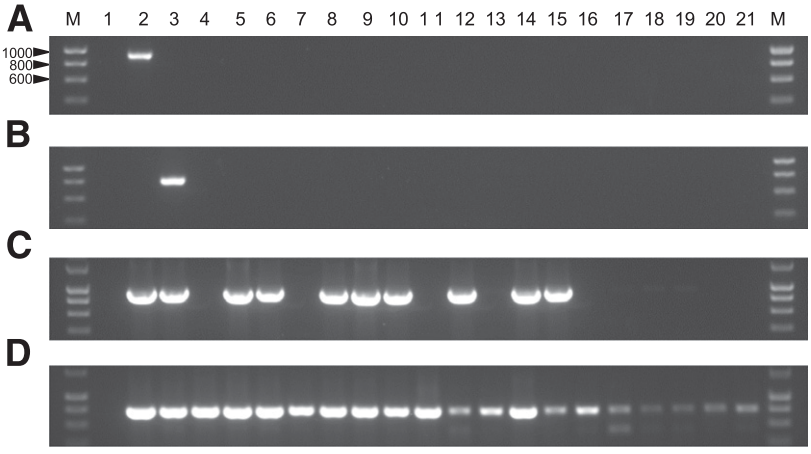


**Fig. 4.** Diagnostic polymerase chain reaction for traceability of *Bacillus amyloliquefaciens* CECT 8237 and CECT 8238. **A** and **B,** Amplicons obtained with a specific pair of primers designed on the unique genomic areas of CECT 8238 or CECT 8237, respectively. **C,** The *bmyB* gene was detected in all of the strains that were producers of the iturins family of lipopeptides (bacillomycin or iturin). **D,** A partial sequence of the housekeeping gene *rpoA* was amplified in all of the strains. Lane 1, Water sample without DNA template; lane 2, CECT 8238; lane 3, CECT 8237; lane 4, *B. subtilis* subsp. *subtilis* 168; lane 5, *B. subtilis* UMAF8561; lane 6, *B. subtilis* UMAF8562; lane 7, *B. subtilis* UMAFBiA758; lane 8, *B. subtilis* UMAF6619; lane 9, *B. subtilis* UMAF1605; lane 10, *B. subtilis* UMAF1610; lane 11, *B. amyloliquefaciens* BGSC10A1; lane 12, *B. amyloliquefaciens* BGSC 10A3; lane 13, *B. amyloliquefaciens* BGSC 10A5T; lane 14, *B. amyloliquefaciens* FZB42; lane 15, *B. amyloliquefaciens* BGSC 10A18; lane 16, *B. flexus* CIP 106928T; lane 17, *B. thuringiensis* subsp. *kurstaki* CECT 4454; lane 18, *B. cereus* ATCC 14579; lane 19, *B. cereus* UMAF8564; lane 20, *Paenibacillus polymyxa* CECT 155, and lane 21, *Brevibacillus laterosporus* CECT 15. M = molecular weight marker HyperLadder 1Kb (Bioline). Numbers on the left are the size of the amplicons in base pairs.

strains (Fig. 4C). Finally, the pair of primers that targeted the *rpoA* gene resulted in a product in all of the strains (Fig. 4D).

### *B. amyloliquefaciens* CECT 8237 and CECT 8238 have singularities in cellular communication and biofilm formation.

It is assumed that the good performance of a biocontrol agent relies on its effective colonization and persistence (Romero et al. 2004; Zeriouh et al. 2014). An important contributor to this bacterial fitness is the formation of biofilms. Within a biofilm, cells are more protected from external aggressions than are planktonic cells, which is due to an extracellular matrix made of exopolysaccharides, proteins, or nucleic acids that provides the entire community with structural robustness (Flemming and Wingender 2010). An important aspect for the establishment and growth of a bacterial biofilm is motility, and both strains display swimming motility, an independent cell movement in liquid medium, and swarming, a multicellular movement driven by the coordinated action of *swrA*, *swrB*, and *swrC* and the lipopeptide surfactin (Supplementary Table S4) (Kearns et al. 2004). Indeed, we have previously demonstrated that the long-term persistence of these strains on melon leaves and efficient biocontrol activity both rely on the formation of biofilms, which appear to be dependent on the action of surfactin (Zeriouh et al. 2014). In agreement with these findings, we identified all of the genes described until now to participate in formation of biofilms, specifically, i) regulatory genes such as *abrB*, *sigW*, *sinI*, *sinR*, and *spo0A* (Vlamakis et al. 2013) and ii) structural genes for the synthesis of the components of the extracellular matrix, including the 15-gene exopolysaccharide operon (*epsA-O*), the operon involved in the formation of TasA amyloid-like protein (*tapA-sipW-tasA*) and the hydrophobin protein BslA (Hobley et al. 2013; Kearns et al. 2005; Romero et al. 2010). We were not able to identify differentiable features between our strains and the *B. amyloliquefaciens* type strain subsp. *plantarum* FZB42, which harbored all of these genes and are almost identical at amino acid sequence level (Supplementary Fig. S3, black boxes). However, the colony architecture was visually different among CECT 8237, CECT 8238, and the type strain FZB42 and the *B. subtilis* type strains 168 and NCBI 3610 (Fig. 5), which led us to speculate about subtle differences in other complementary factors or regulatory pathways that might influence the final phenotypes.

An important aspect of biofilm formation and other physiological responses in bacteria is the cellular communication, and the best-studied is quorum sensing (QS). In *Bacillus* spp., the operon *comQXPA* constitutes a QS system that controls the expression of several genes dedicated to competence or production of surfactin, among other genes (Dogsa et al. 2014), a lipopeptide identified as a self-trigger of biofilm formation (López et al. 2009; Zeriouh et al. 2014). In terms of microbial ecology, this bacterial communication system plays an important role in defining putative interactions and cooperation between members of the same pherotype, which means that they share the same QS type (Dogsa et al. 2014; Stefanic et al. 2012). Unsurprisingly, we identified these loci in CECT 8237 and CECT 8238 and confirmed that their amino acid sequences are slightly conserved with FZB42; *comQ* and *comX*, in particular, retrieved notoriously low levels of similarity. Thus, we characterized the possible diversity of the *comQXPA* locus along several strains corresponding to the *B. subtilis* group, including *B. amyloliquefaciens* plant-associated and non–plant associated strains (Fig. 6). At a glance, it can be observed that the *B. subtilis* group predominantly exhibits overlapping *comQ-comX*, that the *B. amyloliquefaciens* group contains a majority of strains that have independent loci, and that a few *B. amyloliquefaciens* strains show overlapping *comQ-comX-comP* (Fig. 6). However, it was noteworthy that FZB42 clusters together with the *B. subtilis* strains and that the CECT 8237 and CECT 8238 strains fall into divergent pherotypes within the *B. amyloliquefaciens* group. This functional diversification into different pherotypes permits strains to communicate when they belong to the same pherotype but not across different pherotypes; thus, we might assume that this polymorphism emerges as an adaptability response (Dogsa et al. 2014; Stefanic et al. 2012). In the case of CECT 8237 and CECT 8238, these differences might be justified by the fact that they inhabit the same niche (Oslizlo et al. 2015). Therefore, harboring different pherotypes could be beneficial to coexist and avoid interfering gene expression control by QS signals (Tortosa et al. 2001).

In addition to the ComX pheromone, the extracellular Phr peptides regulate relevant biological processes in *Bacillus* spp., such as sporulation, synthesis of antibiotics, and biofilm formation. Each Phr peptide is intimately associated with a cognate intracellular regulator Rap protein that is directly suppressed by the expression of those peptides (Gallego del Sol and Marina 2013). *B. subtilis* encodes 11 Rap proteins (RapA to RapK) and eight Phr peptides (PhrA, PhrC, PhrE, PhrF, PhrG, PhrH, PhrI, and PhrK), which inhibit the response regulators of a diverse two-component regulatory system (Gallego del Sol and Marina 2013). Additional Rap-Phr modules have been found in plasmids of diverse wild *Bacillus* isolates, and examples are Rap60-Phr60 in the plasmid pTA1060, RapP-PhrP in the plasmid pBS32, or RapQ-PhrQ in the plasmid pBSG3 (Yang et al. 2015). Considering the relevance of these *rap-phr* operons in the modulation of biological responses in *Bacillus* spp., we decided to search for the distribution of all of these genes in the genome sequences of CECT 8237, CECT 8238, and other *Bacillus* strains. Most of the *rap* genes and their corresponding antagonistic *phr* were identified, but we did find exceptions. For example, *rapG* is absent in both CECT 8237 and CECT 8238. Strain CECT 8237 appears to have two copies of the gene *rapH*, as represented by *rapH1* and *rapH2*, but not the cognates *phr*. The tandem *rapI-phrI* is only present in strain CECT 8238, an expected finding, given that the presence of the mobile element ICEBs1 is only in this strain (Auchtung et al. 2005). Finally, *rapK-phrK* is present in CECT 8237 and CECT 8238 but is absent in FZB42, and an additional aspartate phosphatase

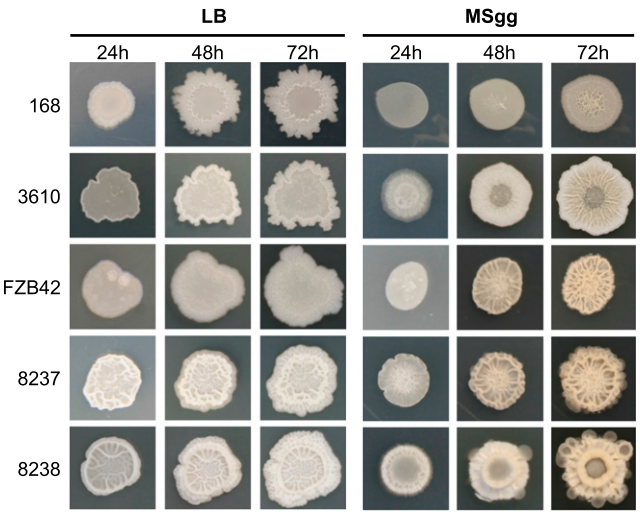


**Fig. 5.** *Bacillus* strains differ in the morphological features of biofilms. Biofilm formation was evaluated as colony morphology in Luria-Bertani (LB) or minimal medium MSgg (minimal salts glycerol glutamate) agar after incubation at 30°C. The strains are *B. subtilis* subsp. *subtilis* 168, *B. subtilis* subsp. *subtilis* NCIB 3610, *B. amyloliquefaciens* subsp. *plantarum* FZB42, *B. amyloliquefaciens* CECT 8237, and *B. amyloliquefaciens* CECT 8238.

annotated as RapX but not the cognate PhrX was identified in both strains (i.e., FZB42 and other *B. amyloliquefaciens* strains but not *B. subtilis*). These findings prove that the regulation system *rap-phr* characterized in *B. subtilis* is not as similar as anticipated for *B. amyloliquefaciens* species. These differences are quite striking and open a wide range of possibilities to examine in future studies to reveal the real implication and effect of these Rap proteins in the developmental program concluding in the formation of biofilms.

### *B. amyloliquefaciens* CECT 8237 and CECT 8238 are triggers of the immune response and growth of plants.

We demonstrated in separate studies that *B. amyloliquefaciens* CECT 8237 and CECT 8238 contribute to plant health by triggering plant defense machinery, known as ISR, and promoting plant growth (García-Gutiérrez et al. 2012, 2013). In these studies, we proved the relevant role of the lipopeptide surfactin in triggering the ISR in melon plants (García-Gutiérrez et al. 2013). Studies in other pathosystems have reported that, as surfactin, the lipopeptide fengycin may also mediate communication with plants, eliciting the ISR in addition to their well-known surfactant or antimicrobial activity (Cawoy et al. 2015; Ongena et al. 2007; Zeriouh et al. 2014). Besides these specific ISR triggers, the cumulative study of diverse biocontrol agents permits a number of generic bacterial attributes to be tentatively implicated in this biocontrol activity (Table 2). Thus, it was not surprising to find in the genomes of both strains the genes involved in the synthesis of these molecules that act as elicitors of the nonspecific basal plant immunity and are known as microbe-associated molecular patterns (MAMPs). Flagellin proteins, which


**Fig. 6.** *Bacillus amyloliquefaciens* CECT 8237 and CECT 8238 belong to different pherogroups, based on the analysis of the competence loci *comQXPA*. The sequences of the competence genes *comQ*, *comX*, *comP*, and *comA* were refined and manually tested for accuracy and concatenated to build a neighbor-joining tree, using MEGA 5 bootstrap values (10,000 replicates), which are shown on the branches. The topology was identical for the trees produced by the minimum evolution and maximum parsimony methods. The sequences from all of the strains used were extracted from published genome sequences. Black triangles represent no overlapping; black circles, *comQ* and *comX* overlaps; the white circle, *comP* and *comA* overlaps; black rectangles, *comQ*, *comX*, and *comP* overlaps; the white rectangle, *comX*, *comP*, and *comA*. Arrows indicate the location of the CECT 8237, CECT 8238, and FZB42 strains in the tree.

are essential in motility, are also known as triggers of the innate plant immune response against potential pathogens. Other known MAMPs present in these strains are encoded by the operon *tuaA-tagO*, which participates in the synthesis of teichuronic acid, a basic component of the cell walls of gram-positive bacteria, and the elongation factor *tufA* genes (Boller and Felix 2009; Kierul et al. 2015).

In a previous study, we demonstrated that both CECT 8237 and CECT 8238 strains are producers of the auxin phytohormone indole-3-acetic acid (García-Gutiérrez et al. 2012), a molecule known to promote plant growth. Similarly, we found the genes of the tryptophan-dependent pathway implicated in the synthesis of this phytohormone (Table 2). In both genomes, we identified the genes implied in the synthesis of acetoin (*alsS* and *alsD*) or 2,3-butanediol (*bdhA*) (Table 2; Supplementary Fig. S4), two bacterial organic volatile compounds involved in this bacteria-plant communication (Farag et al. 2013). We first attempted to detect the production of both volatiles using a qualitative analysis. The Voges-Proskauer test proved that CECT 8237 and CECT 8238 produce acetoin (reddish color of the medium) as well as the indicator strains *B. subtilis* 168, NCIB3610, or *B. amyloliquefaciens* FZB42 (Kierul et al. 2015; Nicholson 2008). The analysis of bacterial supernatants in thin-layer chromatography (TLC) plates revealed a spot with the retention factor (Rf) similar to a 30 mM butanediol standard solution as the control (Nicholson 2008). Next, we evaluated the kinetics of production of both molecules over time, using a gas chromatography–mass spectrometry (MS) analysis of bacterial supernatants, as described below (Supplementary Fig. S5, characteristic chromatogram and mass spectra corresponding to each molecule). It was interesting that *B. subtilis* 168, NCIB3610, or *B. amyloliquefaciens* FZB42 produced more acetoin than did butanediol, whereas the opposite occurred in CECT 8237 and CECT 8238 (Fig. 7A). In addition, the amount of butanediol was surprisingly higher (one order of magnitude) in CECT 8237 and CECT 8238 than in the rest of the strains. In general, the kinetics were similar in these experimental conditions. The amount of both volatiles in the supernatants were picked at 24 h of growth, with the exceptions of FZB42, which accumulates more acetoin at 36 h, and CECT 8237, which accumulates more acetoin at 12 h and butanediol at 36 h. As anticipated from these findings, all of the strains promoted the formation of heavier and more abundant roots of melon seeds than the untreated seeds (Fig. 7B and C). However, no clear differences could be determined between treatments. Based on these observations, we hypothesize that these molecules are important in the PGP activity of CECT 8237 and CECT 8238, as previously observed in other *Bacillus* strains (Farag et al. 2013). However, this finding needs to be confirmed with further analysis.

In addition to the production of plant growth promoters, beneficial bacteria may contribute to plant health as biofertilizers (Pérez-García et al. 2011). Indeed, in the genomes of both strains, we found the genes *yusV* and *yclQ*, putatively involved in iron mobilization, and the group of genes orthologous to *yvrCBA* in *B. subtilis* 168, which is possibly related to vitamin B12 transport (Table 2). Finally, as mentioned above, the *phy* gene that encodes 3-phytase, an enzyme thoroughly reported as a key element in biofertilization, was also identified. Under phosphate-starvation conditions, this enzyme contributes to the phytate-associated phosphorus available in soil and helps eliminate chelate-forming phytate (Fasimoye et al. 2014).

### *B. amyloliquefaciens* CECT 8237 and CECT 8238 are biological factories of antimicrobial compounds.

Secondary metabolites are important and versatile weapons that bacteria use to fight other microbes and are thus highly valuable in plant protection against microbial pathogens (Chen et al. 2009; Pérez-García et al. 2011; Yuan et al. 2012). We experimentally demonstrated the production of lipopeptides,

**Table 2.** Genes involved in the beneficial contribution of *Bacillus amyloliquefaciens* CECT 8237 and CECT 8238 to plant health

| Genes | Protein | Functionality | CECT 8237 % ID[z] | CECT 8238 % ID[z] |
|---|---|---|---|---|
| Bacterial target molecules for general plant immune response | | | | |
| *flgK* | Flagellin HAP1 | Involved in elicitation of plant basal defense | 99 | 99 |
| *fliD* | Flagellin HAP2 | Involved in elicitation of plant basal defense | 95 | 94 |
| *Hag* | Flagellin HAG | Involved in elicitation of plant basal defense | 93 | 90 |
| *tufA* | Elongation factor EF-Tu | Involved in elicitation of plant basal defense | 99 | 100 |
| *tuaA-tagO* | Operon involved in teichuronic acid/ lipopolysaccharide biosynthesis | Involved in elicitation of plant basal defense | 99 | 99 |
| Biofertilization | | | | |
| *yvrC* | Putative iron-binding protein | Putative iron availability | 99 | 99 |
| *yvrB* | Putative iron permease | Putative iron availability | 99 | 99 |
| *yvrA* | Putative ABC transport system ATP-binding protein | Putative iron availability | 98 | 99 |
| *yusV* | Putative iron(III) ABC transport ATPase component | Putative iron availability | 99 | 99 |
| *Phy* | Phytase | Phosphate availability | 99 | 98 |
| *yclQ* | Ferrichrome ABC transporter | Transport/binding proteins and lipoproteins | 99 | 99 |
| Phytostimulation and induced systemic resistance | | | | |
| *dhaS* | Putative indol 3-acet-aldehyde dehydrogenase | Involved in Trp-dependent indole 3-acetic acid (IAA) synthesis | 99 | 99 |
| *ysnE* | Putative IAA-acetyl-transferase | Involved in Trp-dependent IAA synthesis | 98 | 98 |
| *yhcX* | Nitrilase | Involved in Trp-dependent IAA synthesis | 99 | 99 |
| *alsD* | α-Acetolactate decarboxylase | Synthesis of acetoin | 98 | 99 |
| *alsS* | Acetolactate synthase | Synthesis of acetoin | 99 | 99 |
| *alsR* | LysR transcription regulator | Regulator of the *alsDS* operon | 99 | 99 |
| *bdhA* | Acetoin reductase/ 2,3-butanediol dehydrogenase | Synthesis of 2,3-butanediol | 99 | 99 |

[z] Percent identity (ID) compared with *B. amyloliquefaciens* FZB42. The percentage of coverage is 100%, except for the *hag* gene, which is 70 or 81%, respectively, in CECT 8237 or CECT 8238.

surfactin, iturins, and fengycins (Romero et al. 2007a). Genome data mining using the Anti-Smash platform confirmed these findings and revealed the detection of operons for the production of i) bacillibactin siderophore (Miethke et al. 2006), ii) dipeptide bacilysin (Steinborn et al. 2005), iii) macrolactin (Yuan et al. 2012), iv) bacillaene (Butcher et al. 2007), and v) difficidin (Chen et al. 2009). High-pressure liquid chromatography (HPLC) –electrospray (ESI)-MS analysis of the methanolic extracts obtained from cell-free supernatants of both strains grown in Landy broth (Chen et al. 2009) revealed the presence of bacilysin, bacillaene, and dihydrobacillaene (Fig. 8, left charts). The same analysis of the methanolic extracts from the cultures in the GA medium (Chen et al. 2009) showed traces of bacillibactin, difficidin, and macrolactin (Fig. 8, right charts; Supplementary Fig. S6).

In addition, our analysis showed the presence of additional gene clusters that are hypothetically involved in the synthesis of unknown or uncharacterized secondary metabolites (Fig. 9). Interestingly, these novel gene clusters were contained in the unique and atypical regions that were previously described for these strains (Fig. 3), indicating that they are nonconserved among the *Bacillus* genus. In the CECT 8237 strain, we only found a complete gene cluster present in the recently sequenced genome of *B. amyloliquefaciens* LFB112 that was likely involved in the synthesis of a novel nonribosomal peptide synthetase (Fig. 9A) (Cai et al. 2014). In the CECT 8238 strain, we found a group of genes with high probability to synthesize an unknown nonribosomal peptide synthetase (Fig. 9B), which correlated with AR3. This region has been previously reported in *B. amyloliquefaciens* subsp. *plantarum* NAU-B3, YAU B9601-Y2, and *B. amyloliquefaciens* Y2 strains, which are all associated with plants (He et al. 2012). The functionality and structures of the molecules potentially synthesized by these hypothetical gene clusters have not yet been demonstrated.

In addition to antimicrobial activity against bacterial and fungal pathogens, a nascent research field in biological control is the evaluation of potential beneficial bacteria against plant pathogenic nematodes. The wide distribution of serine cuticle-degrading proteases in the *Bacillus* genus suggests that they might be essential in controlling nematode populations in the soil. Indeed, we found two genes in our analysis, *nprE*, which codes an extracellular neutral protease precursor, and *aprE*, which codes a putative cuticle-degrading protease precursor (Lian et al. 2007).

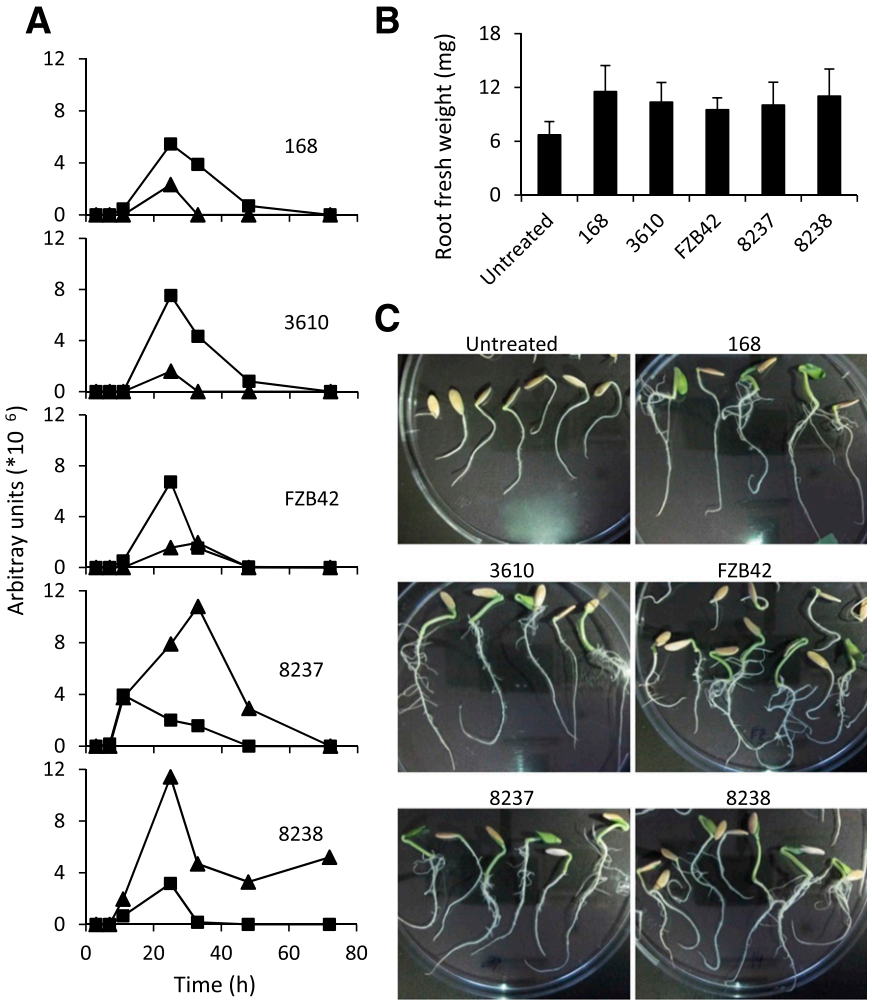


**Fig. 7.** *Bacillus amyloliquefaciens* CECT 8237 and CECT 8238 produce the plant-growth promoter volatile compounds acetoin and 2,3-butanediol. **A,** Gas chromatography–mass spectrometry analysis of bacterial supernatants was used to determine the kinetics of the production of acetoin (black squares) and 2,3-butanediol (black triangles) by *Bacillus* strains grown in Luria-Bertani medium supplemented with 1% glucose at 30°C. **B,** Estimation of the root weight indicates that bacilli induce the growth of the root system of melon seeds. **C,** Representative pictures of the experiment at 7 days shows more abundant and developed roots in melon seeds treated with bacilli compared with untreated seeds. Untreated = nontreated control, 168 = *B. subtilis* subsp. *subtilis* 168, 3610 = *B. subtilis* subsp. *subtilis* NCIB 3610, FZB42 = *B. amyloliquefaciens* FZB42, 8237 = *B. amyloliquefaciens* CECT 8237, and 8238 = *B. amyloliquefaciens* CECT 8238.

**The biocontrol activity of *B. amyloliquefaciens* CECT 8237 and CECT 8238 is supported by their genetic attributes.**

The genome sequences of *B. amyloliquefaciens* CECT 8237 or CECT 8238 are largely shared with other *B. amyloliquefaciens* strains related to plants, including the *B. amyloliquefaciens* subsp. *plantarum* FZB42 reference strain. However, we have shown genetic peculiarities to illustrate the battery of secondary metabolites that might contribute to their outstanding suppressive skills in the plant phyllosphere. To test this idea, we compared the antagonistic activity of *B. amyloliquefaciens* CECT 8237 or CECT 8238 and FZB42 and that of *B. subtilis* 168 and NCIB3610, which are strains that produce different combinations of antimicrobials against *Pectobacterium carotovorum*, the bacterial pathogen responsible for soft rot disease, and *Podosphaera xanthii*, the fungal pathogen responsible for cucurbit powdery mildew disease. In general, *B. amyloliquefaciens* strains performed better than did any of the *B. subtilis* strains in both pathosystems (Fig. 10; Table 3). The protection provided by the CECT 8237 or CECT 8238 strains against fungal (68 and 73% reduction in severity, respectively) and bacterial diseases (61 and 53% reduction in severity, respectively) was slightly better than but not significantly different from that of FZB42 (Table 3). The *B. subtilis* NCIB 3610 strain showed restricted biocontrol activity against both diseases, and *B. subtilis* 168 failed to control *P. carotovorum*, although it provided some protection against *Podosphaera xanthii* (a reduction of 39% of the symptoms). These findings seem to indicate the potential of these strains to produce antimicrobial compounds. Indeed, the *B. subtilis* strains included in this analysis do not produce the polyketides macrolactin and difficidin or lipopeptides of the iturin family, which are important bacterial factors in the control of these diseases (Yuan et al. 2012; Zeriouh et al. 2011). In addition, the strain *B. subtilis* subsp. *subtilis* 168 is not able to synthesize the lipopeptide surfactin, due to a mutation of the phosphopantheinyl transferase *sfp*, which could explain its limited capability to control both diseases as compared with *B. subtilis* NCIB 3610 (Marahiel 2009). In contrast, the *B. amyloliquefaciens* strains that provide efficient control of diseases produce a wide variety of secondary metabolites.

In conclusion, we sequenced and assembled the genomes of two biological control agents, *B. amyloliquefaciens* CECT 8237 and CECT 8238, and created ad hoc scripts to efficiently subtract the information related to their biocontrol skills. Our analysis confirmed that these strains, similar to other *B. amyloliquefaciens* strains associated with plants, have specific genetic features that are absent in industrially relevant strains and would explain their improved performance in these habitats. The consumption of diverse carbon sources and the production of secondary metabolites prevail as the most outstanding differences that ensure an ecological advantage over competitors. In addition, we found attributes of CECT 8237 and CECT 8238 that might explain their high potential as multifaceted biocontrol agents, namely, i) the production of secondary metabolites, which might have double functionality in direct antagonism of plant pathogens and possible competitors or in bacterial fitness and adaptability, ii) multicellularity (biofilm formation and motility) that provides support to colonization and persistence on plant surfaces, and iii) production of phytohormones and MAMPs that trigger plant growth and ISR. Finally, there are "atypical regions" in the genome of both strains that represent new and exciting areas to investigate additional genetic features with potential implications in bacterial fitness and contributions to plant health.

## MATERIALS AND METHODS

**Bacterial strains, growth conditions, and DNA preparation.**

All strains were grown in Luria-Bertani medium (LB) at 37°C overnight with agitation, except *Brevibacillus laterosporus*

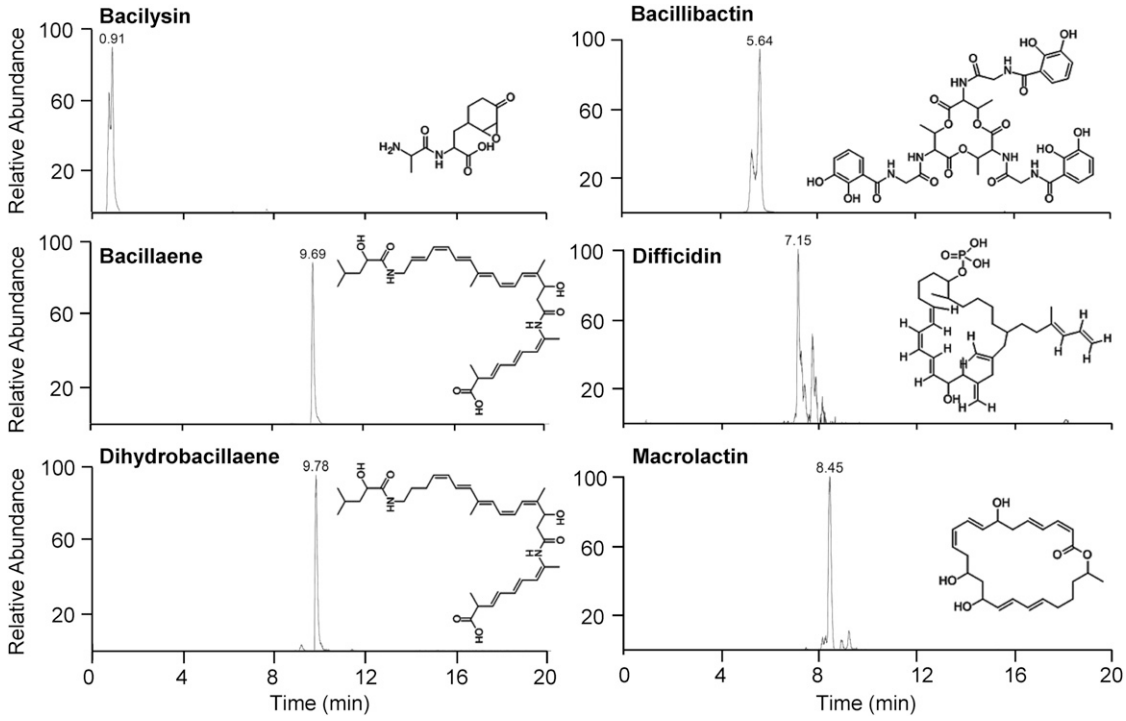


**Fig. 8.** *Bacillus amyloliquefaciens* CECT 8237 and CECT 8238 produce a variety of known secondary metabolites. High-performance liquid chromatography–electrospray mass spectrometry analysis of methanolic extracts from the cell-free supernatants of cultures in Landy medium for 40 h at 30°C revealed the presence of bacilysin, bacillaene, and dihydrobacillaene. The analysis of methanolic extracts from the cell-free supernatants of cultures in gibberellic acid medium for 24 h at 30°C revealed the production of bacillibactin, difficidin, and macrolactin. The chemical structures of the molecules are included (insets).

CECT 15, which was grown at 28°C. For biofilm experiments, 2 µl of the starting culture was spotted onto minimal medium or LB agar plates and was incubated without agitation at 30°C (Romero et al. 2010).

**Genome sequencing and assembling.**

Chromosomal DNA was isolated using the Jet-Flex genomic DNA purification commercial kit (Genomed Laboratories), and genome sequencing and assembling were performed at the Beijing Genomic Institute (BGI, Shenzhen, China). A library of randomly sheared DNA fragments 0.5 to 2 kb in size was subjected to Illumina GA II (Solexa) sequencing. All of the generated reads were qualitatively assessed before assembling with SOAP de novo. Primer walking and PCR amplification were used to fill the remaining gaps and to solve misassembled regions caused by repetitive sequences.

The genome sequences were deposited in National Center for Biotechnology Information (NCBI) GenBank under accession numbers CP006058 (CECT 8237) and CP006960 (CECT 8238).

**Genomic data and annotation.**

The assembled genomes of CECT 8237 and CECT 8238 were submitted to the NCBI Prokaryotic genome annotation pipeline for automatic annotation and were manually reviewed. Gene locations and protein products were generated from the above annotation (ASN.1 file), using the script "asn2all" belonging to the NCBI Toolkit. Genome sequences of the *Bacillus* genus strains used in this work were downloaded from the NCBI complete bacterial genome repository. Annotation of COG categories were computed by aligning the set of predicted protein sequences against the COG PSSM of NCBI's Conserved Domain database using rpsblast. Only hits with an E value ≤ 0.00001 were retained.

**Circular genome visualization.**

Circular layouts were generated using Circos (Krzywinski et al. 2009).

**Phylogenetic analyses.**

For these analyses, *Bacillus* sp. strains selected included the closely related species *B. amyloliquefaciens*, *B. subtilis*, *B. atrophaeus*, and *B. licheniformis* and several representatives of other species such as *B. thuringiensis* and *B. megaterium*. The comparison of partial sequences corresponding to several housekeeping genes, a sporulation gene, and some sigma factor genes was performed. All of these sequences were refined and adjusted to a consensus size by ContigExpress, which belongs to Vector NTI Advance 10 (Life Technologies) and was concatenated in the following order: *nusA-rpoA-dnaA-rpoB-gyrA-gyrB-rpoC-spoVG-sigW-sigH-sigB*. Multiple alignments were conducted with ClustalW and a phylogenetic tree was constructed in MEGA 5, using the neighbor-joining method, based on a pairwise distance matrix with the Tamura-Nei nucleotide substitution model (Tamura et al. 2011).

A similar procedure was used in the comparative study with the competence loci *comQXPA*. The sequences of these proteins were separately refined and adjusted to a consensus size. Further, they were concatenated as they are naturally organized in the genome (*comQ-comX-comP-comA*) and

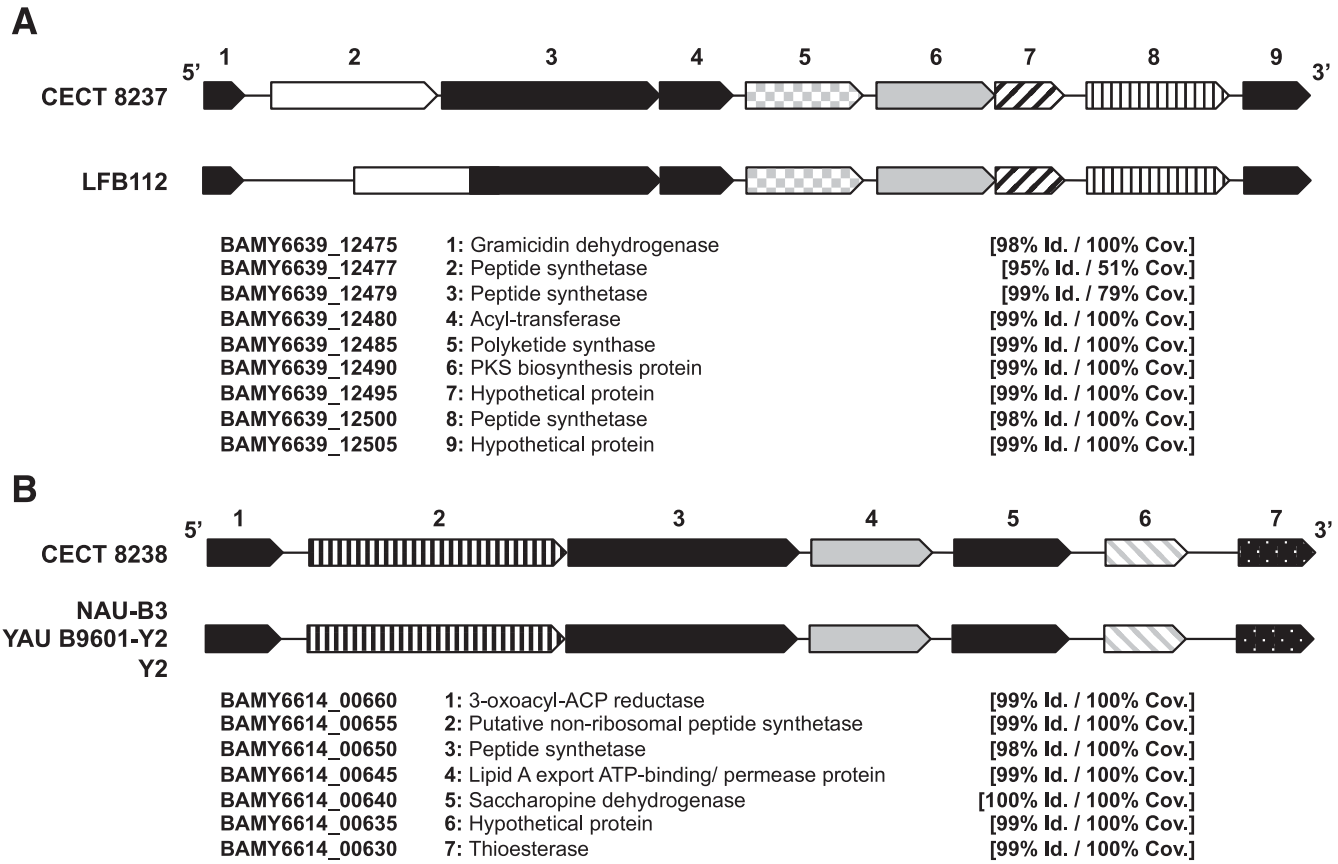


**Fig. 9.** *Bacillus amyloliquefaciens* CECT 8237 and CECT 8238 have putative gene clusters dedicated to production of novel secondary metabolites. **A,** A gene cluster putatively involved in the synthesis, modification, and transport of a new nonribosomal peptide detected in CECT 8237 and *B. amyloliquefaciens* LFB112. **B,** A gene cluster possibly involved in the synthesis, modification, and transport of an uncharacterized nonribosomal peptide and found in CECT 8238 and the *B. amyloliquefaciens* strains NAU-B3, YAU B9601, and Y2.

were aligned to build a phylogenetic tree using the MEGA 5 software and the neighbor-joining method, based on a pairwise distance matrix with the Tamura-3-parameter nucleotide substitution model. The topology of the trees was evaluated by the bootstrap resampling method with 10,000 replicates.

For the construction of phylogeny, only plant-associated *B. amyloliquefaciens* strains (Table 1) were considered. Then, CDS ubiquitously present in such strains and absent in industrial *B. amyloliquefaciens* strains were selected. This group initially contained 24 CDS, but eight were excluded based on their divergent lengths (more than 20 amino acids than the same CDS in the rest of the strains). For each strain, we concatenated the rest of the 16 ubiquitous CDS and used MEGA 5 to align them and build a phylogenetic tree basing on a neighbor-joining method as in the above trees.

**Comparative genome analysis.**

To evaluate the genetic differences between *B. amyloliquefaciens* strains corresponding to plant-associated and non–plant associated representatives (Table 1), the pan-genome of both groups was determined using an iterative BLASTp approach (E value ≤ $5 \times 10^{-7}$). Starting with the whole set of predicted CDS from CECT 8237, we performed iterative BLASTp versus all the plant-associated *B. amyloliquefaciens* genomes sequentially. The genome of CECT 8238 was used for the first iteration. During each iteration, the CDS missing (BLASTp nonhits) from the current version of the pan-genome were identified and added to the pan-genome before the next strain was considered (Baltrus et al. 2011). A similar procedure was carried out to determine the pan-genome corresponding to non–plant associated strains. The CDS of each pan-genome were compared with each other using BLASTp (E value ≤ $5 \times 10^{-7}$), and those only present in either pan-genome were retained. Using this method, it was possible to identify all of the CDS contained in at least one of the group of strains but not in the other.

For the detection of atypical regions in CECT 8237 and CECT 8238 strains nonconserved in the *Bacillus* genus, windows of at least seven consecutive genes with low conservation coverage were retained. First, the conservation measure (CM) for a given gene product in a specific strain with respect to the studied strain (CECT 8237 or CECT 8238, as it corresponds) was calculated. Then, a matrix was generated consisting of the CM values corresponding to all genes of a given strain. In order to evaluate the conservation coverage that is associated to a given gene product belonging to the studied strain (CECT 8237 or CECT 8238), the average of all CM for that gene was estimated. Finally, these data are used to select windows of at least seven consecutive genes whose conservation coverage is under the 30th percentile.

**Prediction of protein domains and events of horizontal transference of DNA.**

Prediction of biosynthesis gene clusters dedicated to production of secondary metabolites was performed with the antiSMASH software 2.0 (Blin et al. 2013). The prediction of putative horizontal gene transfer events was performed using the Alien Hunter software (Vernikos and Parkhill 2006), and potential prophage loci were detected with the Prophage Finder tool (Bose and Barber 2006). To evaluate the trinucleotide pattern for each chromosome, the distribution of all 64 trinucleotides was determined for the whole DNA sequence and for 2-kb subwindows. Then, the $\chi^2$ statistic on the difference between the trinucleotide composition of each window and that of the whole chromosome were computed. Large values for $\chi^2$ in a given window denote different trinucleotide composition from the rest of the chromosome. The probability values were computed assuming uniform distribution of the DNA composition along the genome. Considering the level of error of this assumption, high $\chi^2$ values should be interpreted as indicators of unusual regions on the chromosome that require further investigation.

**PCR analysis for bacterial traceability.**

DNA was obtained from bacteria grown in LB medium for 12 h at 37°C, using the UltraClean microbial DNA isolation commercial kit (Mo Bio Laboratories). The pair of primers CECT 8237-Fw (GGCAGAACAAGAGCAATC) and CECT 8237-Rv (GTCCATGTGAGTGAAATCC) were designed on the loci *BAMY6639_17480* of *B. amyloliquefaciens* CECT 8237, and CECT 8238-Fw (ATTGCCTTTTGGATGATTCG) and CECT 8238-Rv (TCAAGTGGATTTTTGGGAGA) on the loci *BAMY6614_00315* of *B. amyloliquefaciens* CECT 8238. The pair of primers Bm-Fw (GCGATTTGTATGCCTATTT TACA) and Bm-Rv (GCCGTCATACAATTGAATCAGTT)

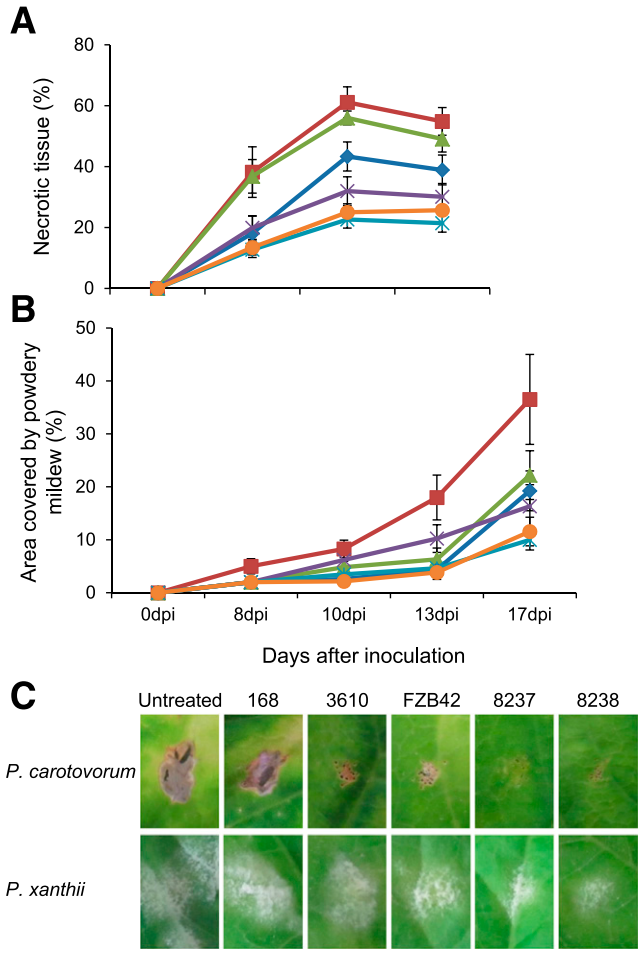


**Fig. 10.** *Bacillus amyloliquefaciens* CECT 8237 and CECT 8238 are more efficient biocontrol agents than other *Bacillus* strains. The biocontrol traits of different strains were tested on melon leaves against **A,** the plant pathogenic bacteria *Pectobacterium carotovorum* subsp. *carotovorum* and **B,** the powdery mildew fungus *Podosphaera xanthii*. The melon leaves were first treated with washed cells of the biocontrol agent and, later, were inoculated with the plant pathogens. The progress of the symptoms was evaluated as the percentage of necrotic leaf induced by *P. carotovorum* or the percentage of leaf area covered by powdery mildew. untreated = non-treated control (red square), 168 = *B. subtilis* subsp. *subtilis* 168 (green triangle), 3610 = *B. subtilis* subsp. *subtilis* NCIB 3610 (blue diamond), FZB42 = *B. amyloliquefaciens* FZB42 (purple X), 8237 = *B. amyloliquefaciens* CECT 8237 (blue X), and 8238 = *B. amyloliquefaciens* CECT 8238 (orange circle). **C,** Representative pictures of each treatment were taken 10 or 17 days after inoculation with *P. carotovorum* or *Podosphaera xanthii*, respectively.

were designed on the *bmyB* loci of the operon involved in the synthesis of the lipopeptide bacillomycin. The pair of primers Rp-Fw (GCGTGGATATGGTACTAC) and Rp-Rv (CTTCAAGTGATTTGCGTCC) amplified a fragment of the housekeeping gene *rpoA*. Standard PCR was performed using the Go Taq Flexi DNA polymerase with 1 to 2 ng of DNA in a final reaction volume of 25 µl, according to the manufacturer's instructions (Promega). The PCR reaction cycle was performed at 94°C for 2 min, followed by a 35-cycle amplification program (94°C for 1 min, 61°C for 1 min, and 72°C for 1 min) and a final extension cycle at 72°C for 7 min. The primers that partially amplified *rpoA* (a housekeeping gene present in all strains) or *bmyB* (gene involved in the biosynthesis of bacillomycin) were included in these studies.

**Detection of volatile compounds and secondary metabolites.**

The detection of acetoin was conducted using the Voges-Proskauer test (Nicholson 2008). The bacterial strains were grown in LB supplemented with 1% glucose at 37°C with slight agitation. After the addition of reagents, the culture turned red to indicate the production of acetoin. The production of 2,3-butanediol was followed using the protocol previously described (Nicholson 2008). Briefly, cells were grown in the same conditions described above. After 24 h of growth, the cells were removed by centrifugation and filtration, and 30 µl of the supernatants were developed in TLC precoated silica gel plates using n-hexane-ethyl acetate (1:5) as the mobile phase. The presence of 2,3-butanediol was revealed with a Cerium-ammonium-molybdate, CAM solution (40 g of ammonium pentamolybdate, 1.6 g of cerium (IV) sulfate, and 800 ml of diluted sulphuric acid [1:9, with water, vol/vol]) and was incubated in an oven at 150°C. A standard of 2,3-butanediol migrated at an Rf of 0.5 under these conditions. The identification and quantification of 2,3-butanediol was further performed on a Thermo Scientific GC Ultra coupled to a DSQ simple quadrupole. The gas chromatograph is equipped with a DB 5 column (15 m × 25 mm, at a film thickness of 0.25 µm). The operating conditions included a helium flow rate of 1.2 ml/min and an injection port temperature of 250°C. The temperature of the column was programmed for 1 min isothermal at 4°C, increased by 10°C per minute for 2 min and 25°C per minute for 2 min.

For the detection of secondary metabolites, CECT 8237 and CECT 8238 were grown in Landy broth or GA medium (Chen et al. 2009). The cultures in Landy medium were grown 12 or 40 h at 30°C and 150 rpm. After centrifugation for 40 min at 11,000 rpm and 4°C, the supernatants were extracted by solid phase extraction in a Merck LiChrolut RP-18 cartridge. After binding and subsequent washing steps with MilliQ water (5-bed volume), the metabolites were eluted with methanol (2-bed volume), were dried under a vacuum, and were resuspended in 100 µl of methanol. The cultures in GA broth were grown at 30°C and 165 rpm for 24 h, and the supernatants were collected after centrifugation for 40 min at 11,000 rpm. The samples were extracted three times with ethyl acetate, were evaporated, and were dissolved in 100 µl of methanol. The resulting samples were analyzed by reverse-phase HPLC (Dionex 3000, Thermo C18, 50 × 2.1 mm, 2.6 µm, Accucore RP-MS) coupled to a ESI-MS (Orbitrap; Q-Exactive, Thermo) in positive mode. The temperature in the HPLC was held constant at 30°C during the experiment. The run was performed with a flow rate of 0.2 ml/min and molecules were eluted in a binary solvent system (solvent A: $CH_3CN$; solvent B: water-0.1% formic acid) as follows: 80% B for 1 min, followed by a 7-min gradient from 80% B to 10% B for 15 min and a subsequent 5-min gradient from 10% B to 80% B for 3 min.

**Promotion of root growth and biological control assays.**

Melon seedlings were used to evaluate the potential of bacterial strains to promote root growth (García-Gutiérrez et al. 2012). The strains were grown at 28°C for 12 h. The cells were pelleted by centrifugation, were washed in water twice, and were resuspended in water to a final cell density of $10^8$ CFU/ml. The melon seed were then soaked in the bacterial suspensions for 1 h and were incubated in a growth chamber for 7 days at 25°C with a 16-h light photoperiod. The roots were split and the weight measured in a precision balance.

Biological control assays were conducted on detached melon leaves (*Cucumis melo* cv. Rochet) using the double petri plate system described elsewhere (Romero et al. 2004, 2007b; Zeriouh et al. 2011). Briefly, cell suspensions of *Bacillus* strains were sprayed on melon leaves. After the whole leaf surface was completely air dried, the leaves were inoculated with cell suspensions ($10^8$ CFU/ml) of the phytopathogenic bacteria *Pectobacterium carotovorum* subsp. *carotovorum* NCPPB 2349 or conidial suspensions ($10^5$ spores/ml) of the cucurbit powdery mildew fungus *Podosphaera xanthii* 2086. After inoculation, the leaves were maintained at 25°C with a 16-h photoperiod. Disease severities were evaluated according to a 0 to 3 scale of specific values based on the diameter of chlorotic and necrotic symptoms for the bacterial disease and the percentage of the leaf area covered with fungal biomass for powdery mildew (Romero et al. 2004, 2007b; Zeriouh et al. 2011).

**Statistical analysis.**

The data were analyzed using SPSS v.20.0 software (SPSS Inc.). A one-way analysis of variance was applied, and the means of each treatment were separated by Fisher's least significant difference test ($P = 0.05$).

**Table 3.** The severity of disease symptoms caused by *Pectobacterium carotovorum* subsp. *carotovorum* and *Podosphaera xanthii* on melon leaves after treatments with vegetative cells of *Bacillus subtilis* or *B. amyloliquefaciens* strains

| | ***P. carotovorum* subsp. *carotovorum*** | | ***Podosphaera xanthii*** | |
|---|---|---|---|---|
| **Treatments** | **Severity[y]** | **Reduction[z]** | **Severity** | **Reduction** |
| Untreated | 55 a | – | 37 a | – |
| *B. subtilis* 168 | 49 ab | 11 | 22 b | 39 |
| *B. subtilis* 3610 | 39 bc | 29 | 19 b | 47 |
| *B. amyloliquefaciens* FZB42 | 30 cd | 45 | 16 bc | 55 |
| *B. amyloliquefaciens* CECT 8237 | 21 d | 61 | 12 c | 68 |
| *B. amyloliquefaciens* CECT 8238 | 26 d | 53 | 10 c | 73 |

[y] Disease severity was assessed for *P. carotovorum* as a percentage of necrotic tissue according to a 0 to 3 scale of symptoms and for *Podosphaera xanthii* as a percentage of leaf area covered by powdery mildew. Values followed by the same letter in each column are not significantly different at $P = 0.05$, according to Fisher's least significant difference test.

[z] Percentage of disease reduction achieved by treatments in regard to values of disease severity on untreated leaves.

## ACKNOWLEDGMENTS

We thank M. L. Pola for help in LC/ESI-MS/MS and GC/MS analyses. We also thank S. de Weert (Koppert B. V.) for critical reading of the manuscript and H. Zeriouh for helpful suggestions and discussion. This work was supported by grants from Plan Nacional de I+D+I of Ministerio de Economía y Competitividad (AGL2010-21848-CO2-01 and AGL-2012-31968) and Incentivos a Proyectos de Excelencia de la Junta de Andalucía (P10-AGR-5797), both co-financed by Fundo Europeu de Desenvolvimento Regional (FEDER) funds (European Union). This work was also supported by grant from Koppert B.V. (8.06/60.4086). D. Romero is funded by the program Ramón y Cajal (RyC-2011-080605), C. Magno is the recipient of a Ph.D. fellowship associated to the Project P10-AGR 5797 from Junta de Andalucía, and P. Martínez-García is supported by the Campus de Excelencia Internacional Andalucía Tech.

## LITERATURE CITED

Alvarez-Martinez, C. E., and Christie, P. J. 2009. Biological diversity of prokaryotic type IV secretion systems. Microbiol. Mol. Biol. Rev. 73: 775-808.

Auchtung, J. M., Lee, C. A., Monson, R. E., Lehman, A. P., and Grossman, A. D. 2005. Regulation of a *Bacillus subtilis* mobile genetic element by intercellular signaling and the global DNA damage response. Proc. Natl. Acad. Sci. U.S.A. 102:12554-12559.

Baltrus, D. A., Nishimura, M. T., Romanchuk, A., Chang, J. H., Mukhtar, M. S., Cherkis, K., Roach, J., Grant, S. R., Jones, C. D., and Dangl, J. L. 2011. Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. PLoS Pathog. 7:e1002132.

Blin, K., Medema, M. H., Kazempour, D., Fischbach, M. A., Breitling, R., Takano, E., and Weber, T. 2013. antiSMASH 2.0—A versatile platform for genome mining of secondary metabolite producers. Nucleic Acids Res. 41:W204-W212.

Boller, T., and Felix, G. 2009. A renaissance of elicitors: Perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. Annu. Rev. Plant Biol. 60:379-406.

Borriss, R., Chen, X.-H., Rueckert, C., Blom, J., Becker, A., Baumgarth, B., Fan, B., Pukall, R., Schumann, P., Spröer, C., Junge, H., Vater, J., Pühler, A., and Klenk, H.-P. 2011. Relationship of *Bacillus amyloliquefaciens* clades associated with strains DSM 7T and FZB42T: A proposal for *Bacillus amyloliquefaciens* subsp. *amyloliquefaciens* subsp. nov. and *Bacillus amyloliquefaciens* subsp. *plantarum* subsp. nov. based on complete genome sequence comparisons. Int. J. Syst. Evol. Microbiol. 61:1786-1801.

Bose, M., and Barber, R. D. 2006. Prophage Finder: A prophage loci prediction tool for prokaryotic genome sequences. Silico Biol. Gedrukt 6:223-227.

Butcher, R. A., Schroeder, F. C., Fischbach, M. A., Straight, P. D., Kolter, R., Walsh, C. T., and Clardy, J. 2007. The identification of bacillaene, the product of the PksX megacomplex in *Bacillus subtilis*. Proc. Natl. Acad. Sci. U.S.A. 104:1506-1509.

Cai, J., Liu, F., Liao, X., and Zhang, R. 2014. Complete genome sequence of *Bacillus amyloliquefaciens* LFB112 isolated from Chinese herbs, a strain of a broad inhibitory spectrum against domestic animal pathogens. J. Biotechnol. 175:63-64.

Carvalhais, L. C., Dennis, P. G., Fan, B., Fedoseyenko, D., Kierul, K., Becker, A., von Wiren, N., and Borriss, R. 2013. Linking plant nutritional status to plant-microbe interactions. PLoS ONE 8:e68555.

Cawoy, H., Debois, D., Franzil, L., De Pauw, E., Thonart, P., and Ongena, M. 2015. Lipopeptides as main ingredients for inhibition of fungal phytopathogens by *Bacillus subtilis/amyloliquefaciens*. Microbial Biotechnol. 8:281-295.

Chen, X. H., Scholz, R., Borriss, M., Junge, H., Mögel, G., Kunz, S., and Borriss, R. 2009. Difficidin and bacilysin produced by plant-associated *Bacillus amyloliquefaciens* are efficient in controlling fire blight disease. J. Biotechnol. 140:38-44.

Dintner, S., Heermann, R., Fang, C., Jung, K., and Gebhard, S. 2014. A sensory complex consisting of an ATP-binding cassette transporter and a two-component regulatory system controls bacitracin resistance in *Bacillus subtilis*. J. Biol. Chem. 289:27899-27910.

Dogsa, I., Choudhary, K. S., Marsetic, Z., Hudaiberdiev, S., Vera, R., Pongor, S., and Mandic-Mulec, I. 2014. ComQXPA quorum sensing systems may not be unique to *Bacillus subtilis*: A census in prokaryotic genomes. PLoS ONE 9:e96122.

Earl, A. M., Losick, R., and Kolter, R. 2007. *Bacillus subtilis* genome diversity. J. Bacteriol. 189:1163-1170.

Farag, M. A., Zhang, H., and Ryu, C.-M. 2013. Dynamic chemical communication between plants and bacteria through airborne signals: Induced resistance by bacterial volatiles. J. Chem. Ecol. 39:1007-1018.

Fasimoye, F. O., Olajuyigbe, F. M., and Sanni, M. D. 2014. Purification and characterization of a thermostable extracellular phytase from *Bacillus licheniformis* PFBL-03. Prep. Biochem. Biotechnol. 44:193-205.

Flemming, H.-C., and Wingender, J. 2010. The biofilm matrix. Nat. Rev. Microbiol. 8:623-633.

Gallego del Sol, F., and Marina, A. 2013. Structural basis of Rap phosphatase inhibition by Phr peptides. PLoS Biol. 11:e1001511.

García-Gutiérrez, L., Romero, D., Zeriouh, H., Cazorla, F., Torés, J., de Vicente, A., and Pérez-García, A. 2012. Isolation and selection of plant growth-promoting rhizobacteria as inducers of systemic resistance in melon. Plant Soil 358:201-212.

García-Gutiérrez, L., Zeriouh, H., Romero, D., Cubero, J., de Vicente, A., and Pérez-García, A. 2013. The antagonistic strain *Bacillus subtilis* UMAF6639 also confers protection to melon plants against cucurbit powdery mildew by activation of jasmonate- and salicylic acid-dependent defence responses. Microb. Biotechnol. 6:264-274.

He, P., Hao, K., Blom, J., Rückert, C., Vater, J., Mao, Z., Wu, Y., Hou, M., He, P., He, Y., and Borriss, R. 2012. Genome sequence of the plant growth promoting strain *Bacillus amyloliquefaciens* subsp. *plantarum* B9601-Y2 and expression of mersacidin and other secondary metabolites. J. Biotechnol. 164:281-291.

Hobley, L., Ostrowski, A., Rao, F. V., Bromley, K. M., Porter, M., Prescott, A. R., MacPhee, C. E., van Aalten, D. M. F., and Stanley-Wall, N. R. 2013. BslA is a self-assembling bacterial hydrophobin that coats the *Bacillus subtilis* biofilm. Proc. Natl. Acad. Sci. U.S.A. 110: 13600-13605.

Kearns, D. B., Chu, F., Branda, S. S., Kolter, R., and Losick, R. 2005. A master regulator for biofilm formation by *Bacillus subtilis*. Mol. Microbiol. 55:739-749.

Kearns, D. B., Chu, F., Rudner, R., and Losick, R. 2004. Genes governing swarming in *Bacillus subtilis* and evidence for a phase variation mechanism controlling surface motility. Mol. Microbiol. 52:357-369.

Kierul, K., Voigt, B., Albrecht, D., Chen, X.-H., Carvalhais, L. C., and Borriss, R. 2015. Influence of root exudates on the extracellular proteome of the plant growth-promoting bacterium *Bacillus amyloliquefaciens* FZB42. Microbiology 161:131-147.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S. J., and Marra, M. A. 2009. Circos: An information aesthetic for comparative genomics. Genome Res. 19:1639-1645.

Lian, L. H., Tian, B. Y., Xiong, R., Zhu, M. Z., Xu, J., and Zhang, K. Q. 2007. Proteases from *Bacillus*:A new insight into the mechanism of action for rhizobacterial suppression of nematode populations. Lett. Appl. Microbiol. 45:262-269.

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., Lin, D., Lu, L., and Law, M. 2012. Comparison of next-generation sequencing systems. J. Biomed. Biotechnol. 2012:251364.

Liu, Y., Lai, Q., Dong, C., Sun, F., Wang, L., Li, G., and Shao, Z. 2013. Phylogenetic diversity of the *Bacillus pumilus* group and the marine ecotype revealed by multilocus sequence analysis. PLoS ONE 8:e80097.

López, D., Fischbach, M. A., Chu, F., Losick, R., and Kolter, R. 2009. Structurally diverse natural products that cause potassium leakage trigger multicellularity in *Bacillus subtilis*. Proc. Natl. Acad. Sci. U.S.A. 106:280-285.

Marahiel, M. A. 2009. Working outside the protein-synthesis rules: Insights into non-ribosomal peptide synthesis. J. Pept. Sci. 15:799-807.

Miethke, M., Klotz, O., Linne, U., May, J. J., Beckering, C. L., and Marahiel, M. A. 2006. Ferri-bacillibactin uptake and hydrolysis in *Bacillus subtilis*. Mol. Microbiol. 61:1413-1427.

Moreno Switt, A. I., den Bakker, H. C., Cummings, C. A., Rodriguez-Rivera, L. D., Govoni, G., Raneiri, M. L., Degoricija, L., Brown, S., Hoelzer, K., Peters, J. E., Bolchacova, E., Furtado, M. R., and Wiedmann, M. 2012. Identification and characterization of novel *Salmonella* mobile elements involved in the dissemination of genes linked to virulence and transmission. PLoS ONE 7:e41247.

Nicholson, W. L. 2008. The *Bacillus subtilis ydjL* (*bdhA*) gene encodes acetoin reductase/2,3-butanediol dehydrogenase. Appl. Environ. Microbiol. 74:6832-6838.

Ongena, M., Jourdan, E., Adam, A., Paquot, M., Brans, A., Joris, B., Arpigny, J.-L., and Thonart, P. 2007. Surfactin and fengycin lipopeptides of *Bacillus subtilis* as elicitors of induced systemic resistance in plants. Environ. Microbiol. 9:1084-1090.

Oslizlo, A., Stefanic, P., Vatovec, S., Beigot Glaser, S., Rupnik, M., and Mandic-Mulec, I. 2015. Exploring ComQXPA quorum-sensing diversity and biocontrol potential of *Bacillus* spp. isolates from tomato rhizoplane. Microb. Biotechnol. 8:527-540.

Pérez-García, A., Romero, D., and de Vicente, A. 2011. Plant protection and growth stimulation by microorganisms: Biotechnological applications of bacilli in agriculture. Curr. Opin. Biotechnol. 22:187-193.

Romero, D., Aguilar, C., Losick, R., and Kolter, R. 2010. Amyloid fibers provide structural integrity to *Bacillus subtilis* biofilms. Proc. Natl. Acad. Sci. U.S.A. 10:2230-2234.

Romero, D., de Vicente, A., Olmos, J. L., Dávila, J. C., and Pérez-García, A. 2007a. Effect of lipopeptides of antagonistic strains of *Bacillus subtilis* on the morphology and ultrastructure of the cucurbit fungal pathogen *Podosphaera fusca*. J. Appl. Microbiol. 103:969-976.

Romero, D., de Vicente, A., Rakotoaly, R. H., Dufour, S. E., Veening, J.-W., Arrebola, E., Cazorla, F. M., Kuipers, O. P., Paquot, M., and Pérez-García, A. 2007b. The iturin and fengycin families of lipopeptides are key factors in antagonism of *Bacillus subtilis* toward *Podosphaera fusca*. Mol. Plant-Microbe Interact. 20:430-440.

Romero, D., Pérez-García, A., Rivera, M. E., Cazorla, F. M., and de Vicente, A. 2004. Isolation and evaluation of antagonistic bacteria towards the cucurbit powdery mildew fungus *Podosphaera fusca*. Appl. Microbiol. Biotechnol. 64:263-269.

Stefanic, P., Decorosi, F., Viti, C., Petito, J., Cohan, F. M., and Mandic-Mulec, I. 2012. The quorum sensing diversity within and between ecotypes of *Bacillus subtilis*. Environ. Microbiol. 14:1378-1389.

Steinborn, G., Hajirezaei, M.-R., and Hofemeister, J. 2005. *bac* genes for recombinant bacilysin and anticapsin production in *Bacillus* host strains. Arch. Microbiol. 183:71-79.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28:2731-2739.

Tortosa, P., Logsdon, L., Kraigher, B., Itoh, Y., Mandic-Mulec, I., and Dubnau, D. 2001. Specificity and genetic polymorphism of the *Bacillus* competence quorum-sensing system. J. Bacteriol. 183:451-460.

Vernikos, G. S., and Parkhill, J. 2006. Interpolated variable order motifs for identification of horizontally acquired DNA: Revisiting the *Salmonella* pathogenicity islands. Bioinformatics 22:2196-2203.

Vlamakis, H., Chai, Y., Beauregard, P., Losick, R., and Kolter, R. 2013. Sticking together: Building a biofilm the *Bacillus subtilis* way. Nat. Rev. Microbiol. 11:157-168.

Xu, X. M., and Jeger, M. J. 2013. Combined use of two biocontrol agents with different biocontrol mechanisms most likely results in less than expected efficacy in controlling foliar pathogens under fluctuating conditions: a modeling study. Phytopathology 103:108-116.

Yang, Y., Wu, H.-J., Lin, L., Zhu, Q. Q., Borriss, R., and Gao, X.-W. 2015. A plasmid-born Rap-Phr system regulates surfactin production, sporulation and genetic competence in the heterologous host, *Bacillus subtilis* OKB105. Appl. Microbiol. Biotechnol. 99:7241-7252.

Yuan, J., Li, B., Zhang, N., Waseem, R., Shen, Q., and Huang, Q. 2012. Production of bacillomycin- and macrolactin-type antibiotics by *Bacillus amyloliquefaciens* NJN-6 for suppressing soilborne plant pathogens. J. Agric. Food Chem. 60:2976-2981.

Zeriouh, H., de Vicente, A., Pérez-García, A., and Romero, D. 2014. Surfactin triggers biofilm formation of *Bacillus subtilis* in melon phylloplane and contributes to the biocontrol activity. Environ. Microbiol. 16:2196-2211.

Zeriouh, H., Romero, D., García-Gutierrez, L., Cazorla, F. M., de Vicente, A., and Pérez-Garcia, A. 2011. The iturin-like lipopeptides are essential components in the biological control arsenal of *Bacillus subtilis* against bacterial diseases of cucurbits. Mol. Plant-Microbe Interact. 24:1540-1552.

## AUTHOR-RECOMMENDED INTERNET RESOURCES

National Center for Biotechnology Information (NCBI) complete bacterial genome repository: ftp://ftp.ncbi.nlm.nih.gov/genomes/Bacteria

NCBI Conserved Domain database: http://www.ncbi.nlm.nih.gov/cdd

NCBI Prokaryotic genome annotation pipeline: http://www.ncbi.nlm.nih.gov/genome/annotation_prok

NCBI Toolkit: http://www.ncbi.nlm.nih.gov/toolkit