


# Sporulation is dispensable for the vegetable-associated life cycle of the human pathogen *Bacillus cereus*

María Luisa Antequera-Gómez,<sup>1</sup> Luis Díaz-Martínez,<sup>1</sup> Juan Antonio Guadix,<sup>2,3</sup> Ana María Sánchez-Tévar,<sup>2,3</sup> Sara Sopeña-Torres,<sup>1</sup> Jesús Hierrezuelo,<sup>1</sup> Hung K. Doan,<sup>4</sup> Johan H.J. Leveau,<sup>4</sup> Antonio deVicente<sup>1</sup> and Diego Romero<sup>1</sup> 

<sup>1</sup>Departamento de Microbiología, Instituto de Hortofruticultura Subtropical y Mediterránea "La Mayora", Universidad de Málaga-Consejo Superior de Investigaciones Científicas (IHSM-UMA-CSIC), Universidad de Málaga, Bulevar Louis Pasteur 31 (Campus Universitario de teatinos), Málaga, 29071, Spain.

<sup>2</sup>Departamento de Biología Animal, Facultad de Ciencias, Universidad de Málaga - IBIMA, Campus de Teatinos s/n, Málaga, 29071, Spain.

<sup>3</sup>Centro Andaluz de Nanomedicina y Biotecnología (BIONAND), Junta de Andalucía, Universidad de Málaga, C/ Severo Ochoa 35, Campanillas (Málaga), 29590, Spain.

<sup>4</sup>Department of Plant Pathology, University of California, Davis, CA, USA.

## Summary

***Bacillus cereus* is a common food-borne pathogen that is responsible for important outbreaks of food poisoning in humans. Diseases caused by *B. cereus* usually exhibit two major symptoms, emetic or diarrheic, depending on the toxins produced. It is assumed that after the ingestion of contaminated vegetables or processed food, spores of enterotoxigenic *B. cereus* reach the intestine, where they germinate and produce the enterotoxins that are responsible for food poisoning. In our study, we observed that sporulation is required for the survival of *B. cereus* in leaves but is dispensable in ready-to-eat vegetables, such as endives. We demonstrate that vegetative cells of *B. cereus* that are originally impaired in sporulation but not biofilm formation are able to reach the intestine and cause severe disorders in a murine model. Furthermore, our findings**

emphasise that the number of food poisoning cases associated with *B. cereus* is underestimated and suggest the need to revise the detection protocols, which are based primarily on spores and toxins.

## Introduction

A major problem in food safety is the contamination of fresh and stored produce with human pathogens (Carlin *et al.*, 2000; Burnett and Beuchat, 2001), which may provoke food poisoning outbreaks (Berget *et al.*, 2010). *Bacillus cereus* is a natural inhabitant of soils and is able to live in association with plants. Members of this group of bacteria are also responsible for food poisoning outbreaks through consumption of raw vegetables and fruits or processed food (Bottone, 2010; Elhariry, 2011). Food poisoning caused by *B. cereus* is classified into two main categories, emetic and diarrheic, depending on the battery of toxins produced (Granum and Lund, 1997). An emetic poisoning is associated with the production of a dodecadepsipeptide called cereulide, (Agata *et al.*, 1994; Agata *et al.*, 1995b) a remarkably heat- and acid-stable toxin typically detected in high concentrations in processed food contaminated with *B. cereus* (Ehling-Schulz *et al.*, 2004). Diarrheic poisoning is known to occur after the ingestion of food contaminated with enterotoxigenic *B. cereus* spores which, upon germination in the intestine, produce heat-labile enterotoxins (McKillip, 2000; Berthold-Pluta *et al.*, 2015). A group of three primary enterotoxins are involved in diarrheic poisoning: haemolysin BL (Hbl), non-haemolytic enterotoxin (Nhe) and cytotoxin K (CytK) (Lund *et al.*, 2000; Stenfors Arnesen *et al.*, 2008).

*Bacillus cereus* spores are highly adhesive and resistant to stress, which ensures their persistence in a variety of environments, under various sanitation measures and in ready-to-eat or processed food (Gao *et al.*, 2015). Under more favourable environmental conditions, spores readily germinate to vegetative cells and begin new life cycles (Wiencek *et al.*, 1990). In addition to spores, contamination with *B. cereus* in the food industry is attributable to the formation of biofilms, in which bacterial communities assemble in response to a variety of signals that enable microorganisms to adapt to different environments (Danhorn and Fuqua, 2007). In the proposed life cycle of *B. cereus*, these microbes live as

Received 9 February, 2021; revised 18 March, 2021; accepted 3 April, 2021.

For correspondence. E-mail diego\_romero@uma.es; Tel./Fax +34951953057.

*Microbial Biotechnology* (2021) 14(4), 1550–1565  
doi:10.1111/1751-7915.13816

saprophytes in soil, move to plants and later enter human-related environments (hospitals, the food industry or human bodies). Spores or vegetative cells rotate depending on the life cycle stage, and they are proposed to remain as individual cells or become embedded in biofilms.

In this study, we investigated the relevance of spores and biofilms of *B. cereus* to the persistence of these species on infrequently investigated plants. The comparative analysis of a group of isolates (Table 1) involved in food poisoning enabled us to propose that sporulation is a major ecological trait in the phylloplane (i.e. on leaf surfaces) but that sporulation is dispensable in ready-to-eat vegetables. Additionally, we demonstrate that intrinsic genetic defects in sporulation do not diminish the pathogenicity of vegetative cells of *B. cereus* in a murine model.

## Results

### *Sporulation permits B. cereus to persist on melon leaves but is dispensable on endive*

We investigated the ability of a group of *B. cereus* strains isolated from different outbreaks of food poisoning (Table 1) to colonize and persist on leave of two different plant species, melon and endive. Two differentiated behaviours of *B. cereus* strains were observed on and in melon leaves: seven strains (AH187,

ATCC 14579, DSM 4282, DSM 4313, DSM 14729, UMAF8564 and 6A27) were capable of persisting at  $10^4$ – $10^6$  CFUs per gram of leaf and had a spore content of 20–40% (Fig. S1), and two strains (DSM 8438 and DSM 2302) which the total CFU and spores counts per gram of leaf were below the detection threshold (Fig. S1). In endive, however, all the strains persisted at  $10^4$ – $10^6$  CFU per gram of leaf. The percentages of sporulation were similar to those recorded on melon leaves with the exception of DSM 2302, which did not show any sign of sporulation. With these findings, we compared the population dynamics of strain AH187 (a representative of the group of persister and sporulating strains), DSM 8438 and DSM 2302 (non-persister on melon leaf and non-sporulating strains) on both plant leaves (Fig. 1). The population size of AH187 progressively decreased and stabilized at  $10^6$  CFU per gram of melon leaf 8 days after inoculation; however, DSM 8438 and DSM 2302 showed a sharp decrease within the first 4 days and became undetectable after 8 days of inoculation (Fig. 1). The AH187 population dynamics on endive leaves were similar to those on melon leaves, and DSM 8438 and DSM 2302 stabilized at  $10^4$ – $10^5$  CFU per gram of endive leaves after 2 DAI (Fig. 1). The sporulation rate of AH187 and DSM 8438 increased starting at 4–6 HAI, and no spores were detected for DSM 2302. Similar results for the persister group were recorded on cucumber leaves and fruit, and DSM 8438 and DSM 2302 did not sporulate in any of the studied cucumber organs (Fig. S1).

**Table 1.** *Bacillus cereus* strains used in this study.

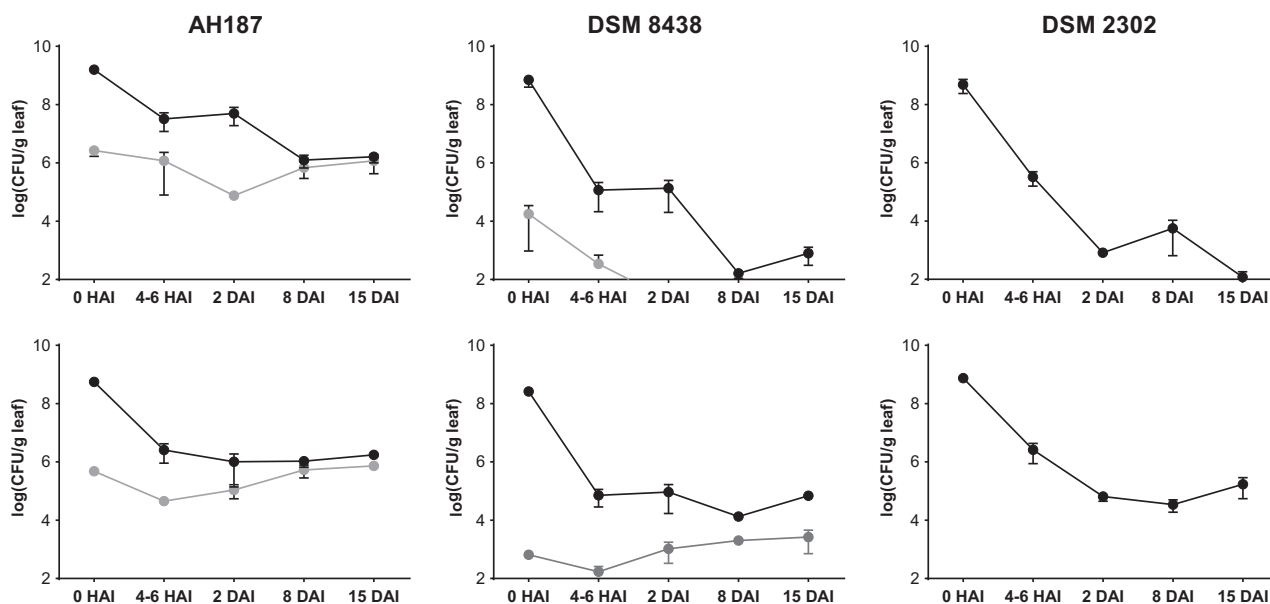
Strains	Origin	Toxins	Reference <sup>a</sup>
AH187	Emetic outbreak (Ehling-Schulz <i>et al.</i> , 2005)	Nhe/	
	Cereulide	BGSC	
ATCC 14579	Cowshed (Ceuppens <i>et al.</i> , 2013)	Hbl/Nhe/ CytK	CECT
DSM 4282	Diarrheal outbreak (Ehling-Schulz <i>et al.</i> , 2005)	Hbl/Nhe/ CytK	DSMZ
DSM 4313	Diarrheal outbreak (Ehling-Schulz <i>et al.</i> , 2005)	Hbl/Nhe/ CytK	DSMZ
DSM 14729	Unknown	Hbl/Nhe/ CytK	DSMZ
UMAF8564	Melon rhizosphere (García-Gutiérrez <i>et al.</i> , 2012)	Hbl/Nhe	UMA
6A27	Diarrheal outbreak (Ehling-Schulz <i>et al.</i> , 2005)	Hbl/Nhe	BGSC
DSM 8438	Diarrheal outbreak (Ehling-Schulz <i>et al.</i> , 2005)	Nhe/ CytK	DSMZ
DSM 2302	Foodstuff (Crovadore <i>et al.</i> , 2016)	Hbl/Nhe/ CytK	DSMZ

Origin indicates the source where the strains were isolated. Toxins are the main enterotoxin or emetic toxin produced by each strain. Reference corresponds to the culture collection where the strains were obtained.

a. CECT (Spanish Type Culture Collection). BGSC (Bacillus Genetic Stock Center). DSMZ (German Collection of Microorganisms and Cell Cultures). UMA (University of Malaga).

### *B. cereus DSM 2302 is genetically impaired in sporulation but capable of biofilm formation*

We sequenced and compared the genomes of all the strains used in this study to define the genetic defects for the inability of DSM 2302 to sporulate, and we searched for genetic traits defining a plant-associated lifestyle. Multilocus sequence analysis using 4 conserved genes (16S rRNA, *gyrB*, *rpoB* and *rpoD*) demonstrated that AH187, DSM 2302 and DSM 8438 belonged to completely separate phylogenetic branches (Fig. S2). This finding enabled us to surmise that the strains that do not share specific niches employ different evolutionary strategies and that their evolutionary distance might be related to their different lifestyles. A deeper study of the pangenome of the nine *B. cereus* strains (Fig. S3A) identified a total of 48 844 coding DNA sequences (CDSs) in the nine *B. cereus* strains. Of these total CDSs, 30 746 were highly conserved across all nine genomes and comprised the core genome of these *B. cereus* strains. The core genes were clustered into 3 326 orthologous groups (OGs), the additional genes (see material and methods for details) were clustered



**Fig. 1.** Sporulation is essential for the persistence of *B. cereus* on melon leaves but not on endive leaves. CFU counts per gram of melon and endive leaf after inoculation (0 h (HAI) 4–6 h (HAI), 2 days, 8 days and 15 days (DAI)) with *B. cereus* strains AH187, DSM 8438 or DSM 2302 strains. The top plots correspond to melon leaf counts, and the bottom plots correspond to endive counts. The black line shows the counts of total CFUs on either melon or endive leaves. The grey line shows the CFU counts corresponding to spores on melon and endive leaves. The detection limit of the technique is  $10^2$  CFU  $g^{-1}$  leaf. Error bars are the standard error of three or more replicates.

into 3 325 Ogs, and the unique genes of the nine strains comprised 3050; therefore, all CDSs were clustered into 9 702 Ogs. Notably, AH187 and DSM 2302 had the largest number of unique genes, exhibiting a total of 444 and 414 exclusive sequences, respectively, which supported our hypothesis on the specialization of these strains to different ecological niches (Table S3). Further analysis of the core genome by 10 different combinations of the sequences of the nine studied strains demonstrated that the number of genes did not decrease with the increase in the number of genomes added to the study (Fig. S3B). In addition, the pangenome of the studied strains was open, indicating that each sequenced genome contributed new genes to the pangenome but never reached an exact number of genes (Fig. S3B). A more specific comparison of DSM 2302 with AH187 showed an uneven distribution of the genes or the absence of several complete regions, including operons (Fig. S3C), as well as widely conserved regions in both strains.

Consistent with the apparent phenotypic defect in sporulation exhibited by DSM 2302, 36 genes involved in the sporulation programme were absent (Table 2, Fig. 2D) (Illing and Errington, 1991; Smith and Youngman, 1993; Tovar-Rojo *et al.*, 2002). To confirm this genotype, we analysed the intrinsic ability of the strains to sporulate in the spore-promoting medium DSM (Fig. 2B). All the strains that could persist over melon

leaves showed high sporulation rates in this medium (Fig. 2B, grey bars), but no sign of sporulation was observed in cultures of DSM 2302. Contrast-phase microscopy analysis of single cells confirmed the presence of refringent endospores within AH187 cells but neither endospores nor mother cell-released spores in cultures of DSM 2302 (Fig. 2B).

In *Bacillus*, sporulation and biofilms are known to contribute to the persistence of *Bacillus* and other species in different habitats, and these two cellular destinies are under the control of the master regulator *spo0A* (Fig. 2C). We disregarded a hypothetical collateral interruption of the biofilm regulatory pathway in DSM 2302, given that most of the lacking genes related to sporulation were structural and belonged to stages of the sporulation programme downstream of phosphorelay (Table 2). In addition, we identified the two major regulators of the expression of biofilm-related genes *sinR* and *sinI* along with the gene cluster *sipW*-to-*calY*, known to be essential for the assembly of the extracellular matrix (ECM) (Fig. S4B). Another indispensable component of the ECM is exopolysaccharides (EPSs). Two genomic regions containing EPS-related genes were found in DSM 2302: (i) one similar to the so-called *eps1* of the type strain ATCC 14579 and orthologous to the *epsA-O* of *B. subtilis*. This region was also found in strain AH187 which, notably, varied in the genes of the central core of the region (Fig. S4C). (ii) One region homologous to the

**Table 2.** Genes absent in DSM 2302 strain compared with AH187 strain.

Gene	Product
BCAH187_A0483	SpolIM; stage II sporulation protein M <sup>a,b</sup>
BCAH187_A0633	Sporulation sigma factor-processing peptidase <sup>b</sup>
BCAH187_A0925	Spore germination protein AC
BCAH187_A1089	Acid-soluble spore protein <sup>b</sup>
BCAH187_A1120	YjcZ family sporulation protein <sup>a,b</sup>
BCAH187_A1125	YjcZ family sporulation protein <sup>a,b</sup>
BCAH187_A1216	YjcZ family sporulation protein <sup>a,b</sup>
BCAH187_A1409	Spore-specific protein <sup>b</sup>
BCAH187_A2125	Spore permease <sup>b</sup>
BCAH187_A2217	AbrB family transcriptional regulator
BCAH187_A2515	Spo0E; stage 0 sporulation regulatory protein <sup>a,b</sup>
BCAH187_A2546	Exosporium protein H
BCAH187_A2674	Spore germination protein gerPA/gerPF <sup>b</sup>
BCAH187_A2711	YjcZ family sporulation protein <sup>a,b</sup>
BCAH187_A2919	CotX; spore coat protein X
BCAH187_A2920	CotW; spore coat protein W <sup>b</sup>
BCAH187_A2921	Spore coat protein X
BCAH187_A3146	CotF; spore coat protein F
BCAH187_A3147	Yqcl/YcgG family protein <sup>a,b</sup>
BCAH187_A3151	Small acid-soluble spore protein A (major alpha-type SASP)
BCAH187_A3165	Spore germination protein GerAC
BCAH187_A3166	Spore germination protein A2
BCAH187_A3167	Spore germination protein GerAA <sup>a,b</sup>
BCAH187_A3201	Putative spore germination B3 GerAC family protein
BCAH187_A3200	GerKB; spore germination protein KB <sup>a</sup> . Spore permease <sup>b</sup>
BCAH187_A3202	Spore germination protein KA
BCAH187_A3600	GerSB; Spore germination protein
BCAH187_A4093	YjcZ family sporulation protein <sup>a,b</sup>
BCAH187_A4323	SpoIIIAA; stage III sporulation protein AA
BCAH187_A4629	Germination protein gerE
BCAH187_A5258	Exosporium leader peptide-containing protein <sup>b</sup>
BCAH187_A5308	SpoVAD; stage V sporulation protein AD
BCAH187_A5309	SpoVAE; stage V sporulation protein AE
BCAH187_C0228	SpoVAD; stage V sporulation protein AD
BCAH187_C0229	SpoVAE; stage V sporulation protein AE
BCAH187_E0003	Sporulation protein <sup>a,b</sup>

Gene corresponds to AH187 locus absent in DSM 2302 strain. Product indicates the protein associated to each gene.

a. The annotation of the product corresponds to orthologous genes in other *Bacillus* species.

b. The annotation of the product corresponds to GO/KO term or Pfam domain function.

so-called *eps2* region of ATCC 14579, which was absent in strain AH187 (Fig. S4D). According to this genetic information, DSM 2302 was able to form biofilms in static cultures, although phenotypically differentiated from AH187. Colonies of DSM 2302 stained with Congo Red (known to be related to the product of the *eps2* region in the type strain ATCC 14579) and adhered to the walls and bottom of the wells but failed to form pellicles in the air-liquid interphase of static liquid cultures. Strain AH187 formed colonies that stained slightly with Congo Red, formed pellicles in the air-liquid interphase and adhered to the wells in static liquid cultures. (Fig. S4A). The pangenomic analysis of DSM 2302 additionally showed several genes involved in the assembly of the

ECM, polysaccharide and putrescine biosynthesis, all of which are putative structural integrant of the biofilm programme (Table S4).

Further genomic analysis demonstrated the presence of genes related to a plant-associated lifestyle in the genome of AH187 but absent in DSM 2302. These genes included several response regulators; ABC transporters hypothetically involved in exoprotein production, competence or sporulation; a regulator inducible by reactive oxygen species (ROS); a putative endoglucanase involved in plant cell wall degradation; diverse surface proteins and important enzymes related to plant nitrogen sources, such as asparagine, glutamine or proline (Table S5). In contrast, DSM 2302 had more genes related to human disorders, bacterial invasion of epithelial cells, bacterial extracellular infection, cell wall surface anchoring, cell adhesion and internalization, and antimicrobial resistance in the genome (Fig. 2A), most of which were observed in the related food-borne bacteria (Roux *et al.*, 2005; Franciosa *et al.*, 2009).

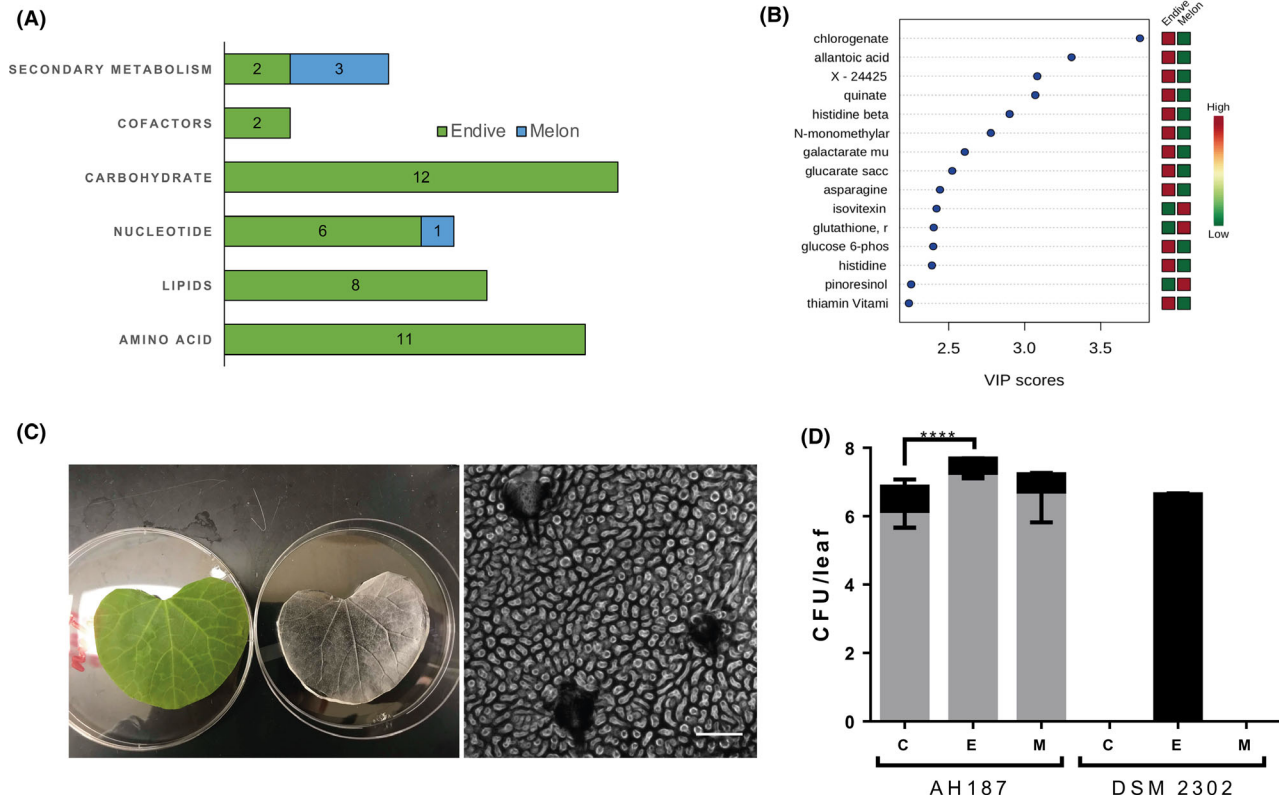
In summary, the defect of sporulation, the lack of genes related to the metabolism of nitrogen sources related to plants and the overrepresentation of genes associated with pathogenic interactions with humans are consistent with the low survival of DSM 2302 on plants.

#### *Metabolic nature of endive enables the persistence of vegetative cells of B cereus*

In addition to the lack of sporulation, the different fitness levels of DSM 2302 on melon or endives leaves raised the question of the major influence of the host. With in-depth metabolomic analysis of leaf tissue, the chemical profiles of both plant species demonstrated significant differences in metabolic families and metabolite concentrations (Fig. S5). Carbohydrates, amino acids, lipids and nucleotides were largely observed in the endive (Fig. 3A). More specifically, chlorogenate, allantoic acid, quinate, galactarate and glucarate were clearly the most representative endive metabolites. However, secondary metabolites of flavonoids (such as isovitexin), glutathione, pinoresinol and siderophores were the most frequently observed metabolites in melon leaves (Fig. 3B).

These findings along with the bacterial population dynamics showed above suggested that endive leaves were more conducive to bacterial growth, while melon leaves represented a more aggressive and stressful environment for bacteria. To confirm this hypothesis, we studied the bacterial behaviour in PDMS artificial surfaces that mirrored the topology of melon leaves (Fig. 3C) but were supplemented with a solution of quinate, galactarate and saccharate (5:1:1), the major components detected in endive. AH187 persisted on melon artificial leaves alone or supplemented with endive metabolites; however, DSM 2302 only persisted





**Fig. 3.** Differences in the nutrient content of the plant host impact the survival of non-sporulating *B. cereus* DSM 2302.

A. Major metabolites grouped by superfamilies found in melon and endive leaves.

B. Important metabolites identified by PLS-DA. The coloured boxes on the right indicate the relative concentration of the corresponding metabolite in each studied surface (red indicates a higher relative concentration, and green indicates a lower relative concentration of the corresponding metabolite).

C. The picture on the left shows a melon leaf and its corresponding replica of PDMS (artificial leaf surface). The image on the right is a micrograph of the surface of an artificial melon leaf. Scale bar corresponds to 100  $\mu$ m.

D. Total CFU (black bars) and spore counts (grey bars) per artificial leaf surface: without metabolites (C) with the majority endive metabolites (E) and the majority flavonoid (isovitexin) found in melon leaves (M) 15 days after inoculation with *B. cereus* AH187 and DSM 2302 strains. The black bar corresponds to the total CFUs per artificial leaf surface, and the grey bars are the counts of CFUs corresponding to spores. The detection limit of the technique is  $10^2$  CFU/g leaf. Error bars are the standard error of three replicates. Statistical analysis is the result of the comparison of the DSM 2302 strain against AH187. \*( $P < 0.05$ ), \*\*( $P < 0.01$ ), \*\*\*( $P < 0.001$ ), \*\*\*\*( $P < 0.0001$ ).

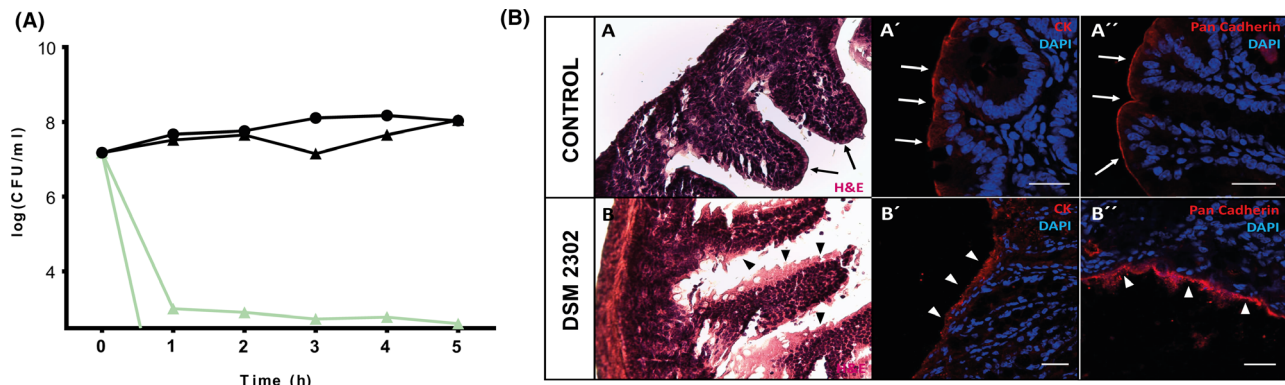
mice as indirect evidence of the presence of this strain in the guts. Haematoxylin–eosin staining showed continuity of the epithelium in the villi and crypts of the intestinal microvilli in untreated mice (Fig. 4B). However, the mice inoculated with DSM 2302 showed disruption of the epithelium. Similar results were obtained in immunohistochemical analyses using anti-Ck or anti-pan-cadherin antibodies and confocal microscopy (Fig. 4B).

These results enable us to conclude that vegetative cells of DSM 2302 (which lack of the acid resistance cereulide toxin) are capable of passing through the stomach and reaching the gut, subsequently causing illness in the host.

## Discussion

Various studies have previously reported the relevance of *B. cereus* spores in food poisoning caused by the

consumption of contaminated raw food or vegetables (Turnbull, 1976; Schoeni and Wong, 2005). It is assumed that both vegetative cells and spores can reach the intestine, but the survival percentage of spores is considerably higher due to their resistance to stomach conditions (Wijnands *et al.*, 2009; Ceuppens *et al.*, 2012a). Therefore, based on this presumption, the few vegetative cells that survive are insufficient to reach the minimum infective dose required to cause disease ( $10^5$ – $10^8$  CFU) (Stenfors Arnesen *et al.*, 2008). In keeping with this model, it is believed that diarrheic diseases provoked by enterotoxigenic *B. cereus* are associated with spore-contaminating foodstuffs, which germinate in the gut and produce the enterotoxins responsible for the symptoms (Clavel *et al.*, 2007; Ceuppens *et al.*, 2012a; Ceuppens *et al.*, 2012b). In our study, we demonstrate that sporulation is not a determinant ecological trait for



**Fig. 4.** Vegetative cells of DSM 2302 can survive simulating stomach conditions and cause symptoms in mouse intestines. A Counts of total CFU/ml on gastric media (GM) supplemented with JB broth of AH187 (▲) and DSM 2302 (●) strains at 37°C for 5 h. Black lines correspond to pH 4.5 simulating conditions after food intake. Green lines correspond to pH 2 (basal stomach conditions). The detection limit of the technique is  $10^1$  CFU  $ml^{-1}$ . Error bars are the standard error of three replicates. B Comparative histological caecum analysis of intestine sections from control intestine (PBS) and (B) epithelium disruption of intestines of infected mice. Immunohistochemistry and confocal microscopy showed epithelial continuity in the villi and crypts of the intestinal microvilli with CK (A'; arrows) and pan-cadherin (A'', arrows) markers of the control intestine and epithelial disruption in the infected caecums using CK (B'; arrowheads) or pan-cadherin (B'', arrowheads). All nuclei were stained with DAPI for immunofluorescence. Scale Bars (A', A'', B', B'') = 25  $\mu m$ .

the persistence of enterotoxigenic *B. cereus* on ready-to-eat vegetables.

Sporulation and biofilm formation are two ecological traits contributing to the completion of the life cycle of *B. cereus*. Our genomic analysis demonstrated that strain DSM 2302, which is naturally unable to sporulate, lacks sporulation genes belonging to mid-late stages of sporulation. However, the first stages related to the phosphorelay and thus sporulation decision-making, as well as other cellular destinies, such as biofilms, are preserved (Hamon and Lazazzera, 2001; Parashar *et al.*, 2013; Fagerlund *et al.*, 2014; Majed *et al.*, 2016). The inability to persist on/in melon leaves which was observed for both DSM 2302 or a *sigE* mutant of the emetic strain (AH187) that was affected in sporulation but not biofilm formation on melon leaves, indicated a major contribution of sporulation over biofilm formation in facilitating long-term bacterial survival in this ecological niche. Consistent with this, metabolomic analysis demonstrated that melon leaves are scarce in carbon and nitrogen sources and dominated by secondary metabolites, especially flavonoids, known as antibacterials and antibiofilms (Cushnie and Lamb, 2005; Awolola *et al.*, 2014; Matilla-Cuenca *et al.*, 2020). Contrary to this scenario, the endive is rich in nutrients and thus supportive of vegetative growth (Fig. 3B), most likely favouring biofilm formation as the preferred long-term bacterial survival strategy. We hypothesize that similar to the findings that have been reported for *B. subtilis* in the interaction with the rhizosphere, the collection of sensor kinases exposed on bacterial cell surfaces may discriminate the variety of signals and thus the final cellular decision-making (Chen

*et al.*, 2012). Therefore, it remains to be determined how different metabolites found in our analysis are sensed by these *Bacillus* cells and integrated in the sporulation or biofilm pathways.

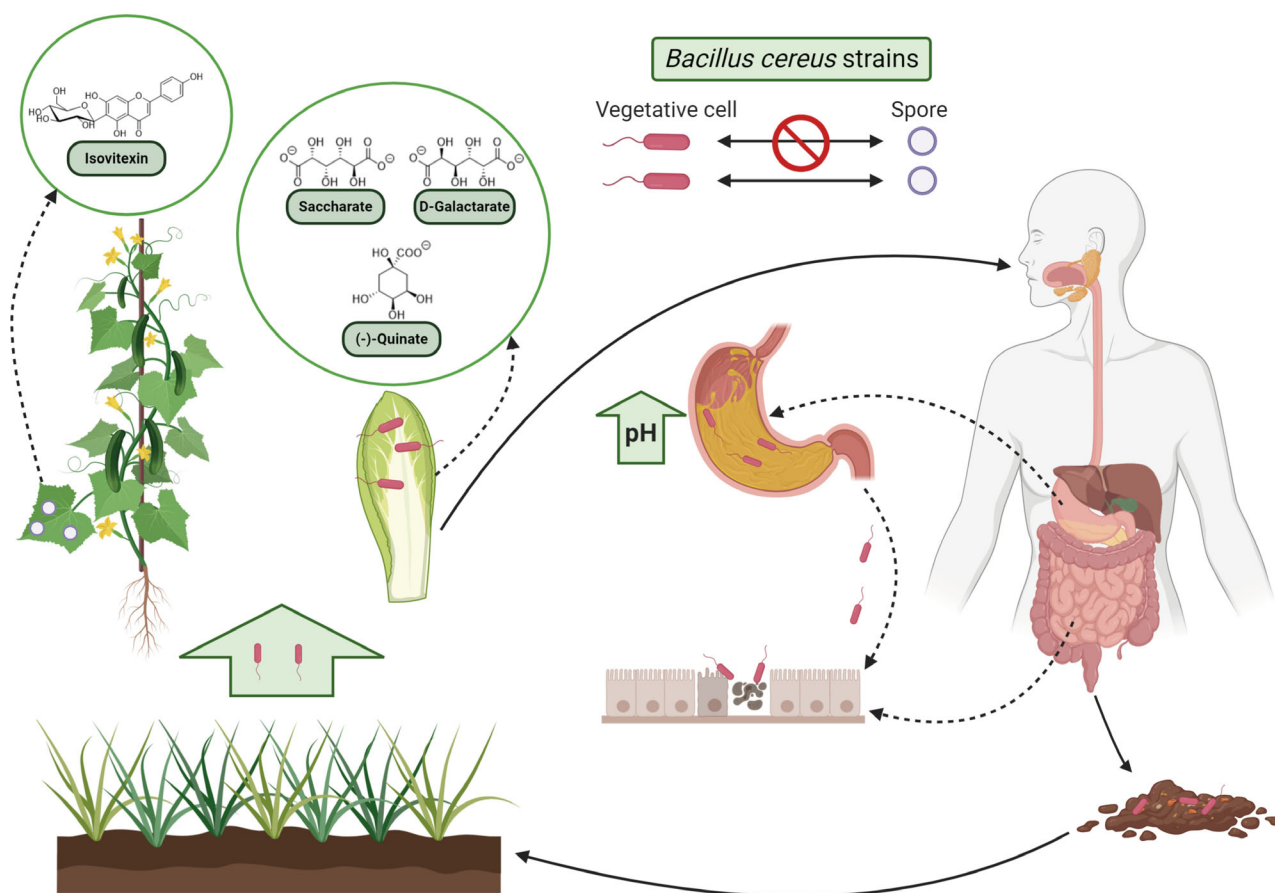
In addition to biofilm formation and sporulation, our genomic analysis showed a group of genes related to the use of plant carbon and nitrogen sources in the emetic strain AH187 (representative of *Bacillus* strain survival on plants), and other persisters strains, but not in DSM 2302. Even though the absence of these genes would be considered an adaptive disadvantage a priori, mutational analysis of these genes in strain AH187 did not indicate that they affect bacterial fitness on leaves (Fig. S6). In alignment with our findings, a previous study also found that genes related to the use of plant carbon sources do not impart an ecological advantage; in fact, the presence of amino acids in food, such as vegetables, could improve bacterial adaptability (Duport *et al.*, 2016). However, considering that only viable single mutants could be tested, we could not determine whether this genetic background may collectively contribute to the survival of bacterial cells on plants, either as single vegetative cells or spores or as communities within biofilms.

It is generally assumed that in *B. cereus* life cycle spores reach the stomach and germinate, giving rise to metabolically active vegetative cells capable of colonizing the gut and producing the toxin responsible for damage to the host (Clavel *et al.*, 2004; Stenfors Arnesen *et al.*, 2008). We have, however, demonstrated that vegetative cells of DSM 2302 are able to overcome acid conditions typically found in the stomach after food

ingestion (pH 4.5) (Fig. 4A) (Clavel *et al.*, 2004) for 5 h, which is sufficient time for the food to move forward from the stomach to the intestine (Camilleri *et al.*, 1989; Dressman *et al.*, 1990). At this stage, two possible but non-exclusive bacterial behaviours are the formation of biofilms associated with the gut and the production of enterotoxins. Genomic analysis of DSM 2302 showed the presence of well-known genes of the biofilm pathway (regulatory and structural elements) and other genes whose products are related to such traits as bacterial cell adhesion to the host tissue, ECM assembly, swarming motility (Sturgill and Rather, 2004; Patel *et al.*, 2006; Flemming and Wingender, 2010; Nwodo *et al.*, 2012) or bacterial invasion of epithelial cells, as previously reported for the human food poisoning pathogens *L. monocytogenes* and *E. coli* (Roux *et al.*, 2005; Franciosa *et al.*, 2009). In addition, DSM 2302 adheres to the bottom of the wells of static cultures, a low oxygen concentration environment that mimics the gut (Fig. S4A). DSM 2302 is genetically determined to produce the three

primary known enterotoxins (Hbl, Nhe and CytK) and the enterotoxins BceT, haemolysin I (cerolysin O), haemolysin II (HlyII), haemolysin III (HlyIII) and enterotoxin CwpFM (EntFM). BceT is lethal against mice, haemolysins are involved in lysing and permeabilizing host cells or intracellular organelles, and these toxins also have haemolytic activity against blood cells. EntFM is involved in bacterial virulence, motility, adhesion to epithelial cells and biofilm formation (Agata *et al.*, 1995a; Tran *et al.*, 2010; Ramarao and Sanchis, 2013). All these findings indicate that DSM 2302 has the necessary arsenal to adhere to epithelial cells of the gut and produce the toxins responsible for food poisoning in humans.

In summary, we identified a strain of *B. cereus* (DSM 2302) naturally impaired in sporulation but still able to survive in certain vegetables and caused illness in mammal hosts (Fig. 5). Based on our findings, we propose that the persistence of this strain on plants relies on the metabolic nature of the host and the ability of the bacteria to form biofilms. Vegetative cells bearing structural



**Fig. 5.** Reconsidering the life cycle of *B. cereus* from ready-to-eat vegetables to humans. Non-sporulating strains could survive over foodstuff products such as vegetables because of the large amount of nutrients compared to plant leaves, which are a more hostile niche. Furthermore, the ingestion of these contaminated foods increases stomach pH, and these conditions allow vegetative cells to survive stomach passage and enter the intestine to provoke human food poisoning. Created with BioRender.com.



elements related to biofilm formation (adhesins or EPSs) and protected by food can cross the stomach of humans and reach the intestine. The biofilm-related structural factors would further ensure either the adhesion of cells to the intestine or the formation of biofilms, as well as the production of enterotoxins that ensure the release of nutrients from the human host. From a clinical perspective, our findings reinforce the presumed underestimation of food poisoning cases caused by *B. cereus* and the need to improve the detection protocols that largely rely on the presence of spores or toxins in food (Soni *et al.*, 2016; Vidic *et al.*, 2020).

## Experimental procedures

### *Bacterial strains and growth conditions*

Nine *B. cereus* isolates from food poisoning outbreaks were used in this study (Table 1). Bacteria were routinely pre-cultured in LB broth (10 g l<sup>-1</sup> tryptone, 5 g l<sup>-1</sup> yeast extract and 5 g l<sup>-1</sup> NaCl) or in 1.5% LB agar. *B. cereus* strains were grown overnight with shaking at 28°C. To study biofilm formation, *B. cereus* strains were cultured for 72 h at 28°C without shaking in TY broth (10 g l<sup>-1</sup> tryptone, 5 g l<sup>-1</sup> yeast extract, 5 g l<sup>-1</sup> NaCl, 10 mM MgSO<sub>4</sub> and 1 mM MnSO<sub>4</sub>) or 1.5% TY agar supplemented with 20 and 10 µg ml<sup>-1</sup> Congo Red and Coomassie Brilliant Blue filtered dyes, respectively (Caro-Astorga *et al.*, 2014). For swarming motility assays, 0.7% TY agar was used (Caro-Astorga *et al.*, 2020). For sporulation experiments, cells were cultured in Difco Sporulation Medium (DSM) (Schaeffer *et al.*, 1965) for 24 h with shaking at 28°C. To simulate *in vitro* gastric passage, the cells were incubated in J broth (JB) (5 g l<sup>-1</sup> peptone, 15 g l<sup>-1</sup> yeast extract, 3 g l<sup>-1</sup> K<sub>2</sub>HPO<sub>4</sub> and 2 g l<sup>-1</sup> filtered glucose) (Clavel *et al.*, 2004).

### *Genome sequencing, assembly and annotation*

Bacterial DNA was isolated using the JetFlex™ Genomic DNA Purification Kit (Thermo Fisher Scientific, Bremen, Germany) following the kit protocol for Gram-positive bacteria. A DNA library was generated by NextSeq 550 (Illumina, United Kingdom) sequencing. The genome assembly was performed using the *B. cereus* type strain (ATCC 14579) as the reference strain of the group with the workflow of Pipeline A5-miseq (Coil *et al.*, 2015). The assembled genomes were annotated with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) version 4.3 algorithm (Angiuoli *et al.*, 2008) and deposited in NCBI GenBank (Table S1). Genomes of strains used in this work (ATCC 14579, AH187 and DSM 2302) were sequenced, and their genomes were obtained from the NCBI complete bacterial genome repository ([\[ncbi.nlm.nih.gov/genome/\]\(http://ncbi.nlm.nih.gov/genome/\)\). This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession \*PHKO00000000\*. The version described in this paper is version \*PHKO00000000\*.](http://www.</a></p>
</div>
<div data-bbox=)

### *Comparative genome, pangenome and phylogenetic analysis*

Genomic analyses for comparison of the strains listed in Table 1 were performed as follows: amino acid sequences for each strain were obtained from the GenBank files using the Perl programming language. Pfam domains (Finn *et al.*, 2016) were identified inside the proteins using a Perl script and an InterProScan algorithm in the InterPro database (Jones *et al.*, 2014). Protein–protein comparisons among the strains of the analysis were performed to obtain the bidirectional best hits (BBH), and orthologous and paralogous proteins were identified using the NCBI BLASTP algorithm. The protein sequences were clustered according to a percentage of identity > 80%, which retained the same Pfam domain with the OrthoMCL algorithm (Li *et al.*, 2003). The protein clusters were classified into three different groups: core (proteins shared among all the studied strains), accessory (proteins shared among all study strains except one or less) and unique (proteins only present in one of the studied strains and not in the others). To compare the function of *B. cereus* strains, GO terms (The Gene Ontology Consortium, 2015) and KO terms (KEGG Orthology) were obtained and analysed using InterProScan. The WEGO (Web Gene Ontology Annotation Plot) tool (<http://wego.genomics.org.cn/>) was used for GO term annotation, and the KAAS server (<https://www.genome.jp/kegg/kaas/>) (Kanehisa *et al.*, 2016a) and BlastKOALA (<https://www.kegg.jp/blastkoala/>) (Kanehisa *et al.*, 2016b) were used for KO term annotation. GO and KO term comparisons and visualization were performed using GraphPad Prism version 6.0 (GraphPad Software, San Diego, CA, USA) and InteractiVenn (Heberle *et al.*, 2015). Multilocus sequence phylogenetic analyses of the nine strains used for the pangenomic analysis, employing *Bacillus subtilis* subsp. *subtilis* 168 as an outgroup reference organism, was performed using the nucleotide sequences of the 16S rRNA, DNA gyrase B subunit *gyrB*, RNA polymerase sigma 70 factor *rpoD* and RNA polymerase beta 70 factor *rpoB* genes (Das *et al.*, 2014). The MUSCLE algorithm included in MEGA7 software was utilized to align the concatenated gene sequences. Next, a maximum likelihood tree was built with the MEGA7 program using the neighbour-joining substitution model, and bootstrap values of each branch were calculated 1000 times. To study the presence/absence of sporulation genes, a comparison between the *B. cereus* AH187 and *B. cereus* DSM 2302 genomes was performed with the Gview server tool

(<https://server.gview.ca>). The *B. cereus* AH187 gff file was obtained from NCBI (ASM2122v1), and the *B. cereus* DSM 2302 fasta file was generated in the pangenome study described elsewhere in this publication. Unique genes were used to identify the Gene Ontology functional categories using comparativeGo (<https://www.comparativego.com>). Sporulation genes from these two strains were selected and compared by blast2 sequencing to obtain the final list of sporulation genes of *B. cereus* AH187 with no matches in *B. cereus* DSM 2302. To determine the possible function of genes annotated as hypothetical proteins or with unknown function, their amino acid sequences were compared with proteins from other organism using the HHPred tool of the Max Planck Institute for Developmental Biology (Tübingen, Germany) (Söding, 2005; Zimmermann *et al.*, 2018). We also searched for orthologous genes in other *Bacillus* species, or we determined the hypothetical function according to the corresponding GO/KO term of the Pfam domain.

#### *Sporulation and image acquisition*

The sporulation assays were performed in DSM medium. One colony of each strain was picked from an LB agar plate and resuspended in DSM broth. Each strain was incubated at 28°C with shaking for 24 h. Bacterial concentrations in CFU ml<sup>-1</sup> were determined by plating 100 µl aliquots of 10-fold serial dilutions onto LB agar. To estimate the proportion of spores, each dilution was heated at 80°C for 10 min and subsequently plated. Microscopy images from DSM cultures were taken on a Nikon Eclipse Ti fluorescence microscope with a Super Plan Fluor objective 100x equipped with a Hamamatsu Orca R2 monochrome 1.3 MP CCD camera. Image processing (for proportion of spores) was performed using ImageJ software.

#### *Construction of a B. cereus sporulation mutants*

*B. cereus* mutants were obtained by electroporation using the pMAD plasmid (Arnaud *et al.*, 2004). To generate mutagenesis constructs, NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs, MA, USA) was utilized following the manufacturer's instructions, and specific primers (Table S2) were employed to amplify the upstream and downstream regions of the target gene. The pMAD vector was digested with SmaI and incubated at 50°C for 1 h in a thermocycler with the upstream and downstream fragments. The total amount of fragments was 0.2 pmol with a 1:2 DNA ratio (vector: insert). The constructions were utilized to transform electrocompetent *B. cereus* cells by electroporation as described previously with several modifications (Silo-Suh *et al.*, 1994; Caro-Astorga *et al.*, 2014). One microgram

of plasmid was incubated with 100 µl of electrocompetent cell suspension on ice for 5–10 min. The mixture of bacteria and DNA was electroporated in a 0.2-cm chilled cuvette using the following electroporation parameters: 1.4 kV voltage, 25 µF capacitance and 200 Ω resistance. The electroporated cells were recovered in LB at 30°C for 5 h and then plated onto LB plates supplemented with erythromycin to select bacterial transformants. To obtain mutants, *B. cereus* transformants were incubated at 40°C without antibiotics to block plasmid replication and force the shuttle vector to integrate into the chromosome via homologous recombination.

#### *Dynamics of bacterial cells on plants*

To evaluate the persistence of *B. cereus* strains on plants, melon (*Cucumis melo* cv. Rochet Panal) (Fitó, Barcelona, Spain) leaves, cucumber (*Cucumis sativus* cv. Marketer) (Fitó, Barcelona, Spain) leaves, cucumber fruit and endive (*Cichorium endivia*) were analysed. Melon and cucumber plants were grown under greenhouse conditions, and fresh cucumber fruit and endive were purchased in a grocery store. Cucumber fruit and endive were rinsed with 1% (w/v) sodium hypochlorite and 96% ethanol in distilled water for 4 min before being inoculated with *B. cereus* cell suspensions. A bacterial suspension (1 ml) at a concentration of 10<sup>8</sup> CFU ml<sup>-1</sup> in distilled water was spread onto the adaxial axis of second and third leaves of melon or cucumber leaves and around cucumber fruit and endive leaves. After bacterial inoculation, the melon and cucumber plants were incubated in a growth chamber at 25°C with a 16-h light/8-h dark photoperiod and 85% relative humidity. Endive and cucumber fruit were incubated at room temperature. Bacterial persistence and sporulation were evaluated 15 days after inoculation. Leaves and cucumber fruit skin were placed individually into sterile plastic stomacher bags with 10 ml of distilled water and homogenized for 3 min in a stomacher homogenizer (Interscience, France). One hundred microlitres of 10-fold serial dilutions of the homogenates were plated onto LB agar plates and incubated at 28°C for 24 h. To estimate the level of spores, serial dilutions were incubated at 80°C for 10 min to kill vegetative cells, and the spores were subsequently plated and incubated overnight at 28°C.

#### *Biofilm and swarming motility assays*

*In vitro* biofilm assays of *B. cereus* strains were performed in TY broth and in TY 1.5% agar supplemented with 20 and 10 µg ml<sup>-1</sup> of Congo Red and Coomassie Brilliant Blue filtered dyes, respectively (Romero *et al.*, 2010; Caro-Astorga *et al.*, 2014). For biofilm assays in TY broth, a colony from a pre-culture in LB agar at 28°C

for each strain was previously resuspended in 1 ml of TY medium; 10 µl of this culture was inoculated into 1 ml of TY broth in the centre wells of a 24-well plate, and the plate was incubated without shaking at 28°C for 72 h. To evaluate bacterial adhesion to abiotic surfaces (i.e. adhesion to the well wall), the biofilm of each well was stained with 1% crystal violet solution (O'Toole *et al.*, 2000) for 3 min and subsequently rinsed several times with distilled water until the excess dye was removed. Finally, the plate was dried on filter paper on a bench. To study biofilm formation on agar plates (colony morphology and staining), 2 µl of the pre-culture described above was inoculated separately on 1.5% TY agar plates with Congo Red and Coomassie Brilliant Blue, and the plates were incubated at 28°C for 72 h. To study swarming motility, 0.7% TY agar plates were utilized. Two microlitres of the pre-culture described above were inoculated on TY plates and incubated at 28°C for 72 h before measuring the diameter of colonies.

#### *Leaf and fruit metabolome analysis*

Metabolomic analysis of melon and endive leaves was performed by Metabolon® (Durham, NC, USA). One hundred milligrams of each tissue (from a pool of leaves) was homogenized and lyophilized. Samples were prepared using the automated MicroLab STAR® system from Hamilton Company (Franklin, MA, USA). Samples were extracted with methanol under vigorous shaking for 2 min to precipitate protein and dissociate small molecules bound to protein or trapped in the precipitated protein matrix followed by centrifugation to recover chemically diverse metabolites. The resulting extract was divided into four fractions: two for analysis by two separate reversed-phase (RP)/UPLC-MS/MS methods using positive ion mode electrospray ionization (ESI), one for analysis by RP/UPLC-MS/MS using negative ion mode ESI, and one for analysis by HILIC/UPLC-MS/MS using negative ion mode ESI. Samples were placed briefly on a TurboVap® (Zymark) to remove the organic solvent. The metabolic analysis was performed by ultra-performance liquid chromatography (UPLC) coupled to a mass spectrometer with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyser operated at 35,000 mass resolution. The four extracts were dried and subsequently reconstituted in methanol. One aliquot was analysed using acidic positive ion conditions, chromatographically optimized for more hydrophilic compounds and subsequently gradient-eluted from a C18 column (Waters UPLC BEH C18-2.1 × 100 mm, 1.7 µm) using water and methanol containing 0.05% perfluoropentanoic acid (PFPA) and 0.1% formic acid (FA). A second aliquot was also analysed using acidic positive ion conditions but was chromatographically optimized for more hydrophobic

compounds. The extract was gradient-eluted from the aforementioned C18 column using methanol, acetonitrile, water, 0.05% PFPA and 0.01% FA and was operated at an organic content that was higher overall. A third aliquot was analysed using basic negative ion optimized conditions using a separate dedicated C18 column. The basic extracts were gradient-eluted from the column using methanol and water but with 6.5 mM ammonium bicarbonate at pH 8. The fourth aliquot was analysed via negative ionization following elution from a HILIC column (Waters UPLC BEH Amide 2.1 × 150 mm, 1.7 µm) using a gradient consisting of water and acetonitrile with 10 mM ammonium formate, pH 10.8. The MS analysis alternated between MS and data-dependent MS<sup>n</sup> scans using dynamic exclusion. The scan range varied slightly between methods but covered approximately 70–1000 m/z. Raw data were extracted, peaks were identified, and data were submitted to QC processing using Metabolon's hardware and software.

To visualize the differences between melon and endive leaves, MetaboAnalyst (<https://www.metaboanalyst.ca/>) was used to generate the heatmap with hierarchical clustering and to determine the most important metabolites through PLS-DA analysis.

#### *Construction of artificial leaf surfaces, bacterial inoculation and incubation*

To produce artificial leaf surfaces, melon leaves were grown as described above. The second and third leaves were taken to mould the leaf surface replica. The process of making artificial leaves was undertaken as previously described (Doan *et al.*, 2020a) and summarized as follows: first, a negative stamp of the adaxial part of the leaf was created. The negative stamp was made with a mixture of 10:1 polydimethylsiloxane (PDMS) and curing agent from a Sylgard 184 PDMS elastomer kit (Dow Chemical, Midland, TX, USA). The mixture was poured over the real leaf previously joined with double-sided adhesive tape on a Petri dish and incubated at 30°C for 17–24 h. When the negative stamp was solidified, it was carefully separated from the real leaf and subsequently exposed to UV light (400 F, 400 W and  $\lambda = 187$ –254 nm) for 1 h to fix the PDMS. After UV exposure, the negative stamp was treated with a mixture of 1000:1 toluene and octadecyltrichlorosilane (ODTS) for 5 min to create a soft cover that facilitated the later separation of the negative stamp from the positive stamp. Next, the negative stamp was rinsed with absolute ethanol for 2 min and dried for at least 10 h at room temperature. To make the positive stamp (artificial leaf), the PDMS mixture was made as described above, and it was poured onto the negative stamp, which was attached to a Petri dish. When the PDMS mixture was solidified, the

positive stamp was carefully separated from the negative stamp. The fabricated artificial leaves were autoclaved and rinsed with a mixture of 1:1000 Triton X-100 in distilled water. Artificial leaves were cut into coupons measuring 25.75 cm in diameter and inoculated with 20- $\mu$ l drops (5 drops per coupon) of a  $10^8$  CFU  $\text{ml}^{-1}$  bacterial suspension in distilled water (Elbing and Brent, 2002; Doan *et al.*, 2020b) or in distilled water supplemented with 0.5% quinate, 0.1% galactarate and 0.1% saccharate or with 1 mM isovitexin to mimic endive and melon leaf metabolite profiles, respectively (obtained from the metabolomic analysis). Inoculated coupons were incubated over moistened filter paper with 10 ml of distilled water into a Petri dish plate or without humidity conditions (without distilled water), allowing the drops to dry on an open Petri dish. All the plates with coupons were incubated for 15 days in a growth chamber as described above for real leaves. After 15 days, the coupons were vortexed for 15 s and sonicated for 10 min in an ultrasonic bath in 10 ml of distilled water in a 50-ml tube (Doan *et al.*, 2020b). Tenfold serial dilutions were made, and 100  $\mu$ l of each dilution was plated on LB agar. To estimate the level of spores, the dilutions were heated at 80°C for 10 min and plated and incubated at 28°C for 24 h to obtain CFU counts.

#### *Gastric passage simulation*

To analyse gastric passage simulation after food ingestion, we employed the protocol described previously (Clavel *et al.*, 2004; Clavel *et al.*, 2007): *B. cereus* strains were incubated in LB agar at 30°C for 24 h. One colony was transferred into a 250-ml flask with 40 ml of J broth (JB) and incubated for 18 h at 30°C and 130 rpm. *B. cereus* cells were added to a 1:1 mixture of JB and gastric media (GM) (4.8 g  $\text{l}^{-1}$  NaCl, 1.56 g  $\text{l}^{-1}$   $\text{NaHCO}_3$ , 2.2 g  $\text{l}^{-1}$  KCl and 0.22 g  $\text{l}^{-1}$   $\text{CaCl}_2$ ) at pH 2 and 4.5 plus 500 U  $\text{l}^{-1}$  pepsin to an initial concentration of  $10^6$ – $10^7$  CFU  $\text{ml}^{-1}$  bacteria and incubated at 37°C for 5 h and 130 rpm to simulate human stomach conditions. Samples were taken hourly, 10-fold serial dilutions were made in 0.2 M sodium phosphate buffer at pH 7 to neutralize the stomach pH, and 100  $\mu$ l of each dilution were plated onto LB agar. For the determination of spores, the dilutions were heated at 80°C for 10 min and plated onto LB agar.

#### *Inoculation of mice with B. cereus and histological analysis of intestines*

CD1 female mice (3 months old) were purchased from the Andalusian Centre for Nanomedicine and Biotechnology (University of Malaga). Each experimental group consisted of six mice housed in cages that were

maintained at a controlled temperature and light cycle; the mice were fed an autoclaved standard balanced conventional diet and a gelatine-only diet 24 h before the infection. Animals were handled in accordance with international regulations for animal welfare. Bacteria were administered by intragastric gavage using a blunt-end needle. Six randomly allocated mice were challenged with 100  $\mu$ l of  $10^8$  CFU of *B. cereus* DSM 2302 in 1x PBS or with 100  $\mu$ l of 1x PBS (Rolny *et al.*, 2014). After 6 h of bacterial administration, mice were anaesthetized with an injection of 200  $\mu$ l of a mix of 10% ketamine, 6% xylazine and 84% water each and sacrificed by cervical dislocation. Samples from the caecum region were aseptically collected and washed in PBS and immediately fixed in 4% formalin overnight (4°C) for histological preparation. To obtain frozen sections, tissues were cryoprotected in 15% and 30% sucrose:PBS solutions, embedded in OCT, snap-frozen in liquid  $\text{N}_2$ -cooled isopentane and sectioned in a cryostat. OCT blocks were cut into 10-mm sections, and some of them were stained with haematoxylin–eosin. At least three caeca were analysed for each treatment.

#### *Immunohistochemistry and confocal microscopy*

Immunofluorescence analyses were performed by blocking non-specific binding sites with 16% sheep serum, 1% bovine serum albumin and 0.5% Triton X-100 in PBS (SBT) and incubating either tissue section slide with primary antibodies (anti-cytokeratin 'DAKO Z0622, 1:100 dilution' and anti-PanCadherin 'Sigma C3678, 1:100 dilution') overnight at 4°C. Next, the samples were washed in PBS (3  $\times$  5 min) and incubated with the secondary antibody (Alexa Fluor® 647 'ImmunoJackson 711-605-152', 1:200 dilution) + DAPI (Sigma D9542, 1:2000 dilution) for 2 h at room temperature. All samples were finally washed in PBS (3  $\times$  5 min), mounted in a 1:1 PBS/glycerol solution and analysed using an SP5 laser confocal microscope (LEICA, Wetzlar, Germany).

#### *Ethics approval and informed consent for mouse experiments*

Experiments were carried out in accordance with the Spanish Institutional Animal Use and Care Committee regulations and with the approval of the Regional Andalusian Government (code A/ES/14/43). The methods were carried out in accordance with the relevant guidelines and regulations.

#### *Statistical analysis*

The statistical analyses were performed using the statistical software GraphPad Prism version 6.0 (GraphPad

Software, CA, USA). For melon and endive leaf experiments, one-way non-parametric ANOVA (Kruskal–Wallis test) was performed with Dunn's multiple comparison tests. For artificial surface experiments, two-way ANOVA (with the Bonferroni test) was applied. The results with *P*-values < 0.05 were considered to be significant. Metabolome data were normalized for internal consistency by processing a constant weight of sample per volume of extraction solvent. No *post hoc* mathematical normalization was imposed on the data. Data were scaled to the median value for each compound, and missing values were imputed with the minimum detected value for that compound. Statistical calculations were performed using natural log-transformed scaled imputed data (table not included).

### Acknowledgements

We would like to thank Josefa Gómez from the Ultrasequencing Unit of the SCBI-UMA for DNA sequencing, Zulema Udaondo from Bio-Ilberis and Rocío Bautista from Bioinformatic Unit of the SCBI-UMA for pangenome and genome analysis support, John Pearson from Bionand for guidance and assistance with microscopy images, María Isabel Castillo from Bionand-UMA for managing and feeding mice and Dr. Atul Parikh from UC Davis for access to his laboratory for the fabrication of PDMS leaf replicants used in this study. This work was supported by grants from European Research Council Starting Grant (BacBio 637971), and Plan Nacional de I + D+i of Ministerio de Economía y Competitividad, Ministerio de Ciencia e Innovación (AGL2016-78662-R and PID2019-107724GB-I00). Johan HJ Leveau was financed by USDA-NIFA grant number 2014-67017-21695. Juan Antonio Guadix acknowledges financial support by grant (PIER-0084-2019) from Proyectos de Investigación en Salud de Junta de Andalucía.

### Funding Information

This work was supported by grants from European Research Council Starting Grant (BacBio 637971), and Plan Nacional de I + D+i of Ministerio de Economía y Competitividad, Ministerio de Ciencia e Innovación (AGL2016-78662-R and PID2019-107724GB-I00). Johan HJ Leveau was financed by USDA-NIFA grant number 2014-67017-21695. Juan Antonio Guadix acknowledges financial support by grant (PIER-0084-2019) from Proyectos de Investigación en Salud de Junta de Andalucía.

### Conflict of interests

The authors declare no competing interests.

### Author contributions

D.R. conceived the study. D.R. and M.L.A.G. designed the experiments. M.L.A.G. performed the main experimental work. M.L.A.G. and L.D.M. analysed and processed the whole genomes. M.L.A.G. and J.H. designed the metabolite assays over artificial leaf surfaces. M.L.A.G., H.K.D. and J.H.J.L. made and design the artificial leaf surfaces protocol. S.S.T. processed the real surfaces for metabolome analysis. A.M.S.T. and J.A.G. performed the mouse assays and analysis. D.R. and M.L.A.G. wrote the manuscript. D.R., M.L.A.G., L.D.M., J.H., J.A.G., J.H.J.L. and A.V. contributed critically to writing the final version of the manuscript.

### References

- Agata, N., Mori, M., Ohta, M., Suwan, S., Ohtani, I., and Isobe, M. (1994) A novel dodecadepsipeptide, cereulide, isolated from *Bacillus cereus* causes vacuole formation in hEp-2 cells. *FEMS Microbiol Lett* **121**: 31–34.
- Agata, N., Ohta, M., Arakawa, Y., and Mori, M. (1995a) The *bceT* gene of *Bacillus cereus* encodes an enterotoxic protein. *Microbiology* **141**: 983–988.
- Agata, N., Ohta, M., Mori, M., and Isobe, M. (1995b) A novel dodecadepsipeptide, cereulide, is an emetic toxin of *Bacillus cereus*. *FEMS Microbiol Lett* **129**: 17–19.
- Angiuoli, S.V., Gussman, A., Klimke, W., Cochrane, G., Field, D., Garrity, G.M., *et al.* (2008) Toward an online repository of standard operating procedures (SOPs) for (meta)genomic annotation. *Omi A J Integr Biol* **12**: 137–141.
- Arnaud, M., Chastanet, A., and Débarbouillé, M. (2004) New vector for efficient allelic replacement in naturally gram-positive bacteria. *Appl Environ Microbiol* **70**: 6887–6891.
- Awolola, G.V., Koorbanally, N.A., Chenia, H., Shode, F.O., and Baijnath, H. (2014) Antibacterial and anti-biofilm activity of flavonoids and triterpenes isolated from the extracts of *Ficus sansibarica* Warb. subsp. *sansibarica* (Moraceae) extracts. *Afr J Tradit Complement Altern Med* **11**: 124–131.
- Berger, C.N., Sodha, S.V., Shaw, R.K., Griffin, P.M., Pink, D., Hand, P., and Frankel, G. (2010) Fresh fruit and vegetables as vehicles for the transmission of human pathogens. *Environ Microbiol* **12**: 2385–2397.
- Berthold-Pluta, A., Pluta, A., and Garbowska, M. (2015) The effect of selected factors on the survival of *Bacillus cereus* in the human gastrointestinal tract. *Microb Pathog* **82**: 7–14.
- Bottone, E.J. (2010) *Bacillus cereus*, a volatile human pathogen. *Clin Microbiol Rev* **23**: 382–398.
- Burnett, S.L., and Beuchat, L.R. (2001) Human pathogens associated with raw produce and unpasteurized juices, and difficulties in decontamination. *J Ind Microbiol Biotechnol* **27**: 104–110.
- Camilleri, M., Colemont, L.j., Phillips, S.f., Brown, M.I., Thomforde, G.m., Chapman, N., and Zinsmeister, A.r. (1989) Human gastric emptying and colonic filling of

- solids characterized by a new method. *Am J Physiol. – Gastrointest Liver Physiol* **257**: G284–G290.
- Carlin, F., Girardin, H., Peck, M.W., Stringer, S.C., Barker, G.C., Martinez, A., *et al.* (2000) Research on factors allowing a risk assessment of spore-forming pathogenic bacteria in cooked chilled foods containing vegetables: a FAIR collaborative project. *Int J Food Microbiol* **60**: 117–135.
- Caro-Astorga, J., Álvarez-Mena, A., Hierrezuelo, J., Guadix, J.A., Heredia-Ponce, Z., Arboleda-Estudillo, Y., *et al.* (2020) Two genomic regions encoding exopolysaccharide production systems have complementary functions in *B. cereus* multicellularity and host interaction. *Sci Rep* **10**: 1000.
- Caro-Astorga, J., Pérez-García, A., de Vicente, A., and Romero, D. (2014) A genomic region involved in the formation of adhesion fibers in *Bacillus cereus* biofilms. *Front Microbiol* **5**: 745.
- Ceuppens, S., Timmery, S., Mahillon, J., Uyttendaele, M., and Boon, N. (2013) Small *Bacillus cereus* ATCC 14579 subpopulations are responsible for cytotoxin K production. *J Appl Microbiol* **114**: 899–906.
- Ceuppens, S., Uyttendaele, M., Hamelink, S., Boon, N., and Van De Wiele, T. (2012a) Inactivation of *Bacillus cereus* vegetative cells by gastric acid and bile during in vitro gastrointestinal transit. *Gut Pathog* **4**: 11.
- Ceuppens, S., Uyttendaele, M., Drieskens, K., Heyndrickx, M., Rajkovic, A., Boon, N., *et al.* (2012b) Survival and germination of *Bacillus cereus* spores without outgrowth or enterotoxin production during in vitro simulation of gastrointestinal transit. *Appl Environ Microbiol* **78**: 7698–7705.
- Chen, Y., Cao, S., Chai, Y., Clardy, J., Kolter, R., Guo, J.-H., and Losick, R. (2012) A *Bacillus subtilis* sensor kinase involved in triggering biofilm formation on the roots of tomato plants. *Mol Microbiol* **85**: 418–430.
- Clavel, T., Carlin, F., Lairon, D., Nguyen-The, C., and Schmitt, P. (2004) Survival of *Bacillus cereus* spores and vegetative cells in acid media simulating human stomach. *J Appl Microbiol* **97**: 214–219.
- Clavel, T., Carlin, F., Dargaignaratz, C., Lairon, D., Nguyen-The, C., and Schmitt, P. (2007) Effects of porcine bile on survival of *Bacillus cereus* vegetative cells and Haemolysin BL enterotoxin production in reconstituted human small intestine media. *J Appl Microbiol* **103**: 1568–1575.
- Coil, D., Jospin, G., and Darling, A.E. (2015) A5-misseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. *Bioinformatics* **31**: 587–589.
- Crovadore, J., Calmin, G., Tonacini, J., Chablais, R., Schnyder, B., Messelhäuser, U., *et al.* (2016) Whole-genome sequences of seven strains of *Bacillus cereus* isolated from foodstuff or poisoning incidents. *Genome Announc* **4**: 4–5.
- Cushnie, T.P.T., and Lamb, A.J. (2005) Antimicrobial activity of flavonoids. *Int J Antimicrob Agents* **26**: 343–356.
- Danhorn, T., and Fuqua, C. (2007) Biofilm formation by plant-associated bacteria. *Annu Rev Microbiol* **61**: 401–422.
- Das, S., Dash, H.R., Mangwani, N., Chakraborty, J., and Kumari, S. (2014) Understanding molecular identification and polyphasic taxonomic approaches for genetic relatedness and phylogenetic relationships of microorganisms. *J Microbiol Methods* **103**: 80–100.
- Doan, H.K., Antequera-Gómez, M.L., Parikh, A.N., and Leveau, J.H.J. (2020a) Leaf surface topography contributes to the ability of *Escherichia coli* on leafy greens to resist removal by washing, escape disinfection with chlorine, and disperse through splash. *Front Microbiol* **11**: 1485.
- Doan, H.K., Ngassam, V.N., Gilmore, S.F., Tecon, R., Parikh, A.N., and Leveau, J.H.J. (2020b) Topography-driven shape, spread, and retention of leaf surface water impacts microbial dispersion and activity in the phyllosphere. *Phytobiomes J* **4**: 268–280.
- Dressman, J.B., Berardi, R.R., Dermentzoglou, L.C., Russell, T.L., Schmaltz, S.P., Barnett, J.L., *et al.* (1990) Upper gastrointestinal (GI) pH in young, healthy men and women. *Pharm Res An Off J Am Assoc Pharm Sci* **7**: 756–761.
- Duport, C., Jobin, M., and Schmitt, P. (2016) Adaptation in *Bacillus cereus*: from stress to disease. *Front Microbiol* **7**.
- Ehling-Schulz, M., Fricker, M., and Scherer, S. (2004) *Bacillus cereus*, the causative agent of an emetic type of food-borne illness. *Mol Nutr Food Res* **48**: 479–487.
- Ehling-Schulz, M., Svensson, B., Guinebretiere, M.-H., Lindbäck, T., Andersson, M., Schulz, A., *et al.* (2005) Emetic toxin formation of *Bacillus cereus* is restricted to a single evolutionary lineage of closely related strains. *Microbiology* **151**: 183–197.
- Elbing, K., and Brent, R. (2002) Media preparation and bacteriological tools. *Curr Protoc Mol Biol Chapter* **1**: 1.1.1–1.1.7.
- Elhairy, H.M. (2011) Attachment strength and biofilm forming ability of *Bacillus cereus* on green-leafy vegetables: cabbage and lettuce. *Food Microbiol* **28**: 1266–1274.
- Fagerlund, A., Dubois, T., Økstad, O. A., Verplaetse, E., Gilois, N., Bennaceur, I., *et al.* (2014) SinR controls enterotoxin expression in *Bacillus thuringiensis* biofilms. *PLoS One* **9**: e87532.
- Finn, R.D., Coghill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., *et al.* (2016) The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res* **44**: D279–D285.
- Flemming, H.C., and Wingender, J. (2010) The biofilm matrix. *Nat Rev Microbiol* **8**: 623–633.
- Franciosa, G., Maugliani, A., Scalfaro, C., Floridi, F., and Aureli, P. (2009) Expression of internalin a and biofilm formation among *Listeria monocytogenes* clinical isolates. *Int J Immunopathol Pharmacol* **22**: 183–193.
- Gao, T., Foulston, L., Chai, Y., Wang, Q., and Losick, R. (2015) Alternative modes of biofilm formation by plant-associated *Bacillus cereus*. *Microbiologyopen* **4**: 452–464.
- García-Gutiérrez, L., Romero, D., Zerriouh, H., Cazorla, F.M., Torés, J.A., 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.
- Granum, P.E., and Lund, T. (1997) *Bacillus cereus* and its food poisoning toxins. *FEMS Microbiol Lett* **157**: 223–228.
- Hamon, M.A., and Lazazzera, B.A. (2001) The sporulation transcription factor Spo0A is required for biofilm development in *Bacillus subtilis*. *Mol Microbiol* **42**: 1199–1209.

- Heberle, H., Meirelles, V.G., da Silva, F.R., Telles, G.P., and Minghim, R. (2015) InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinform* **16**: 169.
- Illing, N., and Errington, J. (1991) The *spoIIIA* operon of *Bacillus subtilis* defines a new temporal class of mother-cell-specific sporulation genes under the control of the  $\sigma^E$  form of RNA polymerase. *Mol Microbiol* **5**: 1927–1940.
- Jones, P., Binns, D., Chang, H.-y., Fraser, M., Li, W., McAnulla, C., *et al.* (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**: 1236–1240.
- Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016a) KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* **44**: D457–D462.
- Kanehisa, M., Sato, Y., and Morishima, K. (2016b) BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* **428**: 726–731.
- Li, L., Stoekert, C.J., and Roos, D.S. (2003) OrthoMCL: IDENTIFICATION of ortholog groups for eukaryotic genomes. *Genome Res* **13**: 2178–2189.
- Lund, T., De Buyser, M.L., and Granum, P.E. (2000) A new cytotoxin from *Bacillus cereus* that may cause necrotic enteritis. *Mol Microbiol* **38**: 254–261.
- Majed, R., Faille, C., Kallassy, M., and Gohar, M. (2016) *Bacillus cereus* biofilms—same, only different. *Front Microbiol* **7**: 1054.
- Matilla-Cuenca, L., Gil, C., Cuesta, S., Rapún-Araiz, B., Žiemytė, M., Mira, A., *et al.* (2020) Antibiofilm activity of flavonoids on staphylococcal biofilms through targeting BAP amyloids. *Sci Rep* **10**: 1–12.
- McKillip, J.L. (2000) Prevalence and expression of enterotoxins in *Bacillus cereus* and other *Bacillus* spp., a literature review. *Antonie Van Leeuwenhoek* **77**: 393–399.
- Nwodo, U.U., Green, E., and Okoh, A.I. (2012) Bacterial exopolysaccharides: functionality and prospects. *Int J Mol Sci* **13**: 14002–14015.
- O'Toole, G., Kaplan, H.B., and Kolter, R. (2000) Biofilm formation as microbial development. *Annu Rev Microbiol* **54**: 49–79.
- Parashar, V., Konkol, M.A., Kearns, D.B., and Neiditch, M.B. (2013) A plasmid-encoded phosphatase regulates *Bacillus subtilis* biofilm architecture, sporulation, and genetic competence. *J Bacteriol* **195**: 2437–2448.
- Patel, C.N., Wortham, B.W., Lines, J.L., Fetherston, J.D., Perry, R.D., and Oliveira, M.A. (2006) Polyamines are essential for the formation of plague biofilm. *J Bacteriol* **188**: 2355–2363.
- Ramarao, N., and Sanchis, V. (2013) The pore-forming haemolysins of *Bacillus cereus*: a review. *Toxins* **5**: 7492–7502.
- Rolny, I.S., Minnaard, J., Racedo, S.M., and Pérez, P.F. (2014) Murine model of *Bacillus cereus* gastrointestinal infection. *J Med Microbiol* **63**: 1741–1749.
- Romero, D., Aguilar, C., Losick, R., and Kolter, R. (2010) Amyloid fibers provide structural integrity to *Bacillus subtilis* biofilms. *Proc Natl Acad Sci USA* **107**: 2230–2234.
- Roux, A., Beloin, C., and Ghigo, J.M. (2005) Combined inactivation and expression strategy to study gene function under physiological conditions: application to identification of new *Escherichia coli* adhesins. *J Bacteriol* **187**: 1001–1013.
- Schaeffer, P., Millet, J., and Aubert, J.P. (1965) Catabolic repression of bacterial sporulation. *Proc Natl Acad Sci USA* **54**: 704–711.
- Schoeni, J.L., and Wong, A.C. (2005) *Bacillus cereus* food poisoning and its toxins. *J Food Prot* **68**: 636–648.
- Silo-Suh, L.A., Lethbridge, B.J., Raffel, S.J., He, H., Clardy, J., and Handelsman, J. (1994) Biological activities of two fungistatic antibiotics produced by *Bacillus cereus* UW85. *Appl Environ Microbiol* **60**: 2023–2030.
- Smith, K., and Youngman, P. (1993) Evidence that the *spoIIM* gene of *Bacillus subtilis* is transcribed by RNA polymerase associated with  $\sigma^E$ . *J Bacteriol* **175**: 3618–3627.
- Söding, J. (2005) Protein homology detection by HMM-HMM comparison. *Bioinformatics* **21**: 951–960.
- Soni, A., Oey, I., Silcock, P., and Bremer, P. (2016) *Bacillus* spores in the food industry: a review on resistance and response to novel inactivation technologies. *Compr Rev Food Sci Food Saf* **15**: 1139–1148.
- Stenfors Arnesen, L.P., Fagerlund, A., and Granum, P.E. (2008) From soil to gut: *Bacillus cereus* and its food poisoning toxins. *FEMS Microbiol Rev* **32**: 579–606.
- Sturgill, G., and Rather, P.N. (2004) Evidence that putrescine acts as an extracellular signal required for swarming in *Proteus mirabilis*. *Mol Microbiol* **51**: 437–446.
- Tettelin, H., Massignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.I., *et al.* (2005) Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial “pan-genome”. *Proc Natl Acad Sci USA* **102**: 13950–13955.
- The Gene Ontology Consortium. (2015) Gene ontology consortium: Going forward. *Nucleic Acids Res.* **43**: D1049–D1056.
- Tovar-Rojo, F., Chander, M., Setlow, B., and Setlow, P. (2002) The products of the *spoVA* operon are involved in dipicolinic acid uptake into developing spores of *Bacillus subtilis*. *J Bacteriol* **184**: 584–587.
- Tran, S.L., Guillemet, E., Gohar, M., Lereclus, D., and Ramarao, N. (2010) CwpFM (EntFM) is a *Bacillus cereus* potential cell wall peptidase implicated in adhesion, biofilm formation, and virulence. *J Bacteriol* **192**: 2638–2642.
- Turnbull, P.C.B. (1976) Studies on the production of enterotoxins by *Bacillus cereus*. *J Clin Pathol* **29**: 941–948.
- Vidic, J., Chaix, C., Manzano, M., and Heyndrickx, M. (2020) Food sensing: detection of *Bacillus cereus* Spores In Dairy Products. *Biosensors* **10**: 15.
- Wiencek, K.M., Klapes, N.A., and Foegeding, P.M. (1990) Hydrophobicity of *Bacillus* and *Clostridium* spores. *Appl Environ Microbiol* **56**: 2600–2605.
- Wijnands, L.M., Pielaat, A., Dufrenne, J.B., Zwietering, M.H., and Van Leusden, F.M. (2009) Modelling the number of viable vegetative cells of *Bacillus cereus* passing through the stomach. *J Appl Microbiol* **106**: 258–267.
- Willenbrock, H., Hallin, P.F., Wassenaar, T.M., and Ussery, D.W. (2007) Characterization of probiotic *Escherichia coli* isolates with a novel pan-genome microarray. *Genome Biol* **8**: 267.
- Zimmermann, L., Stephens, A., Nam, S.-Z., Rau, D., Kübler, J., Lozajic, M., *et al.* (2018) A completely reimplemented

MPI bioinformatics toolkit with a new hHpred Server at its Core. *J Mol Biol* **430**: 2237–2243.

## Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Fig. S1.** Persistence and sporulation of *B. cereus* strains on melon and cucumber leaves, cucumber fruit or endive. CFU counts per gram of leaf fruit 15 days after inoculation with *B. cereus* AH187, ATCC 14579, DSM 4282, DSM 4313, DSM 14729, UMAF8565, 6A27, DSM 8438 and DSM 2302. The black bar plots at the top show the total CFU count per gram of leaf or fruit. The grey bar graphs at the bottom correspond to the spore count per gram of leaf or fruit. The detection limit of the technique is  $10^2$  CFU/g of leaf. Error bars are the standard error of three or more replicates. Statistical analysis is the result of the comparison of all strains against AH187. \*( $P < 0.05$ ), \*\*( $P < 0.01$ ), \*\*\*( $P < 0.001$ ) (Dunn's test). No counts of the DSM 2302 strain appeared in any sporulation assay or on any surface, except endive. No significant counts of DSM 8438 strain appeared over all the surfaces (total CFUs and spores), except endive.

**Fig. S2.** Phylogenetic relation of *B. cereus* AH187, ATCC 14579, DSM 4282, DSM 4313, DSM 14729, UMAF8564, 6A27, DSM 8438 and DSM 2302 strains. Red points indicate the strains that were not able to persist on the melon leaf surface (DSM 8438 and DSM 2302). *B. subtilis* subsp. *subtilis* 168 was used as an out-group. In the analysis, four concatenated genes (16S rRNA, *gyrB*, *rpoD* and *rpoB*) were handled to build a neighbour-joining tree using the MUSCLE algorithm of MEGA7.

**Fig. S3.** Pangenome analysis of the strains AH187, ATCC 14579, DSM 4282, DSM 4313, DSM 14729, UMAF8564, 6A27, DSM 8438 and DSM 2302. A Schematic view of the pangenome analysis of the nine studied strains. The inner layer represents the pangenome (all the analysed genes). Each colour layer corresponds to the genome of one *B. cereus* strain. B Statistical estimation with Willenbock (blue) and Tettelin (red) (Tettelin et al., 2005; Willenbrock et al., 2007). The plot on the top points corresponds to pangenome size according to the number of genomes. In the plot on the bottom, the points correspond to the calculation of the core genome for each genome. C Schematic view of the genomic comparison between strains AH187 and DSM 2302. Genes present in strain AH187 are represented as brown lines throughout the genome, while genes absent in strain DSM 2302 are represented as red lines just above previous. The percentage of CG is represented in black, and the GC skew is represented in purple to locate the intergenic zones and the origins of replication of these strains, respectively.

**Fig. S4.** Comparison of biofilm formation capacity of AH187 and DSM 2302. 72 h in vitro biofilm and swarming assays of AH187 and DSM 2302 strains on TY 1.5% agar supplemented with 20  $\mu\text{g/ml}$  and 10  $\mu\text{g/ml}$  Congo Red and Coomassie Brilliant Blue dyes (A1), TY 0.7% agar (A2) and TY broth (A3). A4 shows bacterial adhesion to an abiotic surface after staining with crystal violet. B *sipW-calY* locus implicated in extracellular matrix synthesis of AH187 and DSM 2302 compared to the *B. cereus* type strain (ATCC 14579). C Region homologous to the *eps1* region of ATCC 14579 involved in exopolysaccharide biosynthesis of AH187 and DSM 2302 strains. Middle grey genes correspond to different genes compared to each other. D Region homologous to the *eps2* region of ATCC 14579 involved in exopolysaccharide biosynthesis found in AH187 and DSM 2302 strains. Grey genes correspond to different genes located in the *eps2* genome region.

**Fig. S5.** Metabolomic profile of melon and endive leaf. Clustering result shown as a heatmap of three replicates of leaves of endive (orange) and melon leaves (green). In the heatmap, red indicates a higher concentration of a metabolite; in contrast, blue indicates a lower concentration of the metabolite.

**Fig. S6.** Persistence of AH187 mutants on melon leaves. CFU (black bar plot at the left) and spore (grey bar plot at the right) counts of AH187 mutants of the genes listed in Table S5 on melon leaves 15 days after bacteria inoculation. The detection limit of the technique is  $10^2$  CFU/g of leaf. Error bars are the standard error of three or more replicates. Statistical analysis is the result of the comparison of all the mutants against AH187 wild-type. \*( $P < 0.05$ ), \*\*( $P < 0.01$ ), \*\*\*( $P < 0.001$ ) (Dunn's test). *sigE* mutant counts fell below the detection limit.

**Table S1.** Sequenced *B. cereus* strains deposited in Genbank. Sequenced *B. cereus* strains deposited in Genbank.

**Table S2.** Primers used to construct mutants of *B. cereus* AH187.

**Table S3.** Pangenome analysis of *B. cereus* strains. Core indicates the proteins shared among all the studied strains, accessory shows the proteins shared among all studied strains except one or less, and unique enumerates the proteins only present in one of the studied strains and not in the others. Differences in core proteins correspond to paralogous genes.

**Table S4.** GO and KO terms, obtained from the pangenome analysis that are present in *B. cereus* DSM 2302 strains and absent in AH187 strain. GO and terms associated with DSM 2302 strain with their corresponding function and DSM 2302 genes related with each GO/KO term.

**Table S5.** Genes absent in *B. cereus* DSM 2302 and DSM 8438. Genomic comparison among persister and non-persister strains reveal genes non present in DSM 8438 and DSM 2302 strains.