J Ind Microbiol Biotechnol (2015) 42:925–937
DOI 10.1007/s10295-015-1612-y

GENETICS AND MOLECULAR BIOLOGY OF INDUSTRIAL ORGANISMS

Whole-genome sequencing of *Bacillus subtilis* XF-1 reveals mechanisms for biological control and multiple beneficial properties in plants

Shengye Guo[1] · Xingyu Li[1] · Pengfei He[1] · Honhing Ho[2] · Yixin Wu[1] · Yueqiu He[1]

Received: 5 November 2014 / Accepted: 25 March 2015 / Published online: 10 April 2015

Abstract *Bacillus subtilis* XF-1 is a gram-positive, plant-associated bacterium that stimulates plant growth and produces secondary metabolites that suppress soil-borne plant pathogens. In particular, it is especially highly efficient at controlling the clubroot disease of cruciferous crops. Its 4,061,186-bp genome contains an estimated 3853 protein-coding sequences and the 1155 genes of XF-1 are present in most genome-sequenced *Bacillus* strains: 3757 genes in *B. subtilis* 168, and 1164 in *B. amyloliquefaciens* FZB42. Analysis using the Cluster of Orthologous Groups database of proteins shows that 60 genes control bacterial mobility, 221 genes are related to cell wall and membrane biosynthesis, and more than 112 are genes associated with secondary metabolites. In addition, the genes contributed to the strain's plant colonization, bio-control and stimulation of plant growth. Sequencing of the genome is a fundamental step for developing a desired strain to serve as an efficient biological control agent and plant growth stimulator. Similar to other members of the taxon, XF-1 has a genome that contains giant gene clusters for the non-ribosomal synthesis of antifungal lipopeptides (surfactin and fengycin), the polyketides (macrolactin and bacillaene), the siderophore bacillibactin, and the dipeptide bacilysin. There are two synthesis pathways for volatile growth-promoting compounds. The expression of biosynthesized antibiotic peptides in XF-1 was revealed by matrix-assisted laser desorption/ionization-time of flight mass spectrometry.

Keywords Genome sequence · *Bacillus subtilis* XF-1 · Bio-control · PGPR · NRPs

Introduction

The gram-positive bacterium *Bacillus subtilis* is the workhorse of a wide range of industrial processes and a model organism for biological research. It competes with pathogens for the nutrients around plants, triggers the plant defense system against pathogen invasion [36, 43], and secretes antimicrobial substances that inhibit the growth of pathogenic bacteria through the successful colonization of the rhizosphere and the plant surface [4]. As member of a family of plant growth-promoting rhizobacteria (PGPR), *Bacillus subtilis* has great potential for increasing the yields and the bio-control of diseases for a wide range of crops. For example, *B. subtilis* can control *Fusarium verticillioides* at the root level [10], suppress the damping-off of tomato seedlings caused by *Rhizoctonia solani* [1], inhibit *Eutypa lata* growth on grapevines [22], and form substances to induce plant resistance [36].

Bacillus subtilis XF-1 (thereafter XF-1) is a bio-control strain which was isolated from the rhizosphere soil of Chinese cabbage (*Brassica pekinensis*) infected by *Plasmodiophora brassicae* in Yunnan. For the control of clubroot disease of cruciferous crops, the patent biological agent (CN 101416641 A) derived from XF-1 has been applied to almost 5,000,000 acres in China. This strain also showed high suppression effect on 21 fungal pathogens belonging to oomycetes, ascomycetes, basidiomycetes

Electronic supplementary material The online version of this article (doi:10.1007/s10295-015-1612-y) contains supplementary material, which is available to authorized users.

✉ Yueqiu He
ynfh2007@163.com

[1] Faculty of Agronomy and Biotechnology, Yunnan Agricultural University, Kunming 650201, China

[2] Department of Biology, State University of New York, New Paltz, NY 12561, USA

and deuteromycetes [75]. Compared with another plant growth-promoting rhizobacterium: *Pseudomonas* [15, 39], relatively little is known about how different kinds of genes affect the life style and regulate the bacterium–plant interaction of *B. subtilis.* Here, we present an analysis of the XF-1 genome in comparison to *B. subtilis* subsp. *subtilis* str.168 and other *Bacillus* strains, by highlighting the genes that may contribute to its antimicrobial action and plant-associated life style.

Materials and methods

Whole-genome shotgun sequencing and gene prediction

Sequencing of the XF-1 genome was performed at the Beijing Genomic Institute (BGI, Shenzhen, China) using a combination of randomly sheared libraries with inserts in the 0.5–2 kb size range. Genomic DNA sequencing was performed using an Illumina GA II system (Solexa), and all generated reads were qualitatively assessed and the vector sequences were trimmed before assembly using the Short Oligonucleotides Alignment Program (SOAP) de novo. Primer-walking and PCR amplification, with the guidance from *B. subtilis* subsp. *subtilis* str. 168 genomic sequences, were used to fill the remaining gaps and to solve the mis-assembled regions generated by the high degree of repetitive sequences. After the independent de novo assembly for each sample independently, the metagene [53], which uses di-codon frequencies estimated by the GC content of the obtained sequence, was used to predict the entire range of ORFs based on the XF-1 genome. The predicted ORFs were then aligned using BLAST [35]. A gene pair with greater than 90 % identity and an aligned length covering over 90 % of the shorter gene was grouped together. The groups with shared genes were then merged, and the longest ORF in each merged group was used to represent the group, and the other members of the group were considered as redundancy. Therefore, the non-redundant gene sets from all the predicted genes were organized by excluding the redundancies, and the ORFs that were less than 100-bp long were filtered. The predicted ORFs were translated into protein sequences using the NCBI genetic codes [17].

Genome analysis and annotation

Putative coding sequences (CDSs) were identified again using Glimmer3 [18]. All identified CDSs were manually reviewed, and false CDSs were removed [3]. The remaining CDSs were then subjected to automatic functional annotation via BLAST with an e-value setting of 1e-5, and a search against publicly available genomic data. The tRNA genes were identified using tRNAscan-SE, and the rRNA genes were confirmed using BLAST on the existing rDNA database with an e-value setting of 1e-5; repeating DNA sequences were identified using Tandem Repeat Finder (TRF) (http://www-is.biotoul.fr/) with the MaxPeriod set to 2000. The transposons and IS sequences were predicted by Repeatmasker. The genomic islands were predicted by Island Viewer (http://www.pathogenomics.sfu.ca/island-viewer/query.php). Detailed information is summarized in Table S1. Genome annotation was performed using the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP), and the GenBank NR database, whereas the KEGG [32] and COG [65] databases were employed for BLASTP identification [66]. The critical standard for gene annotations was based either on genes sharing 80 % identity for 80 % of their length, or on synteny (at least 3 genes in synteny between genomes), e.g., sharing 60 % identity and 80 % of their length between *B. subtilis* 168 with *B. amyloliquefaciens* FZB42 genes.

Construction and comparative analysis of orthologous and paralogous families

Orthologous and paralogous genes were grouped by comparing the XF-1 proteins and coding sequences with the information from 17 strains of *B. subttlis* downloaded for NCBI. Eighteen publicly available genomes of *Bacillus* were obtained using OrthoMCL DB (http://www.orthomcl.org/cgi-bin/OrthoMclWeb.cgi) and MUSCLE (http://www.ebi.ac.uk/Tools/msa/muscle/): *B. subtilis* XF-1 (NC_020244), *B. subtilis* subsp. *spizizenii* TU-B-10 (NC_016047), *B. subtilis* subsp. *spizizenii* W23 (NC_014479), *B. subtilis* subsp. *natto* BEST195 (NC_017196), *B. subtilis* subsp. *subtilis* RO-NN-1 (NC_017195), BSn5 (NC_014976), BSP1 (NC_019896), and 168 (NC_000964), *B. subtilis* QB928 (NC_018520), *B. amyloliquefaciens* Y2 (NC_017912), *B. amyloliquefaciens* FZB42 (NC_009725), and DSM 7 (NC_014551), *B. licheniformis* DSM 13 (NC_006322), *B. cereus* ATCC 14579 (NC_004722), *B. thuringiensis serovarkonkukian* 97-27 (NC_005957), *B. anthracis* str. Ames (NC_003997), *B. pumilus* SAFR-032 (NC_009848), *B. coagulans* 2-6 (NC_015634). Gene families and their phylogenic relationship were constructed using MEGA 4.2 (http://www.megasoftware.net/mega4/m_test_reliab.html).

MALDI-TOF–MS analysis

All solvents and reagents were of analytical grade or the highest grade available. XF-1 (CGMCC NO. 2357), which was stored in our laboratory, was streaked on LB (Luria–Bertani) agar plates to obtain separated colonies. Cells from a single bacterial colony grown on LB agar for 72 h at 30 °C were transferred to a target spot of

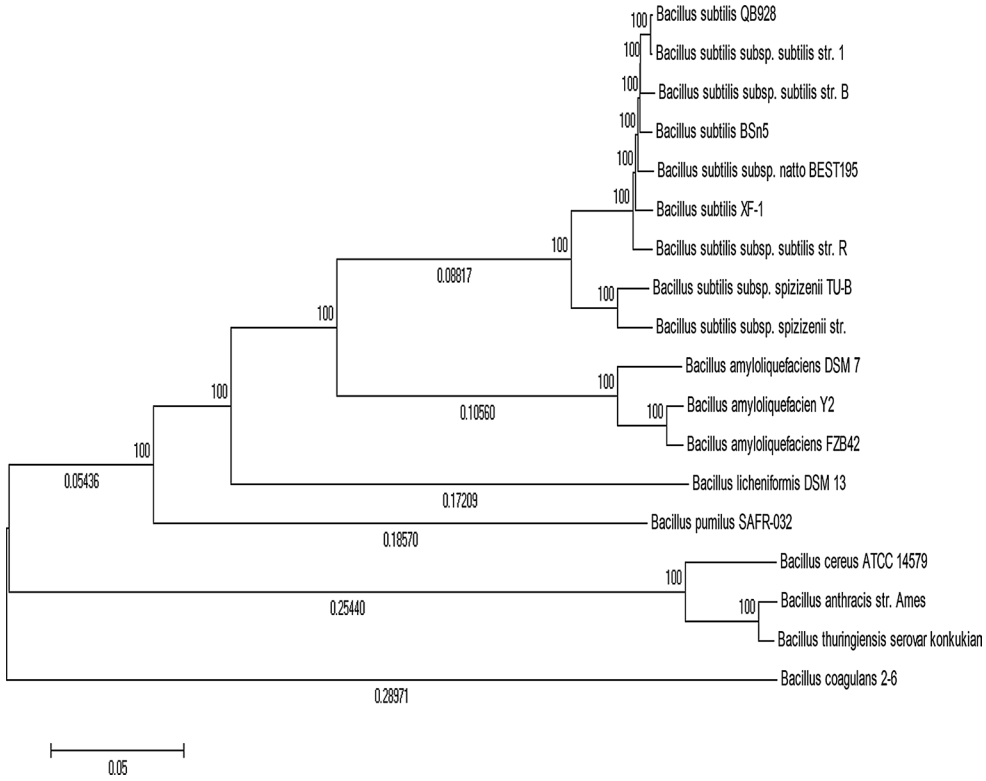

Fig. 1 Construction of the phylogenetic tree of various strains of *Bacillus* spp. The tree was reconstructed from the core genomes of various isolates of *Bacillus* spp. with *B. coagulans* 2–6 (NC_015634) serving as the out group. The *bar* (0.05) indicates substitutions per amino acid within the coding regions of the core genome

a steel target plate using a disposable loop, overlaid with 0.5 μL of an α-cyano-4-hydroxycinnamic acid (CHCA) matrix, and air-dried for 1–2 min at 24–27 °C. The sample mass fingerprints were obtained using a Bruker Autoflex Speed MALDI-TOF–MS (Bruker Daltonics, Billerica, MA, USA), with detection in the linear, positive mode at a laser frequency of 30 Hz and within a mass range of 100–3000 Da. The acceleration voltage was 20 kV, and the extraction delay time was 200 ns.

Results

General features of the genome

The complete sequence of *B. subtilis* XF-1 has been deposited in the NCBI GenBank (Accession no.CP004019). Without the plasmid in the XF-1 genome, the circular chromosome size (4,061,186-bp) is smaller than the genome sizes of both *B. subtilis* 168 and *B. amyloliquefaciens* FZB42, which are the most closely related strains in the genus *Bacillus* (Figs. 1, S1, S2). With 43.8 % GC, 108 tandem repeats, 84 transposons, 77 types of tRNA and 9 clusters of rDNA, the whole genome was predicted to contain 3853 CDSs (coding sequences) (Fig. 2). Ten genomic islands are spread throughout the XF-1 genome (Table 1). BLAST analysis with the available genome-sequenced strains in genus *Bacillus* revealed that 1155 genes of XF-1 belong to the core in most *Bacillus* strains, with 3757 genes found in *B. subtilis* 168, and 1164 in *B. amyloliquefaciens* FZB42 (Fig. 3).

After manual annotation, 2722 CDSs (65 %) were assigned to putative biological functions, whereas 1131 CDSs (36 %) were treated as hypothetical proteins with unknown function. For the CDSs with unassigned functions, 791 CDSs (24 %) corresponded to conserved hypothetical proteins, whereas 340 CDSs (12 %) had no homology to any previously reported sequences. Using the COG module of the MaGe system, 3357 CDSs (81.8 %) were assigned to one or more COG functional classes; and 60 genes were associated with the control of bacterial mobility, and 221 were genes related to the cell wall and membrane biosynthesis, whereas 112 genes were involved in the production of secondary metabolites (Table 2). The genome encodes numerous pathways for the utilization of plant-derived molecules, including proteins for controlling the swarming of bacteria on a plant surface and enzymes for forming architecturally complex biofilms and producing abundant antibiotics and plant growth substances. These genes, which will be described in detail later, are beneficial to the XF-1 strain's plant colonization and growth promotion, and the protection of its plant host against pathogens.

Several features of the origin of replication of XF-1 were similar to those of the corresponding regions in the genus *Bacillus*. Co-localization of three genes (*rpm*H, *dna*A, C663_0002, and *dna*N C663_0003), the GC

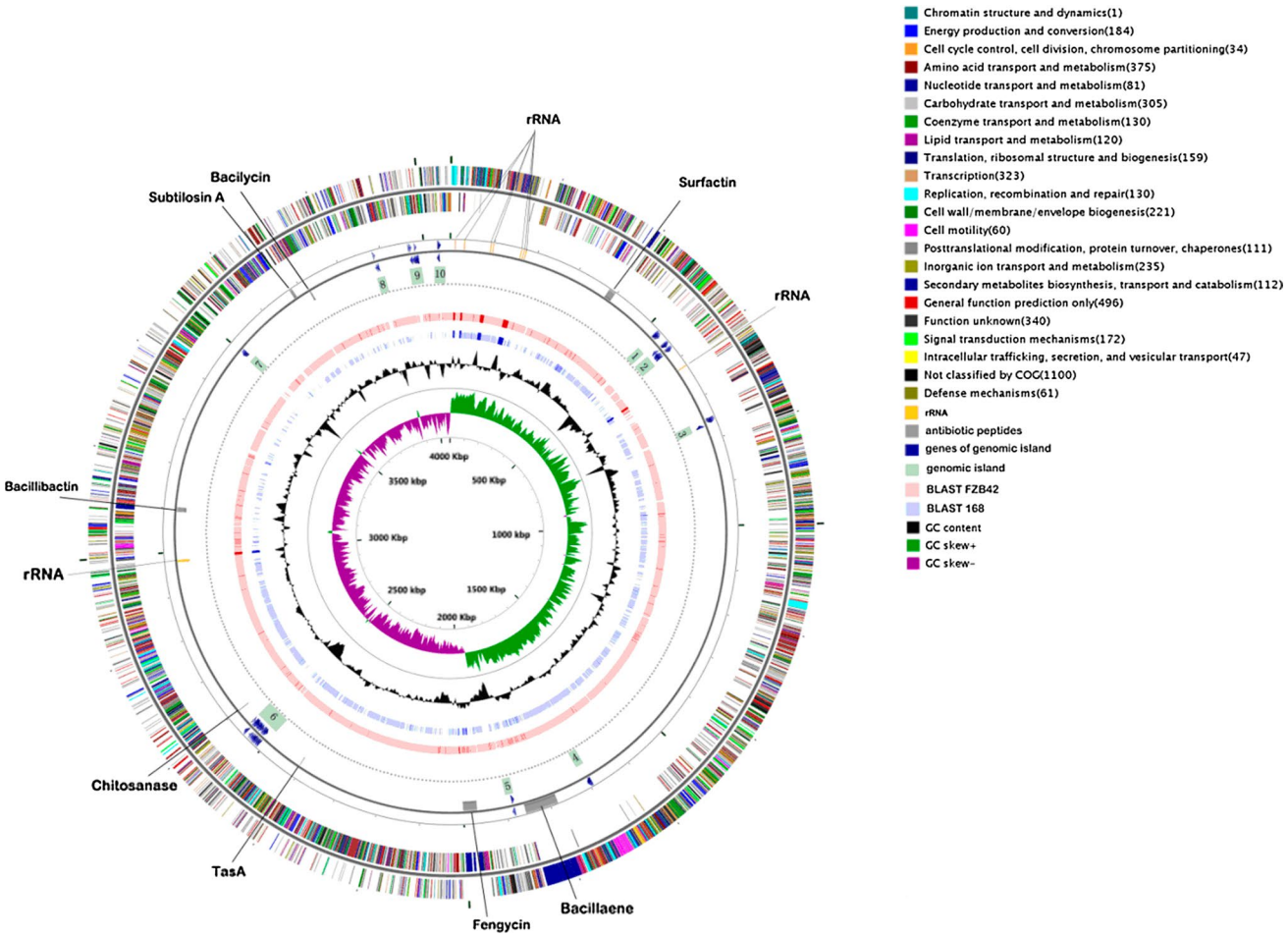

Fig. 2 Circular representation of the *Bacillus subtilis* XF-1 genome for several specific genomic features. *Outermost circle* (1st): all genes are *color* coded according to their functions (see *top right*); 2*nd circle*: rRNA genes (*yellow*), giant gene clusters devoted to the synthesis of secondary metabolites (*gray*), and genes of genomic islands (*blue*); 3*rd circle*: 10 genomic islands (*light green*) numbered 1–10; 4*th circle*: orthologs of FZB42 (*pink*); 5*th circle*: orthologs of 168 (*light blue*); 6*th circle*: GC content (*black*); 7*th circle*: GC skew+ (*green*); 8*th circle*: GC skew− (*violet*); 9*th circle*: scale (bps). GC views were prepared using CGView Server V1.0 (http://wishart.biology.ualberta.ca/cgview/) (color figure online)

Table 1 Prediction of genomic islands in *Bacillus subtilis* XF-1

Gene start	Gene end	Size	GI Prediction program
524293	534238	9945	Predicted by multiple methods (SIGI-HMM)
554480	573751	19271	Predicted by multiple methods (IslandPath-DIMOB)
752051	763294	11243	Predicted by multiple methods (SIGI-HMM)
1694317	1700420	6103	Predicted by multiple methods (IslandPath-DIMOB)
1879630	1885248	5618	Predicted by multiple methods (SIGI-HMM)
2513750	2556692	42942	Predicted by multiple methods (SIGI-HMM) (IslandPath-DIMOB)
3497242	3505660	8418	Predicted by multiple methods (SIGI-HMM)
3880746	3885766	5020	Predicted by multiple methods (SIGI-HMM)
3964998	3983345	18347	Predicted by multiple methods (SIGI-HMM) (IslandPath-DIMOB)
4030931	4036889	5958	Predicted by multiple methods (SIGI-HMM)

nucleotides skew [(G−C)/(G + C)] analysis, and the presence of multiple *dna*A-boxes and AT-rich sequences were common to the *dna*A and *rpm*H genes. The replication termination site, which lies roughly opposite to the origin of replication, was localized at 2.18 megabases by GC skew analysis. The maintenance and stability of the genetic

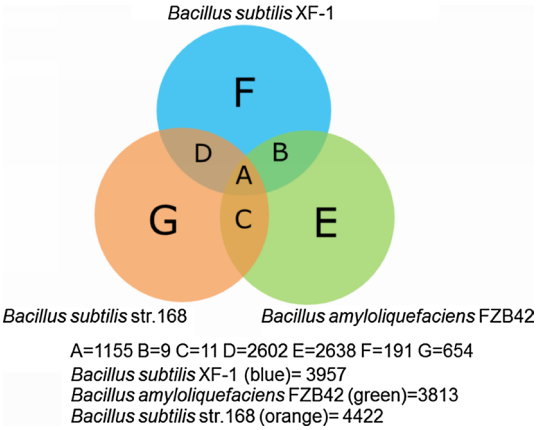

Fig. 3 Venn diagram for *Bacillus subtilis* XF-1, str. 168 and *B. amyloliquefaciens* FZB42. The numbers of gene families and gene clusters in the families are indicated for every species. The intersections between species indicate the number of shared gene families, whereas the number of unique families is shown in the species-specific areas. All three *Bacillus* strains have the same gene (*A*). XF-1 and FZB42 share the common genes ($A + B$). 168 and FZB42 share the common genes ($A + C$). 168 and XF-1 share the common genes ($A + D$)

Table 2 General features of the *Bacillus subtilis* XF-1 genome

General traits	Chromosome
size (bp)	4061186
G + C content	43.8 %
Assigned function (including putative)	3750
tRNA	77
rRNA	9
Chromatin structure and dynamics	1
Energy production and conversion	184
Cell cycle control, cell division, chromosome partitioning	34
Amino acid transport and metabolism	375
Nucleotide transport and metabolism	81
Carbohydrate transport and metabolism	305
Coenzyme transport and metabolism	130
Lipid transport and metabolism	120
Translation, ribosomal structure and biogenesis	159
Transcription	323
Replication, recombination and repair	130
Cell wall/membrane/envelope biogenesis	221
Cell motility	60
Posttranslational modification, protein turnover, chaperones	111
Inorganic ion transport and metabolism	235
Secondary metabolites biosynthesis, transport and catabolism	112
General function prediction only	496
Function unknown	340
Signal transduction mechanisms	172
Intracellular trafficking, secretion, and vesicular transport	47
Defense mechanisms	61

information are important to all kingdoms of life and depend on numerous types of DNA repair systems; these systems include excision repair system (*yxl*J C663_3775, *ung* C663_3705, *mut*M C663_2753, *yfh*Q C663_0885, *nfo* C663_2395, *mfd* C663_0056, *pcrA* C663_0685), the mismatch repair system (*mut*S C663_1747, *mut*L C663_1748, *xse*A C663_2313, *xse*B C663_2312), the non-homologous end-joining system (*yko*U C663_1379, *yko*V C663_1380) and the homologous repairing system (Table 3). Other repair systems include *uvs*E (C663_3620), which repairs UV radiation-induced cyclobutane pyrimidine dimerization, and *spl*A (C663_1435) along with *spl*B (C663_1436) which are spore photoproduct lyases for repairing UV-induced DNA damage and are conserved in the genus *Bacillus*. The DNA double-strand checking was performed by *dis*A (C663_0090), which has diadenylate cyclase activity for cyclic diadenosine phosphate (c-di-AMP) synthesis [47, 72]. These genes may contribute to the genomic stability and cell survival of the strain, especially in harsh soil conditions.

Plant colonization by XF-1

It is widely recognized that efficient colonization on plant roots is a key step for plant growth promotion and disease suppression by PGPR [67]. Our research results show that the XF-1 strain can colonize in the rhizosphere, on the root surface and within root tissues of a variety of plants, including Chinese cabbage, tomato, tabacco and maize [44, 75]. The first step for XF-1 colonization reaches the surface of the plant roots, either by passive movement in water fluxes or by active, flagella-propelled swimming. XF-1 has one flagella biosynthesis operon (*fla/che*) and two stator elements, *mot*AB and *mot*PS; this situation is similar to that found in most species of *Bacillus*, for example FZB42 [71]. The activation and expression of genes required for flagellar motion are required to ensure cell motility (*hag* C663_3399, *fli*D C663_3420, *fli*P C663_1681 and *flg*M C663_1677) [11, 21, 23] and chemotaxis (*che*C C663_1691, *che*D C663_1692, *che*V C663_1445) in response to root exudates [34, 37, 56] in many *Bacillus* strains. It is conceivable that XF-1 may actively reach the plant root surface by inducing the expression of *fla/che* operons during the colonization process [50]. The next step of colonization is to establish the interaction of the plant and microbe by non-specific binding or a firm anchorage, which would rely on a variety of cell surface-associated materials that enable bacterial biofilm formation.

Table 3 Genes of DNA repair systems in *Bacillus subtilis* XF-1

Bacillus subtilis XF-1 ORF	Gene name	Product
C663_2306	*rec*N	ATPase involved in DNA repair
C663_2409	*rec*O	Recombinational DNA repair protein (RecF pathway)
C663_2250	*din*B2	Nucleotidyl transferase/DNA polymerase involved in DNA repair
C663_3402	*uvr*B	Helicase subunit of the DNA excision repair complex
C663_1436	*spl*B	DNA repair photolyase
C663_0885	*yfh*Q	A/G-specific DNA glycosylase
C663_1435	*spl*A	TRAP-like transcriptional regulator
C663_1379	*yko*U	Predicted eukaryotic-type DNA primase
C663_0090	*dis*A	Predicted nucleic-acid-binding protein (contains the HHH domain)
C663_3620	*uvs*E	UV damage repair endonuclease
C663_3705	*ung*	Uracil DNA glycosylase
C663_0005	*rec*F	Recombinational DNA repair ATPase (*Rec*F pathway)
C663_2646	*rad*C	DNA repair proteins
C663_2703	*mut*SB	Mismatch repair ATPase (*Mut*S family)
C663_2753	*mut*M	Formamidopyrimidine-DNA glycosylase
C663_1747	*mut*S	Mismatch repair ATPase (*Mut*S family)
C663_1748	*mut*L	DNA mismatch repair enzyme (predicted ATPase)
C663_1004	*yha*Z	DNA alkylation repair enzyme
C663_1088	*sbc*D	DNA repair exonuclease
C663_0056	*mfd*	Transcription-repair coupling factor (superfamily II helicase)
C663_0427	*mut*T	NTP pyrophosphohydrolases including oxidative damage repair enzymes
C663_2239	*nud*F	NTP pyrophosphohydrolases including oxidative damage repair enzymes
C663_2395	*nfo*	Endonuclease IV
C663_0685	*pcr*A	Superfamily I DNA and RNA helicases
C663_2313	*xse*A	Exonuclease VII, large subunit
C663_2312	*xse*B	Exonuclease VII small subunit
C663_3775	*yxl*J	3-Methyladenine DNA glycosylase
C663_0824	*yfj*P	3-Methyladenine DNA glycosylase/8-oxoguanine DNA glycosylase
C663_3366	*yvc*I	NTP pyrophosphohydrolases including oxidative damage repair enzymes
C663_2911	*ytk*D	NTP pyrophosphohydrolases including oxidative damage repair enzymes
C663_1380	*yko*V	Uncharacterized conserved protein

The biofilm not only strengthens the interaction between plants and PGPR but also provides the plant root system with a protective barrier against attack by pathogenic organisms. Extracellular polymeric substances (EPS) are the dominant biofilm components on the root surface, and it is densely packed with communities of microorganisms. The EPS produced by *Bacillus* strains exhibit a wide range of chemical compositions: neutral polysaccharides, charged polymers, amphiphilic molecules and proteins with various molecular sizes ranging from 0.57 to 128 kDa [54]. It is conceivable that the expression of 16 genes in the XF-1 genome clustered as an operon (*eps*A-O, C663_3317–C663_3303) was likely inhibited by *sin*R (C663_2344) and activated by *sin*I (C663_2343) whereas *pgc*A (C663_0954) might play an important role in biofilm formation with a putative function as either a phosphoglucomutase or phosphomannomutase [6, 33, 42, 48]. The main charged EPS is poly-g-glutamate, which is synthesized by the machinery system encoded by the *pgs*BCA (C663_1735) operon; this system presumably involved in cell–surface interactions via the absorption of other charged molecules [73]. The third type of EPS molecule is surface-active lipopeptides, which are involved in biofilm formation via suppression of both bacterial and fungal pathogens. The fourth form of EPS contains extracellular proteins that are completely dissociated from the cell and are found free in the surrounding medium or within the exopolymeric matrix. They include enzymes for the degrading plant polymers (e.g., proteins, polysaccharides and nucleic acids, etc.) and for biofilm formation and architecture. Three distinct pathways exist in the genome for exporting these proteins from the cytoplasm to the surrounding environment [8]. Most

genes for the secretory pathway, such as Sec-SRP encoding genes (*sec*A C663_3415, *ffh* C663_1642, *sec*G C663_3234, *sec*E C663_0101, *sec*D C663_2605) and the genes for the twin-arginine translocation (*tat*AC C663_1825, *tat*AD C663_0254, *tat*AYC663_0581, *tat*CD C663_0255, *tat*CY C663_0582), have been found in the XF-1 genome [26]. We predicted 268 genes encoding proteins with putative secretion signals (Table S1).

Synergistic interactions for plant growth

The disease defense mechanism of a plant host is generally regarded as resulting from the direct action of plant hormone synthesis, which increases the availability and assimilation of plant mineral nutrients and indirect actions, via disease prevention or suppression [24, 29, 31]. *Bacillus subtilis* XF-1 does not have the ability to fix atmospheric nitrogen. However, it contains genes encoding molybdenum cofactor biosynthesis protein A (*moa*A, C663_3566), which may be a relic of a nitrogen-fixing gene cluster or a cofactor for nitrogen assimilation [5]. The genes for nitrate reduction *nar*IJHG (C663_3626–C663_3629), the nitrate transporter *nar*K (C663_3633), their putative regulator gene *arf*M (C663_3630) and *fnr* (C663_3632), the nitrite transporter *nas*A (C663_0323) and *nas*B–E (C663_0322-C663_0319) [49] are found in the XF-1 genome along with an amino acid transporter (*nrg*A, C663_3545) and its regulator (*nrg*B, C663_3546); *nas*B–E is predicted to be involved in nitrite transport and reduction [74].

Three genes (*ykq*A C663_1493, *ktr*A C663_2960, *yug*O C663_2990) were predicted as potassium transporters in XF-1 [28]. For phosphate, there are at least 12 genes related to its uptake, starvation regulation and organic phosphate hydrolysis. We have used the phytase gene (inositol-phosphate phosphatase, *ppn*K C663_2800) as a molecular marker for the primary selection of plant growth-promoting bacteria. Other mineral-related genes operate via the synthesis of siderophores to produce antibiotics with antagonistic effects and to deprive iron from the surrounding microorganisms of iron as described previously [19, 20, 27, 52]. The genes for the transportation of magnesium (*mgt*E C663_1369), manganese (*mnt*H C663_0431), cobalt (*cor*A C663_2358) and ferrochrome (*yvd*K C663_3343) were predicted to function either in the uptake of nutrients or in the de-toxification of heavy metal ions for both the bacteria and the plant host (Table 4).

Growth-stimulating volatile compounds, including acetoin and 2,3-butanediol, are emitted by some of the most efficient plant growth-promoting rhizobacteria, including *Bacillus* spp., *Pseudomonas* spp. and *Enterobacter* spp. [16, 55, 57, 70]. We found two such synthetic pathways in XF-1. The first pathway involves *als*S (C663_3496), which encodes acetolactate synthase; *als*D (C663_3495), which encodes acetolactate decarboxylase; and *als*R (C663_3497), which regulates these genes. Acetolactate synthase catalyzes the condensation of two pyruvate molecules into acetolactate, and acetolactate decarboxylase converts decarboxylated acetolactate into acetoin. The second pathway was predicted using the *ilv*H (C663_2673) and *ilv*C (C663_2672) genes, which are parts of the leucine, valine, and isoleucine biosynthesis pathway [2]. The reduction of acetoin to 2,3-butanediol is thought to be catalyzed by unknown enzymes as a result of the reverse reaction of dehydrogenation with NADH as a hydrogen and electron donor. In the XF-1 genome, 6 alcohol dehydrogenases were found, but their functions require further investigation (Table 4). The acetoin and 2,3-butanediol synthesis systems in PGPR can shift carbohydrate catabolism from acidic products, lactate, toxic compounds or ethanol, to neutral volatile compounds with growth-promoting ability. This type of shift is beneficial for strain colonization and the establishment of the plant–symbiotic relationship especially under the low O_2 partial pressure in the rhizosphere. Substances known to be involved in plant growth promotion, including putrescine and spermidine, are often produced by a large variety of PGPRs [9]. *Spe*A (C663_1507), the gene encoding arginine decarboxylase; *spe*B (C663_3654), which encodes agmatinase; and *spe*E (C663_3655), which encodes spermidine synthase in the arginine degradation pathway, were found in XF-1 genome [59, 64]. These genes may transform amino acids to plant growth-promoting substances, which contribute to the utilization of this strain in agriculture.

Non-ribosomal biosynthesized antibiotic peptides (NRPs)

NRPs exist in many bacteria and fungi [69]. The synthesis enzymes and non-ribosomal peptide synthetases, which are composed of multi-modularly arranged catalytic domains, catalyze peptide formation in multiple repeated steps [60]. Non-ribosomally generated amphipathic lipopeptide antibiotics with condensed beta-hydroxyl or beta-amino fatty acids are widespread in *Bacillus* spp. and vary in the length of the fatty acid chains and amino acid composition, resulting in remarkably different molecules. In general, NRPs in XF-1 are grouped into two families: surfactin and fengycin.

Surfactin is a potent biosurfactant and shows hemolytic, antimicrobial and antiviral activities via alteration of membrane integrity [30]. Comparative analysis of the genomic sequence of *B. amyloliquefaciens* FZB42, a surfactin-producing strain, with XF-1 showed that *srf*AA (C663_0341), *srf*AB (C663_0342), *srf*AC (C663_0343) and *srf*AD (C663_0344) are highly similar to those of FZB42 (*srf*A-*srf*D), and are likely involved in the non-ribosomal synthesis of surfactin. *sfp* [C663_0349, phosphopantetheinyl

Table 4 Genes for promoting plant growth and the assimilation of minerals in *Bacillus subtilis* XF-1

Bacillus subtilis XF-1 ORF	Gene name	Product
C663_3566	*moaA*	Molybdenum cofactor biosynthesis protein A
C663_3626	*narI*	Nitrate reductase gamma subunit
C663_3627	*narJ*	Nitrate reductase delta subunit
C663_3628	*narH*	Nitrate reductase beta subunit
C663_3629	*narG*	Nitrate reductase alpha subunit
C663_3633	*narK*	Nitrate/nitrite transporter
C663_3630	*arfM*	Transcriptional regulator
C663_3632	*fnr*	cAMP-binding proteins—catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases
C663_0323	*nasA*	Nitrate/nitrite transporter
C663_0319	*nasE*	Ferredoxin subunits of nitrite reductase and ring-hydroxylatingdioxygenases
C663_0320	*nasD*	NAD(P) H-nitrite reductase
C663_0321	*nasC*	Anaerobic dehydrogenases, typically selenocysteine-containing
C663_0322	*nasB*	NAD(P) H-nitrite reductase
C663_3545	*nrgA*	Ammonia permease
C663_3546	*nrgB*	Nitrogen regulatory protein PII
C663_2800	*ppnK*	Inositol-phosphate phosphatase
C663_2358	*corA*	Magnesium and cobalt transport protein
C663_1369	*mgtE*	Magnesium transporter
C663_0431	*mntH*	Manganese transport protein
C663_2923	*mntD*	Manganese ABC transporter
C663_2924	*mntC*	Manganese ABC transporter (membrane protein)
C663_2925	*mntB*	Manganese ABC transporter (ATP-binding protein) system ATP-binding protein
C663_2926	*mntA*	Manganese ABC transporter (membrane protein)
C663_1493	*ykqA*	Potassium uptake protein
C663_2960	*ktrA*	Potassium uptake protein
C663_2990	*yugO*	Predicted NAD-binding protein 2 (probable Kef-type transporter subunit); voltage-gated potassium channel
C663_3496	*als*S	Thiamine pyrophosphate-requiring enzymes [acetolactate synthase, pyruvate dehydrogenase (cytochrome), glyoxylatecarboligase, phosphonopyruvate decarboxylase]
C663_3495	*als*D	Alpha-acetolactate decarboxylase
C663_3497	*als*R	Transcriptional regulator
C663_2673	*ilv*H	Acetolactate synthase, small (regulatory) subunit
C663_2672	*ilv*C	Ketol-acid reductoisomerase

transferase (PPTase)] and *ycz*E (C663_0350, an integral membrane protein with unknown function) are essential for the non-ribosomal synthesis of lipopeptides and the synthesis of polyketides.

With the structural properties of cyclization, branching and unusual amino acid constituents, fengycin, which specifically suppresses the growth of filamentous fungi, is probably biosynthesized by five NRPSs, *pps*A–*pps*E (C663_1894–C663_1887) with high homology to *B. amyloliquefaciens* FZB42 [41, 61, 62, 68]. All proteins of *pps*A, *pps*D (C663_1888) and *pps*E are of similar size, containing modules that activate and ligate the first six amino acid residues of fengycin. It is postulated that *pps*B (C663_1893) may be a three-module enzyme that activates Pro, Gln and Tyr in positions 7–9 of fengycin, whereas *pps*C (C663_1890) may contribute to the C-terminal Ile and the closing of internal ring via the formation of macrolactone between the carboxyl group of Ile and the free hydroxy group of Tyr at position 3. The *deg*Q (C663_3033) gene, which is a regulatory protein, can promote the expression of *sfp* to increase fengycin production. Fengycin biosynthesis is presumably initiated similarly to the synthesis of surfactin [13]. The cells of *P. brassicae* were shown to be lysed by fengycin-type cyclopeptides (FTCPs) extracted and purified from XF-1. These compounds may partially account for the antifungal activity against *P. brassicae* [50].

Five gene fragments with high amino acid identity (more than 98 %) to the *dhb* FBECA, which directs the

synthesis of the siderophore bacillibactin (2,3-dihydroxybenzoyl-Gly-Thr) of FZB42, were detected in the XF-1 strain [14, 38]. Similar to fengycin synthesis, synthesis of the siderophore bacillibactin is dependent on a functional PPTase gene *sfp* and its regulatory gene *ycz*E. Bacilysin (L-alanyl-[2,3-epoxycyclohexanone-4]-L-alanine), which is another compound with activity against a broad spectrum of bacteria and *Candida albicans*, is a dipeptide antibiotic containing the N-terminal l-Ala residue at the C-terminal of the non-proteinogenic amino acid l-anticapsin. The synthesis of bacilysin is catalyzed by enzymes encoded by the *bac*ABCDE gene cluster [13], and the gene cluster of XF-1 was identified to be co-linear to FZB42.

Synthesis of polyketides

As major tertiary metabolites in bacteria, polyketides have been found to have many bioactivities, including the suppression of pathogenic organisms, enhancement of the plant immunity and inhibition of tumor growth. In *Bacillus*, polyketides are mainly synthesized by type I polyketide synthases (PKSs); the synthesis starts with acyl-CoA precursors and is followed by decarboxylative Claisen–Schmidt condensation. The biosynthetic pathway is similar to the pathway for the non-ribosomally synthesized peptides, which is catalyzed by giant multi-module enzymes [69]. The genome of XF-1 contains the *pks*A–S (C663_1751-C663_1772) locus with a remarkable size of 72.9 kb, a PKS system, and gene clusters encoding type I polyketide synthases similar to those found in the FZB42 strain [41]. Macrolactin is the polyketide biosynthesis product of the *pks* cluster. The XF-1 genome also harbors another giant gene cluster which may be involved in the polyketide synthesis of bacillaene. This cluster spans nearly 76 kb (*bae*B–S, from 1,783,092 to 1,859,586 bp).

Result of matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass analysis of biosynthesized low-molecular-weight antibiotic peptides

In the MALDI-TOF mass spectra, several groups of related mass signals were observed that represent protonated as well as mono- and dialkali-ion adducts. The corresponding compounds were identified by comparison with MS data previously obtained from both *B. amyloliquefaciens* Y2 and *B. amyloliquefaciens* FZB42 [12, 25, 41] and attributed to fengycin ($[M + H, Na]^+ = 1449.693$ and 1499.754; 1527.768), surfactin ($[M + Na, K]^+ = 1030.616$; 1044.629; 1058.651 and 1074.617; 1102.612), bacillaene ($[M + H]^+ = 581.282$), and macrolactin ($[M + K]^+ = 441.099$) (Fig. 4).

It is possible that the XF-1 strain might produce a "cocktail" of antibiotics for improved competition and survival, and the antibiotics described above are most likely only part of the total antibiotic spectrum. Genes involved in antibiotic biosynthesis and the expression of enzymes for polymer degradation of pathogenic organisms would contribute to the suppression of pathogenic organisms.

Discussion

The genomic sequencing and annotation of XF-1 provide a useful platform for studying the mechanisms governing plant growth promotion, protection and colonization. The genus *Bacillus* is known for forming metabolically inactive spores that are highly resistant to environmental stresses. However, in addition to the spores, genetically similar *Bacillus* spp. have numerous distinct cell and biofilm types with different abilities for producing antibiotics and enzymes to suppress the growth of pathogenic organisms under various physiological conditions [46]. These cell types, which are convertible, are controlled by three major transcriptional factors, *Deg*U, *Com*A, and *Spo*0A, with the phosphorylated forms (*Deg*U-P, *Com*A-P and *Spo*0A-P) as the active forms [45]. The XF-1 genome showed that both *Deg*U-p (C663_3435) and *Com*A-p (C663_1118) are likely to be linked to their cognate histidine kinases, *Deg*S and *Com*P (C663_3030), respectively, whereas *Spo*0A-P (C663_1845), mainly maintained by the phosphorylation of *Spo*0F and *Spo*0B to produce *Spo*0F-P and *Spo*0B-P, is probably activated by several histidine kinases (*kin*A C663_1442, *kin*B C663_3005, *kin*C C663_1492, *kin*D C663_1407, *kin*E C663_1393) and deactivated by response-regulator aspartyl-phosphate phosphatases (*rap*A C663_1275, *rap*B C663_3565, *rap*C C663_0370, *rap*D C663_3533, *rap*H C663_0711) [46]. *Deg*U-P induces the expression of the genes for cell movement (e.g., formation of flagella protein) and genes for the secretory machinery responsible for the production and secretion of enzymes for polysaccharide, protein and lipid degradation [40]. *Com*A-P induces the pathway leading to the production of antibiotics [45], killing neighboring microbes and promoting the uptake of hydrolyzed nutrients. *Com*A-P also promotes competence, which prepares cells to take up foreign DNA, leading to horizontal gene transfer [58]. *Spo*0A-P promotes the sporulation process of the XF-1 strain [63], allowing the cells to survive in harsh conditions, such as a totally depleted food supply and extreme stress. Prior to plant colonization, most XF-1 cells would become motile to seek out the plant host. Successful attachment and colonization of the root are accomplished by the production of bacterial enzymes to synthesize the extracellular matrix using available carbon sources derived from the degradation of pectin,

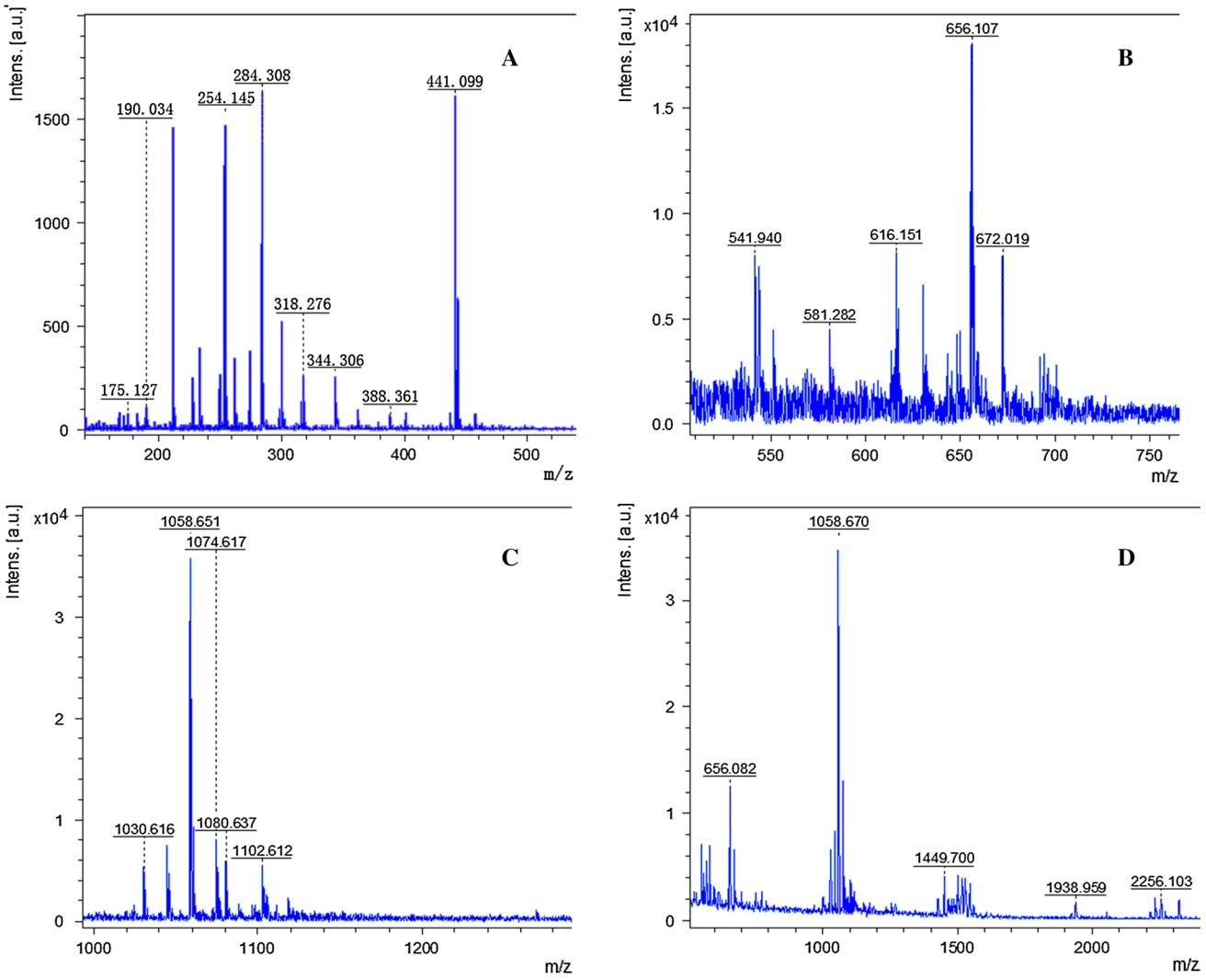

Fig. 4 MALDI-TOF mass spectra of an extract of a lyophilized culture filtrate obtained by the cultivation of *Bacillus subtilis* XF-1 in Luria–Bertani medium. Mass range between 400 and 600 showing the mass peaks of the polyketide products bacillaene and macrolactin (**a**, **b**); mass range between $m/z = 1000$ and 1200 showing the mass peaks of the surfactin lipopeptide family (**c**); mass range between $m/z = 1400$ and 1550 showing the mass peaks of the fengycin lipopeptide family (**d**)

cellulose and hemicelluloses of the plant host. The diversity of hydrolases for polymer degradation controlled by *Deg*U-P in XF-1 is consistent with this hypothesis.

The control of clubroot and other soil-borne pathogens by XF-1 could be primarily attributed to its occupation of the soil space, resulting in the depletion of available organic and inorganic compounds needed for plant pathogen propagation, the production of antibiotics, such as lipopeptides and polyketides, and the degradation of the cell wall of fungi and bacteria by ribosomally synthesized enzymes. From a genomic point of view, XF-1 contains several gene clusters for the synthesis of three types of EPS, the major component of the biofilm that can effectively block *Plasmodiophora* spores from coming in contact with plant roots. Another set of gene clusters produces different types of antimicrobial substances (e.g., both NRPs and polyketides and other low-molecular weight polypeptides, such as surfactin and fengycin, macrolactin and bacillaene, and bacillibactin and bacilysin) for the suppression of pathogenic microorganisms. In addition, carbohydrate utilizing proteins or enzymes produced by XF-1 are probably responsible for the degradation of the polysaccharides of fungi and bacteria: chitosanase (*csn*, C663_2522, EC: 3.2.1.132), xylanase/chitin deacetylase (*yxk*H, C663_3795, EC:3.5.1.41), *N*-acetylglucosaminidase (*lyt*D, C663_3472, EC:3.2.1.52) and exo-glucosaminidase (*lyt*G, C663_2964). Consistent with other fungi, *Plasmodiophora* has a spore wall whose major component is chitin [7, 51]. The carbohydrate utilizing genes produced by XF-1 suppress the pathogen infection by directly degrading the cell components of

the pathogen. Further researches using a combination of transcriptomics, proteomics, metabolomics and mutagenesis of *Bacillus subtilis* XF-1 will help to elucidate the exact functions of the putative genes and pathways.

Acknowledgments We thank the Ministry of Agriculture, China, and the Department of Science and Technologies, Yunnan Province, China, for their financial support through the Special Fund for Agro-scientific Research in the Public Interest (201003029) and the Natural Science Foundation of Yunnan Province (2008CC024).

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Standards This article does not contain any studies with human participants or animals performed by any of the authors. Informed consent was obtained from all individual participants included in the study.

References

1. Asaka O, Shoda M (1996) Biocontrol of *Rhizoctonia solani* damping-off of tomato with *Bacillus subtilis* RB14. Appl Environ Microbiol 62(11):4081–4085
2. Atkins T, Prior R, Mack K, Russell P, Nelson M, Oyston P, Dougan G, Titball R (2002) A mutant of *Burkholderia pseudomallei*, auxotrophic in the branched chain amino acid biosynthetic pathway, is attenuated and protective in a murine model of melioidosis. Infect Immun 70(9):5290
3. Badger JH, Olsen GJ (1999) CRITICA: coding region identification tool invoking comparative analysis. Mol Biol Evol 16(4):512
4. Bais HP, Fall R, Vivanco JM (2004) Biocontrol of *Bacillus subtilis* against infection of *Arabidopsis* roots by *Pseudomonas syringae* is facilitated by biofilm formation and surfactin production. Plant Physiol 134(1):307–319
5. Bird C, Wyman M (2003) Nitrate/nitrite assimilation system of the marine picoplanktonic cyanobacterium *Synechococcus* sp. strain WH 8103: effect of nitrogen source and availability on gene expression. Appl Environ Microbiol 69(12):7009–7018
6. Branda SS, Gonzalez-Pastor JE, Dervyn E, Ehrlich SD, Losick R, Kolter R (2004) Genes involved in formation of structured multicellular communities by *Bacillus subtilis*. J Bacteriol 186(12):3970–3979
7. Buczacki S, Moxham SE (1983) Structure of the resting spore wall of *Plasmodiophora brassicae* revealed by electron microscopy and chemical digestion. T Brit Mycol Soc 81(2):221–231
8. Candela T, Fouet A (2006) Poly-gamma-glutamate in bacteria. Mol Microbiol 60(5):1091–1098
9. Cassán F, Maiale S, Masciarelli O, Vidal A, Luna V, Ruiz O (2009) Cadaverine production by *Azospirillum brasilense* and its possible role in plant growth promotion and osmotic stress mitigation. Eur J Soil Biol 45(1):12–19
10. Cavaglieri L, Orlando J, Rodriguez M, Chulze S, Etcheverry M (2005) Biocontrol of *Bacillus subtilis* against *Fusarium verticillioides* in vitro and at the maize root level. Res Microbiol 156(5):748–754
11. Chen L, Helmann JD (1994) The *Bacillus subtilis* sigma D-dependent operon encoding the flagellar proteins FliD, FliS, and FliT. J Bacteriol 176(11):3093–3101
12. Chen XH, Vater J, Piel J, Franke P, Scholz R, Schneider K, Koumoutsi A, Hitzeroth G, Grammel N, Strittmatter AW (2006) Structural and functional characterization of three polyketide synthase gene clusters in *Bacillus amyloliquefaciens* FZB 42. J Bacteriol 188(11):4024–4036
13. Chen XH, Koumoutsi A, Scholz R, Schneider K, Vater J, Süssmuth R, Piel J, Borriss R (2009) Genome analysis of *Bacillus amyloliquefaciens* FZB42 reveals its potential for biocontrol of plant pathogens. J Biotechnol 140(1–2):27–37
14. Chen XH, Koumoutsi A, Scholz R, Eisenreich A, Schneider K, Heinemeyer I, Morgenstern B, Voss B, Hess WR, Reva O, Junge H, Voigt B, Jungblut PR, Vater J, Süssmuth R, Liesegang H, Strittmatter A, Gottschalk G, Borriss R (2007) Comparative analysis of the complete genome sequence of the plant growth–promoting bacterium *Bacillus amyloliquefaciens* FZB42. Nat Biotechnol 25(9):1007–1014
15. Compant S, Duffy B, Nowak J, Clément C, Barka EA (2005) Use of plant growth-promoting bacteria for biocontrol of plant diseases: principles, mechanisms of action, and future prospects. Appl Environ Microbiol 71(9):4951–4959
16. Cox GN, Kusch M, Edgar RS (1981) Cuticle of *Caenorhabditis elegans*: its isolation and partial characterization. J Cell Biol 90(1):7–17
17. Delcher AL, Bratke KA, Powers EC, Salzberg SL (2007) Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23(6):673
18. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL (1999) Improved microbial gene identification with GLIMMER. Nucleic Acid Res 27(23):4636
19. Duckworth OW, Bargar JR, Sposito G (2009) Coupled biogeochemical cycling of iron and manganese as mediated by microbial siderophores. Biometals 22(4):605–613
20. Duffy BK, Défago G (1999) Environmental factors modulating antibiotic and siderophore biosynthesis by *Pseudomonas fluorescens* biocontrol strains. Appl Environ Microbiol 65(6):2429–2438
21. Fan B, Carvalhais LC, Becker A, Fedoseyenko D, von Wirén N, Borriss R (2012) Transcriptomic profiling of *Bacillus amyloliquefaciens* FZB42 in response to maize root exudates. BMC Microbiol 12(1):116
22. Ferreira J, Matthee F, Thomas A (1991) Biological control of *Eutypa lata* on grapevine by an antagonistic strain of *Bacillus subtilis*. Phytopathology 81(3):283–287
23. Fredrick KL, Helmann JD (1994) Dual chemotaxis signaling pathways in *Bacillus subtilis*: a sigma D-dependent gene encodes a novel protein with both CheW and CheY homologous domains. J Bacteriol 176(9):2727–2735
24. Hayat R, Ali S, Amara U, Khalid R, Ahmed I (2010) Soil beneficial bacteria and their role in plant growth promotion: a review. Ann Microbiol 60(4):579–598
25. He P, Hao K, Blom J, Rückert C, Vater J, Mao Z, Wu Y, Hou M, He P, He Y (2013) 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(2):281–291
26. Hecker M, Völker U (2004) Towards a comprehensive understanding of *Bacillus subtilis* cell physiology by physiological proteomics. Proteomics 4(12):3727–3750
27. Heymann P, Gerads M, Schaller M, Dromer F, Winkelmann G, Ernst JF (2002) The siderophore iron transporter of *Candida albicans* (Sit1p/Arn1p) mediates uptake of ferrichrome-type siderophores and is required for epithelial invasion. Infect Immun 70(9):5246–5255
28. Holtmann G, Bakker EP, Uozumi N, Bremer E (2003) KtrAB and KtrCD: two K^+ uptake systems in *Bacillus subtilis* and their role in adaptation to hypertonicity. J Bacteriol 185(4):1289–1298
29. Huber DM (1980) The role of mineral nutrition in defense. In: Horsfall JG, Cowling EB (eds) Plant Disease. AcademicPress, New York, pp 381–406

30. Jacques P (2011) Surfactin and other lipopeptides from *Bacillus* spp. In: Soberón-Chávez G (ed) Biosurfactants. Springer, Berlin, pp 57–91
31. Jones JD, Dangl JL (2006) The plant immune system. Nature 444(7117):323–329
32. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T (2008) KEGG for linking genomes to life and the environment. Nucleic Acid Res 36(suppl 1):480–484
33. Kearns DB, Chu F, Branda SS, Kolter R, Losick R (2004) A master regulator for biofilm formation by *Bacillus subtilis*. Mol Microbiol 55(3):739–749
34. Kearns DB, Chu F, Rudner R, Losick R (2004) Genes governing swarming in *Bacillus subtilis* and evidence for a phase variation mechanism controlling surface motility. Mol Microbiol 52(2):357–369
35. Kent WJ (2002) BLAT-the BLAST-like alignment tool. Genome Res 12(4):656
36. Kilian M, Steiner U, Krebs B, Junge H, Schmiedeknecht G, Hain R (2000) FZB24® *Bacillus subtilis*–mode of action of a microbial agent enhancing plant vitality. Pflanzenschutz-Nachrichten Bayer 1(1):72–93
37. Kirby JR, Kristich CJ, Saulmon MM, Zimmer MA, Garrity LF, Zhulin IB, Ordal GW (2001) CheC is related to the family of flagellar switch proteins and acts independently from CheD to control chemotaxis in *Bacillus subtilis*. Mol Microbiol 42(3):573–585
38. Kloepper JW, Leong J, Teintze M, Schroth MN (1980) Enhanced plant growth by siderophores produced by plant growth-promoting rhizobacteria. Nature 286(5776):885–886
39. Kloepper JW, Ryu C-M, Zhang S (2004) Induced systemic resistance and promotion of plant growth by *Bacillus* spp. Phytopathology 94(11):1259–1266
40. Kobayashi K (2007) Gradual activation of the response regulator DegU controls serial expression of genes for flagellum formation and biofilm formation in *Bacillus subtilis*. Mol Microbiol 66(2):395–409
41. Koumoutsi A, Chen X-H, Henne A, Liesegang H, Hitzeroth G, Franke P, Vater J, Borriss R (2004) Structural and functional characterization of gene clusters directing nonribosomal synthesis of bioactive cyclic lipopeptides in *Bacillus amyloliquefaciens* strain FZB42. J Bacteriol 186(4):1084–1096
42. Lazarevic V, Soldo B, Médico N, Pooley H, Bron S, Karamata D (2005) *Bacillus subtilis* α-phosphoglucomutase is required for normal cell morphology and biofilm formation. Appl Environ Microbiol 71(1):39–45
43. Leclère V, Béchet M, Adam A, Guez JS, Wathelet B, Ongena M, Thonart P, Gancel F, Chollet-Imbert M, Jacques P (2005) Mycosubtilin overproduction by *Bacillus subtilis* BBG100 enhances the organism's antagonistic and biocontrol activities. Appl Environ Microbiol 71(8):4577–4584
44. Liu Q, Xiong G, Mao Z, Wu Y, He Y (2012) Analyses for the colonization ability of *Bacillus subtilis* XF-1 in the rhizosphere. Acta phytophylacica sinica 39:425–430
45. Lopez D, Kolter R (2010) Extracellular signals that define distinct and coexisting cell fates in *Bacillus subtilis*. FEMS Microbiol Rev 34(2):134–149
46. Lopez D, Vlamakis H, Kolter R (2009) Generation of multiple cell types in *Bacillus subtilis*. FEMS Microbiol Rev 33(1):152–163
47. Marsin S, McGovern S, Ehrlich SD, Bruand C, Polard P (2001) Early steps of *Bacillus subtilis* primosome assembly. J Biol Chem 276(49):45818
48. Marvasi M, Visscher PT, Casillas Martinez L (2010) Exopolymeric substances (EPS) from *Bacillus subtilis* : polymers and genes encoding their synthesis. FEMS Microbiol Lett 313(1):1–9
49. Merrick M, Edwards R (1995) Nitrogen control in bacteria. Microbiol Rev 59(4):604–622
50. Mirel DB, Lauer P, Chamberlin MJ (1994) Identification of flagellar synthesis regulatory and structural genes in a sigma D-dependent operon of *Bacillus subtilis*. J Bacteriol 176(15):4492–4500
51. Moxham SE, Buczacki S (1983) Chemical composition of the resting spore wall of *Plasmodiophora brassicae*. T British Mycol Soc 80(2):297–304
52. Neilands J (1995) Siderophores: structure and function of microbial iron transport compounds. J Biol Chem 270(45):26723–26726
53. Noguchi H, Park J, Takagi T (2006) MetaGene: prokaryotic gene finding from environmental genome shotgun sequences. Nucleic Acid Res 34(19):5623
54. Omoike A, Chorover J (2004) Spectroscopic study of extracellular polymeric substances from *Bacillus subtilis*: aqueous chemistry and adsorption effects. Biomacromolecules 5(4):1219–1230
55. Ping L, Boland W (2004) Signals from the underground: bacterial volatiles promote growth in *Arabidopsis*. Trends Plant Sci 9(6):263–266
56. Rao CV, Glekas GD, Ordal GW (2008) The three adaptation systems of *Bacillus subtilis* chemotaxis. Trends Microbiol 16(10):480–487
57. Ryu CM, Farag MA, Hu C-H, Reddy MS, Wei HX, Paré PW, Kloepper JW (2003) Bacterial volatiles promote growth in *Arabidopsis*. P Natl A Sci 100(8):4927–4932
58. Schultz D, Wolynes PG, Jacob EB, Onuchic JN (2009) Deciding fate in adverse times: sporulation and competence in *Bacillus subtilis*. P Natl A Sci 106(50):21027
59. Sekowska A, Bertin P, Danchin A (1998) Characterization of polyamine synthesis pathway in *Bacillus subtilis* 168. Mol Microbiol 29(3):851–858
60. Stein T (2005) *Bacillus subtilis* antibiotics: structures, syntheses and specific functions. Mol Microbiol 56(4):845–857
61. Steller S, Vater J (2000) Purification of the fengycin synthetase multienzyme system from *Bacillus subtilis* b213. J Chroma Togr B: Biomed Sci Appl 737(1):267–275
62. Steller S, Vollenbroich D, Leenders F, Stein T, Conrad B, Hofemeister J, Jacques P, Thonart P, Vater J (1999) Structural and functional organization of the fengycin synthetase multienzyme system from *Bacillus subtilis* b213 and A1/3. Chem Biol 6(1):31–41
63. Stragier P, Losick R (1996) Molecular genetics of sporulation in *Bacillus subtilis*. Annu Rev Genet 30(1):297–341
64. Tabor CW, Tabor H (1984) Polyamines. Annu Rev Biochem 53(1):749–790
65. Tatusov RL, Galperin MY, Natale DA, Koonin EV (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acid Res 28(1):33–36
66. Tatusova TA, Madden TL (2006) BLAST 2 sequences, a new tool for comparing protein and nucleotide sequences. FEMS Microbiol Lett 174(2):247–250
67. Timmusk S, Grantcharova N, Wagner EGH (2005) *Paenibacillus polymyxa* invades plant roots and forms biofilms. Appl Environ Microbiol 71(11):7292
68. Tosato V, Albertini AM, Zotti M, Sonda S, Bruschi CV (1997) Sequence completion, identification and definition of the fengycin operon in *Bacillus subtilis* 168. Microbiology 143(11):3443–3450
69. Walsh CT (2004) Polyketide and nonribosomal peptide antibiotics: modularity and versatility. Science 303(5665):1805
70. Weilharter A, Mitter B, Shin MV, Chain PS, Nowak J, Sessitsch A (2011) Complete genome sequence of the plant growth-promoting endophyte *Burkholderia phytofirmans* strain PsJN. J Bacteriol 193(13):3383–3384
71. Werhane H, Lopez P, Mendel M, Zimmer M, Ordal G, Márquez-Magaña L (2004) The last gene of the *fla/che* operon in *Bacillus*

subtilis, *ylxL*, is required for maximal σ^D function. J Bacteriol 186(12):4025–4029

72. Witte G, Hartung S, Buttner K, Hopfner KP (2008) Structural biochemistry of a bacterial checkpoint protein reveals diadenylate cyclase activity regulated by DNA recombination intermediates. Mol Cell 30(2):167–178
73. Wolff S, Antelmann H, Albrecht D, Becher D, Bernhardt J, Bron S, Büttner K, van Dijl JM, Eymann C, Otto A (2007) Towards the entire proteome of the model bacterium *Bacillus subtilis* by gel-based and gel-free approaches. J Chroma Togr B 849(1–2):129–140
74. Wray L, Atkinson MR, Fisher SH (1994) The nitrogen-regulated *Bacillus subtilis* nrgAB operon encodes a membrane protein and a protein highly similar to the *Escherichia coli* glnB-encoded PII protein. J Bacteriol 176(1):108–114
75. Xiong G, Zhao G, Fan C, He Y (2009) Identification and fungistatic effect of a biocontrol strain. J Yunnan Agr U 24:190–194