Current Genetics (2022) 68:449–466
https://doi.org/10.1007/s00294-022-01240-4

ORIGINAL ARTICLE

Genomic insights into the diversity of non-coding RNAs in *Bacillus cereus* sensu lato

Kátia B. Gonçalves[1] · Renan J. Casarotto Appel[1] · Laurival A. Vilas Bôas[1] · Priscilla F. Cardoso[1] · Gislayne T. Vilas Bôas[1]

Received: 19 December 2021 / Revised: 20 March 2022 / Accepted: 30 March 2022 / Published online: 12 May 2022

Abstract

Bacillus cereus sensu lato is a group of bacteria of medical and agricultural importance in different ecological niches and with controversial taxonomic relationships. Studying the composition of non-coding RNAs (ncRNAs) in several bacterial groups has been an important tool for identifying genetic information and better understanding genetic regulation towards environment adaptation. However, to date, no comparative genomics study of ncRNA has been performed in this group. Thus, this study aimed to identify and characterize the set of ncRNAs from 132 strains of *Bacillus cereus*, *Bacillus thuringiensis* and *Bacillus anthracis* to obtain an overview of the diversity and distribution of these genetic elements in these species. We observed that the number of ncRNAs differs in the chromosomes of the three species, but not in the plasmids, when species or phylogenetic clusters were compared. The prevailing functional/structural category was Cis-reg and the most frequent class was Riboswitch. However, in plasmids, the class Group II intron was the most frequent. Also, nine ncRNAs were selected for validation in the strain *B. thuringiensis* 407 by RT-PCR, which allowed to identify the expression of the ncRNAs. The wide distribution and diversity of ncRNAs in the *B. cereus* group, and more intensely in *B. thuringiensis*, may help improve the abilities of these species to adapt to various environmental changes. Further studies should address the expression of these genetic elements in different conditions.

Keywords *Bacillus cereus* (sensu lato) · Diversity and distribution non-coding RNA · Comparative genomics · Bioinformatics

Introduction

Bacillus cereus sensu lato (*sl*) is a group of bacteria of medical and agricultural importance. The species of this bacterial group colonize diverse hosts and have different ecological niches. The three most studied species of the group, *Bacillus cereus* sensu stricto (*Bc*), *Bacillus thuringiensis* (*Bt*), and *Bacillus anthracis* (*Ba*), are known for their pathogenic potential. *B. cereus* is considered an opportunistic pathogen due to food contamination (Ehling-Schulz et al. 2006). *B. thuringiensis* has been widely used in the biological control of insects due to the production of different toxins, mainly Cry proteins (Schnepf et al. 1998). *B. anthracis* is considered a pathogenic species to animals, acting even on humans (Pilo and Frey 2011). In these three species, genes encoding virulence and pathogenicity factors are often located in plasmids, including the toxins and capsule components of *B. anthracis*, Cry proteins in *B. thuringiensis*, and cereulide synthesis in emetic strains of *B. cereus* (Schnepf et al. 1998; Ehling-Schulz et al. 2006; Pilo and Frey 2011; Adang et al. 2014).

Several studies classify *B. cereus s.l.* as a complex phylogenetic group (Tourasse et al. 2011; Okinaka and Keim 2016). Multilocus sequence typing (MLST) data divide the group into seven clusters (I-VII), which show important genetic and host association differences (Guinebretière et al. 2008; Patiño-Navarrete and Sanchis 2017). Strains of *B. thuringiensis* and *B. cereus* are located mainly in clusters III and IV (Guinebretière et al. 2008), while strains of *B. anthracis* are found exclusively in cluster III. Thus, although they may belong to the same phylogenetic group,

Communicated by Michael Polymenis.

✉ Gislayne T. Vilas Bôas
gvboas@uel.br

[1] Depto Biologia Geral, Universidade Estadual de Londrina, Londrina, Brazil

B. anthracis strains form a distant subcluster from commercially used bioinsecticidal *B. thuringiensis* strains (Raymond and Federici 2017). However, due to the complexity of the *B. cereus* group, taxonomic relationships among species remain a subject of debate (Bazinet 2017). Recently, the average nucleotide identity genome species threshold (92.5 ANI) was proposed (Carroll et al. 2020).

Although multiple tools have been used to study taxonomic relationships among *B. cereus* group species (Guinebretière et al. 2008; Van der Auwera et al. 2013; Liu et al. 2015; Patiño-Navarrete and Sanchis 2017), non-coding RNAs (ncRNAs), which can act as regulatory factors for several genetic and ecological traits (Hör et al. 2018), have never been used. These elements are organized in a database of RNA families (Rfam) according to sequence homology (Kalvari et al. 2018). Furthermore, families of ncRNAs are grouped into classes (sRNA, Leader, Riboswitch, etc.) and functional categories (Gene, Cis-reg, and Intron) through functional and structural similarities (Nieselt and Herbig 2013). Another criterion used for the classification of ncRNAs is related to size, being separated into small non-coding RNA (sRNA), which are frequently found in bacteria (Nobukazu et al. 2013), and Long non-coding RNA (lncRNA) (Amin et al. 2019), which are uncommon elements in bacteria (Harris and Breaker 2018).

In general, ncRNAs can play important roles in controlling cellular processes in bacteria, such as quorum sensing and biofilm formation (Lo 2018), and acting in gene regulation, metabolism, and physiological adaptations (Zur Bruegge et al. 2017; Harris and Breaker 2018; Amin et al. 2019). They also exhibit different modes of action, such as acting on gene expression by base-pairing with the target messenger RNA (mRNA), as in the case of sRNAs (Gripenland et al. 2010; Nitzan et al. 2017; Svenningsen, 2018). Recent studies show that sRNAs can also directly affect mRNA stability without affecting translation by recruiting degradation machinery or interfering with their action (Durand et al. 2017). In the *B. cereus* group, studies to date have described the occurrence and functional role of some specific ncRNAs, such as Riboswitches and miRNA, for example (Xu et al. 2015; Tang et al. 2016; Sajid et al. 2018; Nie et al. 2019), but have not analyzed the diversity of these elements.

Therefore, this study describes the first comparative genomic analysis of ncRNAs in the *B. cereus* group aiming to understand the distribution and diversity of these elements in the bacteria of the group, and thus, contributing to their taxonomy/phylogenetic framework discussions. Our results demonstrate that the three main species of the group have a rich profile of ncRNAs, that a great number and diversity of these genetic elements are shared among the three species, and that *B. thuringiensis* has the highest number of elements, which may help improve the abilities of these species to adapt to various environmental changes.

Materials and methods

The pipeline of tools used in this study is described in Fig. 1.

Genome mining

The *B. cereus*, *B. thuringiensis*, and *B. anthracis* strains with the complete genome sequences available in the Gold "Genomes Online Database" (Mukherjee et al. 2019) (https://gold.jgi.doe.gov/) and the NCBI "National *Center for Biotechnology Information*" (Johnson et al. 2008) (https://www.ncbi.nlm.nih.gov/genome) databases up to 09/15/2017 were selected for this study. In cases of duplicate genomes with the same nomenclature, both genomes were kept in the analyses due to variations observed in replicon size and the number of plasmids and ncRNAs analyzed in later steps.

In silico identification and classification of ncRNAs

The complete genome sequences were used for ncRNA search and annotation using the *cmsearch* tool contained in the Infernal program package, "INFERence of RNA ALignment" (Nawrocki et al. 2009) (http://eddylab.org/infernal/) version 1.1.2. Covariance models were extracted from the RNA families database, Rfam (Kalvari et al. 2018) (http://rfam.xfam.org/) version 13.0 (downloaded in 09/2017). In total, 2,686 families and 191 covariance models were used, restricted to families already identified as being in *Bacillales*, following criteria previously established by Nawrocki (2014). The following criteria or quality filters were employed for investigation: Only sequences that achieved an E-value of 0.01 were used; the tRNA, tmRNA, 5.8S rRNA, and 5S rRNA families (with no regulatory role) were removed from the analysis because they were not the focus of the study; and the lowest-scoring overlaps identified by *cmsearch*, with the cmsearch_tblout_deoverlap tool version 0.02 (https://github.com/nawrockie/cmsearch_tblout_deoverlap), were removed.

To visualize the distribution of the identified ncRNAs among the genomes, the jVenn tool (Bardou et al. 2014) (http://jvenn.toulouse.inra.fr/app/index.html) was used and edited with the online tool (https://miro.com/). Regarding their occurrence, ncRNA families were classified into three categories: i) common: present in all three species; ii) shared: present in two species and iii) exclusive: present in a single species.

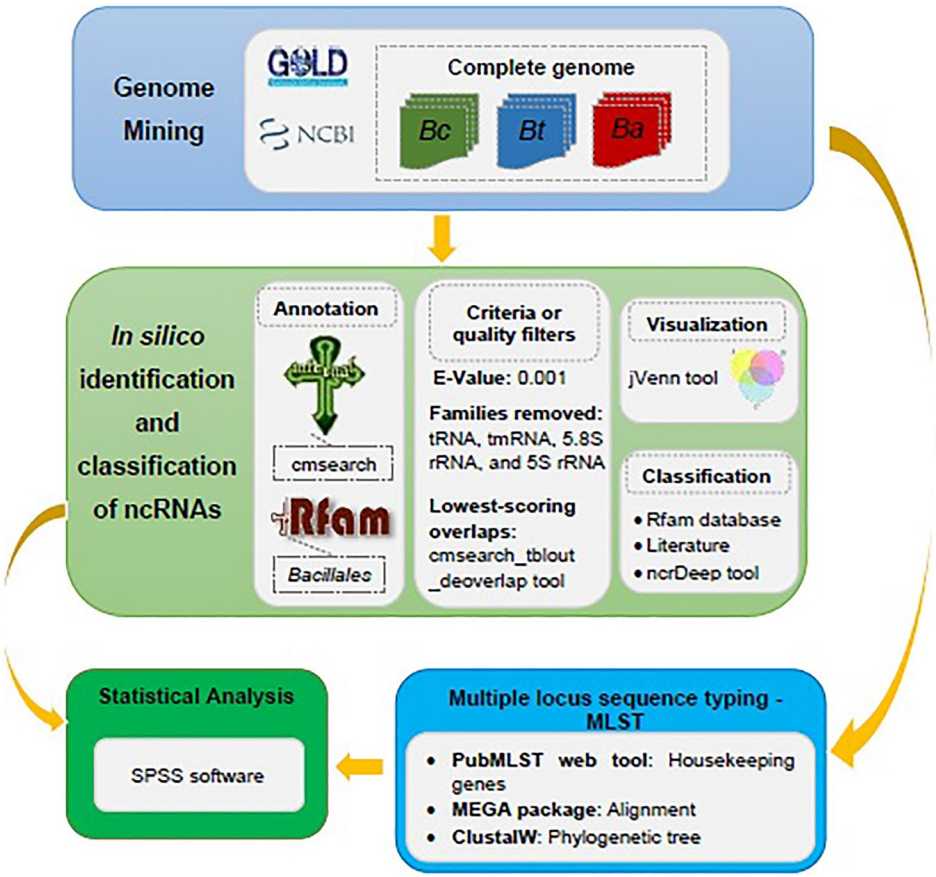

Fig. 1 Pipeline of the tools used in this study

The identified families of ncRNAs were classified into functional/structural categories, classes, and subclasses as described in the Rfam database or literature. The ncrDeep tool was used to identify families whose class information was not described (Chantsalnyam et al. 2020) (http://nsclbio.jbnu.ac.kr/tools/ncRDeep/).

Finally, we separated the ncRNAs according to size, unrelated to the functional/structural categories or classes, thus, we considered as sRNAs the families with up to 500 bp and lncRNAs over 500 bp. For this, we used the Rfam database (Nawrocki et al. 2015).

Statistical analysis

To investigate the relationship between the number of ncRNA families and the size (bp) of chromosomes and plasmids, a normality test (*Shapiro–Wilk*) and correlation analysis (*Spearman's*) were performed. The Kruskal–Wallis test was performed to analyze the variance among the three species under study. To analyze the differences between strains in Cluster III and strains in Cluster IV of the MLST, the Mann–Whitney test was performed. These analyses were carried out using SPSS version 25 software.

Multiple locus sequence typing, MLST

The partial sequences of seven housekeeping genes (*glpF*, *gmk*, *ilvD*, *pta*, *pur*, *pycA*, and *tpi*) from all selected genomes were retrieved via the PubMLST web tool (Jolley et al. 2018) (https://pubmlst.org/), using the MLST strategy described by (Priest et al. 2004) for the *B. cereus* group. The sequences were concatenated and the resulting sequences from each genome were aligned using the MEGA package version 7.0.26 (Kumar et al. 2016) (https://www.megasoftware.net/). The ClustalW program and neighbor-joining algorithm were used to construct a phylogenetic tree, maintaining the default parameters. Visualization of the phylogenetic tree was performed with the web tool iTOL version 4 (Letunic and Bork 2019) (https://itol.embl.de/).

Genome alignment

The plasmid sequences (*Bt* YBT-1520 plasmid pBMB293, *Bt* YWC2-8 plasmid pYWC2-8–1, *Bt* Bc601 plasmid pBTBC2, *Bt* IS5056 plasmid pIS56-285, *Bt* CT-43 plasmid pCT281, *Bt* HD-29 plasmid pBMB267, *Bt* HD-1 plasmid pBMB299, and *Bt* YBT-1518 plasmid pBMB0232) were compared to the Basic Local Alignment Search Tool (Blastn), using the "align two or more sequences" option (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

The program Mauve (Darling et al. 2004) was used within the Geneious Prime® 2020.1.2 toolkit (Restricted) to align the genome of *Bt* YBT-1518 with the genomes of phylogenetically closed strains (*Bt* MYBT1846 and *Bt* 407). The genomic location of Group II introns between the strains was plotted using the ClicoFS tool (http://clicofs.codoncloud.com/).

Evaluation of ncRNA expression

Strains and growth conditions

The expression of ncRNA genes was evaluated in the *B. thuringiensis* var. *thuringiensis* 407 Cry- strain (Vilas-Bôas et al. 1998). For DNA extraction, the strain was grown at 30 °C in a liquid Luria–Bertani (LB) medium. For RNA extraction, the strain was grown in LB medium, also at 30 ºC, in triplicate and monitored until it reached $O.D_{600}$ of around 0.8, corresponding to the end of the logarithmic phase (Gilois et al. 2007).

Selection of ncRNA genes and primer design for validation

We evaluated the expression of some ncRNA genes identified in the researched genomes. We selected genes with low copy numbers (between 1 and 6), with a minimum size of 80 nt, that had been identified in the genome of *Bt407* Cry^- (the strain used for the gene expression assays), that were present in strains from different clusters of the phylogenetic tree, and that were not found inside genic/ORF regions. The genomic coordinates of these genes in the reference strain (*Bt407* Cry^-) were identified using the INFERNAL program and used, along with the U-gene program version 1.17.0 (Okonechnikov et al. 2012), to extract the sequences of each element selected for validation. The sequences of the primers and their properties, such as melting temperature and size of the amplified product, are listed in Online Resource Table 3. The primers were synthesized by Invitrogen™.

DNA extraction

DNA samples from the *B. thuringiensis* 407 Cry- strain were extracted using the Wizard® Genomic DNA Purification kit (Promega), following the protocol for Gram-positive bacteria, with adaptations. Thus, a treatment with 120 µL of lysozyme (10 mg/mL) was added at 37 ºC for 1 h in the step prior to cell lysis. At the end of the extraction procedure, the DNA was rehydrated with 20 µL of DNA Rehydration Solution.

RNA extraction and cDNA synthesis

Total RNA extraction was performed using the lllustra™ RNAspin Mini RNA kit (GE Healthcare illustra™) with adaptations. Initially, centrifuged cell samples were homogenized with 350 µL of lysis buffer, 3.5 µL of β-mercaptoethanol, and glass beads (2 mm) to lyse the cells through vortex agitation. The following steps were performed in accordance with the kit protocol, including a DNAse digestion. RNA was eluted by adding 50 µL of RNAse-free water.

The cDNA was obtained by reverse transcription of total RNA in a 20 µL reaction using the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems™), according to the manufacturer's recommendations. The quality of the RNA and cDNA was analyzed using a Nanodrop 2000 spectrophotometer.

PCR amplification from cDNA

For validation, the obtained cDNA was used to amplify each of the selected ncRNAs. The reactions were carried out in an Applied Biosystems™ A24811 thermal cycler with the QIAGEN TopTaq Master Mix kit following the manufacturer's specifications and using the primers presented in Online Resource Table 3. PCR products were analyzed on 1% agarose gel electrophoresis, stained with SYBR Safe (Invitrogen™), and visualized in an L-PIX EX photo documentation system (Loccus Biotechnology) with a K9-100L molecular weight marker (1 kb) from Kasvi. Water (negative control) and DNA (positive control) were used as controls of the PCR reactions. In standardization reactions, PCR products were not found when the RT-PCR reactions was run with samples from RNA extraction that were not reversely transcribed (data not shown).

Results

Genomic aspects and in silico identification of ncRNAs

The mining of genomic sequences enabled the selection of 132 complete genomes, with 44 being *B. cereus*, 40 being B. *thuringiensis*, and 48 being *B. anthracis*. Chromosome size varied slightly among species, with *B. anthracis* having the

lowest chromosome variation (from 5.20 to 5.25 Mb) and *B. thuringiensis* having the highest variation (from 5.21 to 6.00 Mb). The GC content of the chromosomal sequences was about 35%, and no differences were observed among the three species (Online Resource Table 4).

Figure 2A graphically presents the chromosomal size (Mb) and number of chromosomal ncRNAs of each strain studied; they are represented by spots. The data from the three species enabled us to identify three distinct clusters. The first one, positioned around 5.20 Mb, was composed by three strains of *B. cereus* and all strains of *B. anthracis*, indicating that this last species showed little variation in chromosome size and number of ncRNAs identified. The second cluster, positioned between 5.20 and 5.60 Mb, was composed of most *B. cereus strains* grouped with some *B. thuringiensis* strains, which demonstrates that these species showed more variation in number of ncRNAs and chromosome size as compared to *B. anthracis*. The third cluster, composed exclusively of *B. thuringiensis* strains, positioned between 5.60 and 6.00 Mb, demonstrates that some strains of this species showed greater chromosome size than the other evaluated strains. Four strains were not included in the clusters because they presented divergent chromosome sizes. One example is strain *Bt* YBT-1518, which had the largest chromosome among all studied (6.00 Mb) and the highest number of ncRNAs identified (332).

Distribution analysis between chromosome size and the number of ncRNAs was calculated by the Shapiro–Wilk normality test, which showed that the data did not show a normal distribution ($W = 0.778$, p value < 0.05, and $W = 0.543$, p value < 0.05, respectively). Spearman's test indicated a high positive correlation (Rho $= 0.811$, p value < 0.01 and Linear $R^2 = 0.524$) between number of ncRNAs and size (in bp) of chromosomes. This reveals that as the size of a chromosome increases, the number of identified ncRNAs increases as well.

The number of identified chromosomal ncRNAs varied among strains (Table 1). The Kruskal–Wallis test demonstrated that the number of ncRNAs differed significantly ($p < 0.001$) among the three species. Thus, the results of this study demonstrated that *B. anthracis* strains showed the smallest average chromosome size and the smallest average of ncRNAs. Similarly, *B. thuringiensis* strains presented both the largest average chromosome size and the highest average of chromosomal ncRNAs (Table 1), which may be due to the existence of *B. thuringiensis* strains with chromosome enlargement (Fig. 2A).

Analysis of the plasmid content of the 132 selected genomes identified 386 plasmids. Plasmids were not found in all strains of the three species, but were greater among *B. thuringiensis* strains, which also showed the highest average of plasmids per strain (Table 1). The average genome increase caused by the presence of plasmids was also higher among *B. thuringiensis* strains (12%), reaching 24% in *Bt* MC28 and *Bt* IS5056 strains (Online Resource Table 4). Among *B. cereus* strains, the presence of the plasmids led to an average increase of 6% in genome size, with a maximum of 14% in *Bc* 03BB108 and *Bc* C1L; in *B. anthracis*, the average increase was 5%, with a maximum of 6% in *Ba* 14RA5914 (Online Resource Table 4). The GC content among plasmids of the three species was similar: *B. cereus* plasmids averaged 33.21%, *B. thuringiensis* 33.28%, and *B. anthracis* 32.77%, as expected for species with a common ancestral origin (Nishida 2012).

Among the *B. anthracis* strains, ncRNAs were identified in all plasmids, while in the other two species, ncRNAs were identified in 58% and 61% of the plasmids of *B. cereus* and *B. thuringiensis,* respectively (Table 1). The Kruskal–Wallis test showed no significant difference in the number of ncRNAs identified in the plasmids of the three species ($p = 0.107$).

The Shapiro–Wilk test for all 386 plasmids evaluated in this study demonstrated that the size in bp and the number of ncRNAs did not follow a normal distribution ($W = 0.798$, p value < 0.05, and $W = 0.639$, p value < 0.05, respectively), while Spearman*'s* test demonstrated a moderate and significant correlation between plasmid size and the number of ncRNAs (Rho $= 0.625$, p value < 0.01). Figure 2B illustrates that many plasmids follow the trend of the straight line, while others are located distantly, indicating a moderate linear relationship between the variables.

In addition, some plasmids (black circle) showed a high number of ncRNAs compared to other plasmids of the same size (from 150 to 350 kb). Blastn alignment demonstrated that these plasmids are similar to each other because they showed identity values between 87.07 and 100.00 and coverage ranging from 58 to 99% (Online Resource Table 5), except for plasmid pBMB0232 from the *B. thuringiensis* strain YBT-1518, which obtained the lowest coverage among the strains (0% to 2%). Analysis of these plasmids identified only Group II Introns, indicating that the presence of this genetic element precedes the dispersal and diversification events of these plasmids among *B. thuringiensis* strains.

Figure 2B highlights two other plasmids from *B. thuringiensis* strains, which showed the highest amounts of ncRNAs identified in a single plasmid. The plasmid pBMB0233 with 240.661 kb, from strain YBT-1518, has 24 ncRNAs, as well as plasmid pYC1 with 761.374 kb, from strain YC-10, has 33 ncRNAs. Furthermore, the bottom of Fig. 2B reveals some large *B. cereus* plasmids that have no ncRNAs or only one ncRNA—for example, the unique plasmid from *B. cereus* strain FM1 with 402.61 kb and plasmid pRML01 from *Bc* strain HN001 with 435.42 kb. Sequence alignment of these plasmids by Blastn showed low identity and coverage values (data not shown), indicating that these plasmids did not have the same genetic origin.

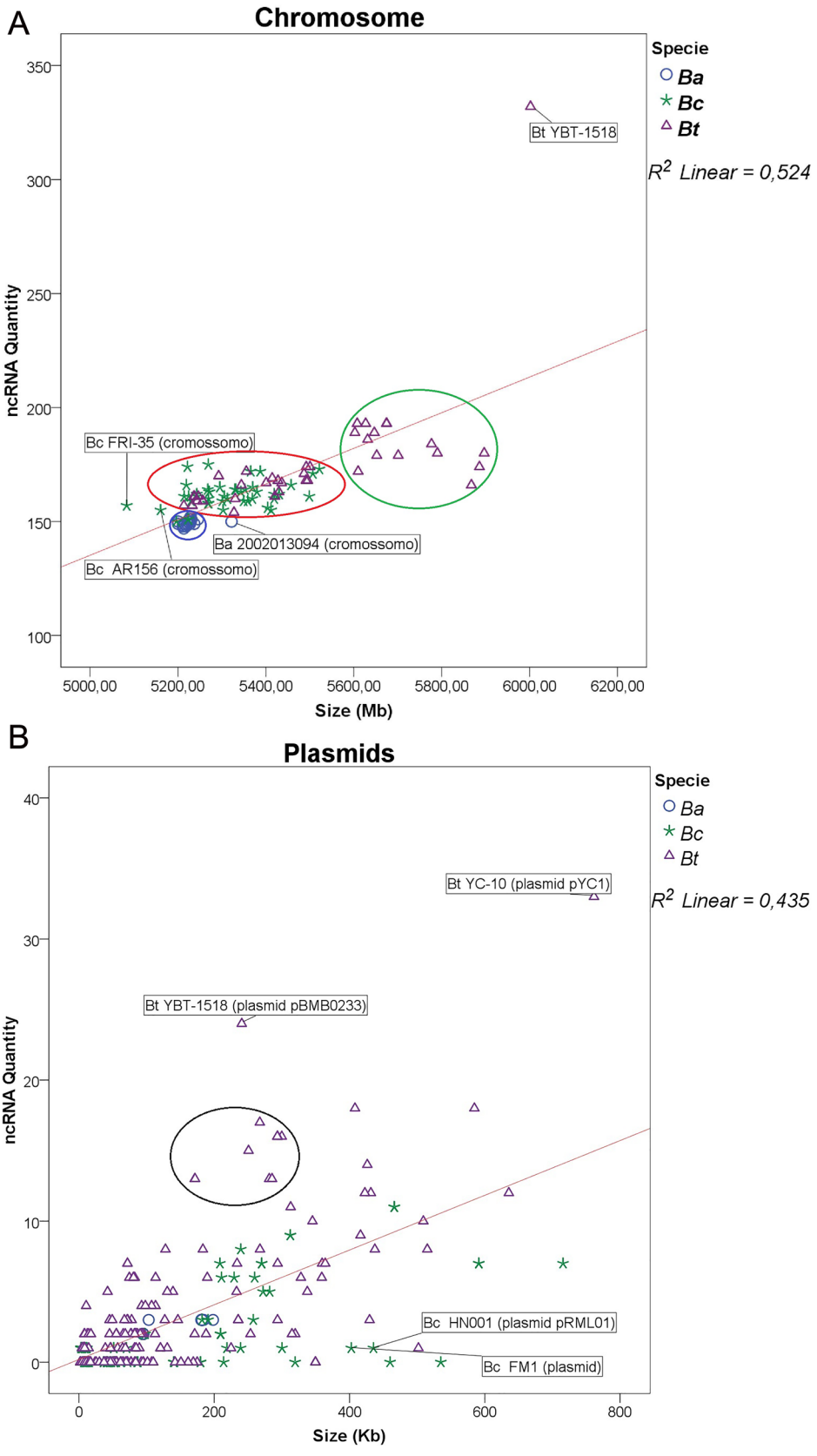

Fig. 2 Scatter plots representing the correlation between replicon size and the number of ncRNAs identified in the three species. Each spot corresponds to a replicon of a *B. cereus s.l.* strain. The red lines represent the Coefficients of Determination R^2. **A** Chromosomes–the blue circle represents the chromosome clustering of the *B. anthracis* strains with some strains of *B. cereus*. The red circle represents the chromosome cluster of the *B. cereus* and *B. thuringiensis* strains. The green circle represents the chromosome cluster of the *B. thuringiensis* strains. **B** Plasmids–the black circle represents a cluster of plasmids with high numbers of identified ncRNAs

Table 1 Genetic data and presence of chromosomal and plasmid ncRNA in *B. cereus*, *B. thuringiensis* and *B. anthracis*

	Bc	*Bt*	*Ba*
Number of strains	44	40	48
Number of strains with plasmids (%)	38 (86%)	39 (98%)	33 (69%)
Number of plasmids	89	238	59
Average plasmids per strain	2.02	5.95	1.22
Chromosomes with ncRNA (%)	44 (100%)	40 (100%)	48 (100%)
Plasmids with ncRNA (%)	52 (58%)	144 (61%)	59 (100%)
Average number of chromosomal ncRNA per strain (smaller and larger quantity)	161.8 (150–175)[a]	175.7 (154–332)[b]	150.4 (147–152)[c]
Average number of plasmidial ncRNA per strain (smaller and larger quantity)[A]	3.2 (1-11)[d]	4.5 (1–33)[d]	2.5 (2–3)[d]
Number of ncRNA chromosomal families	51[e]	58[f]	40[g]
Number of ncRNA plasmidial families (h)	18	23	4
Total number of ncRNA families	55	64	43

Kruskal-Wallis Test = Different letters in each line correspond to significant differences in values ($p < 0.05$)

[A]Only plasmids with ncRNA identified

Different letters ([a], [b], [c], [d], [e], [f], [g]) in each line correspond to significant differences invalues ($p < 0.05$)

Identification and occurrence of ncRNA families

One of the objectives of this study was to verify the occurrence of ncRNA families in *B. cereus s.l.* We identified 65 families of ncRNAs among the three species, which were assigned as described in the Rfam database. *B. thuringiensis* is the species with the highest number of families, a total of 64, while B. *cereus* presented 55 families and *B. anthracis* presented 43 families (the least) of ncRNAs (Online Resource Fig. 6, Table 1). Through the Kruskal–Wallis test, we identified that the difference in the number of families between species was significant for the chromosomes ($p < 0.001$), while there was no difference for the plasmids ($p = 0.051$) (Table 1).

Most of the families (43 of 64) were classified as common (Online Resource Fig. 6). In addition, *B. cereus* and *B. thuringiensis* strains shared 11 families. However, *B. anthracis* strains did not share ncRNA families with the other two species separately (Online Resource Fig. 6 and Online Resource Table 6). In addition, 10 families exclusive to *B. thuringiensis* strains, and one family exclusive to *B. cereus* strains, were identified (Online Resource Fig. 6 and Online Resource Table 6). Together, these data indicate that ncRNA acquisition occurred both at the common ancestor of the three species (ncRNA families classified as common) and after separation from *B. anthracis* (ncRNA families shared between *B. cereus* and *B. thuringiensis*). Moreover, after the separation between the latter two species, *B. thuringiensis* strains continued to acquire ncRNAs at a higher frequency than *B. cereus* strains, as they had a higher number of unique families.

Among the common families, most are located on chromosomes such as the SurA, BtsR1, and RsaE families, which are related to sporulation (Silvaggi et al. 2006; Irnov et al. 2010), defense and pathogenicity (Peng et al. 2018), oxidative stress (Durand et al. 2017; Marincola et al. 2019), and virulence (Geissmann et al. 2009; Guillet et al. 2013). The chromosomal copy number per strain was highly variable among the families (Table 2 and Online Resource Table 6), but the copy number of each family was similar among the three species. In plasmids, the occurrence of the ncRNA families was always a single copy or a low copy number (Table 2). A few families located both on chromosomes and on plasmids were also identified (T-box, BsrC, crcB, cspA, Intron_gpI, RsaE, SAM, SurA, ykok, Intron_gpII, group-II-D1D4-5, yidF, and L21_leader) (Online Resource Table 6). A large proportion of these families are present on the chromosomes of most strains, whereas on plasmids they were found in few strains, except for the group-II-D1D4-5 and Intron_gpII families, which were identified in high copy number on a larger number of *B. thuringiensis* plasmids.

Of the 11 families shared by *B. cereus* and *B. thuringiensis*, six were found exclusively on chromosomes (5_ureB_sRNA, Histone, rli38, rsaJ, sau-5971, and group-II-D1D4-4). Three were found exclusively on plasmids (Bacillus-plasmid, CRISPR-DR43, and rliF) and two were found on both chromosomes and plasmids (epsC and rli40) (Online Resource Table 6). In addition, all families of ncRNAs classified as shared were found in a single copy or a low copy number on chromosomes and plasmids (Table 2).

Of the families identified exclusively in a single species, the one exclusive to *B. cereus* (tsr25) was detected only on the chromosome of a single strain, sequenced twice, *Bc* E33L (CP000001.1 and CP009968.1). Meanwhile, from the 10 families exclusive to *B. thuringiensis* (Online Resource Table 6), five were identified only on chromosomes (ncr1175, PyrG_leader, rliI, RNAI, and ykkC- ykkD), two

Table 2 Distribution of ncRNA families identified in *B. cereus s.l*

Category functional/structural[a]	Classes[b]	SubClass[c]	ncRNA family[d]	Rfam[e]	Length ncRNA (nt)[f]	Copy variation/total strain[g]					
						Chromosome			Plasmids		
						Bc	*Bt*	*Ba*	*Bc*	*Bt*	*Ba*
Gene	CRISPR		CRISPR-DR43	RF01352	29				0–1/1	0–1/2	
	miRNA		mir-10	RF00104	75					0-1/1	
	Ribozyme[i]		Bacteria_large_SRP	RF01854	264	0–1/43	1/40	1/48			
	Ribozyme		RNaseP_bact_b	RF00011	366	1/44	1/40	1/48			
	rRNA[l]		LSU_rRNA_bacteria	RF02541	2925	11–16/44	12–16/40	10–11/48			
	rRNA[l]		SSU_rRNA_bacteria	RF00177	1533	11–16/44	11–15/40	10–11/48			
	sRNA		5_ureB_ncRNA	RF02514	292	0–1/3	0–2/11				
	sRNA[k]		6S	RF00013	184	2/44	2/40	2/48			
	sRNA		Bacillus-plasmid	RF01691	61	0-1/19	0-2/48				
	sRNA		BsrC	RF01410	86	1–3/43	2–4/40	2/48	0–1/1	0–1/6	
	sRNA		BsrF	RF01411	113	1/43	1/40	1/48			
	sRNA		BtsR1	RF02888	51	1/44	1/40	1/48			
	sRNA		ncr1175	RF02450	110		1/1				
	sRNA		ncr1241	RF02451	63	1/44	1–2/40	1/48			
	sRNA		ncr1575	RF02452	215	1–2/44	1–2/40	2/48			
	sRNA		rli38	RF01470	150	0–1/2	0–1/4				
	sRNA		rli40	RF01472	263	0–1/12	0–1/13		0–2/2		
	sRNA	Antisense	rliD	RF01494	327	0–1/32	0–1/26	1/48			
	sRNA	Antisense	rliE	RF01459	222	0–1/10	0–1/5	1/48			
	sRNA		rliF	RF01476	220				0–1/1	0–1/1	
	sRNA		rliI	RF01487	239		0–1/1				
	sRNA	Antisense	RNAI	RF00106	102		0–1/1				
	sRNA		RsaE	RF01820	112	1/44	1–2/40	1/48	0–1/2	0–1/5	
	sRNA		RsaJ	RF01822	287	0–1/6	0–1/1				
	sRNA		sau-5971	RF02397	96	0–1/5	0–1/7				
	sRNA		SR1	RF02376	210	2–3/44	2–3/40	2/48			
	sRNA		SurA	RF02377	278	0–2/19	0–2/17	1/48	0–1/2	0–1/2	
	sRNA	Antitoxin	ToxI	RF02519	34				0–1/1	0–3/19	0–1/28
	sRNA		tsr25	RF02578	181	0–1/2					

Table 2 (continued)

Category functional/structural[a]	Classes[b]	SubClass[c]	ncRNA family[d]	Rfam[e]	Length ncRNA (nt)[f]	Copy variation/total strain[g]					
						Chromosome			Plasmids		
						Bc	*Bt*	*Ba*	*Bc*	*Bt*	*Ba*
Cis-reg	Leader		L10_leader	RF00557	139	1/44	0–1/39	1/48			
	Leader		L19_leader	RF00556	42	1/44	1/40	1/48			
	Leader		L20_leader	RF00558	127	1–2/44	1/40	1/48			
	Leader		L21_leader	RF00559	78	1/44	1/40	1/48		0–1/1	
	Leader		PyrG_leader	RF02371	78		0–1/2				
	Leader		T-box	RF00230	224	38–47/44	44–48/40	43–45/48	0–1/1	0–1/2	0–2/28
	Leader		ylbH	RF00516	122	1–2/44	1–3/40	1–2/48			
	Riboswitch		c-di-GMP-I	RF01051	87	1–6/44	1–6/40	1/48			
	Riboswitch		Cobalamin	RF00174	191	1/44	1/40	1/48			
	Riboswitch		crcB	RF01734	63	1–2/44	1–2/40	1/48	0–1/3	0–1/1	
	Riboswitch[k]	Thermoregulator	cspA	RF01766	428	6–7/44	6–7/40	6/48	0–1/1	0–2/6	
	Riboswitch[i]		epsC	RF01735	117	0–1/29	0–2/37		0–1/1	0–1/1	
	Riboswitch		FMN	RF00050	140	2/44	2–3/40	2/48			
	Riboswitch		glmS	RF00234	171	1/44	1/40	1/48			
	Riboswitch		Glycine	RF00504	94	1/44	1/40	1/48			
	Riboswitch[i]		Histone3	RF00032	46	0–1/15	0–2/17				
	Riboswitch		Lysine	RF00168	183	4–5/44	4–5/40	4/48			
	Riboswitch[k]		pan	RF01749	93	1/44	1/40	1/48			
	Riboswitch		PreQ1	RF00522	45	2/44	2/40	2/8			
	Riboswitch		Purine	RF00167	102	6–8/44	6–8/40	6/48			
	Riboswitch[k]		PyrR	RF00515	126	3–4/44	3–4/40	3/48			
	Riboswitch		SAM	RF00162	108	16–20/44	16–19/40	17/48	0–1/2	0–1/4	
	Riboswitch		TPP	RF00059	105	6–8/44	6–8/40	7/48			
	Riboswitch		ydaO-yuaA	RF00379	136	3–4/44	4/40	4/48			
	Riboswitch[j]		yjdF	RF01764	101	0–2/40	1/40	1/48	0–1/1		
	Riboswitch		ykkC-yxkD	RF00442	108		0–1/1				
	Riboswitch		ykoK	RF00380	169	1–4/44	1–8/40	1/48	0–2/4	0–2/11	

Table 2 (continued)

Category functional/structural[a]	Classes[b]	SubClass[c]	ncRNA family[d]	Rfam[e]	Length ncRNA (nt)[f]	Copy variation/total strain[g]					
						Chromosome			Plasmids		
						Bc	*Bt*	*Ba*	*Bc*	*Bt*	*Ba*
	Riboswitch		yybP-ykoY	RF00080	175	1–2/44	1–2/40	1/48			
Intron	Group I introns		Intron_gpI	RF00028	251	0–5/38	0–5/36	1–3/48	0–1/1	0–2/6	
	Group II introns		Intron_gpII	RF00029	93	0–8/25	0–16/24		0–6/25	0–20/80	0–2/31
	Group II introns		group-II-D1D4-1	RF01998	84		0–1/4			0–6/5	
	Group II introns		group-II-D1D4-2	RF01999	118					0–1/2	
	Group II introns		group-II-D1D4-3	RF02001	171		0–1/4			0–7/6	
	Group II introns		group-II-D1D4-4	RF02003	138	0–1/1	0–5/5			0–1/5	
	Group II introns		group-II-D1D4-5	RF02004	200	0–6/20	0–7/21		0–5/25	0–10/63	0–1/31
	Group II introns		group-II-D1D4-6	RF02005	228		0–11/10			0–3/25	

Color gradient (gray) represents the variation between the maximum and minimum amount of copies identified. Lighter shades indicate smaller quantities and darker shades indicate larger quantities. *Bt B. thuringiensis, Bc B. cereus, Ba B. anthracis*

The YBT-1518 strain has 52 copies of the Intron_gpII element and 53 copies of the group-II-D1D4-1 element and 53 copies of group-II-D1D4-3.

[a]Functional/structural category according to Rfam.

[b]Classification of ncRNA by Rfam.

[c]Subclass according to Rfam.

[d]ncRNA name (Rfam).

[e]Rfam Identification.

[f]ncRNA size (nt).

[g]Maximum and minimum copy number identified on chromosome and plasmid of the three species.

[i]ncrDeep: epsC; Histone3; Bacteria_large_SRP.

[j](Li et al. 2016)

[k](Geissler et al. 2021)

[l]The ncRNAs SSU_rRNA_bacteria and LSU_rRNA_bacteria were the only families considered lncRNA in this study

were identified only on plasmids (group-II-D1D4-2 and mir-10), and three were present on chromosomes and plasmids (group-II-D1D4-1, group-II-D1D4-3, and group-II-D1D4-6) (Online Resource Table 6). Most of these exclusive families also showed a limited ability to spread in the group, occurring in few phylogenetically related strains and in a single copy, except for group-II-D1D4-1, group-II-D1D4-2, group-II-D1D4-3, and group-II-D1D4-6, which occurred in a higher percentage of strains and in more copies than the other exclusive families (Table 2 and Online Resource Table 6).

In this study, we separated the ncRNAs according to size, unrelated to the functional/structural categories or classes, thus, we considered as sRNAs the families with up to 500 bp and lncRNAs over 500 bp. Thus, 63 families were classified as sRNAs and their copy numbers varied from 0 to 48 among strains (Table 2). Two families (SSU_rRNA_bacteria and LSU_rRNA_bacteria) were classified as lncRNA and identified with 10–16 copies per strain. Therefore, most of the identified families belong to the sRNA class, as previously described in other bacterial groups (Mars et al. 2016; Wolf et al. 2018), also confirming the low occurrence of lncRNAs for the *B. cereus* group.

Functional/structural categories and classes of ncRNAs

Through the Rfam database and literature (Li et al. 2016; Fiannaca et al. 2017; Harris and Breaker 2018; Amin et al. 2019; Lee et al. 2019; Drecktrah et al. 2020; Geissler et al. 2021), all identified families were distributed into three functional/structural categories, nine classes, and three subclasses (Fig. 3 and Table 2). Each functional/structural category was investigated for relative abundance. Thus, the functional/structural category with the highest abundance in *B. cereus s.l.* was Cis-reg with 67.78%, followed by Gene with 25.28% and Intron with the lowest abundance of 6.94% (Fig. 3). Among the classes, Riboswitch, Leader, rRNA, sRNA, and Group II introns accounted for between 37.83% and 5.85% of the ncRNA molecules, while the classes with the lower abundance were Group I introns, ribozyme, miRNA, and CRISPR, reaching up to 1.09% each (Fig. 3).

The occurrence of ncRNA classes varied between chromosomes and plasmids. Thus, the Leader, Riboswitch, and sRNA were identified mainly on chromosomes, whereas the Group II introns class was identified on both chromosomes and plasmids (Table 2). In addition, the number of families identified in *B. cereus s.l.* for each class was investigated. The sRNA presented 23 families that can regulate the expression of target genes at the post-transcriptional level (Sajid et al. 2018; Sridhar and Gayathri 2019). The Riboswitch presented 21 families; this class contains ncRNAs capable of controlling gene expression at the transcriptional

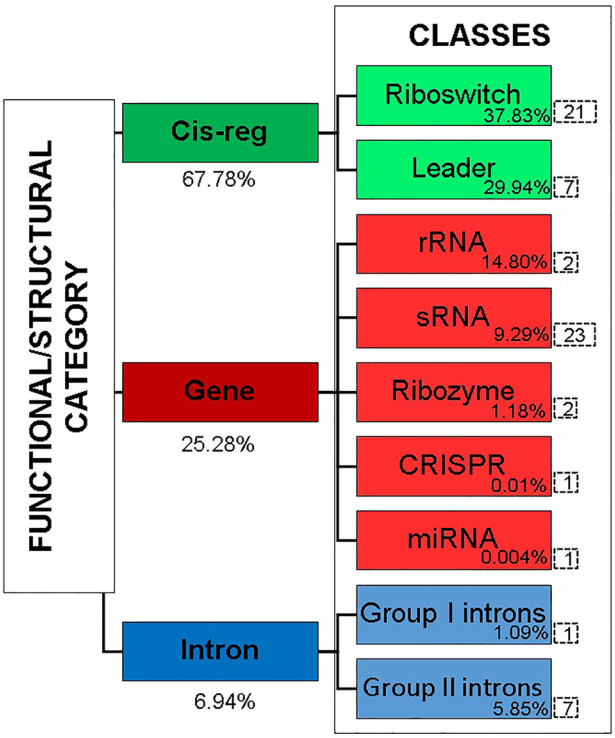

Fig. 3 Functional/structural categories and classes of ncRNAs identified in chromosomes and plasmids of *B. cereus*, *B. thuringiensis*, and *B. anthracis*. *Boxes with dotted borders represent the number of families of each class

and translational levels (Hör et al. 2018) and, then, to respond to environmental changes. With seven families, the Leader class is composed of ncRNA families that regulate translation (Irnov et al. 2010) and interact with metabolism-related proteins (Xia et al. 2012). The Group II introns contained seven families; these elements are ribozymes that act as mobile genetic elements, predominantly as retroelements (Harris and Breaker 2018; Toro et al. 2018; Fayad et al. 2019).

Among the classes found to have a lower incidence, ribozyme and rRNA were composed of two families, whose function is to catalyze reactions, such as RNA splicing, RNA cleavage, and protein synthesis (Moore and Steitz 2002; Lehmann and Schmidt 2003; Skilandat and Sigel 2013). The CRISPR, miRNA, and Group I introns are represented by a single family (Fig. 3, Table 2). CRISPR can provide defense mechanisms against viruses and may be inactivated in some strains (Grissa et al. 2007; Zheng et al. 2020), miRNA has already been described as acting in insect resistance to *B. thuringiensis* (Xu et al. 2015), while Group I introns have the same function as elements of the Group II introns (Harris and Breaker 2018).

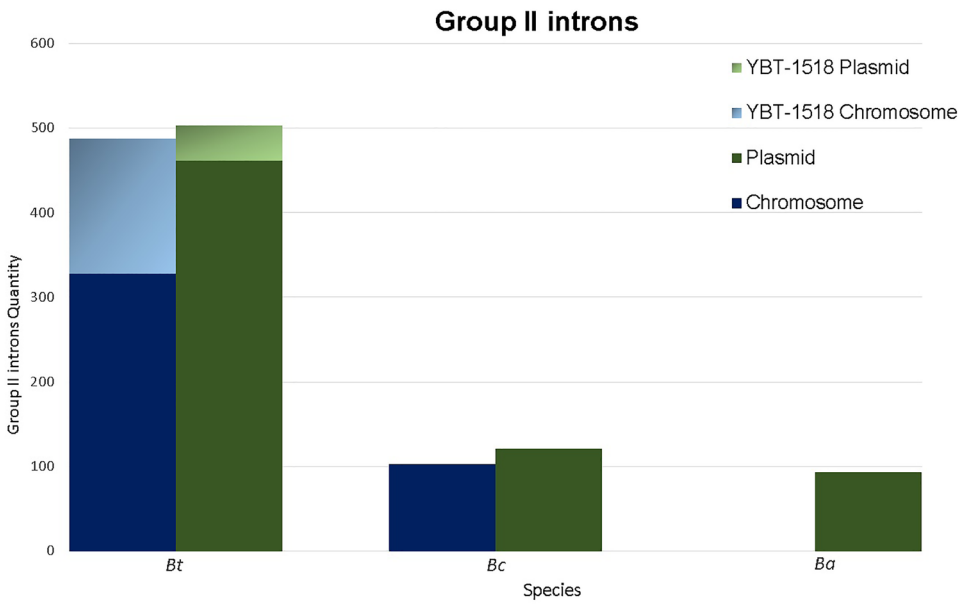

Fig. 4 Quantity of class Group II introns on the chromosomes and plasmids. Compared to the quantity identified in the *B. thuringiensis* strain YBT-1518

Group II introns

Since the class of Group II introns (Group-II-D1D4 and Intron_gpII) correspond to 77% of the ncRNAs identified in *B. thuringiensis* plasmids and to 72% and 62% in *B. cereus* and *B. anthracis* plasmids, respectively, it was further analyzed in both chromosomes and plasmids. The sum of these elements found on the chromosomes of all *B. thuringiensis* strains was 487 copies, 103 copies on all *B. cereus* strains, and no copies on the chromosomes of *B. anthracis*. The Kruskal–Wallis test for the number of Group II introns in the chromosomes showed a significant difference between *B. cereus, B. anthracis* and *B. thuringiensis* species ($p < 0.001$). This result highlights the great number of these elements in *B. thuringiensis* strains, especially on the strain *Bt* YBT-1518 that showed about one-third (32.6%) of the Group II introns from this species. Hence, a total of 159 copies of Group II introns were identified on the chromosome of the *Bt* YBT-1518 (Fig. 4) and when this strain was compared to the three species, the difference was statistically significant ($p < 0.001$).

Among the plasmids, 503 copies of Group II introns were identified in *B. thuringiensis* strains, while 121 copies were identified in *B. cereus*, and 93 copies were identified in *B. anthracis* (Fig. 4). The Kruskal–Wallis test for the number of Group II introns in the plasmids also showed a difference between the three species—*B. cereus*, *B. thuringiensis*, and *B. anthracis* ($p < 0.001$). However, we observed that *B. cereus* and *B. thuringiensis* exhibit no significant difference regarding the copy number of Group II introns ($p = 0.495$), while the pairwise analysis showed that *B. anthracis* is different from the other two species (*Ba-Bc* $p = 0.011$; *Ba-Bt* $p < 0.001$). Similarly, to the chromosomes, the plasmids of the strain *Bt* YBT-1518 showed a remarkable number of Group II introns, corresponding to 10% of the elements found in all *B. thuringiensis* plasmids. Therefore, *Bt* YBT-1518 showed 42 copies in four plasmids (three copies in plasmid pBMB0229, three copies in pBMB0230, 13 copies in pBMB0232, and 23 copies in pBMB0233). In addition, when *Bt* YBT-1518 was compared to *B. cereus*, *B. thuringiensis*, and *B. anthracis*, the Kruskal–Wallis test demonstrated a statistically significant difference ($p = 0.004$), indicating that the distribution of these elements in this strain is significantly different from the distribution in the plasmids of the three species.

From the first analysis, concerning the correlation among the replicons and the number of copies of ncRNAs, the strain *Bt* YBT1518 has already shown a divergent profile of ncRNAs content, with the highest number of chromosomal ncRNAs identified among all the evaluated strains and with a higher number of ncRNAs in the plasmids than other strains. Given the results for Group II introns, we observe that these elements are responsible for the higher number of ncRNAs in this strain. Due to this atypical profile, the statistical analyses between the species were also performed excluding *Bt* YBT1518. However, the result of the analysis of variance did not change, so we retained *Bt* YBT1518 in the overall comparative analyses.

Therefore, the genomic context of these elements was analyzed by aligning the chromosomal sequence of *Bt* strain YBT-1518 with the chromosomal sequences of *Bt* strain MYBT1846 and *Bt* 407 (Online Resource Fig. 7 A), which are closely related phylogenetically. Thus, on the chromosome of *Bt* strain MYBT1846, four copies of the Group II class introns were identified, whereas in *Bt* strain 407, six copies were identified (Online Resource Fig. 7 A). Elements

belonging to the class Group II introns were detected both in highly conserved regions and in regions with low conservation among the three genomes (Online Resource Fig. 7 A). This result indicates that the acquisition and spread of these elements in strain *Bt* YBT-1518 took place after the evolutionary divergence of the three strains. This statement is also true for plasmid analysis, as these elements were observed in the plasmids of *Bt* strains MYBT1846 and *Bt* 407, although in smaller quantities (Online Resource Fig. 7 B), as compared to the occurrence in the plasmids of *Bt* strain YBT-1518.

Correlation between ncRNAs and phylogeny

Due to the complex taxonomic and phylogenetic relationships among *B. cereus*, *B. thuringiensis*, and *B. anthracis*, an MLST tree containing all strains of this study was constructed and classified into clusters according to Guinebretière et al. (2008). Moreover, we include data about the occurrence of ncRNAs in chromosomes and plasmids, aiming to obtain a profile of the distribution of ncRNAs in correlation with the phylogenetic and ecological classification of these strains.

Analyzing the phylogenetic tree (Additional file 1), one can see that *B. cereus* and *B. thuringiensis* strains are distributed all over the clusters, while *B. anthracis* strains form a highly clonal subcluster that remains separate from the other species, as described previously (Rasko et al. 2005; Tourasse et al. 2011; Bazinet 2017). Cluster IV contains *B. thuringiensis* strains, widely used as a bioinsecticide, and environmental strains of *B. cereus*. On the other hand, pathogenic *B. cereus* strains and *B. anthracis* strains, in addition to some *B. thuringiensis* strains, are part of cluster III.

Concerning ncRNA, the families common to the three species were present in most of the strains in all clusters and represent the majority of ncRNA families. Meanwhile, most of the ncRNA families shared between *B. cereus* and *B. thuringiensis* showed a limited ability to spread, being shared mainly by a few phylogenetically related strains (Additional file 1), except for Histone3, epsC, and rli40 families, which were found in a large number of strains in both clusters (Online Resource Table 6). The exclusive families also had a limited ability to spread in the group, as they were not identified in more than one strain and presented a single copy in the genomes in which they were found, except for the families group-II-D1D4-1, group-II-D1D4-2, group-II-D1D4-3, and group-II-D1D4-6, which are found in many strains and have a higher copy number. Analyzing this information was essential to selecting ncRNAs for use in the PCR expression validation step.

The comparative analysis of ncRNA abundance on chromosomes revealed differences between MLST cluster III and cluster IV. The Mann–Whitney test showed that the distribution of the number of families ($p = 0.000$) and the number of ncRNAs ($p = 0.000$) was different between the two clusters. However, in the analysis of ncRNA abundance in plasmids, the Mann–Whitney test showed no difference between clusters III and IV for either the number of families ($p = 0.890$) or the number of ncRNAs ($p = 0.644$).

Validation

For the in vitro validation, nine families were selected based on phylogenetic distribution, the abundance of each family in the tree species, and the number of copies in the genomes. These families belong to the three functional and structural categories (Genes, Cis-Reg, and Intron). The description of the families, as well as the primers used, are presented in Online Resource Table 3. All ncRNAs were amplified from the cDNA of *B. thuringiensis* var. *thuringiensis* 407 Cry- from the end of the exponential phase with the size corresponding to the amplicon (Fig. 5, Online Resource Table 3). This demonstrates their presence in the RNA samples of the tested strain and confirms their expression.

Discussion

Bacteria of the *Bacillus cereus* group colonize several ecological niches and infect different hosts. The phylogeny of this group is complex, with polyphyletic clades and paraphyletic species (Guinebretière et al. 2008; Tourasse et al. 2011; Liu et al. 2015; Bazinet 2017; Raymond and Federici 2017). Recently, a new taxonomic nomenclature, based on an average nucleotide identity genome species threshold (92.5 ANI) was proposed (Carroll et al. 2020). Although ncRNAs have several important physiological roles (Nitzan et al. 2017; Harris and Breaker 2018; Svenningsen, 2018), the role of these genetic elements in this bacterial group is based on studies of specific interest (Xu et al. 2015; Lu et al. 2018; Nie et al. 2019). Thus, in this study, we describe the prevalence, diversity, and genomic localization of ncRNAs in a high number of *B. cereus, B. thuringiensis,* and *B. anthracis* strains, and validate the expression of some ncRNAs in *Bt* strain 407.

The in silico strategy allowed us to identify ncRNA sequences in the chromosomes of all strains studied, with variation in the average number of chromosomal ncRNAs among the three species. The evaluation of the set of selected strains demonstrates that *B. anthracis* strains showed no variation in chromosome size besides presenting the smallest chromosomes within the group—a result in agreement with the clonality of this species (Rasko et al. 2005; Kolstø et al. 2009). *B. cereus* strains, in contrast, showed variation in chromosome size, with a gain in genetic material, although this variation was greater among *B. thuringiensis* strains,

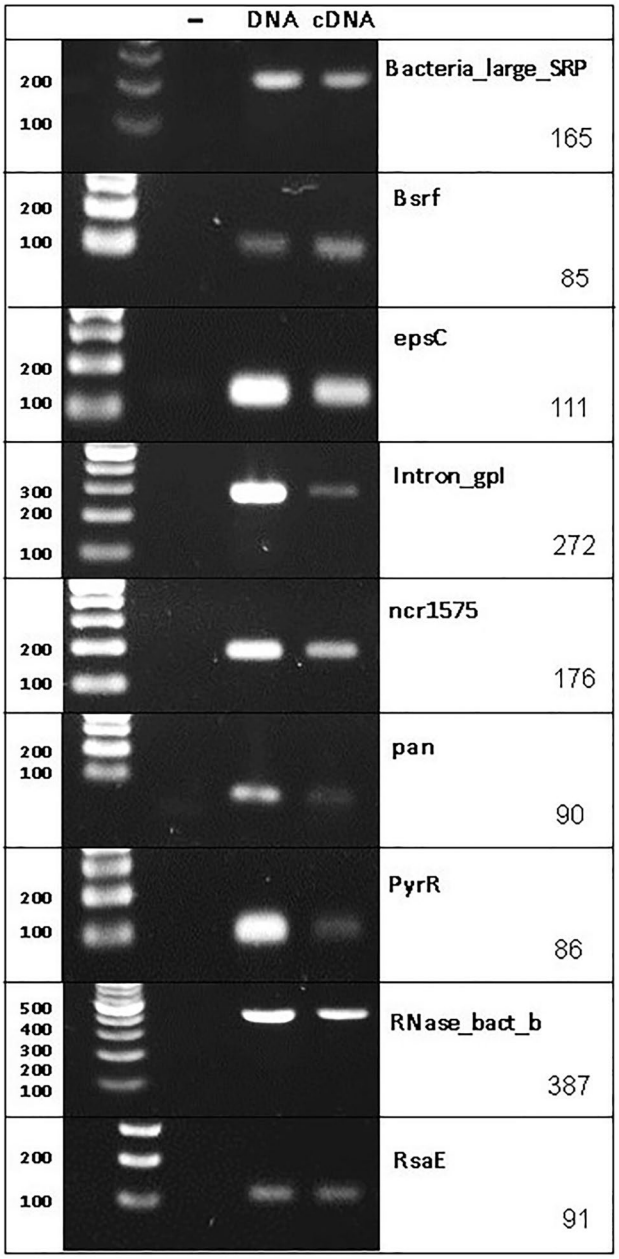

Fig. 5 Validation of the ncRNAs families (Bacteria_large_SRP, BsrF, epsC, Intron_gpI, ncr1575, pan, PyrR, RNaseP_bact_b and RsaE) by RT-PCR. The ncRNAs were checked on agarose gel for presence in DNA and cDNA of the *Bt* 407 Cry- strain, in a culture with $O.D._{600}$ around 0.8. The molecular weight markers are indicated on the left of the image and correspond to the size in bp, and the amplicon on the right

as the latter species presented the strains with the largest chromosomes detected, as already observed in other studies (Cardoso et al. 2019; Fayad et al. 2019). Thus, in *B. thuringiensis*, some strains showed more important genetic material gain than the other strains evaluated in this study. The results regarding the number of ncRNAs in the chromosomes of *B. thuringiensis* strains reflect these differences in the chromosomal profile of the species due to the positive correlation between replicon size and number of ncRNAs. In addition, the chromosome of *B. thuringiensis* showed the highest number of families and the highest absolute number of ncRNAs, followed by *B. cereus* and, finally, *B. anthracis*.

Strains of *B. anthracis* exhibit low genetic diversity and are believed to have diverged from an ancestral *B. cereus* through the acquisition of two large plasmids (pXO1 and pXO2) that contain genes encoding for characteristic virulence factors (Pilo and Frey 2011; Moayeri et al. 2015). Thus, the plasmid content is very similar among their strains (Online Resource Table 6) (et al. 2000; Helgason; Rasko et al. 2005;). Furthermore, the use of the in silico strategy to search for ncRNAs in plasmids demonstrated that among *B. anthracis* strains, all analyzed plasmids had ncRNA (Table 1), with minimal variation with respect to the number or variety of families identified. This result, again, agrees with the fact that this species is highly clonal. However, among *B. cereus* and *B. thuringiensis* strains, plasmid content is highly variable. Thus, the results of this study demonstrated that the *B. thuringiensis* species shows a higher percentage of plasmid-bearing strains, as well as a higher number of plasmids per strain; as a consequence, the presence of these plasmids allows for a greater average increase in genome size (Zhong et al. 2011; Patiño-Navarrete and Sanchis 2017; Fayad et al. 2019). In contrast to what was seen for *B. anthracis*, the presence of ncRNA is not a common feature of all *B. cereus* and *B. thuringiensis* plasmids, which can be explained by the wide plasmid network in these species, especially in *B. thuringiensis*. However, the frequency of plasmids carrying ncRNAs is higher in *B. cereus* and *B. thuringiensis*, and the presence of plasmid ncRNAs (in absolute numbers) is higher in *B. thuringiensis*. This large and diverse plasmid network may have been responsible for the lack of statistical difference in the number of ncRNA among the three species in the analysis of variance.□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

Concerning the results of ncRNA families in plasmids, we observed a large number of Group II introns on the plasmids of all three species, especially in *B. thuringiensis*. This class of ncRNAs consists of elements characterized as ribozymes that act as mobile genetic elements (Harris and Breaker 2018; Toro et al. 2018; Fayad et al. 2019). Therefore, the abundance of these elements in *B. cereus* group strains reinforces the role of horizontal genetic transfer (HGT) in genome plasticity, evolution, and the acquisition of novel ncRNAs, as already observed for *Streptococcus agalactiae* (Wolf et al. 2018).

Of all the strains with Group II introns analyzed in this study, *B. thuringiensis* strain YBT-1518 deserves special attention. Our results showed that this strain has a different

profile of ncRNAs as compared to all other evaluated strains due to the significant increase in the number of copies of Group II introns. However, the sum of the size of these elements (16.774 kb) were not the main responsible for the increase in the genome size of this strain (chromosome of 6.00 Mb). Furthermore, the presence of these transposable elements in a high copy number did not cause the loss of toxicity, which is active against nematode larvae (Zheng et al. 2016). This toxicity results mainly from the activity of the product of three genes, i.e., one located on the chromosome (*cry5Ba2*) and two located on the plasmid pBMB0228, *app6Aa1* (formerly *cry6Aa2*) and *xpp55Aa1* (formerly *cry55Aa1*) (Wang et al. 2014; Crickmore et al. 2020).

Regarding the number of ncRNAs that this study identified, we demonstrated the wide prevalence of ncRNAs in the three species studied, especially in *B. thuringiensis,* which presented a higher number of ncRNA copies and families than the other species. The results presented here, along with previous results from the research group, demonstrate that *B. thuringiensis* has accumulated different metabolic regulators, both on the chromosome, as in the case of ncRNAs (as demonstrated here), and on plasmids, as can be observed for ncRNAs and Rap-Phr quorum sensing systems (Cardoso et al. 2019). Rap-phr systems regulate essential processes such as sporulation in bacteria of the *B. cereus* group (Fazion et al. 2018; Cardoso et al. 2019) and were identified in all studied strains, being found in a higher abundance in plasmids of *B. thuringiensis* than of the other species of the group. Moreover, the accumulation of these systems in *B. thuringiensis* plasmids may favor a fine regulation of bacterial metabolism to environmental variations (Cardoso et al. 2020). These data are in agreement with the already described plasticity of the *B. thuringiensis* genome (Gillis et al. 2018), enabling this bacterium to better adapt to the environment, possibly due to its complex ecological niche (Argôlo-Filho and Loguercio 2014; Mishra et al. 2017).

Few studies have conducted broad genomic analysis of ncRNAs in different bacterial strains or groups. However, it is not possible to perform an absolute numerical comparison between these studies and our results because, in addition to interspecific differences, this variation can be attributed to variations in bioinformatics tools that predict and identify ncRNAs. For example, in the annotation study performed by (Geissler et al. 2021), using the Rfam tool, the authors identified 230 ncRNAs in *Bacillus subtilis* strain 168; however, the number of ncRNAs annotated in this strain changed with other tools: RefSeq = 212; BsubCyc = 183; Dar et al. Riboswitches = 82, and Nicolas et al. Predictions = 153 (Geissler et al. 2021).

However, the results of the comparative analysis of ncRNAs in other bacterial groups highlight the importance of these genetic elements for the adaptation of bacteria to the environment. Thus, Wolf et al. (2018), in a previous work developed by our research group, verified a great variation of ncRNAs in *S. agalactiae* strains, which were distributed in three clusters, corresponding to different hosts (Cluster 0: fish; Cluster 1: mammals; and Cluster 2: fish and mammals). In addition, they observed a correlation between the number of RNA families and the origin (cluster) of the strains, demonstrating that ncRNAs may be involved with the bacterial-host interaction in pathogenic species.

Other studies have observed that sRNAs can be conserved between related species. Chen et al. (2011) observed, in the genus *Clostridium*, that many sRNAs are conserved among physiologically related species, mainly for pathogenic species as compared to solvogenic and cellulolytic species. They even showed a correlation between the number of sRNAs and genome size for pathogenic species, hypothesizing that the amount of sRNAs is related to the organism's physiology (Chen et al. 2011). Similar results were observed in the study by Wurtzel (2012), which compared the transcriptome of *Listeria monocytogenes* and *Listeria innocua*, revealing both species-specific and conserved sRNAs. Several of these elements were conserved at the genomic level, but were expressed in only one of the species. This suggests that there is a difference between pathogenic and non-pathogenic species, which indicates that regulation by ncRNAs may be an evolutionary strategy allowing species to rapidly adapt to environmental changes (Wurtzel et al. 2012). The aspects that these different studies observed can be assumed to pertain to the *B. cereus* group, whose diversity (number of copies and families) may contribute to the establishment of such diverse ecological niches observed among species and strains even if no species-specific ncRNAs have been identified.

Besides the *in silico* analysis of ncRNAs within the *B. cereus s.l.*, we evaluated the expression of nine families of ncRNAs by RT-PCR; all of them were validated. These results are essential because they confirm the occurrence of these elements previously identified by computational methods. Also, besides being present in the genome, they are transcribed, which leads to a better understanding of these elements in this group of bacteria.

By analyzing the composition, distribution, and number of ncRNAs in *B. anthracis, B. cereus*, and *B. thuringiensis* strains, it was possible to obtain an overview of the diversity of these RNAs in this bacterial group. First, we can highlight the wider analysis of ncRNA families in bacteria that had not been classified so far. Furthermore, our results contribute to the discussion of the phylogenetic question of the *B. cereus* group, as they confirm previous studies indicating that: a) the three species are closely related (high number of common ncRNA families); b) *B. anthracis* is a clonal species (conserved profile of families and number of ncRNAs) and easily differentiated from *B. cereus* and *B. thuringiensis*; and c) *B. thuringiensis* has the genome with greater plasticity

(with more chromosomal ncRNAs and unique families) and *B. cereus* is in the intermediate position between the two other species. Finally, the acquisition and maintenance of genetic elements that present regulatory profiles in the genomes of *B. cereus* strains, especially those identified as *B. thuringiensis,* may lead to advantages regarding their ability to adapt, multiply, and disseminate in different environments. Further studies should address the expression of these genetic elements in different conditions.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00294-022-01240-4.

Acknowledgements The authors are grateful to the Universidade Estadual de Londrina and to the institutions that financed this paper. Gonçalves, K.B. was supported by fellowships from CAPES—finance code 001, grant number 88882.448512/2019-01. We would like to thank the reviewers who helped to improve this manuscript.

Funding Capes, 88882.448512/2019-01, Kátia Gonçalves.

Declarations

Competing interests The authors have not disclosed any competing interests.

References

Adang MJ, Crickmore N, Jurat-Fuentes JL (2014) Diversity of *Bacillus thuringiensis* crystal toxins and mechanism of action. In: Dhadialla TS, Gill SS (eds) Advances in insect physiology. Elsevier, Amsterdam, Netherlands, pp 39–87

Amin N, McGrath A, Chen Y-PP (2019) Evaluation of deep learning in non-coding RNA classification. Nat Mach Intell 1:246–256. https://doi.org/10.1038/s42256-019-0051-2

Argôlo-Filho RC, Loguercio LL (2014) *Bacillus thuringiensis* is an environmental pathogen and host-specificity has developed as an adaptation to human-generated ecological niches. Insects 5:62–91. https://doi.org/10.3390/insects5010062

Bardou P, Mariette J, Escudié F et al (2014) Jvenn: an interactive Venn diagram viewer. BMC Bioinformatics 15:1–7. https://doi.org/10.1186/1471-2105-15-293

Bazinet AL (2017) Pan-genome and phylogeny of *Bacillus cereus* sensu lato. BMC Evol Biol 17:1–16. https://doi.org/10.1186/s12862-017-1020-1

Cardoso P, Fazion F, Perchat S et al (2020) Rap-phr systems from plasmids pAW63 and pht8-1 act together to regulate sporulation in the *Bacillus thuringiensis* serovar kurstaki HD73 strain. Appl Environ Microbiol 86:1–16. https://doi.org/10.1128/AEM.01238-20

Carroll LM, Wiedmann M, Kovac J (2020) Proposal of a taxonomic nomenclature for the *Bacillus cereus* group which reconciles genomic definitions of bacterial species with clinical and industrial phenotypes. Mbio 11:1–15. https://doi.org/10.1101/779199

Chantsalnyam T, Lim DY, Tayara H, Chong KT (2020) ncRDeep: non-coding RNA classification with convolutional neural network. Comput Biol Chem 88:107364. https://doi.org/10.1016/j.compbiolchem.2020.107364

Chen Y, Indurthi DC, Jones SW, Papoutsakis E (2011) Small RNAs in the genus *Clostridium*. Mbio 2:1–11. https://doi.org/10.1128/mBio.00340-10

Crickmore N, Berry C, Panneerselvam S et al (2020) A structure-based nomenclature for *Bacillus thuringiensis* and other bacteria-derived pesticidal proteins. J Invertebr Pathol. https://doi.org/10.1016/j.jip.2020.107438

Darling ACE, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14:1394–1403. https://doi.org/10.1101/gr.2289704

Cardoso PdF, Perchat S, Vilas-Boas LA et al (2019) Diversity of the Rap-Phr quorum-sensing systems in the *Bacillus cereus* group. Curr Genet 65:1367–1381. https://doi.org/10.1007/s00294-019-00993-9

Drecktrah D, Hall LS, Brinkworth AJ et al (2020) Characterization of 6S RNA in the Lyme disease spirochete. Mol Microbiol 113:399–417. https://doi.org/10.1111/mmi.14427

Durand S, Braun F, Helfer AC et al (2017) sRNA-mediated activation of gene expression by inhibition of 5'-3' exonucleolytic mRNA degradation. Elife 6:1–23. https://doi.org/10.7554/eLife.23602

Ehling-Schulz M, Fricker M, Grallert H et al (2006) Cereulide synthetase gene cluster from emetic *Bacillus cereus*: structure and location on a mega virulence plasmid related to *Bacillus anthracis* toxin plasmid pXO1. BMC Microbiol 6:1–11. https://doi.org/10.1186/1471-2180-6-20

Fayad N, Kallassy Awad M, Mahillon J (2019) Diversity of *Bacillus cereus* sensu lato mobilome. BMC Genomics 20:1–11. https://doi.org/10.1186/s12864-019-5764-4

Fazion F, Perchat S, Buisson C et al (2018) A plasmid-borne Rap-Phr system regulates sporulation of *Bacillus thuringiensis* in insect larvae. Environ Microbiol 20:145–155. https://doi.org/10.1111/1462-2920.13946

Fiannaca A, La Rosa M, La Paglia L et al (2017) nRC: non-coding RNA Classifier based on structural features. BioData Min 10:1–18. https://doi.org/10.1186/s13040-017-0148-2

Geissler AS, Anthon C, Alkan F et al (2021) Bsgatlas: a unified *Bacillus subtilis* genome and transcriptome annotation atlas with enhanced information access. Microb Genomics 7:000524. https://doi.org/10.1099/mgen.0.000524

Geissmann T, Chevalier C, Cros MJ et al (2009) A search for small noncoding RNAs in *Staphylococcus aureus* reveals a conserved sequence motif for regulation. Nucleic Acids Res 37:7239–7257. https://doi.org/10.1093/nar/gkp668

Gillis A, Fayad N, Makart L et al (2018) Role of plasmid plasticity and mobile genetic elements in the entomopathogen *Bacillus thuringiensis* serovar israelensis. FEMS Microbiol Rev 42:829–856. https://doi.org/10.1093/femsre/fuy034

Gilois N, Ramarao N, Bouillaut L et al (2007) Growth-related variations in the *Bacillus cereus* secretome. Proteomics 7:1719–1728. https://doi.org/10.1002/pmic.200600502

Gripenland J, Netterling S, Loh E et al (2010) RNAs: regulators of bacterial virulence. Nat Rev Microbiol 8:857–866. https://doi.org/10.1038/nrmicro2457

Grissa I, Vergnaud G, Pourcel C (2007) The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinformatics 8:1–10. https://doi.org/10.1186/1471-2105-8-172

Guillet J, Hallier M, Felden B (2013) Emerging functions for the *Staphylococcus aureus* RNome. PLoS Pathog 9:1–13. https://doi.org/10.1371/journal.ppat.1003767

Guinebretière MH, Thompson FL, Sorokin A et al (2008) Ecological diversification in the *Bacillus cereus* Group. Environ Microbiol 10:851–865. https://doi.org/10.1111/j.1462-2920.2007.01495.x

Harris KA, Breaker RR (2018) Large noncoding RNAs in bacteria. Microbiol Spectr 6:515–526. https://doi.org/10.1128/microbiolspec.rwr-0005-2017

Helgason E, Økstad OA, Dominique A et al (2000) *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*–one species on the

basis of genetic evidence. Appl Environ Microbiol 66:2627–2630. https://doi.org/10.1128/AEM.66.6.2627-2630.2000

Hör J, Gorski SA, Vogel J (2018) Bacterial RNA biology on a genome scale. Mol Cell 70:785–799. https://doi.org/10.1016/j.molcel.2017.12.023

Irnov I, Sharma CM, Vogel J, Winkler WC (2010) Identification of regulatory RNAs in *Bacillus subtilis*. Nucleic Acids Res 38:6637–6651. https://doi.org/10.1093/nar/gkq454

Johnson M, Zaretskaya I, Raytselis Y et al (2008) NCBI BLAST: a better web interface. Nucleic Acids Res 36:5–9. https://doi.org/10.1093/nar/gkn201

Jolley KA, Bray JE, Maiden MCJ (2018) Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res 3:1–20. https://doi.org/10.12688/wellcomeopenres.14826.1

Kalvari I, Argasinska J, Quinones-Olvera N et al (2018) Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. Nucleic Acids Res 46:D335–D342. https://doi.org/10.1093/nar/gkx1038

Kolstø AB, Tourasse NJ, Økstad OA (2009) What sets *Bacillus anthracis* apart from other *Bacillus* species? Annu Rev Microbiol 63:451–476. https://doi.org/10.1146/annurev.micro.091208.073255

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874

Lee H, Zhang Z, Krause HM (2019) Long noncoding RNAs and repetitive elements: junk or intimate evolutionary partners? Trends Genet 35:892–902. https://doi.org/10.1016/j.tig.2019.09.006

Lehmann K, Schmidt U (2003) Group II introns: structure and catalytic versatility of large natural ribozymes. Crit Rev Biochem Mol Biol 38:249–303. https://doi.org/10.1080/713609236

Letunic I, Bork P (2019) Interactive tree of life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 47:256–259. https://doi.org/10.1093/nar/gkz239

Li S, Hwang XY, Stav S, Breaker RR (2016) The ydjF riboswitch candidate regulates gene expression by binding diverse azaaromatic compounds. RNA 22:530–541. https://doi.org/10.1261/rna.054890.115

Liu Y, Lai Q, Göker M et al (2015) Genomic insights into the taxonomic status of the *Bacillus cereus* group. Sci Rep 5:1–11. https://doi.org/10.1038/srep14082

Lu X, Liu SF, Yue L et al (2018) Epsc involved in the encoding of exopolysaccharides produced by *Bacillus amyloliquefaciens* FZB42 act to boost the drought tolerance of *Arabidopsis thaliana*. Int J Mol Sci 19:1–18. https://doi.org/10.3390/ijms19123795

Marincola G, Wencker F, Ziebuhr W (2019) The many facets of the small non-coding RNA RsaE (RoxS) in metabolic niche adaptation of gram-positive bacteria. J Mol Biol 431:4684–4698. https://doi.org/10.1016/j.jmb.2019.03.016

Mars RAT, Nicolas P, Denham EL, van Dijl JM (2016) Regulatory RNAs in *Bacillus subtilis*: a gram-positive perspective on bacterial RNA-mediated regulation of gene expression. Microbiol Mol Biol Rev 80:1029–1057. https://doi.org/10.1128/MMBR.00026-16

Mishra PK, Bisht SC, Ruwari P et al (2017) Genetic diversity and functional characterization of endophytic *Bacillus thuringiensis* isolates from the North Western Indian Himalayas. Ann Microbiol 67:143–155. https://doi.org/10.1007/s13213-016-1244-0

Moayeri M, Leppla SH, Vrentas C et al (2015) Anthrax pathogenesis. Annu Rev Microbiol 69:185–208. https://doi.org/10.1146/annurev-micro-091014-104523

Moore PB, Steitz TA (2002) The involvement of RNA in ribosome function. Nature 418:229–235. https://doi.org/10.1038/418229a

Mukherjee S, Stamatis D, Bertsch J et al (2019) Genomes online database (GOLD) vol 7: updates and new features. Nucleic Acids Res 47:D649–D659. https://doi.org/10.1093/nar/gky977

Nawrocki EP (2014) Annotating functional RNAs in genomes using infernal. In: Gorodkin J, Ruzzo WL (eds) RNA sequence, structure, and function: computational and bioinformatic methods. Springer Humana Press, NY, pp 163–197

Nawrocki EP, Kolbe DL, Eddy SR (2009) Infernal 1.0: inference of RNA alignments. Bioinformatics 25:1335–1337. https://doi.org/10.1093/bioinformatics/btp157

Nawrocki EP, Burge SW, Bateman A et al (2015) Rfam 12.0: updates to the RNA families database. Nucleic Acids Res 43:D130–D137. https://doi.org/10.1093/nar/gku1063

Nie P, Chen C, Yin Q et al (2019) Function of miR825 and miR825* as negative regulators in *Bacillus cereus* AR156-elicited systemic resistance to *Botrytis cinerea* in *Arabidopsis thaliana*. Int J Mol Sci 20:1–21. https://doi.org/10.3390/ijms20205032

Nieselt K, Herbig A (2013) Non-coding RNA, classification. In: Dubitzky W, Wolkenhauer O, Cho K-H, Yokota H (eds) Encyclopedia of systems biology. Springer, NY, pp 1532–1534

Nishida H (2012) Evolution of genome base composition and genome size in bacteria. Front Microbiol 3:1–3. https://doi.org/10.3389/fmicb.2012.00420

Nitzan M, Rehani R, Margalit H (2017) Integration of bacterial small RNAs in regulatory networks. Annu Rev Biophys 46:131–148. https://doi.org/10.1146/annurev-biophys-070816-034058

Nobukazu N, Tatsuhiko S, Gota K (2013) Translational control by small RNAs, bacteria. In: Dubitzky W, Wolkenhauer O, Cho KH, Yokota H (eds) Encyclopedia of systems biology. Springer, pp 2289–2292

Okinaka RT, Keim P (2016) The phylogeny of *Bacillus cereus* sensu lato. Microbiol Spectr 4:1–12. https://doi.org/10.1128/microbiolspec.tbs-0012-2012

Okonechnikov K, Golosova O, Fursov M et al (2012) Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28:1166–1167. https://doi.org/10.1093/bioinformatics/bts091

Patiño-Navarrete R, Sanchis V (2017) Evolutionary processes and environmental factors underlying the genetic diversity and lifestyles of *Bacillus cereus* group bacteria. Res Microbiol 168:309–318. https://doi.org/10.1016/j.resmic.2016.07.002

Peng D, Luo X, Zhang N et al (2018) Small RNA-mediated cry toxin silencing allows *Bacillus thuringiensis* to evade *Caenorhabditis elegans* avoidance behavioral defenses. Nucleic Acids Res 46:159–173. https://doi.org/10.1093/nar/gkx959

Pilo P, Frey J (2011) *Bacillus anthracis*: Molecular taxonomy, population genetics, phylogeny and patho-evolution. Infect Genet Evol 11:1218–1224. https://doi.org/10.1016/j.meegid.2011.05.013

Priest FG, Barker M, Baillie LWJ et al (2004) Population structure and evolution of the *Bacillus cereus* group. J Bacteriol 186:7959–7970. https://doi.org/10.1128/JB.186.23.7959-7970.2004

Rasko DA, Altherr MR, Han CS, Ravel J (2005) Genomics of the *Bacillus cereus* group of organisms. FEMS Microbiol Rev 29:303–329. https://doi.org/10.1016/j.femsre.2004.12.005

Raymond B, Federici BA (2017) In defense of *Bacillus thuringiensis*, the safest and most successful microbial insecticide available to humanity–a response to EFSA. FEMS Microbiol Ecol 93:1–8. https://doi.org/10.1093/femsec/fix084

Sajid M, Geng C, Li M et al (2018) Whole-genome analysis of *Bacillus thuringiensis* revealing partial genes as a source of novel cry toxins. Appl Enviromental Microbiol 84:1–15. https://doi.org/10.1128/AEM.00277-18

Schnepf E, Crickmore N, Van Rie J et al (1998) *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol Mol Biol Rev 62:775–806. https://doi.org/10.1128/MMBR.62.3.775-806.1998

Silvaggi JM, Perkins JB, Losick R (2006) Genes for small, noncoding RNAs under sporulation control in *Bacillus subtilis*. J Bacteriol 188:532–541. https://doi.org/10.1128/JB.188.2.532

Skilandat M, Sigel RKO (2013) Ribozymes. In: Maloy S, Hughes K (eds) Brenner's encyclopedia of genetics, 2nd edn. Elsevier, Amsterdam, Netherlands, pp 254–258

Sridhar J, Gayathri M (2019) Transcriptome based identification of silver stress responsive sRNAs from *Bacillus cereus* ATCC14579. Bioinformation 15:474–479. https://doi.org/10.6026/97320630015474

Svenningsen SL (2018) Small RNA-based regulation of bacterial quorum sensing and biofilm ormation. Microbiol Spectr 6(4). https://doi.org/10.1128/microbiolspec.RWR-0017-2018.PMID: 30003870

Tang Q, Yin K, Qian H et al (2016) Cyclic di-GMP contributes to adaption and virulence of *Bacillus thuringiensis* through a riboswitch-regulated collagen adhesion protein. Sci Rep 6:1–12. https://doi.org/10.1038/srep28807

Toro N, Martínez-Abarca F, Molina-Sánchez MD et al (2018) Contribution of mobile group II introns to *Sinorhizobium meliloti* genome evolution. Front Microbiol 9:1–8

Tourasse NJ, Helgason E, Klevan A et al (2011) Extended and global phylogenetic view of the *Bacillus cereus* group population by combination of MLST, AFLP, and MLEE genotyping data. Food Microbiol 28:236–244. https://doi.org/10.1016/j.fm.2010.06.014

Van der Auwera GA, Feldgarden M, Kolter R, Mahillon J (2013) Whole-genome sequences of 94 environmental isolates of *Bacillus cereus* sensu lato. Genome Announc 1:1–4. https://doi.org/10.1128/genomeA.e00380-13

Vilas-Bôas G, Vilas-Bôas L, Lereclus D, Arantes O (1998) *Bacillus thuringiensis* conjugation under environmental conditions. FEMS Microbiol Ecol 25:369–374. https://doi.org/10.1016/S0168-6496(98)00005-1

Wang P, Zhang C, Guo M et al (2014) Complete genome sequence of *Bacillus thuringiensis* YBT-1518, a typical strain with high toxicity to nematodes. J Biotechnol 171:1–2. https://doi.org/10.1016/j.jbiotec.2013.11.023

Wolf IR, Paschoal AR, Quiroga C et al (2018) Functional annotation and distribution overview of RNA families in 27 *Streptococcus agalactiae* genomes. BMC Genomics 19:1–11. https://doi.org/10.1186/s12864-018-4951-z

Wurtzel O, Sesto N, Mellin JR et al (2012) Comparative transcriptomics of pathogenic and non-pathogenic *Listeria* species. Mol Syst Biol 8:1–14. https://doi.org/10.1038/msb.2012.11

Xia L, Xia W, Li S et al (2012) Identification and expression of small non-coding RNA, L10-Leader, in different growth phases of *Streptococcus mutans*. Nucleic Acid Ther 22:177–186. https://doi.org/10.1089/nat.2011.0339

Xu LN, Ling YH, Wang YQ et al (2015) Identification of differentially expressed microRNAs between *Bacillus thuringiensis* Cry1Ab-resistant and -susceptible strains of *Ostrinia furnacalis*. Sci Rep 5:1–15. https://doi.org/10.1038/srep15461

Zheng Z, Zheng J, Zhang Z et al (2016) Nematicidal spore-forming bacilli share similar virulence factors and mechanisms. Sci Rep 6:1–9. https://doi.org/10.1038/srep31341

Zheng Z, Zhang Y, Liu Z et al (2020) The CRISPR-Cas systems were selectively inactivated during evolution of *Bacillus cereus* group for adaptation to diverse environments. ISME J 14:1479–1493. https://doi.org/10.1038/s41396-020-0623-5

Zhong C, Peng D, Ye W et al (2011) Determination of plasmid copy number reveals the total plasmid DNA amount is greater than the chromosomal DNA amount in *Bacillus thuringiensis* YBT-1520. PLoS ONE 6(1):e16025. https://doi.org/10.1371/journal.pone.0016025

Zur Bruegge J, Einspanier R, Sharbati S (2017) A long journey ahead: long non-coding RNAs in bacterial infections. Front Cell Infect Microbiol 7:1–6. https://doi.org/10.3389/fcimb.2017.00095