**Standards in
Genomic Sciences**

EXTENDED GENOME REPORT   **Open Access**

Draft genome sequence of *Bacillus amyloliquefaciens* subsp. *plantarum* strain Fito_F321, an endophyte microorganism from *Vitis vinifera* with biocontrol potential

Cátia Pinto[1,2], Susana Sousa[1], Hugo Froufe[1], Conceição Egas[1,3], Christophe Clément[2], Florence Fontaine[2] and Ana C Gomes[1,3*]

Abstract

Bacillus amyloliquefaciens subsp. *plantarum* strain Fito_F321 is a naturally occurring strain in vineyard, with the ability to colonise grapevine and which unveils a naturally antagonistic potential against phytopathogens of grapevine, including those responsible for the *Botryosphaeria* dieback, a GTD disease. Herein we report the draft genome sequence of *B. amyloliquefaciens* subsp. *plantarum* Fito_F321, isolated from the leaf of *Vitis vinifera* cv. Merlot at Bairrada appellation (Cantanhede, Portugal). The genome size is 3,856,229 bp, with a GC content of 46.54% that contains 3697 protein-coding genes, 86 tRNA coding genes and 5 rRNA genes. The draft genome of strain Fito_F321 allowed to predict a set of bioactive compounds as bacillaene, difficidin, macrolactin, surfactin and fengycin that due to their antimicrobial activity are hypothesized to be of utmost importance for biocontrol of grapevine diseases.

Keywords: Genome sequencing, *Bacillus amyloliquefaciens* subsp. *plantarum*, Fito_F321 strain, Grapevine-associated microorganism, Biocontrol, Endophytic microorganism

Introduction

Bacillus amyloliquefaciens is a species from the genus *Bacillus*, genetically and phenotypically related to *B. subtilis*, *B. vallismortis*, *B. mojavensis*, *B. atrophaeus*, *B. methylotrophicus*, *B. siamensis*, *B. velezensis*, *B. licheniformis*, and *B. pumilus*, which altogether form the *B. subtilis* group [1–9]. Taxonomic problems involving the species *B. velezensis*, *B. amyloliquefaciens* subsp. *plantarum*, *B. methylotrophicus* and *B. oryzicola* had been recently reported [10]. In order to avoid this taxonomic misunderstanding, a more recent study proposed *B. amyloliquefaciens* subsp. *plantarum* as a later heterotypic synonym of *B. velezensis*, based on phylogenomic analysis [10]. Another study also reinforced that *B. amyloliquefaciens*, *B. velezensis* and *B. siamensis* should be kept as singular species across their clade however, and due to their close relationship, these species should be included in the "operational group *B. amyloliquefaciens*" within the *B. subtilis* group [11].

B. amyloliquefaciens is ubiquitously distributed, Gram-positive, rod-shaped, aerobic and endospore-forming bacteria. Together with other different *Bacillus* species from the *Bacillus subtilis* group, *B. amyloliquefaciens* has been reported to develop beneficial relationships with plants by promoting growth, improving resistance to environmental stress or having important biological activities for plant diseases control [12–14]. These species produce a variety of antimicrobial compounds, such as bacteriocins, antifungal compounds such as lipopeptides, namely iturins and fengycins, and siderophores [15, 16]. Given its biocontrol potential, aligned with its physiological characteristics, such as UV light and heat resistant spores, long shelf life [17] and their advantageous characteristics for formulation, this microorganism is an environmental-friendly alternative to agrochemicals. Indeed, some of *B. amyloliquefaciens*

* Correspondence: acgomes@cnc.uc.pt
[1]Biocant - Biotechnology Innovation Center, Cantanhede, Portugal
[3]Center for Neurosciences and Cell Biology (CNC), Faculty of Medicine, University of Coimbra, Polo I, 1st floor, Rua Larga, 3004-504 Coimbra, Portugal
Full list of author information is available at the end of the article

strains are commercially available as biological control agents or generic plant growth promoters [18, 19].

Altogether these characteristics prompted us to explore the *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321, a naturally occurring strain in vineyards that we have isolated from grapevine leaves in the Bairrada appellation - Portugal. In this study, we have obtained the draft genome sequence of *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321, analysed it and compared it with known genome sequences of representative related species, to gain knowledge on the genes involved in plant interaction with grapevine, as well as the genes conferring antimicrobial activity, and thus to evaluate the potential of this strain for further viticulture and agronomic applications.

Organism information
Classification and features
Strain Fito_F321 was isolated from *Vitis vinifera* cv. Merlot at Bairrada appellation – Cantanhede, Portugal during the 2012 vine cycle. The samples collection was authorized by the private owner, who is fully acknowledged in this paper, and no specific permissions were required for this activity. Briefly, leaf tissues were homogenised in a sterile saline solution (0.85% NaCl) with a sterile pestle. The bacterial isolates were then obtained after plating the homogenised leaves on PDA medium and incubation for 24 h at 28 °C. Sub-cultures were then carried out on the same culture medium until obtaining pure colonies that were further assigned to an isolation code. Microscopic analysis showed that strain Fito_F321 is a Gram-positive and rod shape microorganism (Fig. 1). The classification and general features of strain Fito_F321 are listed in Table 1.

Strain Fito_F321 was taxonomically identified by combining the analysis of the 16S rRNA gene sequence using
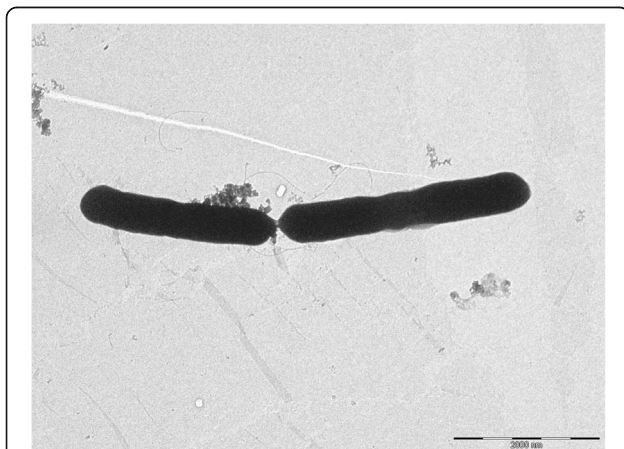

Fig. 1 Transmission electron micrograph of Bacillus amyloliquefaciens subsp. plantarum strain Fito_F321. Bar: 2 μm

Table 1 Classification and general features of Bacillus amyloliquefaciens subsp. *plantarum* strain Fito_F321, according to the MIGS recommendations [69]

MIGS ID	Property	Term	Evidence code[a]
	Classification	Domain *Bacteria*	TAS [70]
		Phylum *Firmicutes*	TAS [71–73]
		Class *Bacilli*	TAS [74, 75]
		Order *Bacillales*	TAS [72, 76]
		Family *Bacillaceae*	TAS [72, 77]
		Genus *Bacillus*	TAS [72, 78]
		Species *Bacillus amyloliquefaciens*	TAS [1, 79]
		Strain: Fito_F321	
	Gram stain	Gram-positive	IDA
	Cell shape	Rod-shaped	IDA
	Motility	Motile	NAS
	Sporulation	Spore-forming	NAS
	Temperature range	unreported	
	Optimum temperature	28 °C	IDA
	pH range; Optimum	6–9, 6.5	IDA
	Carbon source	Organic carbon compounds	NAS
MIGS-6	Habitat	Leaf, grapevine	IDA
MIGS-6.3	Salinity	0–6% (*w/v*); salt tolerant	IDA
MIGS-22	Oxygen requirement	Aerobic	NAS
MIGS-15	Biotic relationship	free-living	IDA
MIGS-14	Pathogenicity	Non-pathogen	NAS
MIGS-4	Geographic location	Cantanhede, Portugal	IDA
MIGS-5	Sample collection	2012	IDA
MIGS-4.1	Latitude	40°19′40.11″N	
MIGS-4.2	Longitude	8°32′59.54″O	
MIGS-4.4	Altitude	90 m	

[a]Evidence codes – IDA: Inferred from Direct Assay; TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [80]

both SILVA database [20] and EzBioCloud [21], and by genome comparisons. In SILVA the 16S rRNA sequence of strain Fito_F321 showed 99% of similarity to *B. amyloliquefaciens* subsp. *plantarum* AS43.3 (CP003838) and to *Bacillus amyloliquefaciens* subsp. *plantarum* SQR9 (CP006890), both non-type strains. A last updated available, reclassified these two strains as *Bacillus velezensis*. In the other hand, results obtained from EzBioCloud showed a 99.93% similarity of strain Fito_F321 to *B. velezensis* CR-502 (type strain). Given these results, the 16S rRNA gene sequence of strain Fito_F321 and

other representative related and type strains species available on GenBank [22] were then selected for phylogenetic analysis (Fig. 2). The phylogenies were generated using the Neighbor-Joining method [23] and evolutionary distances were computed by the Maximum Composite Likelihood method [24] with 1000 bootstrap replicates. Phylogenetic analysis was conducted in MEGA 7.0 [25]. Phylogenetic analysis of the 16S rRNA revealed that strain Fito_F321 is positioned in the same group as *B. amyloliquefaciens* subsp. *plantarum* FZB42, *B. siamensis* PD-A10 and *B. methylotrophicus* CBMB205 and is closest to *B. amyloliquefaciens* subsp. *amyloliquefaciens* DSM7 and *B. velezensis* CR-502. To overcome the difficulties of strain Fito_F321 classification, a comparison of genome sequences between type and other strains of both *B. amyloliquefaciens* and *B. velezensis* species was performed according to the methodology proposed by Chun et al. [26] and is fully presented in the section Comparisons with other genomes. Overall, our results showed that strain Fito_F321 is closer to *B. amyloliquefaciens* subs. *plantarum* FZB42, with a DDH estimate of 85.90% (> 70%) and an ANI similarity of 98.40% (≥95–96%), than to *B. amyloliquefaciens* subsp. *amyloliquefaciens* DSM7 (DDH estimate of 55.30% and ANI similarity of 94.15%). Thus, and according to this data, strain Fito_F321 was classified as a *B. amyloliquefaciens* subsp. *plantarum*.

Extended feature descriptions

The physiological and biochemical features of *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321

Fig. 2 Phylogenetic tree highlighting the taxonomic relation of *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321(●) based on 16S rDNA amplicon within the Bacillus clade. Only type strains are included. The GenBank accession numbers are shown in parentheses. Sequences were aligned using ClustalW 1.6 [81]. The phylogenetic tree was constructed by using the Neighbor-Joining method [23] and evolutionary distances were computed by the Maximum Composite Likelihood method [24] within MEGA 7.0 [25]. There were a total of 1380 positions in the final dataset. Numbers at the nodes are bootstrap values calculated from 1000 replicates; only bootstrap values > 70 are indicated

were analysed to explore the mechanisms behind its antagonistic potential, namely its ability to produce hydrolytic enzymes, presence of siderophores and phosphate solubilization. The tolerance to pH and salinity conditions were also tested. All tests were performed in triplicate. Given the enzymatic production, the amylase, cellulase, lipase, pectinase, protease and urease activity were screened under in vitro conditions by using specific culture media. Results were expressed as positive activity, when a clear halo around strain colony was observed, and the enzymatic index (EI) was calculated through the ration between the average diameter of the degradation halo (clear zone) and the average diameter of the colony growth. The strain Fito_F321 was able to produce all enzymes under in vitro conditions except ureases. Amongst them, cellulases had the higher enzymatic index (10.50 ± 0.20), followed by pectinases (5.44 ± 0.39). This strain was also able to produce siderophores and to solubilise phosphate under in vitro conditions. Overall, these phenotypic features are of high interest, since they are intimately involved in the biocontrol action. Further, this strain was able to grow between pH 6.0–9.0, with an optimal growth at pH 6.5, and grew under up to 6% NaCl. Interestingly, the morphology of Fito_F321 colonies was altered with salt concentration, and colonies became smaller with increasing NaCl concentration in the culture media. It is recognised that excess of soil salinity reduces both plant growth and yield thus, salt tolerant strains may confer plant tolerance against these abiotic stresses [27].

Genome sequencing information
Genome project history

B. amyloliquefaciens subsp. *plantarum* strain Fito_F321 was selected for sequencing as a part of an ongoing project that focuses on the deep characterization of the grapevine-associated microorganisms and their natural antagonistic potential. Thus, its specific antagonistic activity against important grapevine pathogens, such as grey mould or grapevine trunk diseases, together with its physiological and biochemical unique features such as the ability to growth on a range of pHs and salinity conditions, the production of siderophores, the phosphate solubilisation and the high enzymatic activity, were the drivers for its sequencing.

Sequencing of the wild-type *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321 genome was performed at Biocant, Portugal and the draft genome sequencing project has been deposited at DDBJ/ENA/GenBank under the Bioproject PRJNA360208, Biosample ID SAMN06205151 and the accession number MSYT00000000. The version described in this paper is version MSYT01000000. A summary of the project is shown in Table 2.
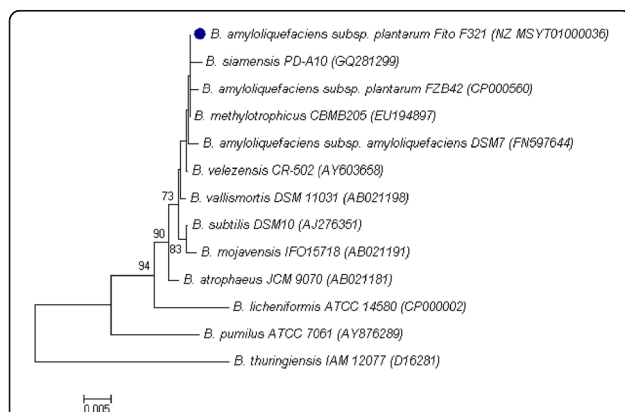
Table 2 Project information

MIGS ID	Property	Term
MIGS 31	Finishing quality	Draft-genome
MIGS-28	Libraries used	Rapid Library Preparation Method GS FLX+ Series XL+
MIGS 29	Sequencing platforms	GS FLX Titanium Sequencing Kit XL+
MIGS 31.2	Fold coverage	41X
MIGS 30	Assemblers	GS Assembler, version 2.9
MIGS 32	Gene calling method	Prodigal, GenePRIMP
	Locus Tag	BVY13
	Genbank ID	MSYT00000000
	Genbank Date of Release	05/01/2018
	GOLD ID	–
	BIOPROJECT	SAMN06205151
MIGS 13	Source Material Identifier	Fito_F321
	Project relevance	Biocontrol, Grapevine, GTD

Growth conditions and genomic DNA preparation

B. amyloliquefaciens subsp. *plantarum* strain Fito_F321 was grown in Luria-Agar medium at 28 °C for 24 h. The genomic DNA was extracted by using the Wizard Genomic DNA Purification kit (Promega, Madison, USA), following the standard protocol for Gram- positive bacteria. The DNA integrity was checked by 0.8% agarose gel electrophoresis, the concentration was determined by using Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher Scientific) and quality assessed with NanoDrop spectrophotometer (Thermo Scientific, USA). Prior to genome sequencing, the quality of the isolated DNA and the molecular identity was confirmed by the sequencing of the 16S rRNA gene.

Genome sequencing and assembly

A DNA library was built from 1 mg of high-quality genomic DNA. Briefly, genomic DNA was fragmented by nebulization and the sequencing adaptors ligated to create double stranded DNA libraries. After quality assessment by using high sensitivity DNA analysis kit (Agilent Technologies) and library titration with KAPA library quantification kit (Kapa Biosystems), the final genome fragments were pyrosequenced in the GS FLX+ system (Roche, 454 Life Sciences), using GS FLX Titanium Sequencing Kit XL+ at Biocant (Cantanhede, Portugal). The sequencing reads were assembled with the GS Assembler, version 2.9 (Roche, 454 Life Sciences) using the default parameters. The sequencing produced 285,879 reads with an average length of 580 bases. The final assembly yielded – 54 contigs, a genome coverage of 41-fold and generated a genome of 3.86 Mb (Table 2).

Genome annotation

The structural and functional annotations were performed using the Prokaryotic Genome Prediction pipeline [28]. Prediction of non-coding RNA genes and miscellaneous features were performed with the PGP pipeline by using tRNAscan-SE [29], RNAMMer [30] and PILERCR [31]. Coding sequences were predicted with Prodigal [32] and automatically corrected by PGP pipeline based on the GenePRIMP algorithm [33]. Functional annotation of protein coding genes was carried out under Prokaryotic Genome Prediction pipeline in InterProScan [34] against Pfam database [35], TIGRFAM [36], Hamap [37], PIRSF [38], PRINTS [39], SMART [40], SUPERFAMILY [41], ProSite [42] databases and RPS-BLAST against Clusters of Orthologous Groups (COG) database [43]. The product name of the identified coding sequences (CDSs) was assigned by using Pfam database, TIGRFAM and COG annotation [44]. The CDSs that were not assigned to a specific product with these databases were named as hypothetical proteins.

Genome properties

The genome statistics are provided in Table 3, and genome visualisation was performed on Artemis version 16.0.0 [45]. The draft genome sequencing of *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321 was distributed across 54 contigs with an estimated genome size of 3,856,229 bp and an average of GC content of 46.53%. The genome analysis showed that Fito_F321 strain' genome contained 3657 protein coding genes predicted, 95 RNAs and without any CRISP elements. The predicted protein encoding genes showed a total length of 3,424,790 bp which represents 88.81% of the total

Table 3 Genome statistics

Attribute	Value	% of Total[a]
Genome size (bp)	3,856,229	100
DNA coding (bp)	3,424,790	88.81
DNA G + C (bp)	1,794,204	46.53
DNA scaffolds	54	–
Total genes	3846	100
Protein coding genes	3657	95.09
RNA genes	95	2.47
Pseudo genes	94	2.44
Genes in internal clusters	NA	–
Genes with function prediction	2790	72.54
Genes assigned to COGs	2697	70.12
Genes with Pfam domains	3241	84.27
Genes with signal peptides	2,48	6.45
Genes with transmembrane helices	2500	65.00
CRISPR repeats	0	0.00

[a]The total is based on either the size of genome in base pairs or the total number of genes in the predicted genome

genome size. Of these, 2697 proteins were assigned to a COG functional category across 20 categories (Table 4). The majority of protein-coding genes were assigned as *function unknown* (264 proteins) and *general function prediction only* (306 proteins), which all together represents 15.59% of the protein encoding genes (Table 4). The proteins not assigned in COGs (960 proteins) represent 26.25% and the *amino acid transport* (269 proteins), *transcription* (227 proteins) and *carbohydrate transport and metabolism* (191 proteins) were the followed categories with 7.36%, 6.21% and 5.22%, respectively. Interestingly, the *defense mechanisms* included 43 protein-coding genes, which represent about 1% of the annotated genome, and included β-lactamase (class C), multi-drug efflux pumps as ATP-binding cassette (ABC) transport and the multidrug and toxic compound extrusion (matE), antimicrobial peptides (AMPs) and lanthionine synthetase component C-like protein (LANCL).

Insights from the genome sequence

A total of 111 metabolic pathways were identified using the KEGG annotation and included, several metabolism pathways (as alanine, aspartate and glutamate, fructose, mannose, galactose, glutathione, methane, nitrogen, pyruvate, sulphur, tryptophan or starch and sucrose), glycolysis, TCA cycle, fatty acid biosynthesis, glucosinolate biosynthesis, antibiotic biosynthesis (neomycin, kanamycin, gentamicin, puromycin, streptomycin or tetracycline) or degradation pathways of noxious compounds (atrazine, benzoate, bisphenol, dioxin, ethylbenzene, limonene, pinene, naphthalene, polycyclic aromatic hydrocarbon or toluene). In general, and as previously described, the metabolic pathways identified showed that the majority of protein-coding genes are involved in the amino acid metabolism (7.36%), carbohydrate metabolism (5.22%), energy metabolism (4.27%) and lipid metabolism (3.20%).

Plant-bacteria interactions

The genome of *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321 was also analysed for genes contributing directly or indirectly for plant-growth promotion (PGP) and biocontrol activities (Additional file 1: Table S1):

Colonisation, adhesion, and movement of bacteria across plant root

It is recognized that a crucial feature of a successful plant growth promoter microorganism, as well as of a

Table 4 Number of genes associated with general COG functional categories

Code	Value	%age[a]	Description
J	158	4.32	Translation, ribosomal structure and biogenesis
A	0	0.00	RNA processing and modification
K	227	6.21	Transcription
L	97	2.65	Replication, recombination and repair
B	1	0.03	Chromatin structure and dynamics
D	34	0.93	Cell cycle control, Cell division, chromosome partitioning
V	43	1.18	Defense mechanisms
T	105	2.87	Signal transduction mechanisms
M	136	3.72	Cell wall/membrane biogenesis
N	41	1.12	Cell motility
U	40	1.09	Intracellular trafficking and secretion
O	78	2.13	Posttranslational modification, protein turnover, chaperones
C	156	4.27	Energy production and conversion
G	191	5.22	Carbohydrate transport and metabolism
E	269	7.36	Amino acid transport and metabolism
F	78	2.13	Nucleotide transport and metabolism
H	122	3.34	Coenzyme transport and metabolism
I	117	3.20	Lipid transport and metabolism
P	149	4.07	Inorganic ion transport and metabolism
Q	85	2.32	Secondary metabolites biosynthesis, transport and catabolism
R	306	8.37	General function prediction only
S	264	7.22	Function unknown
–	960	26.25	Not in COGs

[a]The total is based on the total number of protein coding genes in the genome

biocontrol agent relies on its competence for plant colonisation, notably at roots level [46]. Overall a colonisation process may involve a plant surface adhesion/attachment and a bacterial biofilm formation [47]. The *B. amyloliquefaciens subsp. plantarum* strain Fito_F321 genome encodes a set of proteins involved in flagella biosynthesis, such as *fliZ* (BVY13_00370), *flgC* (BVY13_14075), *flhF* (BVY13_00340), *flhA* (BVY13_00345), *flhB* (BVY13_00350), *fliR* (BVY13_00355), *fliQ* (BVY13_00360), *fliP* (BVY13_00365) or chemotaxis, namely *cheA* (BVY13_00325), *cheD* (BVY13_00310), *cheV* (BVY13_01575) and *che*W (BVY13_00320). *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321 also displays a swarming motility, which allows a rapid surface colonization [48]. Herein, genes encoding for *swrA* (BVY13_02415), *swrB* (BVY13_00300) and *swrC* (BVY13_18860) were predicted. Overall, the swarming motility requires both flagella biosynthesis and surfactant production [48]. Other genes such as hook-associated proteins - *flgK* (BVY13_02315), *fliD* (BVY13_02345), or *hag flagellin* (BVY13_02340) can be expressed in response to root exudates secreted by plant roots. Bacterial flagellins can interact with host and are involved in elicitation of general plant immune response [49]. Furthermore, this strain may also produce biofilms. Indeed, the sporulation transcription factor *spo0A* (BVY13_05345) was here identified, and has an important role on biofilm formation, by repressing the expression of AbrB [50, 51]. *Spo0A* is essential for surface-adhered cells prior to transition to a three-dimensional biofilm structure [50, 51].

Plant-growth promotion

B. amyloliquefaciens subsp. *plantarum* strain Fito_F321 encodes proteins that enhance the plant growth such as those involved in the biosynthesis of indole-3-acetic acid, a plant auxin. Herein, genes encoding for tryptophan, the main precursor of IAA [52], were identified and include *trp* genes such as *trpA* (BVY13_06245), *trpB* (BVY13_06240) or *trpE* (BVY13_06220). Going forward, the synthesis of volatile compounds, as 2,3-butanediol and acetoin, released by some *Bacillus* strains, may also enhance the plant growth promotion and be involved in the eliciting induce systemic resistance [53]. Herein, a set of genes that catalyse the 2,3-butanediol pathway, such as butanediol dehydrogenase *bdhA* gene (BVY13_17360), acetolactate synthase *als* and *alsD* (BVY13_14285, BVY13_09195) and acetolactate decarboxylase *alsD* (BVY13_14290) were identified. Regarding nitrogen fixation, several *nif* genes were not identified among Fito_F321 genome, though other genes involved in nitrate reduction pathways were predicted. Further, a scaffold protein *nifU* (BVY13_03720) and a cysteine desulfurase *nifS*, which are involved in the Fe-S cluster assembly and required for the activation of nitrogenase, where identified. Another feature of *B. amyloliquefaciens*

subsp. *plantarum* strain Fito_F321 is the *nirK* gene responsible for the nitric oxide synthase, a signalling molecule that protects Gram-positive strains from antibiotics and oxidative stress [54, 55]. Regarding the phosphate solubilisation, no *pqq* genes were predicted for this bacterial strain. These genes encode a pyrroloquinoline quinine, a PGP agent involved in the phosphate solubilisation process [56]. However, Fito_F321 strain displayed a phytase activity (BVY13_15080) that contributes to the subsequent use of phosphorous by the plant. This activity is important for the plant growth under phosphate limitation [57, 58]. These predictions are in agreement with the in vitro results obtained by using the Pikovskaya culture medium, which unveiled the ability of Fito_F321 strain to solubilise phosphate. Another important feature is that this strain encodes an inositol 2-dehydrogenase (BVY13_11530), important for the inositol catabolism. Inositol or other inositol derivatives are end-products of phytate degradation, abundant in the plant rhizosphere and can be use by microorganisms as carbon sources [57, 59].

An indirect PGP effect can also be mediated through the siderophores production. Siderophores are iron (Fe)-specific chelating small molecules secreted by bacteria and have high affinity with ferric ionic from soils and surrounding environments [60], thus increasing the bio-availability of Fe for plants, by promoting its solubilisation. On other hand, the siderophores production by BCAs may also confer a clear competition for the available carbon sources, allowing for plant colonisation in detriment with other microorganisms. The *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321 encodes genes for ABC transporters for iron and iron uptake, which was further supported by the genome analysis using antiSMASH 3.0 [61] that also predicted siderophores.

Biocontrol activity

B. amyloliquefaciens subsp. *plantarum* strain Fito_F321 revealed high potential to produce bioactive secondary metabolites (2.32%) with important biocontrol activities. In agreement with the genome analysis using antiSMASH 3.0 [61], 13 secondary metabolites gene clusters were identified (Additional file 2: Table S2). Amongst them, *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321 encoded 4 polyketide synthases clusters, 4 nonribosomal peptides synthases clusters and 1 hybrid PKS-NRPS clusters. Thus, 3 types of antibacterial polyene PKs can be produced, comprising bacillaene, difficidin, macrolactin and butirosin; 2 types of lipopeptides as fengycin, bacilysin, surfactin; as well as siderophore bacillibactin. In addition, the remaining 4 clusters were predicted to produce secondary metabolites including ladderane, lantipeptide or terpene cyclase, namely a putative squalene-hopene cyclase (Additional file 2: Table S2).

Table 5 Comparative analysis of the genome features of *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321 with both *B. amyloliquefaciens* and *B. velezensis*. Details for each genome was completed according to the information available at NCBI and EzBioCloud

	Strain	GB acession number	Isolation source	Country	Genome size (Mb)	G+C content (%)	Protein-coding sequences	tRNA coding genes	rRNA
	B. amyloliquefaciens subsp. *plantarum* strain Fito_F321	MSYT00000000	Leaves (vineyard)	Portugal	3.86	46.54%	3.697	86	5
Type-strains	*B. amyloliquefaciens* subsp. *plantarum* FZB42	CP000560	Soil (sugar beet)	Germany	3.92	46.50%	3.687	89	31
	B. amyloliquefaciens subsp. *amyloliquefaciens* DSM 7	FN597644	Not available/unknown	Germany	3.98	46.10%	3.870	94	30
	Bacillus velezensis KCTC 13012	LHCC00000000.1	River velez	Spain	4.04	46.30%	3.806	80	9
	Bacillus velezensis NRRL B-41580	LLZC00000000.1	River Velez	Spain	4.03	46.30%	3.790	80	9
Non type-strains	*B. amyloliquefaciens* subsp. *plantarum* SQR9	CP006890	Cucumber rhizosphere	China	4.12	46.10%	3.902	72	21
	Bacillus amyloliquefaciens WS-8	CP018200.1	Soil	China	3.93	46.50%	3.670	86	27
	Bacillus amyloliquefaciens CC178	CP006845.1	Cucumber phyllosphere	South Korea	3.92	46.50%	3.702	86	27
	Bacillus amyloliquefaciens KHG19	CP007242.1	Fermented soybean paste	South Korea	3.95	46.60%	3.698	89	31
	Bacillus amyloliquefaciens Y2	CP003332.1	Wheat rhizosphere	China	4.24	45.90%	4.038	87	31
	B. amyloliquefaciens subsp. *plantarum* AS43.3	CP003838	Surface of a wheat spike	USA	3.96	46.60%	3.669	89	31
	Bacillus amyloliquefaciens UMAF6614	CP006960.1	Not available/unknown	Not available/unknown	4.01	46.50%	3.754	83	27
	Bacillus amyloliquefaciens B15	CP014783.1	Grape skin	China	4.01	46.50%	3.759	90	31
	Bacillus amyloliquefaciens UMAF6639	CP006058.1	Not available/unknown	Not available/unknown	4.03	46.30%	3.741	83	27
	Bacillus amyloliquefaciens S499	CP014700.1	Soil	Democratic Republic of the Congo	3.93	46.60%	3.720	81	24
	Bacillus amyloliquefaciens IT-45	CP004065.1	Not available/unknown	Not available/unknown	3.93	46.60%	3.726	95	30
	Bacillus amyloliquefaciens LFB112	CP006952.1	Chinese herbs	China	3.94	46.70%	3.684	94	32
	Bacillus amyloliquefaciens Y14	CP017953.1	Rhizosphere of peanut	China	3.96	46.40%	3.741	87	27
	Bacillus amyloliquefaciens LM2303	CP018152.1	Wild yak dung	China	3.99	46.70%	3.771	86	27
	Bacillus amyloliquefaciens L-S60	CP011278.1	Soil	China	3.90	46.70%	3.662	91	28
	Bacillus amyloliquefaciens L-H15	CP010556.1	Cucumber seedlings	China	3.91	46.70%	3.666	84	28
	Bacillus amyloliquefaciens MBE1283	CP013727.1	Korean traditional alcoholic beverage	South Korea	3.97	46.50%	3.725	86	27
	Bacillus amyloliquefaciens RD7-7	CP016913.1	Fermented soybean paste	South Korea	3.69	46.30%	3.501	87	27
	Bacillus amyloliquefaciens MT45	CP011252.1	Daqu	China	3.90	46.10%	3.752	81	24
	Bacillus amyloliquefaciens SRCM101267	CP021505.1	Food	South Korea	4.07	45.90%	4.014	87	27

Table 5 Comparative analysis of the genome features of *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321 with both *B. amyloliquefaciens* and *B. velezensis*. Details for each genome was completed according to the information available at NCBI and EzBioCloud (*Continued*)

Strain	GB acession number	Isolation source	Country	Genome size (Mb)	G+C content (%)	Protein-coding sequences	tRNA coding genes	rRNA
B. amyloliquefaciens subsp. *plantarum* strain Fito_F321	MSYT00000000	Leaves (vineyard)	Portugal	3.86	46.54%	3.697	86	5
Bacillus amyloliquefaciens LL3	CP002634.1	Fermented food	South Korea	4.00	45.69%	3.935	72	22
Bacillus amyloliquefaciens TA208	CP002627.1	Soil	China	3.94	45.80%	3.891	70	19
Bacillus amyloliquefaciens XH7	CP002927.1	Not available/unknown	Not available/unknown	3.94	45.80%	3.889	75	22

Antimicrobial resistance

In the meantime, the strain Fito_F321 encodes antimicrobial resistance genes (Additional file 1: Table S1) such as bacitracin (bcr - BVY13_11500), fosfomycin (*fosB* - BVY13_12675) and tetracycline (BVY13_08560) [62]. Regarding bacitracin and fosfomycin resistance genes, these are antibiotics that interfere with peptidoglycan synthesis of the bacterial cell wall [63, 64]. Given the bacitracin, herein multiple genes encoding for ABC transporter system were identified, which are associated with bacitracin resistance. Tetracycline antibiotics inhibit the bacterial ribosome, and thus, protein synthesis [65]. In Fito_F321 strain genome, the resistance to tetracycline occurs via active efflux (BVY13_08560).

Comparisons with other genomes

To further characterize the extent of which *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321 differentiates from other strains, genome comparisons of strain Fito_F321 were carried out with the genomes of four types trains, namely *B. amyloliquefaciens* subsp. *plantarum* FZB42, *B. amyloliquefaciens* subsp. *amyloliquefaciens* DSM7, *B. velezensis* KCTC 13012 and *B. velezensis* NRRL B-41580, and other 23 complete genomes of non-type strains of *B. amyloliquefaciens,* including related species that show ≥98.7% 16S sequence similarity. For this, both GGDC 2.1 web server [66], using the DSMZ phylogenomics pipeline [67] to estimate the DNA-DNA hybridization, and the JSpecies WS web server [68] to estimate the Average Nucleotide Identity through pairwise comparisons of genomes were applied. The estimate DDH was calculated with the formula two at the GGDC website, which is the recommended for draft genomes and the ANI values were calculated using the MUMmer software (ANIm) as described by Richter and Roselló-Móra (2009) [68]. This analysis allowed for the calculation of the intergenomic distances between genomes and the probability of belonging to the same species. The general comparison is shown in Table 5 and the intergenomic distances, through the DDH estimate and ANI, are in Table 6. Given the analysis with type-strains, results have shown that *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321 had a lower distance with *B. amyloliquefaciens* subs. *Plantarum* FZB42 with a DDH estimate of 85.90% and a probability to correspond to the same species of 94.14%. These results were also supported by the ANI analysis where both strains reached a similarity of 98.40%, with 95.22% of the genome aligned. Contrary, *B. amyloliquefaciens* subsp. *amyloliquefaciens* DSM7 was the strain most distant from strain Fito_F321, with a DDH estimate of 55.30% and a probability to correspond to the same species of 35.90%. The same comparative results were performed for non-type strains. Herein, *B. amyloliquefaciens* subsp. *plantarum* SQR9 showed the lower intergenomic distance and the higher similarity with *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321.

Conclusions

In this study, we have characterized the genome of *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321, a

Table 6 Comparative analysis of the *in-silico* genome distances between *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321 withboth *B. amyloliquefaciens* and *B. velezensis*, through the DNA-DNA hybridization (DDH method) and average nucleotide identities (ANI)

| | Strain | GB acession number | DDH method | | | | | ANIm | | |
			Distance	DDH estimate	Probability that DDH > 70% (same species)	Probability that DDH > 79% (same subspecies)	Difference in % G+C	ANIm (%)	Aligned (%)	Aligned (bp)
Type-strains	*B. amyloliquefaciens* subsp. *plantarum* FZB42	CP000560	0.0167	85.90%	94.14%	58.4%	0.06	98.40%	95.22%	3.671.944
	Bacillus velezensis KCTC 13012	LHCC00000000.1	0.0175	85.10%	93.79%	57.14%	0.21	98.37%	95.60%	3.686.632
	Bacillus velezensis NRRL B-41580	LLZC00000000.1	0.0175	85.10%	93.80%	57.17%	0.21	98.37%	95.59%	3.856.229
	B. amyloliquefaciens subsp. *amyloliquefaciens* DSM 7	FN597644	0.0604	55.3%	35.9%	7.84%	0.45	94.15%	88.6%	3.416.734
Non type-strains	*B. amyloliquefaciens* subsp. *plantarum* SQR9	CP006890	0.0114	90.60%	96.01%	66.28%	0.44	98.91%	95.86%	3.696.425
	Bacillus amyloliquefaciens WS-8	CP018200.1	0.0120	90.10%	95.83%	65.43%	0.04	98.90%	96.65%	3.727.194
	Bacillus amyloliquefaciens CC178	CP006845.1	0.0167	85.90%	94.14%	58.39%	0.07	98.40%	95.3%	3.675.098
	Bacillus amyloliquefaciens KHG19	CP007242.1	0.0168	85.70%	94.09%	58.18%	0.09	98.37%	95.81%	3.694.724
	Bacillus amyloliquefaciens Y2	CP003332.1	0.0172	85.30%	93.90%	57.50%	0.68	98.36%	95.09%	3.666.760
	B. amyloliquefaciens subsp. *plantarum* AS43.3	CP003838	0.0173	85.30%	93.89%	57.46%	0.05	98.37%	96.4%	3.717.580
	Bacillus amyloliquefaciens UMAF6614	CP006960.1	0.0174	85.20%	93.82%	57.24%	0.05	98.38%	95.39%	3.678.467
	Bacillus amyloliquefaciens B15	CP014783.1	0.0180	84.70%	93.55%	56.31%	0.07	98.28%	96.51%	3.721.654
	Bacillus amyloliquefaciens UMAF6639	CP006058.1	0.0181	84.60%	93.50%	56.14%	0.19	98.29%	96.35%	3.715.305
	Bacillus amyloliquefaciens S499	CP014700.1	0.0231	80.20%	90.74%	48.15%	0.09	97.74%	95.91%	3.698.339
	Bacillus amyloliquefaciens IT-45	CP004065.1	0.0231	80.20%	90.75%	48.16%	0.08	97.74%	95.69%	3.690.216
	Bacillus amyloliquefaciens LFB112	CP006952.1	0.0233	80.00%	90.61%	47.82%	0.11	97.74%	95.84%	3.692.925
	Bacillus amyloliquefaciens Y14	CP017953.1	0.0233	80.00%	90.62%	47.85%	0.10	97.74%	95.13%	3.668.289
	Bacillus amyloliquefaciens LM2303	CP018152.1	0.0234	80.00%	90.58%	47.75%	0.14	97.74%	95.65%	3.688.426
	Bacillus amyloliquefaciens L-S60	CP011278.1	0.0237	79.70%	90.38%	47.26%	0.12	97.73%	95.84%	3.696.002
	Bacillus amyloliquefaciens L-H15	CP010556.1	0.0237	79.70%	90.35%	47.20%	0.12	97.72%	98.85%	3.696.084
	Bacillus amyloliquefaciens MBE1283	CP013727.1	0.0238	79.60%	90.31%	47.10%	0.00	97.74%	95.21%	3.671.471
	Bacillus amyloliquefaciens RD7-7	CP016913.1	0.0602	55.40%	36.23%	7.93%	0.28	94.14%	87.88%	3.388.859
	Bacillus amyloliquefaciens MT45	CP011252.1	0.0603	55.30%	36.14%	7.90%	0.45	94.13%	90.95%	3.507.122
	Bacillus amyloliquefaciens SRCM101267	CP021505.1	0.0605	55.20%	35.76%	7.81%	0.60	94.15%	88.80%	3.424.154
	Bacillus amyloliquefaciens LL3	CP002634.1	0.0613	54.80%	34.25%	7.42%	0.83	94.08%	88.75%	3.422.356
	Bacillus amyloliquefaciens TA208	CP002627.1	0.0614	54.80%	34.17%	7.39%	0.71	94.08%	88.93%	3.429.250
	Bacillus amyloliquefaciens XH7	CP002927.1	0.0614	54.80%	34.18%	7.40%	0.72	94.07%	88.94%	3.429.758

The nucleotide sequences were download from GenBank and the respective accession numbers are shown in the table. *In-silico* DNA-DNA hybridization (DDH) was calculated by using the Genome-to-Genome Distance Calculator (GGDC 2.1) [67] and ANI values were computed through pairwise genome comparison by using the MUMmer software [68]. Values with grey colour are of a below cut-off (< 95%)

natural grapevine-associated microorganism, which was isolated from grapevine leaves. Given its genomic and physiological characteristics, this microorganism may provide an interesting model to study the plant-microbial interactions and their role in grapevine protection. The intergenomic distances amongst genomes showed that *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321 is highly close to type strain *B. amyloliquefaciens* subs. *plantarum* FZB42, with a DDH estimate value of 85.90% and a ANIm value of 94.14%, and more distant to the type strain *B. amyloliquefaciens subsp. amyloliquefaciens* DSM7.

The predicted gene compounds of *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321 such as bacillaene, difficidin, macrolactin, surfactin, fengycin and siderophore, together with other protein-coding genes herein presented, are of utmost importance for its biocontrol activities and could explain its positive plant-microbial interactions, as well as its role on the natural protection of vineyard. Thus, these gene clusters suggest that the strain Fito_F321 can produce bioactive compounds of biocontrol value, which represents a source of novel bioactive compounds and that may be essential for the grapevine protection in the pursue of a more sustainable viticulture.

Additional files

Additional file 1: Table S1. General overview of genes involved in bacterium-plant interaction in *B. amyloliquefaciens* subsp. *plantarum* strain Fito_F321. (XLSX 11 kb)

Additional file 2: Table S2. Secondary metabolite gene clusters identified. (XLSX 14 kb)

Abbreviations
BCA: Biological control agent; GH: Glycoside hydrolase; GTD: Grapevine Trunk Diseases; PGP: Plant growth promoting

Authors' contributions
CP, FF and ACG designed research. CP performed the experiments and SS did DNA isolation and purification. CP and HF were evolved on the Bioinformatic analysis. Contributed reagents/materials/analysis tools: SS, CE, CC, FF and ACG. Wrote the paper: CP, FF and ACG. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Biocant - Biotechnology Innovation Center, Cantanhede, Portugal. [2]SFR Condorcet - FR CNRS 3417, University of Reims Champagne-Ardenne, Induced Resistance and Plant Bioprotection (RIBP)- EA 4707, BP1039, Cedex 2 51687 Reims, France. [3]Center for Neurosciences and Cell Biology (CNC), Faculty of Medicine, University of Coimbra, Polo I, 1st floor, Rua Larga, 3004-504 Coimbra, Portugal.

References

1. Priest F, Goodfellow M, Shute L, Berkeley R. *Bacillus amyloliquefaciens* sp. nov., nom. rev. Int J Syst Bacteriol. 1987;37:69–71.
2. Fritze D. Taxonomy of the genus *Bacillus* and related genera: the aerobic endospore-forming bacteria. Phytopathology. 2004;94:1245–8.
3. Madhaiyan M, Poonguzhali S, Kwon SW, Sa TM. *Bacillus methylotrophicus* sp. nov., a methanol-utilizing, plant-growth-promoting bacterium isolated from rice rhizosphere soil. Int J Syst Evol Microbiol. 2010;60:2490–5.
4. Sumpavapol P, Tongyonk L, Tanasupawat S, Chokesajjawatee N, Luxananil P, Visessanguan W. *Bacillus siamensis* sp. nov., isolated from salted crab (poo-khem) in Thailand. Int J Syst Evol Microbiol. 2010;60:2364–70.
5. Roberts MS, Nakamura LK, Cohan FM. *Bacillus vallismortis* sp. nov., a close relative of *Bacillus subtilis*, isolated from soil in death valley, in California. Int J Syst Evol Microbiol. 1996;46:470–5.
6. Roberts MS, Nakamura LK, Frederick MCOW. *Bacillus mojavensis* sp. nov., distinguishable from *Bacillus subtilis* by sexual iolation, divergence in DNA sequence, and differences in fatty acid composition. Int J Syst Bacteriol. 1994;44:256–64.
7. Nakamura LK. Taxonomic relationship of black pigmented *Bacillus subtilis* strains and a proposal for *Bacillus atrophaeus* sp. nov. Int J Syst Bacteriol. 1989;39:295–300.
8. Ruiz-García C, Béjar V, Martínez-Checa F, Llamas I, Quesada E. *Bacillus velezensis* sp. nov., a surfactant-producing bacterium isolated from the river Vélez in Málaga, southern Spain. Int J Syst Evol Microbiol. 2005;55:191–5.
9. Nakamura LK, Roberts MS, Cohan FM. Relationship of *Bacillus subtilis* clades associated with strains 168 and W23: a proposal for *Bacillus subtilis* subsp. *subtilis* subsp. nov. and *Bacillus subtilis* subsp. *spizizenii* subsp. nov. Int J Syst Bacteriol. 1999;49:1211–5.
10. Dunlap CA, Kim SJ, Kwon SW, Rooney AP. *Bacillus velezensis* is not a later heterotypic synonym of *Bacillus amyloliquefaciens*; *Bacillus methylotrophicus*, *Bacillus amyloliquefaciens* subsp. plantarum and "*Bacillus oryzicola*" are later heterotypic synonyms of *Bacillus velezensis* based on phylogenomics. Int J Syst Evol Microbiol. 2016;66:1212–7.
11. Fan B, Blom J, Klenk HP, Borriss R. *Bacillus amyloliquefaciens*, *Bacillus velezensis*, and *Bacillus siamensis* form an "operational group *B. amyloliquefaciens*" within the *B. subtilis* species complex. Front Microbiol. 2017;8:1–15.
12. Schnepf E, Crickmore N, Van Rie J, Lereclus D, Baum J, Feitelson J, et al. *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol. Mol. Biol. Rev. 1998;62:775–806.
13. Boriss R. Use of plant-associated *Bacillus* strains as biofertilizers and biocontrol agents in agriculture. In: Maheshwari D, editor. Bact. Agrobiol. Plant growth responses. Berlin, Heidelberg: Springer Berlin Heidelberg; 2011. p. 21–40.
14. Qiao J-Q, Wu H-J, Huo R, Gao X-W, Borriss R. Stimulation of plant growth and biocontrol by *Bacillus amyloliquefaciens* subsp. *plantarum* FZB42 engineered for improved action. Chem Biol Technol Agric. 2014;1:12.
15. Baruzzi F, Quintieri L, Morea M, Caputo L. Antimicrobial compounds produced by *Bacillus* spp . and applications in food. Formatex. 2011;2:1102–11.
16. Sumi CD, Yang BW, Yeo I-C, Hahm YT. Antimicrobial peptides of the genus *Bacillus*: a new era for antibiotics. Can J Microbiol. 2015;61:93–103.
17. McSpadden Gardener BB. Ecology of *Bacillus* and *Paenibacillus* spp. in agricultural systems. Phytopathology. 2004;94:1252–8.
18. Choi S, Jeong H, Kloepper JW, Ryu C. Genome sequence of *Bacillus amyloliquefaciens* GB03, an active ingredient of the first commercial biological control product. Genome Announc. 2014;2:e01092–14.
19. Chen XH, Koumoutsi A, Scholz R, Eisenreich A, Schneider K, Heinemeyer I, et al. Comparative analysis of the complete genome sequence of the plant

Pinto *et al. Standards in Genomic Sciences* (2018) 13:30

Page 11 of 12

growth–promoting bacterium *Bacillus amyloliquefaciens* FZB42. Nat Biotechnol. 2007;25:1007–14.

20. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013;41:D590–6.

21. Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, et al. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int J Syst Evol Microbiol. 2017;67:1613–7.

22. Borriss R, Chen XH, Rueckert C, Blom J, Becker A, Baumgarth B, et al. Relationship of *Bacillus amyloliquefaciens* clades associated with strains DSM 7 T and FZB42 T: a proposal for *Bacillus amyloliquefaciens* subsp. *amyloliquefaciens* subsp. nov. and *Bacillus amyloliquefaciens* subsp. *plantarum* subsp. nov. based on complete gen. Int J Syst Evol Microbiol. 2011; 61:1786–801.

23. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987;4:406–25.

24. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci U S A. 2004;101: 11030–5.

25. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33:1870–4.

26. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, et al. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int J Syst Evol Microbiol. 2018;68:461–6.

27. Shrivastava P, Kumar R. Soil salinity: a serious environmental issue and plant growth promoting bacteria as one of the tools for its alleviation. Saudi J Biol Sci King Saud University. 2015;22:123–31.

28. Egas C, Barroso C, Froufe HJC, Pacheco J, Albuquerque L, da Costa MS. Complete genome sequence of the radiation-resistant bacterium *Rubrobacter radiotolerans* RSPS-4. Stand Genomic Sci. 2014;9:1062–75.

29. Schattner P, Brooks AN, Lowe TM. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. Nucleic Acids Res. 2005;33:686–9.

30. Lagesen K, Hallin P, Rødland EA, Stærfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 2007;35:3100–8.

31. Edgar RC. PILER-CR: fast and accurate identification of CRISPR repeats. BMC Bioinformatics. 2007;8:1–6.

32. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010;11:119.

33. Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, et al. GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. Nat Methods. 2010;7:455–7.

34. Zdobnov EM, Apweiler R. InterProScan - an integration platform for the signature-recognition methods in InterPro. Bioinformatics. 2001;17:847–8.

35. Finn RD, Tate J, Mistry J, Coggill P, Sammut S. The Pfam protein families databse. Nucleic Acids Res. 2008;36:D281–8.

36. Haft DH, Selengut JD, White O. The TIGRFAMs database of protein families. Nucleic Acids Res. 2003;31:371–3.

37. Pedruzzi I, Rivoire C, Auchincloss AH, Coudert E, Keller G, De Castro E, et al. HAMAP in 2013, new developments in the protein family classification and annotation system. Nucleic Acids Res. 2013;41:584–9.

38. Wu CH, Nikolskaya A, Huang H, Yeh LS, Natale DA, Vinayaka CR, et al. PIRSF: family classification system at the protein information resource. Nucleic Acids Res. 2004;32:D112–4.

39. Attwood TK, Coletta A, Muirhead G, Pavlopoulou A, Philippou PB, Popov I, et al. The PRINTS database: a fine-grained protein sequence annotation and analysis resource-its status in 2012. Database. 2012;2012:1–9.

40. Letunic I, Doerks T, Bork P. SMART 7: recent updates to the protein domain annotation resource. Nucleic Acids Res. 2012;40:302–5.

41. Wilson D, Pethica R, Zhou Y, Talbot C, Vogel C, Madera M, et al. SUPERFAMILY - sophisticated comparative genomics, data mining, visualization and phylogeny. Nucleic Acids Res. 2009;37:380–6.

42. Sigrist CJA, Cerutti L, De Castro E, Langendijk-Genevaux PS, Bulliard V, Bairoch A, et al. PROSITE, a protein domain database for functional characterization and annotation. Nucleic Acids Res. 2010;38:161–6.

43. Tatusoy R, Koonin E, Lipman D. A genomic perspective on protein families. Science. 1997;278:631–7.

44. Mavromatis K, Ivanova NN, Chen I-MA, Szeto E, Markowitz VM, Kyrpides NC. The DOE-JGI standard operating procedure for the annotations of microbial genomes. Stand Genomic Sci. 2009;1:63–7.

45. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. Bioinformatics. 2000;16:944–5.

46. Lugtenberg BJJ, Dekkers L, Bloemberg GV. Molecular determinats of rhizosphere colonization by *Pseudomonas*. Annu. Rev. Phytopathology. 2001;39:461–90.

47. Bogino PC, Oliva Mde L, Sorroche FG, Giordano W. The role of bacterial biofilms and surface components in plant-bacterial associations. Int. J Mol Sci. 2013;14:15838–59.

48. Kearns DB, Losick R. Swarming motility in undomesticated *Bacillus subtilis*. Mol Microbiol. 2003;49:581–90.

49. Abramovitch RB, Anderson JC, Martin GB. Bacterial elicitation and evasion of plant innate immunity. Nat Rev Mol cell Biol. 2006;7:601–11.

50. Hamon MA, Lazazzera BA. The sporulation transcription factor Spo0A is required for biofilm development in *Bacillus subtilis*. Mol Microbiol. 2001;42: 1199–209.

51. Hamon M, Stanley N, Britton R, Grossman A, Lazzazzera B. Identification of AbrB-regulated genes involved in biofilm formation by Bacillus subtilis. Mol Microbiol. 2004;52:847–60.

52. Spaepen S, Vanderleyden J. Auxin and plant-microbe interactions. Cold Spring Harb Perspect Biol. 2011;3:a001438.

53. Ryu C-M, M a F, Hu C-H, Reddy MS, Wei H-X, Paré PW, et al. Bacterial volatiles promote growth in Arabidopsis. Proc Natl Acad Sci U S A. 2003;100:4927–32.

54. Gusarov I, Nudler E. NO-mediated cytoprotection: instant adaptation to oxidative stress in bacteria. Proc Natl Acad Sci U S A. 2005;102:13855–60.

55. Sudhamsu J, Crane BR. Bacterial nitric oxide synthases: what are they good for? Trends Microbiol. 2009;17:212–8.

56. Choi O, Kim J, Kim J-G, Jeong Y, Moon JS, Park CS, et al. Pyrroloquinoline quinone is a plant growth promotion factor produced by Pseudomonas fluorescens B16. Plant Physiol. 2007;146:657–68.

57. Idriss EE, Makarewicz O, Farouk A, Rosner K, Greiner R, Bochow H, et al. Extracellular phytase activity of *Bacillus amyloliquefaciens* FZB45 contributes to its plant-growth-promoting effect. Microbiology. 2002;148:2097–109.

58. Richardson AE, Hadobas PA, Hayes JE, O'hara CP, Simpson RJ. Utilization of phosphorus by pasture plants supplied with myo-inositol hexaphosphate is enhanced by the presence of soil micro-organisms. Plant Soil. 2001;229:47–56.

59. van Straaten KE, Zheng H, Palmer DRJ, Sanders D a R. Structural investigation of myo-inositol dehydrogenase from *Bacillus subtilis*: implications for catalytic mechanism and inositol dehydrogenase subfamily classification. Biochem J. 2010;432:237–47.

60. Alexander DB, Zuberer DA. Use of chrome azurol-S reagents to evaluate siderophore production by rhizosphere bacteria. Biol Fertil Soils. 1991;12:39–45.

61. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, et al. AntiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res. 2015;43:W237–43.

62. Lakin SM, Dean C, Noyes NR, Dettenwanger A, Ross AS, Doster E, et al. MEGARes: an antimicrobial resistance database for high throughput sequencing. Nucleic Acids Res. 2017;45:D574–80.

63. Stone KJ, Strominger JL. Mechanism of action of bacitracin: complexation with metal ion and C55-isoprenyl pyrophosphate. Proc Natl Acad Sci U S A. 1971;68:3223–7.

64. Storm DR. Mechanism of bacitracin action: a specific lipid peptide interaction. Ann N Y Acad Sci. 1974;235:387–98.

65. Chopra I, Roberts M. Tetracycline antibiotics : mode of action , applications , molecular biology, and epidemiology of bacterial resistance. Microbiol Mol Biol Rev. 2001;65:232–60.

66. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics. 2013;14:60.

67. Meier-Kolthoff JP, Klenk HP, Göker M. Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. Int J Syst Evol Microbiol. 2014;64:352–6.

68. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A. 2009;106:19126–31.

69. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, et al. The minimum information about a genome sequence (MIGS) specification. Nat. Biotechnol. 2008;26:541–7.

70. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains archaea, bacteria, and Eucarya. Proc Natl Acad Sci U S A. 1990;87:4576–9.

71. Gibbons NE, Murray RGE. Proposals concerning the higher taxa of bacteria. Int J Syst Bacteriol. 1978;28:1–6.

72. Skerman VBD, McGowan V, Sneath PHA. Approved lists of bacterial names. Int J Syst Bacteriol. 1980;30:225–420.

73. Murray R. The higher taxa, or, a place for everything…? In: Holt JG, editor. Bergey's manual of systematic bacteriology, vol. 1. First ed. Baltimore: The Williams and Wilkins Co.; 1984. p. 31–4.

74. Euzéby J. Validation list no. 132. List of new names and new combinations previously effectively, but not validly, published. Int J Syst Evol Microbiol. 2010;60:469–72.

75. Ludwig W, Schleifer KH, Whitman WB. Class I. Bacilli class nov. In: De Vos P, Garrity G, Jones D, Krieg NR, Ludwig W, Rainey FA, Schleifer KH, Whitman WB, (eds). Bergey's manual of systematic bacteriology, Second edition, Volume 3, New York: Springer-Verlag. 2009, p.19–20.

76. Prévot AR. Dictionaire des bactéries pathogènes. In: Hauderoy P, Ehringer G, Guillot G, Magrou J, Prévot AR, Rosset D, et al., editors. 2nd. Paris: Masson et Cie; 1953. p. 1–692.

77. Fischer A. Untersuchungen über bakterien. Jahrbücher für Wissenschaftliche Botanik. 1895;27:1–163.

78. Cohn F. Untersuchungen über bakterien. Beiträge zur Biologie Pflanzen. 1872;1:127–224.

79. Wang LT, Lee FL, Tai CJ, Kuo HP. *Bacillus velezensis* is a later heterotypic synonym of *Bacillus amyloliquefaciens*. Int J Syst Evol Microbiol. 2008;58:671–5.

80. Ashburner M, Ball C, Blake J, Botstein D, Butler H, Cherry M, et al. Gene ontologie: tool for the unification of biology. Nat Genet. 2000;25:25–9.

81. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22:4673–80.