

RESEARCH ARTICLE

# dRNA-Seq Reveals Genomewide TSSs and Noncoding RNAs of Plant Beneficial Rhizobacterium *Bacillus amyloliquefaciens* FZB42

Ben Fan<sup>1\*</sup>, Lei Li<sup>2☯</sup>, Yanjie Chao<sup>2☯</sup>, Konrad Förstner<sup>2</sup>, Jörg Vogel<sup>2</sup>, Rainer Borriss<sup>3</sup>, Xiao-Qin Wu<sup>1\*</sup>

**1** Co-Innovation Center for Sustainable Forestry in Southern China, College of Forestry, Nanjing Forestry University, 210037 Nanjing, China, **2** RNA Biology Group, Institute for Molecular Infection Biology, University of Würzburg, 97080 Würzburg, Germany, **3** Fachgebiet Phytomedizin, Albrecht Daniel Thaer Institut für Agrar- und Gartenbauwissenschaften, Lebenswissenschaftliche Fakultät, Humboldt Universität zu Berlin, 14195 Berlin, Germany

☯ These authors contributed equally to this work.

\* [fanben2000@gmail.com](mailto:fanben2000@gmail.com) (BF); [xqwu2000@gmail.com](mailto:xqwu2000@gmail.com) (XW)



**OPEN ACCESS**

**Citation:** Fan B, Li L, Chao Y, Förstner K, Vogel J, Borriss R, et al. (2015) dRNA-Seq Reveals Genomewide TSSs and Noncoding RNAs of Plant Beneficial Rhizobacterium *Bacillus amyloliquefaciens* FZB42. PLoS ONE 10(11): e0142002. doi:10.1371/journal.pone.0142002

**Editor:** Lennart Randau, Max-Planck-Institute for Terrestrial Microbiology, GERMANY

**Received:** July 13, 2015

**Accepted:** October 15, 2015

**Published:** November 5, 2015

**Copyright:** © 2015 Fan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All detailed TSS information is available from the NCBI Gene Expression Omnibus (GEO) with accession number GSE66681.

**Funding:** The financial support by the National Natural Science Foundation of China (No. 31100081) and Scientific Research Foundation of Nanjing Forestry University for Introduction of High Level Talents (YJ2012-03) is gratefully acknowledged, as well as the financial support given by the Priority Academic Program Development (PAPD) of Jiangsu Higher Education Institutions. This study was further

## Abstract

*Bacillus amyloliquefaciens* subsp. *plantarum* FZB42 is a representative of Gram-positive plant-growth-promoting rhizobacteria (PGPR) that inhabit plant root environments. In order to better understand the molecular mechanisms of bacteria-plant symbiosis, we have systematically analyzed the primary transcriptome of strain FZB42 grown under rhizosphere-mimicking conditions using differential RNA sequencing (dRNA-seq). Our analysis revealed 4,877 transcription start sites for protein-coding genes, identified genes differentially expressed under different growth conditions, and corrected many previously mis-annotated genes. We also identified a large number of riboswitches and *cis*-encoded antisense RNAs, as well as *trans*-encoded small noncoding RNAs that may play important roles in the gene regulation of *Bacillus*. Overall, our analyses provided a landscape of *Bacillus* primary transcriptome and improved the knowledge of rhizobacteria-host interactions.

## Introduction

Rhizobacteria are a heterogeneous group of bacteria directly colonizing or living in the vicinity of plant roots. Many rhizobacteria are beneficial to plants as they stimulate plant growth and/or protect plants from phytopathogens. Rhizobacteria have been developed as microbial fertilizer and biocontrol agents. Understanding the symbiotic relationship between rhizobacteria and their host plants has been a longstanding research interest in environmental microbiology. Among a large variety of plant inhabitants, the Gram-positive rhizobacterium *B. amyloliquefaciens* subsp. *plantarum* FZB42 was extensively used as a model of plant-growth-promoting organisms to study the mechanisms of rhizobacterial biocontrol activity [1–6]. However, while

supported by the Natural Science Foundation of Jiangsu Province (No. BK20151514).

**Competing Interests:** The authors have declared that no competing interests exist.

genomic studies have revealed novel gene clusters and informed the pathways for the synthesis of natural products of this organism, our knowledge of the structure and expression of these genes has been restricted to a series of individual investigations [2, 4, 6–11].

In order to adapt to the fast-changing environment surrounding plant roots, rhizobacteria need a tuneable regulation system that can turn on or off gene expression in a dynamic fashion. To understand the mechanisms of such a gene regulation, it is critical to know where the transcription starts. Global mapping of transcription starting sites (TSSs) can facilitate the identification of promoter regions, *cis*-acting regulatory elements and cognate transcriptional regulators, and thereby helping to design reporter constructs. We here demonstrate the first global TSS map of a rhizobacterial *Bacillus* species using a recently developed differential RNA-seq (dRNA-seq) approach [12]. This approach can differentiate the 5' end status between primary transcripts (5'-ppp) and processed transcripts (5'-p) by exploiting a terminator exonuclease (TEX) that only degrades 5' monophosphorylated RNA. By splitting a total RNA sample and treating each half with or without TEX, primary transcripts will be enriched in the TEX-treated samples due to the selective removal of processed fragments [12]. Then, the TSSs of primary transcripts are identified in a single-nucleotide resolution using next-generation sequencing technology. Similar approaches have been used to generate high-quality transcriptome maps of diverse bacterial species, including Gram-negative species such as *Helicobacter pylori* [12], *Escherichia coli* [13], *Salmonella typhimurium* [14], *Chlamydia pneumoniae* [15], *Xanthomonas campestris* [16], and Gram positive bacteria like *B. subtilis* [17] and *Mycobacterium tuberculosis* [18].

In addition to charting the landscape of expressed protein-coding genes, dRNA-seq also provides a tool for a rapid genome-wide discovery of small noncoding RNAs (sRNAs). Over the last decade, sRNAs have been shown to play important roles in bacterial adaptation to changing environments. To date, sRNAs have been extensively characterized in the Gram-negative model bacteria like *E. coli* or *Salmonella enterica* where their total count is approaching 300 [19–21]. Much fewer sRNAs have been functionally characterized in Gram-positive organisms. It has also been argued that riboregulation in Gram-positive bacteria relies on *cis*-acting elements such as riboswitches rather than *trans*-acting sRNAs, partially because the genetic inactivation of Hfq, an RNA chaperone facilitating basepairing interactions between sRNAs and their target mRNAs [22], seems to have a minor effect in Gram-positive species [23–26], as compared to Gram-negative bacteria. Nonetheless, increasing numbers of sRNAs have been identified in Gram-positive organisms [17, 27–40]. For example, 62 novel sRNA candidates were identified in the first RNA-seq study of *B. subtilis* [17], bringing the number of known sRNAs in this model organism to 80, with 30 of them having been confirmed experimentally [17, 31–40].

A few sRNAs in *B. subtilis* have been characterized in detail [29, 35, 41], and demonstrated as crucial regulators in diverse physiological circuits. For example, the SR1 sRNA (a.k.a. BsrF), which itself is activated by the *B. subtilis* transcriptional regulator CodY, represses the translation of the AhrC transcription activator thereby regulating arginine catabolism [35, 42]; the RnaC/S1022 sRNA modulates the expression of transcriptional regulator AbrB to suppress exponential growth and to promote heterogeneity of *Bacillus* population [41]. Given these established roles of sRNAs in modulating gene expression, we reasoned that sRNAs may also be active in bacilli that live in a complex environment such as the rhizosphere. Here, we present the primary transcriptome of *B. amyloliquefaciens* FZB42 grown under several different conditions. Our dRNA-seq analysis led to an identification of close to 5,000 TSSs in a genome with approx. 3,700 mRNA genes [1]. This improves our knowledge of gene regulation events that enable *Bacillus* to respond to environmental cues. Moreover, we identified and further experimentally validated 21 new sRNAs, whose functions remain elusive. To our knowledge, this is

so far the most extensive transcriptomic study of plant associated *B. amyloliquefaciens*, a group of bacteria with fundamental importance in plant-microbe interaction.

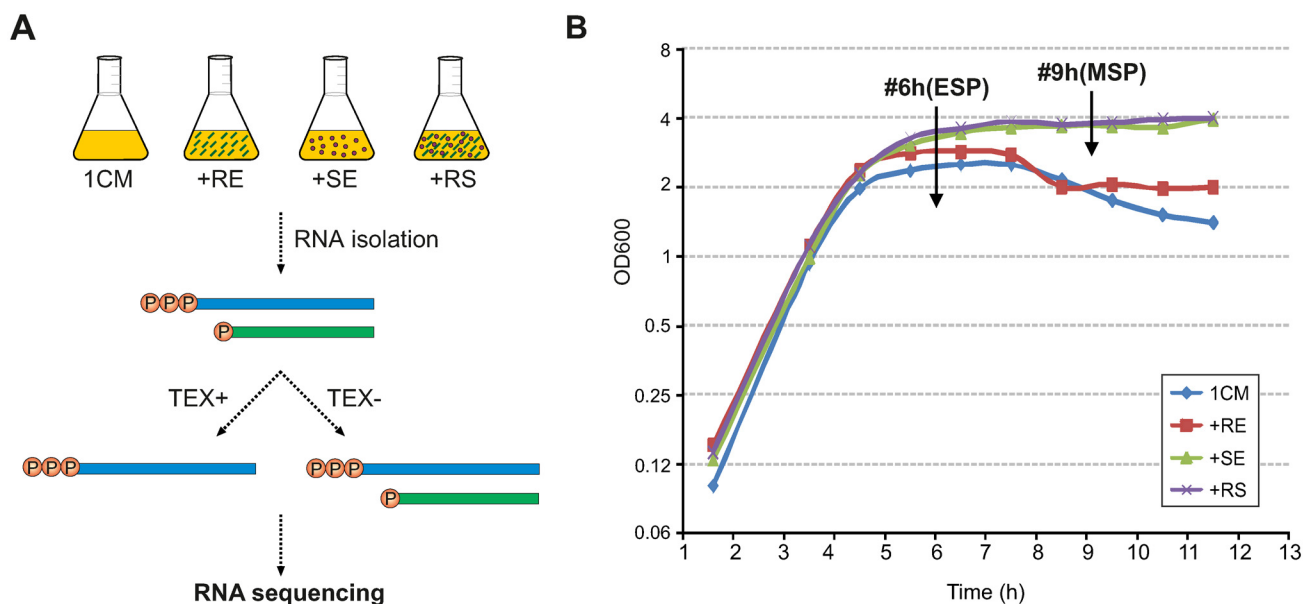
## Materials and Methods

### Root exudates and bacterial cultures

*B. amyloliquefaciens* FZB42 was cultured under four conditions [43]: i) in 1CM medium (1% peptone, 0.05% yeast extract, 0.5% NaCl); ii) 1CM medium supplemented with maize root exudates (RE) of 0.25 mg/ml; iii) 1CM medium supplemented with 10% soil extract (SE) prepared from soil collected from the farmland (with the permission of farmland owner Mrs Constanze Ackermann) in Wuerzburg, Germany; iv) 1CM medium supplemented with both the maize root exudates (0.25 mg/ml) and the soil extract (10%). The maize root exudates were collected from the cultivar “DengHai 11” as described in our previous work [43]. The cultures were incubated at 210 rpm and 28°C and collected at early stationary phase and at middle stationary phase, respectively, for total RNA preparation (Fig 1).

### Total RNA preparation

The bacterial cells harvested were immediately mixed with 0.2V STOP buffer (95% ethanol + 5% phenol). After a quick shake, the cultures were snap-frozen in liquid nitrogen and then stored at -80°C until RNA isolation. To isolate total RNA, the frozen cultures were half-melted on ice and then centrifuged at 5,000rpm for 10 minutes at 4°C. The pellets were suspended with ice-cold Trizol<sup>®</sup> agent and then transferred to a cryo-tube filled with 0.2V sterile glass beads of 0.1 mm in diameter. Subsequently, the mixtures were milled on FastPrep<sup>®</sup>-24 system for twice of 30 seconds at speed of 6.0 m/s with a one-minute interval cooling on ice. The



**Fig 1. Growth conditions of *Bacillus amyloliquefaciens* FZB42.** (A) Schematic diagram of the dRNA-seq workflow of *B. amyloliquefaciens* FZB42. FZB42 was cultured in four media: i) 1CM medium; ii) 1CM medium supplemented with maize root exudates (+RE); iii) 1CM medium supplemented with soil extract (+SE); iv) 1CM medium supplemented with both maize root exudates and soil (+RS). (B) Growth curves of FZB42 growing in the four media. The cultures were collected for total RNA preparation at different time points (#6h & #9h) as indicated by the arrows. TEX: terminator exonuclease; ESP: early stationary phase; MSP: middle stationary phase.

doi:10.1371/journal.pone.0142002.g001

milled cells-containing Trizol<sup>®</sup> agent was continued to isolate total RNA as previously described [44].

### Deep sequencing

The isolated RNA samples were initially quality filtered, and then used to construct cDNA libraries by Vertis Biotechnologie AG (<http://www.vertisbiotech.com>), Germany, as described in the supporting information of reference [12]. Sequencing was performed on a Solexa HiSeq 2000 machine. Raw cDNA reads were quality trimmed and poly-A clipped. The remaining clean reads were mapped to the *B. amyloliquefaciens* FZB42 genome (NC\_009725.1) using READemption (version 0.3.4) with “create”, “align” and “coverage” subcommands. The align subcommand was called with the parameters (`—split -r -p 24 -a 95 -l 12—poly_a_clipping -progress`) [45]. The coverage per nucleotide was calculated, normalized by the number of total number of mapped of reads for each library and visualized in the Integrated Genome Browser [46].

### TSS prediction and differential expression analysis

TSSs were predicted by TSSpredator [47] which use a ratio-based approach to identify the enriched sites in TEX treated samples compared to TEX untreated samples. Based on their locations relative to annotated genes, TSSs were classified as primary TSS, secondary TSS, anti-sense TSS, internal TSS and orphan TSS. The TSS with the strongest expression and within 500 nucleotides upstream of a gene was considered as primary. The TSS that associated with same gene was considered as secondary. Antisense TSS was located internal or within 100 nucleotides of a gene on the antisense strand, while internal TSS was located inside a gene on the sense strand. The other TSS without any annotated gene nearby was orphan TSS. To identify the differential expressed genes under two different conditions (A&B), we defined the differential expressed genes as those passing three filters: i) fold change of Reads Per Kilobase Per Million mapped reads (RPKM) under two conditions (condition A and condition B) was greater than 2 (for simplicity, we refer to this as  $RPKM_{A/B} > 2$  throughout the article), and ii) RPKM of the gene in condition A was greater than 2 (refer to  $RPKM_A > 2$ ), and iii) raw reads of the gene in condition A was greater than 50 (refer to as  $Raw\ reads_A > 50$ ). The RPKM values were calculated from the untreated libraries.

### Radioactive labelling of oligonucleotides

Oligonucleotides were radioactively-labelled at their 5'-OH ends by T4 polynucleotide kinase (T4 PNK) that catalyses transfer of  $\gamma$ -phosphate from  $^{32}P$ -ATP. For this, 40 pmol of oligos were mixed with 4  $\mu$ l [ $\gamma$ - $^{32}P$ ] ATP (10  $\mu$ Ci/ml) and phosphorylation took place by incubation of the mixture with T4-Kinase at 37°C for 30 minutes. The reaction was stopped by heat inactivation at 70°C for 10 minutes.

### sRNA detection by northern blotting

The total RNA samples were separated (10  $\mu$ g/each sample) on 6% PAA 7M urea gel in 1 $\times$ TBE buffer. The samples were denatured at 95°C for 5 minutes and then cooled on ice for another 5 minutes. After running the RNAs were transferred to a positively charged nylon membrane. Finally the RNAs were immobilized on the membrane by cross-linking using UV radiation. The membrane was initially incubated in 15 ml QIAquick hybridization buffer for 1 hour at 42°C, while the radioactively-labelled oligo probes were denatured at 95°C for 5 minutes and then immediately cooled on ice to unfold the secondary structures. Subsequently, the

membrane was hybridized overnight with 1  $\mu$ l denatured oligo probes at 42°C. The membrane was washed three times at 42°C, each for 15 minutes, with 2 $\times$  SSC/0.1% SDS, 1 $\times$  SSC/0.1% SDS, and 0.5 $\times$  SSC/0.1% SDS, respectively. The results were visualized by Typhoon FLA 9500 scanner (GE Healthcare).

## KEGG Pathway Analysis

The differentially expressed genes with various fold changes (logarithmic) were mapped into KEGG pathway using the GSEA method [48], a gene set enrichment analysis based approach for determining the significant pathways. The significant pathways with  $p$ -value smaller than 0.1 were displayed using Pathview [49].

## Data deposition

Raw sequence reads were stored in NCBI Gene Expression Omnibus (GEO) with accession number GSE66681.

## Results

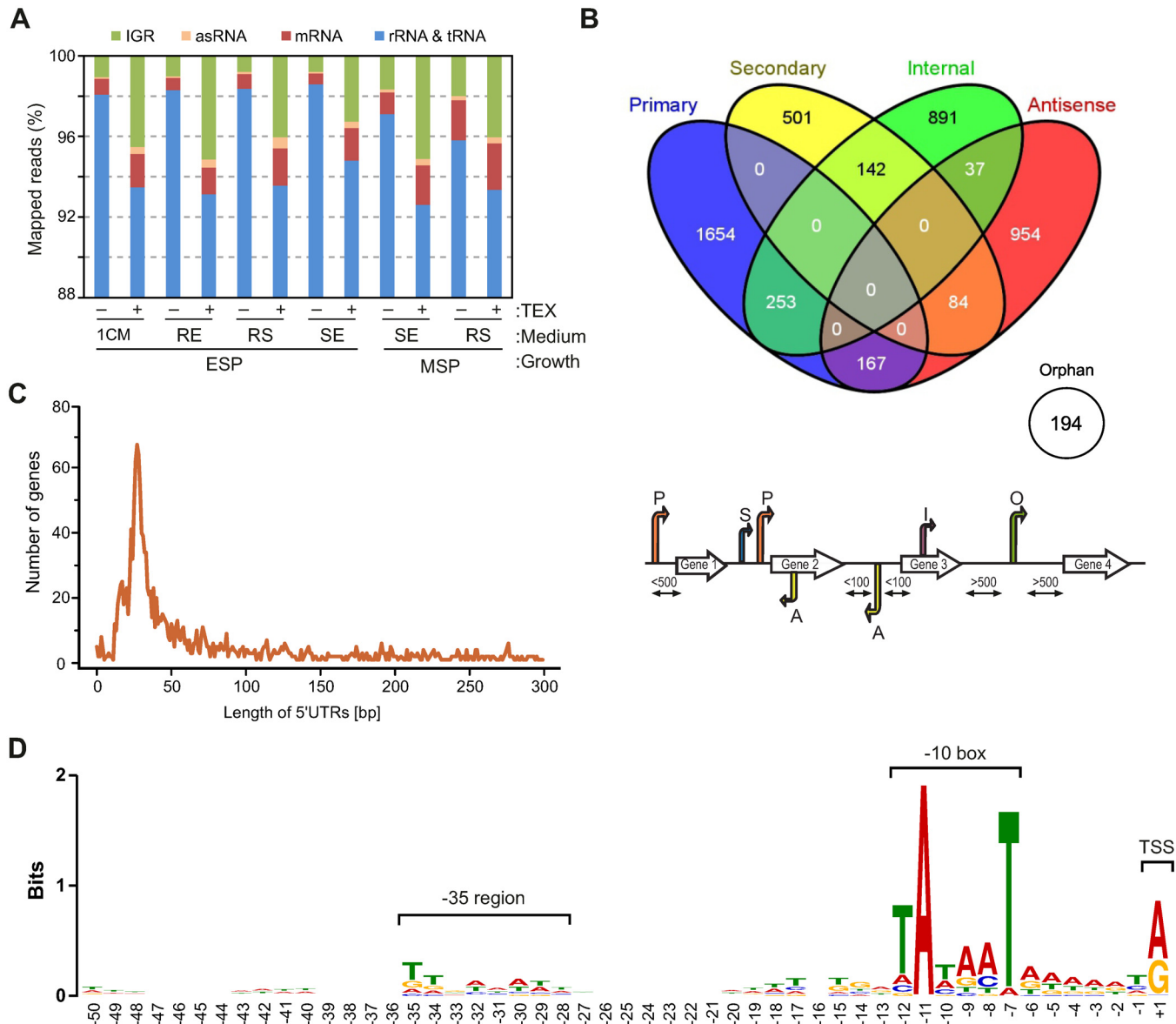
### Genome-wide TSS Analysis

In order to mimic natural plant-root environments and to promote the expression of genes important in rhizosphere ecosystem, we cultivated *B. amyloliquefaciens* FZB42 in 1CM medium supplemented with either maize root exudates (RE), or soil extract (SE), or both (RE +SE, a.k.a. "RS"). We determined the kinetics of bacterial growth under four different conditions and collected bacterial cultures for total RNA isolation at the indicated time points (Fig 1). Based on the growth curves, the sampling points at ~6 hours roughly correspond to the early stationary phase (ESP) for SE and RS samples, and to the middle stationary phase (MSP) for 1CM and RE samples. Since we wanted to identify genes affected by RE at a similar growth stage, two additional samples were further collected at MSP under SE and RS conditions (at ~9 hours), resulting in six bacterial samples in total. For simplicity, we hereafter refer to these samples using medium features (*i.e.*, 1CM, RE, SE, RS) and sampling time (*i.e.*, 6h or 9h) connected with a "#" symbol, *e.g.*, 1CM#6h (S1 Table). Following the published dRNA-seq cDNA synthesis protocol [12], we split each total RNA sample into two aliquots, one of which was treated with terminator exonuclease (TEX). A total of 12 cDNA libraries were constructed and deep sequenced to ~120 million reads using Illumina technology.

After a standard quality check and adaptor removal, cDNA reads were mapped to the FZB42 genome using the READemption pipeline [45]. The number of reads allocated to each gene annotation was quantified and the sum of reads of each RNA class was calculated (S1 Table). We have observed a high mappability for all these cDNA libraries as most reads (>91%) could be mapped to the FZB42 genome. As expected, the TEX-treated (TEX+) libraries contain a lower percentage of housekeeping rRNA/tRNA reads (Fig 2A), arguing that the TEX treatment successfully depleted processed transcripts (mostly rRNA/tRNA). We have detected 4,877 TSSs in total, which were grouped into five different categories according to [12], including 2,074 (43%) primary TSSs, 727 (15%) secondary TSSs, 1,323 (27%) internal TSSs, 1,424 (25%) antisense TSSs, and 194 (4%) orphan TSSs (S2 Table). As expected, some TSSs belong to multiple categories (Fig 2B).

The 2,074 primary TSSs detected in all six conditions were assigned to transcripts for ~56% of the annotated 3,701 genes in the FZB42 genome. We randomly selected 10 genes with experimentally mapped TSSs to verify the reliability of TSS designation. The result showed that our analysis accurately captured all these known TSS, confirming the robustness of our TSS





**Fig 2. General profile of the sequencing data.** Distribution of the mapped reads of various RNAs from each library varying in different culture conditions or terminator exonuclease treatment. 1CM: 1CM medium; +RE: 1CM medium supplemented with maize root exudates (RE); +SE: 1CM medium supplemented with soil extract (RE); +RS: 1CM medium supplemented with both maize root exudates and soil (RS); TEX: terminator exonuclease; ESP: early stationary phase; MSP: middle stationary phase. (B) Venn diagram showing overlaps among various TSS categories. The TSS were designated into the categories as defined in [12]. (C) Length distribution of 5' leader sequence in *B. amyloliquefaciens* FZB42. (D) Motif analysis among upstream of TSSs of *B. amyloliquefaciens* FZB42. Sequences upstream of TSSs were extracted (positions -1 to -50) and common sequence motifs were searched using MEME software. The coordinates give positions relative to TSSs.

doi:10.1371/journal.pone.0142002.g002

annotation (S3 Table). Eighty of the primary TSSs were detected in only one specific condition, predominantly (71, ~90%) in the SE#9h condition (S4 Table). The assigned primary TSSs allowed us to determine the 5' untranslated region (UTR) for 1,958 mRNAs. The majority of mRNAs possess a 5'UTR sequence between 10 to 50 nucleotides in length (Fig 2C), indicating an optimal arrangement to support efficient translation in *Bacilli*, similar to the situation in many Gram-negative bacteria [12, 15, 50]. 445 mRNAs carry a longer 5'UTR sequence (>100 nucleotides) which may harbor potential regulatory elements for post-transcriptional

regulation *in cis* or *in trans*. This number should be taken with some caution, though, since many of these genes (>50%) encode hypothetical proteins whose start codon annotation may need correction. By contrast, 28 mRNAs possess no or a very short 5'UTR (<10 nucleotides) and should be considered leaderless mRNAs (S5 Table).

Sequence analysis of the nucleotide composition for primary TSS revealed a strong selection of purines (A or G) at the +1 site, which is required for efficient transcription initiation by RNA polymerase, at least in many Gram-negative bacteria [51, 52]. Using all identified TSSs, we aligned their upstream sequences to identify potential elements signaling transcription activity. A typical TATA box (TATAAT) was identified at the -10 region, which further supported the high accuracy of our TSS mapping. A weaker motif (TTnAAAAtt) is observed at the -35 region, which is slightly different from the canonical TTGACA sequence in *B. subtilis* (Fig 2D). This may be due to a large assembly of different promoter sequences that are recognized by a variety of sigma factors.

Secondary TSSs are operationally defined as a TSS being located downstream of primary TSS but having fewer cDNAs [12]. Gene expression initiating from secondary TSSs indicates alternative promoters. More than 30% of the total secondary TSSs (727) were detected in each condition, while a higher percentage of 65% and 80% were found in conditions SE#9h and RS#9h, respectively (S6 Table). Excluding the overlap between different conditions, *i.e.*, only considering those genes specifically expressed in one condition, we identified a total of 95 secondary TSSs, which mostly fall in the conditions of SE#9h and RS#9h (28 and 59 TSSs, respectively). The major portion of secondary TSSs unique to condition SE#9h and RS#9h precedes a gene encoding a hypothetical protein (S6 Table), similar to our observation that two thirds of the primary TSSs detected only in SE#9h precede a gene of unknown function. Such information implies that a large part of the expressed genes at middle stationary phase remain elusive with respect to function.

## Differentially expressed genes at two growth phases

Upon entering stationary phase, *Bacillus* species initiate the sporulation gene expression programme to prepare for production of spores. The relevant RNA samples SE#9h and RS#9h were isolated from MSP while samples SE#6h and RS#6h were isolated from ESP. Therefore a comparison of gene transcription between the two time points will allow us to identify sporulation-associated genes differentially expressed in the two different growth stages. For scoring differential gene expression, we focused on transcripts passing our strict abundance filters (see experimental procedures). We identified 208 and 433 genes for the SE and RS conditions, respectively, showing strong up-regulation at MSP compared to ESP. 148 activated genes were common to both the SE and RS conditions. Nearly 40% of the common genes are of unknown function while 27% of them are associated with sporulation (S1 Fig).

Down-regulated genes at MSP were rare (5 genes in SE condition and 23 genes in RS condition) and only three (*spoIIB*, *spoIIE*, and *bpr*) of these overlapped in both conditions. The down-regulation of the three genes is consistent with previous studies [53, 54] where they were reported to drop in expression after sporulation start, roughly corresponded to MSP here.

To further characterize the 151 common genes (148 up-regulated, 3 down-regulated; S7 Table) that may be relevant to the sporulation process, we examined their upstream sequences for conserved binding sites of RNA polymerase sigma factors. Using knowledge from *B. subtilis*, we determined sigma factor binding motifs in 37 genes, 33 of which are known to be regulated by SigK or SigG. Moreover, based on the TSSs identified in this study, we predicted with the online tool [55] binding motifs for another 36 genes, 24 of which are also regulated by SigK or SigG (S7 Table). Activation of SigG in the forespore precedes and is required for the

subsequent activation of SigK in the mother cells. After activation, SigG and SigK are responsible for transcription of a series of genes in the forespores and the mother cells, respectively [56, 57]. No binding motif is predicted for the remaining half of the common genes, partly due to lack of defined TSSs. The high percentage of SigG/SigK-dependent genes (57 out of 73) suggests that the time of #9h sample roughly corresponded to stage III-IV (completion of engulfment and cortex synthesis) of the seven stages of sporulation cycle [58, 59]. This deduction in turn supports the suitability of our above designation the time of #9h as MSP.

## Riboswitches and *Cis*-encoded antisense RNAs

*Cis*-acting regulatory RNA elements encoded within 5' UTRs are prevalent in Gram-positive bacteria [60, 61]. By querying all 5'-UTRs longer than 70 nt against the Rfam database [62, 63], we identified 53 *cis*-encoded RNA elements (S8 Table) which include various conserved riboswitches (TPP, SAM riboswitch, purine riboswitch, glycine riboswitch, FMN riboswitch, preQ1 riboswitch, glmS riboswitch) and other leader sequence such as L10-, L13-, L19-, L20, and L21-leader, T-box leader, *Trp* leader\_2, *ykoK* leader, *ykkC-ykkD* leader [64], pan RNA motif [65], *yjdF* RNA [65], and PyrR binding site [66]. The 53 RNA elements validated most of the *cis*-regulatory riboswitches or leader sequences provided by the Rfam database for *B. amyloliquefaciens*, except *ylbH* leader and *pyrG* leader. The absence of these RNAs is probably due to a low expression of them in our specific growth conditions.

Antisense RNAs (asRNAs) regulate bacterial gene expression *in cis*, often by forming long RNA duplexes with the oppositely encoded mRNAs [67]. Manual inspection identified a total of 136 antisense RNA candidates, 90 of which are located antisense to coding regions (S9 Table). The remaining 46 antisense RNA candidates are opposite to a 5'UTR or 3'UTR (S10 Table). Eight asRNA candidates showed higher expression in the #9h samples as compared to the #6h samples (S11 Table); only two of them are opposite to genes of known function (*yocH* and *sspF*), both of which are specifically expressed at stationary growth phase.

## Gene re-annotation and operon prediction

A group of 46 highly expressed intergenic transcripts were detected. We identified six of them as tRNAs according to annotation in Rfam, including four that were previously annotated to encode hypothetical proteins. We blasted the remaining 40 transcripts against the non-redundant protein database for homologous proteins that have been annotated in other bacterial genomes. The blast results suggest that the 40 transcripts encode small proteins (30–99 aa), only a few of which have been functionally characterized. For example, we have detected *pznA*, predicted to encode a 41-aa protein and recently shown to be involved in synthesis of an antimicrobial peptide [3, 4]. Further, we found eight TSSs within 50 bp downstream of previously annotated start codons. Sequence analysis revealed a new start codon locating at 40–100 bp downstream of their TSSs and a plausible ribosomal binding site (RBS) locating at ~7 bp upstream of the new start codons. Accordingly, we propose to change the annotation for these eight genes, and for two other transcripts bearing an ORF antisense to previously annotated genes (S12 Table). Overall, we propose to add or re-annotate 46 undefined genes to the genome annotation of FZB42 (S13 Table) and to correct 10 gene annotation. In addition, we failed to detect any reliable reads in the coding region of gene *yhaY* (222 bp) may not exist since we failed to detect any reliable reads in its coding region but found a highly expressed transcript on its antisense strand. The expressed transcript is 223 nt in length and contains no ORF, suggesting it could be an sRNA which may silence *yhaY* completely in our tested conditions.

Some of these short transcripts from intergenic regions (IGRs) may have dual functions, acting as both mRNA and regulatory RNA. For example, two of the small protein coding



transcripts (*rho-glpX*, *yrzI-yrhG*) were previously identified as sRNAs [17]. Another two short IGR transcripts (*xylB-pps*, *cypC-yitS*; see S13 Table) containing an ORF are located antisense to the 3'UTR or terminator regions of their neighbor genes, indicating they may also function as an sRNA regulating *in cis* the expression of the neighboring genes.

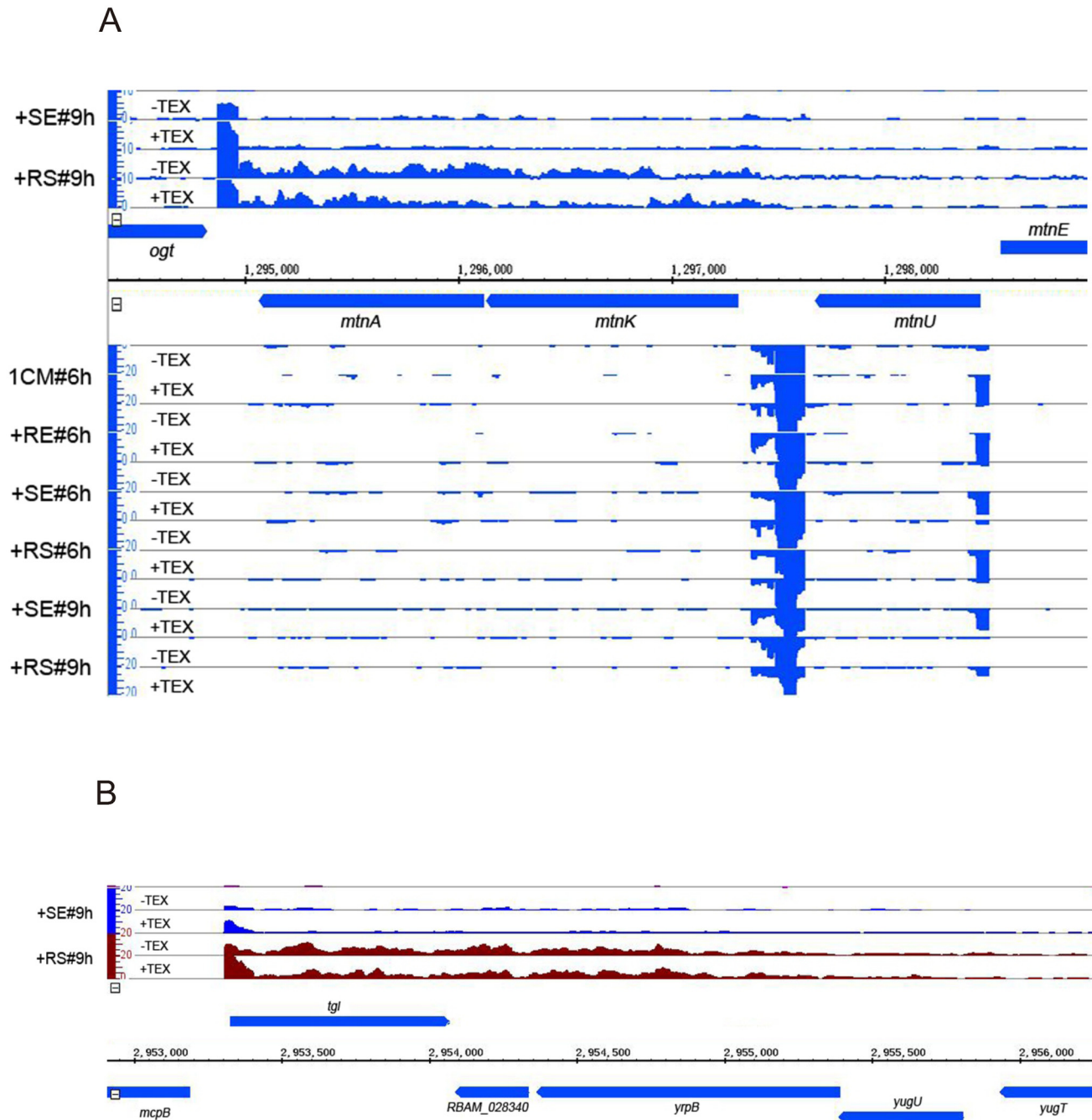
We found 730 operons for the FZB42 genome predicted by the DOOR database [68]. We manually checked the 730 operon loci with our visualized dRNA-seq data and thus validated 78 polycistronic transcripts covering a total of 210 genes (S14 Table). Most of the confirmed operons (53 out of the 78) are bicistronic. Intriguingly, we observed two long asRNA that resemble excludon structures (Fig 3A and 3B). Excludon is a transcriptome organization where divergent genes with mutually exclusive or related functions transcribe long antisense RNAs in order to suppress the expression of the other gene in the same locus [69, 70]. In addition, we found that 13 genes, which are listed within a polycistronic operon by DBTBS [55], were actually transcribed as a monocistronic unit (S15 Table). This indicates that alternative TSSs could be used to initiate their gene transcription in a certain environment. For example, the polycistronic operon of *yhaA-yhfA-yhgC* varied their transcript initiation from different TSSs at ESP (#6h) or MSP (#9h) (S2 Fig).

## Identification of *trans*-encoded small RNAs

sRNAs are known to be encoded by independent genes located within IGRs. To identify such sRNAs in FZB42, we screened over 500 abundant IGR transcripts (with more than 50 raw reads), excluding those extending from or into coding regions of their flanking genes. After further filtering out mis-annotated mRNAs, riboswitches and other leader sequences, or antisense RNAs, as explained above, we determined 86 candidate sRNAs. Thirteen of them were known sRNAs, which include the FZB42 homologs of three ubiquitous RNAs (M1 RNA of RNase P, 6S RNA [there is only one gene copy in strain FZB42], and tmRNA) and of seven experimentally confirmed sRNAs in *B. subtilis* (FsrA, RsaE, BsrG, BsrF, BsrI, polC-ylxS, and ncr1175). Three more potential homologs of *B. subtilis* sRNAs were detected in the IGRs of *ymzA-nrdI*, *yvcI-trxB*, and *yneK-cotM* [17]. By contrast, we failed to detect potential homologs of the BsrC or SurA sRNAs [31, 38]. BsrC was expressed in *B. subtilis* during the vegetative phase but undetectable during the stationary phase [38], which may explain why it was missing in our analysis; SurA, a ~280 nt-long sRNA from the IGR between *yndK* and *yndL* in *B. subtilis* [31], is not conserved in FZB42 probably due to genomic re-arrangement. Though FZB42 encodes a *yndL* homolog, the other flanking gene *yndK* is absent. However, a part of the SurA locus is expressed as a 94 nt-long transcript antisense to the *yndL* mRNA (S3 Fig).

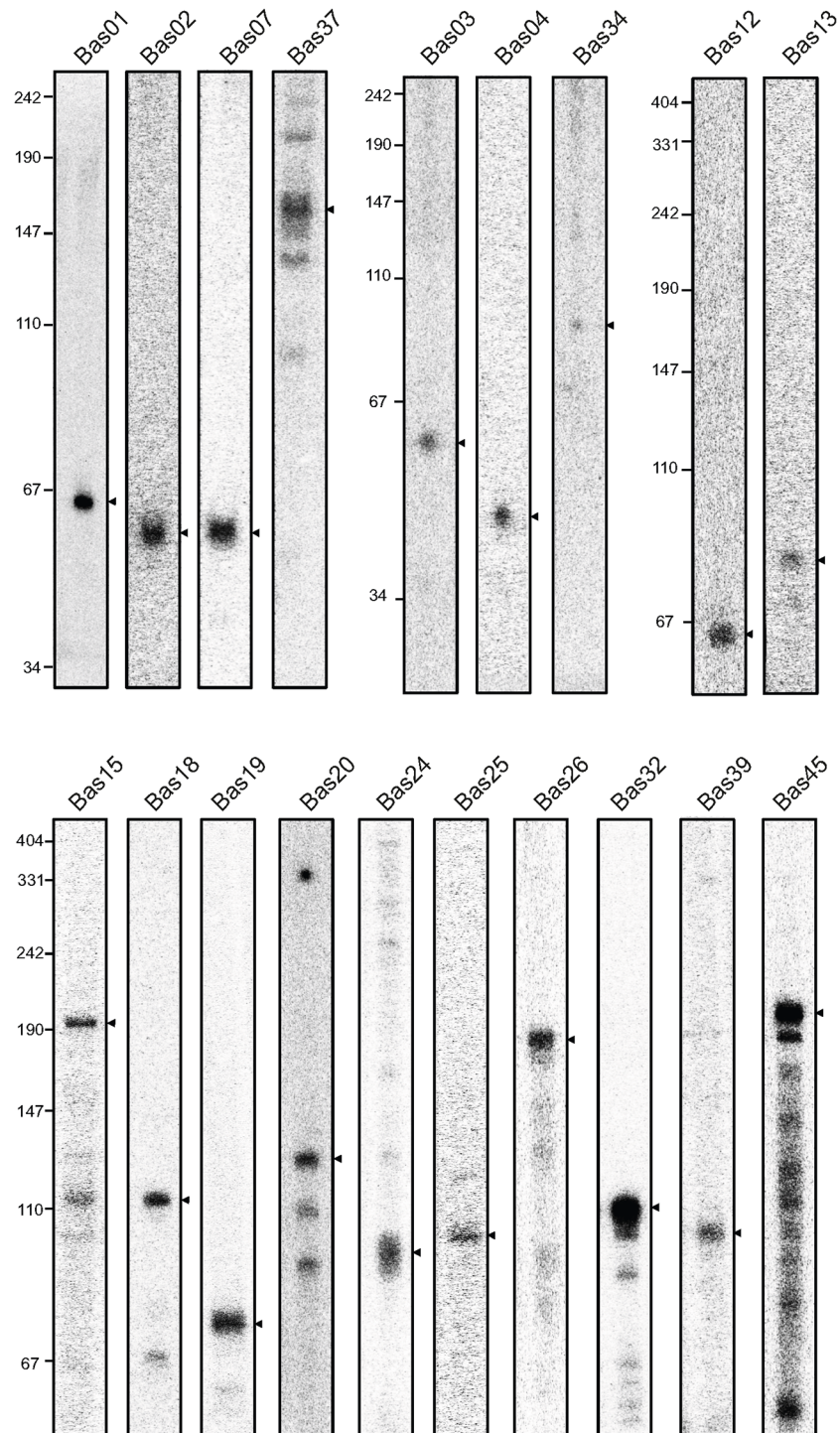
To validate our results, we probed on Northern blots 35 selected sRNA candidates, 19 of which were detected as distinct transcripts (Fig 4 and S16 Table). One of the candidates (Bas28) was also detected but with additional shorter bands, indicating sRNA processing (S4 Fig). A distinct transcript of >500 nt in length was identified in the *trpS-oppA* IGR and designated as Bas35; in this case, the detected transcript much exceeded the size predicted by the RNA-seq data (~95nt) (S5 Fig). Since we failed to find a convincing ORF on the sense strand within the IGR, Bas35 is probably a long noncoding RNA covering 5'-UTR of *oppA*. Intriguingly, Bas35 is located antisense to the T-box leader upstream of *trpS*, suggesting it may possess a regulatory role *in cis*.

The expression profiles of the new sRNAs along growth in different media give valuable hints of how these sRNAs are regulated by rhizosphere environments. Four sRNAs (Bas02, Bas03, Bas24, and Bas39) seemed to be affected by root exudates, and almost a half of all sRNAs (Bas02, Bas04, Bas19, Bas20, Bas23, Bas24, Bas26, Bas32, Bas39) were either positively or negatively regulated by soil extract (S6 Fig). Furthermore, many sRNAs (Bas01, Bas02,



**Fig 3. A subset of operons antisense to known genes.** Two operons identified antisense to a known operon (A) or to known genes (B). cDNA reads in each library without (-TEX) or with (+TEX) terminator exonuclease treatment were mapped to *B. amyloliquefaciens* FZB42 chromosome and visualized by IGB 8.1. Vertical bars indicate cDNA recovery detected in each library and were adjusted to the same scale of 20. The genes indicated above the horizontal coordinates are located on the forward strand while the genes indicated below the horizontal coordinates are located on the reverse strand. Bacterial cultures were sampled for total RNA extraction at two time points (#6h&#9h) from four conditions: i) in 1CM medium; ii) 1CM medium supplemented with the maize root exudates (+RE); iii) 1CM medium supplemented with soil extract (+SE); iv) 1CM medium supplemented with both maize root exudates and soil extract (+RS). The time point #6h corresponded to early stationary phase while #9h corresponded to middle stationary phase.

doi:10.1371/journal.pone.0142002.g003



**Fig 4. Verification of sRNAs in *B. amyloliquefaciens* FZB42.** sRNAs identified by sequencing data were validated by Northern blot. One representative lane was displayed for each sRNA identified. The names of sRNAs and the molecular marker were indicated on the top and on the left respectively. For more details of each sRNA see [S6 Fig](#).

doi:10.1371/journal.pone.0142002.g004

Bas07, Bas12, Bas19, Bas20, Bas26, Bas32 and Bas45) strongly accumulated in MSP (S6 Fig), with several of them showing the highest expression peak at the last sampling time point. These sRNAs may be involved in stationary phase-specific processes such as sporulation. For Bas26 (*yjcA*-RBAM\_011810) (S7 Fig), the intergenic region where its major cDNA reads occurred belongs to 3'UTR region of *yjcA* [71]. Thus we deduced that Bas26 could be a 3'UTR-derived sRNA, a demonstrated phenomenon in Gram-negative bacteria that 3'UTR serve as a genomic reservoir of sRNAs [44].

## Expression of genes involved in bacterium-plant interaction

*B. amyloliquefaciens* FZB42 is a model strain of plant beneficial rhizobacteria. To find out how FZB42 responds to root environment, we compared the gene expression profile from the RS and SE conditions at MSP (#9h). Our analysis revealed 361 genes that were regulated by the presence of maize root exudates. These genes, belonging to various functional categories, were mostly up-regulated by root exudates. For example, two of these genes, *tatCY* and *secY*, encoding components of Tat pathway and Sec pathway respectively, have been promoted in transcription. Their up-regulation indicates that FZB42 increased its secretion activity in the presence of root exudates. Only two genes were down-regulated by the root exudate; one encodes an endospore development protein (*spoIIB*) and the other one (*ywcl*) is of unknown function.

Among the 361 genes, 63 are reportedly implicated in plant-bacterium interactions (S17 Table). Our results supported this implication and, further, provided more information like TSSs, favoring future investigations on them. For example, FZB42 impressively devotes >8.5% of its genome to synthesis of antimicrobial metabolites [1, 4, 6], an important mechanism that many rhizobacteria protect plants from phytopathogens. The operons for three antimicrobial metabolites (bacillomycin D, plantazolicin, and diffidin) showed higher transcription (RPKM>2) than the others. We confirmed the published TSS of *bmyD* [7], but identified the previously unknown TSSs for plantazolicin and diffidin gene clusters (S8 Fig). Further, we identified the TSS of *yczE*. This highly transcribed gene encodes a putative membrane-spanning protein that is essential for the synthesis of diffidin, bacillaene, macrolactin, and bacillomycin D [7] (S8 Fig). In addition, our data showed that bacillomycin D and plantazolicin were expressed mainly in the #6h but not in the #9h samples, while diffidin was expressed at both the #6h and #9h time points.

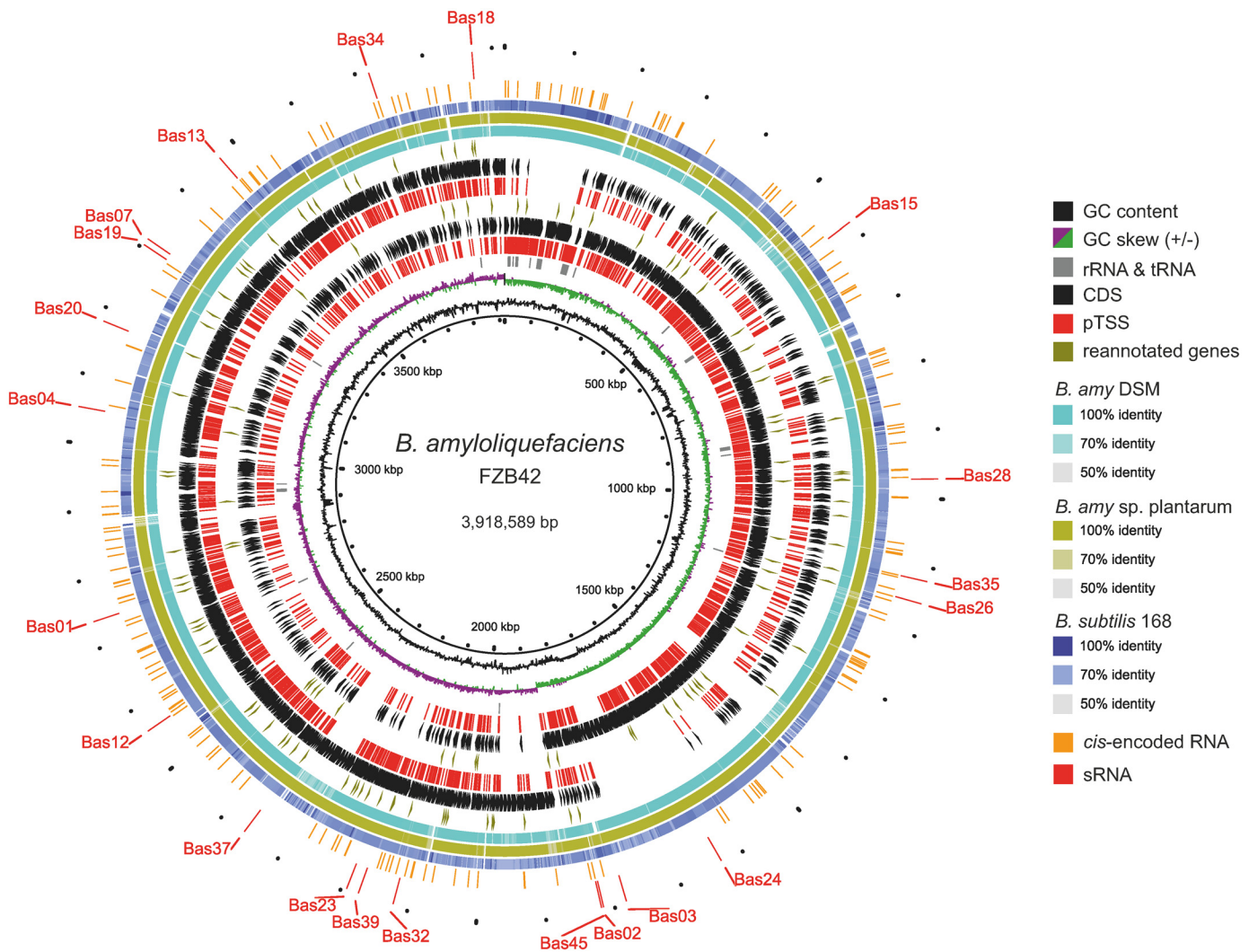
Other highly expressed genes in S17 Table include those involved in degrading plant derived organic substrates, such as celluloses, semicelluloses, chitin, opines, and plant proteins; the induction of these genes reflects the ability of rhizobacteria to utilize plant-derived nutrients which are essential for them to thrive in the rhizosphere. Some genes involved in swarming motility and biofilm formation were also highly expressed (S17 Table). These genes are usually related efficient colonization of plant roots, a prerequisite for rhizobacteria to interplay with plants. A FZB42 unique gene cluster (RBAM\_007740 to RBAM\_007770) with no counterpart in *B. subtilis*, encodes proteins containing a collagen-related GXT structure motif. This cluster, required for biofilm formation and adhesion to plant roots [72], was highly expressed in #9h but not in #6h samples (S8 Fig). Further, some genes involved in response to oxidative stress (e.g., *sodA* and *tpx*) or adaption to atypical stresses (*cspB*, *cspC*, and *cspD*) were also highly induced; these genes may help the rhizobacteria to survive the plant basal defense.

## Discussion

*B. amyloliquefaciens* FZB42 is a Gram-positive representative of rhizobacteria that colonize plant roots and promote plant growth. Understanding its transcriptome structure and



dynamics is critical to study the symbiotic relationship between rhizobacteria and host plants. In this work we have employed dRNA-seq and investigated the primary transcriptome of strain FZB42 under several conditions mimicking plant root environments. Our analyses established a comprehensive transcriptome structure for FZB42 (Fig 5) by determining primary TSSs for 60% of all FZB42 genes, assigning 5'-UTRs and operon structures, and identifying a large number of *cis*-encoded and *trans*-encoded RNAs which may be regulators of gene expression in *B. amyloliquefaciens*. Our analyses at the transcriptome level also improved the genome annotation of FZB42; for example, we proposed 46 new genes and corrected 10 mis-annotated genes (S13 Table).



**Fig 5. Genomic Visualization of various RNAs identified.** The 1st ring from inside depicts genomic size maker; 2nd ring in black indicates DNA GC content; 3rd ring in purple and green indicates GC skew (+/-); 4th ring in grey indicates housekeeping rRNA and tRNAs; 5th and 6th ring indicates primary TSS (in red) and genes located on the leading strand (black arrows clockwise), 7th ring with olive arrows indicates the re-annotated genes on the leading strand, including corrected CDS, leaderless mRNAs and proposed operons; 8th and 9th ring indicates primary TSS (in red) and genes (black arrows counterclockwise) located on the lagging strand; 10th ring with Olive green arrows indicates re-annotated genes on the lagging strand; 11th-12th ring show nucleotide conservation patterns with genomes of *B. amyloliquefaciens* DSM7 (NC\_014551.1), *B. amyloliquefaciens* subsp. plantarum CAU B946 (NC\_016784.1) and *B. subtilis* 168 (NC\_000964.3), respectively; 13th ring in orange depicts all the *cis*-encoded RNAs annotated in this study, including riboswitches and all the antisense RNAs; 14th ring (the most outer-ring in red) depicts experimentally validated sRNAs in this study and their names are labeled.

doi:10.1371/journal.pone.0142002.g005



In contrast with genome annotation, the structure of most bacterial transcriptomes has not been well-defined, mainly because of the lack of high-throughput approaches. Traditionally, bacterial TSSs were determined for a handful of well-studied genes using low-throughput methods such as primer extension or 5'-RACE. These methods are generally time-consuming and come with low resolution due to the gel-based system. The dRNA-seq based genome-wide TSS determination applied here generated a high-resolution transcriptome structure and revealed a complexity of gene organizations in bacterial transcriptome. Many coding genes are transcribed from multiple TSSs and may possess several different UTRs; and many mRNAs have overlapping and antisense transcripts (S3, S4 and S8–S10 Figs). For instance, the *yqzM* RNA encoding a 46-AA conserved hypothetical protein (S13 Table) was found complementary to the 5'UTR of *holA* as well as the 3'UTR of *comEC* (S9 Fig). We suspect that *yqzM* RNA and *holA/comEC* might also serve as antisense RNA regulating the expression of each other. Similarly, a large part of the *ytkA* mRNA (>200 bp) is antisense to another peptide-encoding transcript on the opposite strand (S10 Fig), suggesting another case of potential antisense regulation *in cis*.

Another characteristic of the bacterial transcriptome structure is its high plasticity under different conditions. We note that our current analyses are restricted to stationary phase, which is of high interest because most stress responses and noncoding regulators are activated under this condition. The transcriptome profile is very likely to differ from that in exponential growth, as exemplified by the dynamic expression profile of several sRNAs throughout bacterial growth (S6 Fig). As the sequencing cost continues to drop, full transcriptome analyses of bacterial samples from different growth stages and arrays of stress conditions will become feasible in the future.

In this study, the cultures were sampled from different growth stages and FZB42 was grown in different media supplemented with soil extract and/or maize root exudates in order to mimic rhizosphere environments. While soil extract is important for the growth of soil bacterial [73], root exudates are composed of a wide range of organic compounds and known to be the crucial players mediating the interaction between plants and rhizobacteria [74, 75]. These different conditions allow us not only to elucidate the complexity of FZB42 transcriptome structure as stated above but also to compare their gene expression in response to diverse influential factors. For example, we identified a number of differentially expressed transcripts with respect to growth phase or root exudate. Since our sequence output contains many rRNA/tRNA reads (rRNA depletion would be beneficial for future sequencing analyses to obtain higher reads coverage), we applied a strict filter (see [material and methods](#) part) to select genes with high number of reads for reliable quantification and comparison of gene expression, though this would compromise the detection of genes with low expression levels. The differentially expressed genes we identified form a useful resource meriting our attention for future work, in order to unravel molecular details of growth phase specific physiology or plant-bacteria communication.

Further, relating the expression of these genes with specific environmental or physiologic conditions would facilitate exploration of their function, regulation and relevance. For example, a number of genes up-regulated were enriched in the pathways of histidine (S11 Fig) or arginine metabolism (S11 Fig). In these two pathways, many genes seem upregulated at #9h than at #6h, which suggests enhanced synthesis of L-glutamate and arginine at MSP. This is in accordance with a previous report that L-glutamate and L-arginine are two predominant amino acids in spores of *B. subtilis* [76]. Together, these observations suggest a critical role of glutamate in spore formation, dormancy, or germination. The regulated genes in the two pathways also allow us to trace the synthesis source and flow destination of glutamate and arginine in *B. amyloliquefaciens*, such that production of L-glutamate must derive at least in part from

L-histidine and further originate from phosphoribosyl pyrophosphate. Further, we can reason that glutamate and arginine here probably serve not only for the urea cycle but also as substrates for energy generation. For glutamate, this process seems to be completed by firstly converting L-glutamate to L-proline and then directing proline to TCA via pyruvate. Alternatively, glutamate could be converted into ornithine of urea cycle and integrated into TCA cycle via fumarate, in the course of which NADH were generated for ATP production.

Another example is the enrichment in the pathway of inositol phosphate metabolism. This enrichment suggested that at MSP catabolism of 1D-*myo*-inositol was finally directed into TCA cycle (S11 Fig). Inositol phosphates are abundant in most soils and the *myo* isomer is the most prevalent form occurring in nature [77]. Thus, we inferred that soil extract used in both SE and RS conditions supplied inositol phosphates, which were firstly transformed into *myo*-inositol, to serve as an important energy source at MSP. Besides, we also identified an enrichment of some repressed genes at #9h in the pathway relating to fructose and mannose metabolism (S11 Fig). Most of the genes were involved in the first part, before production of glyceraldehyde-3-phosphate, of glycolysis pathway. The suppression of these genes was probably resulted from the absence of easily decomposed saccharides at MSP, which may also account for inositol phosphate and amino acids were utilized as energy resource as deduced above. Overall, a number of regulated genes in the exemplified pathways display, at least partially, a profile of the potential surviving strategies of FZB42 adapting to rhizosphere.

Although the functions of most identified sRNAs yet to be understood, surging amount of evidences indicate their roles in bacterial adaptation to changing environments and growth conditions [61, 78, 79]. For example, Spot42, one of the first sRNAs identified in *E. coli*, was recently shown to facilitate the metabolism reprogramming when switching to different carbon sources [78]; and the CsrB family of sRNAs regulate the expression of a large number of genes in response to altered nutrient availability [80–83]. Beyond central metabolism, sRNAs are now known involved in many bacterial life styles, including the switch between aerobic to anaerobic growth [84, 85], motility to biofilm formation [86], low population density to high density [87, 88], and saprophytism to virulence [21, 89, 90]. To make a living in the complex rhizosphere environments, bacteria need to possess a potential to utilize plant-derived nutrients, deal with dynamic biotic and abiotic challenges in the niches, colonize plant roots and form biofilm. Similar to the environmental alteration confronted by pathogenic bacteria during the course of infection, it is reasonable to expect that the identified *Bacillus* regulatory RNAs may also coordinate rhizobacterial adaptation to rhizosphere.

Here we have identified a large number of non-coding regulatory RNAs in FZB42. Specifically, we annotated in the FZB42 genome 53 *cis*-encoded riboswitch or leader sequence, 136 asRNAs, and 86 sRNA candidates, a number of which have been validated by independent northern blotting method. Some of the candidates could not be confirmed by independent northern blot analysis probably due to lower expression level and/or lower sensitivity of compared to RNA-seq. The validated sRNAs assemble an array of novel regulators in *Bacillus* worthy of further scrutiny. From their gene locations, most of the sRNAs are probably *trans*-acting sRNAs and act on genes from a distant locus. These sRNAs often directly binds to the 5'UTR of the target mRNAs to modulate translation initiation [61]. Importantly, our global TSS map has enabled a genome-wide annotation of 5'UTRs in FZB42, and this information would be immensely valuable for studying sRNA regulations, identifying their target sites in mRNAs, and establishing sRNA-mRNA interaction networks in the long term.

In summary, we have systemically identified the TSSs and non-coding RNAs of the rhizobacterium FZB42 in this study. Our results provide a valuable resource for gene studies on *B. amyloliquefaciens* and related species, e.g., the model organism *B. subtilis*. Detection of sRNAs in rhizobacteria not only broadens the scope of sRNA world in the model Gram positive

rhizobacteria, but also helps to elucidate the functional insights into sRNAs in some unexplored environmental conditions, e.g., rhizosphere, and in some important physiological behavior, e.g., sporulation.

## Supporting Information

**S1 Fig. Functional classification of the genes up-regulated at middle stationary phase.** The genes specifically up-regulated at middle stationary phase compared to early stationary phase were identified for SE samples and RS samples respectively. The common genes exhibited in both of the two conditions were distributed in various functional classes according to [91]. The inner pie map was the major class while the outer pie map were subclass subordinated to each major group. The percentage and/or number of the genes in each class/subclass are indicated respectively.

(PDF)

**S2 Fig. Sequencing reads of the *yhaA-yhfA-yhgC* operon.** Two operons identified antisense to a known operon (A) or to known genes (B). cDNA reads in each library without (-TEX) or with (+TEX) terminator exonuclease treatment were mapped to *B. amyloliquefaciens* FZB42 chromosome and visualized by IGB 8.1. Vertical bars indicate cDNA recovery detected in each library which were adjusted to the same scale of 20 or 30. The genes indicated above the horizontal coordinates are located on the forward strand while the genes indicated below the horizontal coordinates are located on the reverse strand. Bacterial cultures were sampled for total RNA extraction at two time points (#6h&#9h) from four conditions: i) in 1CM medium; ii) 1CM medium supplemented with the maize root exudates (+RE); iii) 1CM medium supplemented with soil extract (+SE); iv) 1CM medium supplemented with both maize root exudates and soil extract (+RS). The time point #6h corresponded to early stationary phase while #9h corresponded to middle stationary phase.

(PDF)

**S3 Fig. Sequencing reads of the transcript antisense to *yndL*.** cDNA reads in each library without (-TEX) or with (+TEX) terminator exonuclease treatment were mapped to *B. amyloliquefaciens* FZB42 chromosome and visualized by IGB 8.1. Vertical bars indicate cDNA recovery detected in each library which were adjusted to the same scale of 50. The genes indicated above the horizontal coordinates are located on the forward strand while the genes indicated below the horizontal coordinates are located on the reverse strand. Bacterial cultures were sampled for total RNA extraction at two time points (#6h&#9h) from four conditions: i) in 1CM medium; ii) 1CM medium supplemented with the maize root exudates (+RE); iii) 1CM medium supplemented with soil extract (+SE); iv) 1CM medium supplemented with both maize root exudates and soil extract (+RS). The time point #6h corresponded to early stationary phase while #9h corresponded to middle stationary phase.

(PDF)

**S4 Fig. Expression of sRNA Bas28 (*rhaX-yheI*) shown in Northern Blot and RNA-seq result.** Panel A: Northern blot was performed using total RNAs prepared from different sampling time points and different media. The arrow indicates the position where the sRNA was expected to be present according to RNA-seq result. Panel B: cDNA reads in each library without (-TEX) or with (+TEX) terminator exonuclease treatment were mapped to *B. amyloliquefaciens* FZB42 chromosome and visualized by IGB 8.1. Vertical bars indicate cDNA recovery detected in each library which were adjusted to the same scale of 50. The genes indicated above the horizontal coordinates are located on the forward strand while the genes indicated below the horizontal coordinates are located on the reverse strand. Bacterial cultures were sampled

for total RNA extraction at two time points (#6h&#9h) from four media: i) in 1CM medium; ii) 1CM medium supplemented with the maize root exudates (+RE); iii) 1CM medium supplemented with soil extract (+SE); iv) 1CM medium supplemented with both maize root exudates and soil extract (+RS). The time point #6h corresponded to early stationary phase while #9h corresponded to middle stationary phase.

(PDF)

**S5 Fig. Expression of sRNA Bas35 (*trpS-oppA*) shown in RNA-seq result and Northern**

**Blot.** Panel A: Northern blot was performed using total RNAs prepared from different sampling time points and different media. The arrow indicates the position where the sRNA was expected to be present according to RNA-seq result. Panel B: cDNA reads in each library without (-TEX) or with (+TEX) terminator exonuclease treatment were mapped to *B. amyloliquefaciens* FZB42 chromosome and visualized by IGB 8.1. Vertical bars indicate cDNA recovery detected in each library which were adjusted to the same scale of 50. The genes indicated above the horizontal coordinates are located on the forward strand while the genes indicated below the horizontal coordinates are located on the reverse strand. Bacterial cultures were sampled for total RNA extraction at two time points (#6h&#9h) from four media: i) in 1CM medium; ii) 1CM medium supplemented with the maize root exudates (+RE); iii) 1CM medium supplemented with soil extract (+SE); iv) 1CM medium supplemented with both maize root exudates and soil extract (+RS). The time point #6h corresponded to early stationary phase while #9h corresponded to middle stationary phase.

(PDF)

**S6 Fig. Expression of the verified sRNAs along growth of FZB42 in four different media.**

Bacterial cultures were sampled for total RNA extraction at different time points from four media: i) in 1CM medium; ii) 1CM medium supplemented with the maize root exudates (+RE); iii) 1CM medium supplemented with soil extract (+SE); iv) 1CM medium supplemented with both maize root exudates and soil extract (+RS). Northern blots were performed using corresponding oligo probes for each sRNA ([S16 Table](#)). In panel A and panel B, the arrows and the numbers at the end of each arrow indicate time points where cultures were collected. The numbers of RNA samples used for Northern blot in panel C & D corresponded to the culture numbers in panel A, while the numbers of RNA samples used for Northern blot in panel E-H corresponded to culture numbers in panel B.

(PDF)

**S7 Fig. Expression of sRNA Bas26 (*yjca-RBAM\_011810*) shown in RNA-seq result.**

cDNA reads in each library without (-TEX) or with (+TEX) terminator exonuclease treatment were mapped to *B. amyloliquefaciens* FZB42 chromosome and visualized by IGB 8.1. Vertical bars indicate cDNA recovery detected in each library which were adjusted to the same scale of 50. The genes indicated above the horizontal coordinates are located on the forward strand while the genes indicated below the horizontal coordinates are located on the reverse strand. Bacterial cultures were sampled for total RNA extraction at two time points (#6h&#9h) from two conditions: i) 1CM medium supplemented with soil extract (+SE); ii) 1CM medium supplemented with maize root exudates and soil extract (+RS). The time point #6h corresponded to early stationary phase while #9h corresponded to middle stationary phase.

(PDF)

**S8 Fig. Sequencing reads of the genes involved in plant-microbe interaction.**

cDNA reads in each library without (-TEX) or with (+TEX) terminator exonuclease treatment were mapped to *B. amyloliquefaciens* FZB42 chromosome and visualized by IGB 8.1. Vertical bars indicate cDNA recovery detected in each library which were adjusted to the same scale of 50. The genes

indicated above the horizontal coordinates are located on the forward strand while the genes indicated below the horizontal coordinates are located on the reverse strand. Bacterial cultures were sampled for total RNA extraction at two time points (#6h&#9h) from four conditions: i) in 1CM medium; ii) 1CM medium supplemented with the maize root exudates (+RE); iii) 1CM medium supplemented with soil extract (+SE); iv) 1CM medium supplemented with both maize root exudates and soil extract (+RS). The time point #6h corresponded to early stationary phase while #9h corresponded to middle stationary phase.  
(PDF)

**S9 Fig. Expression of *yqzM (hola-comEC)* shown in RNA-seq result.** cDNA reads in each library without (-TEX) or with (+TEX) terminator exonuclease treatment were mapped to *B. amyloliquefaciens* FZB42 chromosome and visualized by IGB 8.1. Vertical bars indicate cDNA recovery detected in each library which were adjusted to the same scale of 50. The genes indicated above the horizontal coordinates are located on the forward strand while the genes indicated below the horizontal coordinates are located on the reverse strand. Bacterial cultures were sampled for total RNA extraction at two time points (#6h&#9h) from four conditions: i) in 1CM medium; ii) 1CM medium supplemented with the maize root exudates (+RE); iii) 1CM medium supplemented with soil extract (+SE); iv) 1CM medium supplemented with both maize root exudates and soil extract (+RS). The time point #6h corresponded to early stationary phase while #9h corresponded to middle stationary phase.  
(PDF)

**S10 Fig. Expressed transcripts occurred in the region of locus *dps-ytkA*.** cDNA reads in each library without (-TEX) or with (+TEX) terminator exonuclease treatment were mapped to *B. amyloliquefaciens* FZB42 chromosome and visualized by IGB 8.1. Vertical bars indicate cDNA recovery detected in each library which were adjusted to the same scale of 50. The genes indicated above the horizontal coordinates are located on the forward strand while the genes indicated below the horizontal coordinates are located on the reverse strand. Bacterial cultures were sampled for total RNA extraction at two time points (#6h&#9h) from four conditions: i) in 1CM medium; ii) 1CM medium supplemented with the maize root exudates (+RE); iii) 1CM medium supplemented with soil extract (+SE); iv) 1CM medium supplemented with both maize root exudates and soil extract (+RS). The time point #6h corresponded to early stationary phase while #9h corresponded to middle stationary phase  
(PDF)

**S11 Fig. A group of significantly regulated genes in #9h samples compared with in #6h samples are known in function involved in histidine metabolism (A), arginine and proline metabolism (B), inositol phosphate metabolism (C), and fructose and mannose metabolism (D).** The regulated genes were mapped in the KEGG pathway and the diagram was accordingly adapted. The products encoded by up-regulated genes at middle stationary phase are highlighted in red while those by down-regulated genes are highlighted in green. RE: *B. amyloliquefaciens* FZB42 was grown in 1CM medium supplemented with the maize root exudates (RE); RS: *B. amyloliquefaciens* FZB42 was grown in 1CM medium supplemented with the maize root exudates (RE) and soil extract (SE).  
(PDF)

**S1 Table. Growth conditions and distribution of sequencing reads mapped to *B. amyloliquefaciens* FZB42 genome in each library.**  
(XLSX)



**S2 Table. Statistics of number and percentage of all categories TSSs detected in various growth condition.**

(XLSX)

**S3 Table. The 10 primary TSSs randomly selected to check robustness of TSSs identification.**

(XLSX)

**S4 Table. Primary TSSs specific to one growth condition.**

(XLSX)

**S5 Table. Leaderless mRNAs detected in *B. amyloliquefaciens* FZB42.**

(XLSX)

**S6 Table. Secondary TSSs specific to middle stationary phase (SE#9h and RS#9h).**

(XLSX)

**S7 Table. The significantly regulated mRNA genes by different growth phase.**

(XLSX)

**S8 Table. Cis-encoded regulatory RNAs identified.**

(XLSX)

**S9 Table. RNAs antisense to the middle region of a gene.**

(XLSX)

**S10 Table. Antisense RNAs locating at an intergenic region.**

(XLSX)

**S11 Table. Regulatory RNAs highly expressed at middle stationary phase.**

(XLSX)

**S12 Table. Re-annotation of mRNA genes.**

(XLSX)

**S13 Table. Annotation of tRNAs and mRNA genes newly identified.**

(XLSX)

**S14 Table. Predicted operons in genome of *B. amyloliquefaciens* FZB42.**

(XLSX)

**S15 Table. The genes expressed from an inner TSS within a polycistronic operon.**

(XLSX)

**S16 Table. Small RNAs detected by Northern Blot.**

(XLSX)

**S17 Table. Genes involved in plant-bacteria interaction.**

(XLSX)

## Acknowledgments

The financial support by the National Natural Science Foundation of China (No. 31100081) and Scientific Research Foundation of Nanjing Forestry University for Introduction of High Level Talents (YJ2012-03) is gratefully acknowledged, as well as the financial support given by the Priority Academic Program Development (PAPD) of Jiangsu Higher Education

Institutions. This study was further supported by the Natural Science Foundation of Jiangsu Province (No.BK20151514). The authors declare there is no conflict of interest.

## Author Contributions

Conceived and designed the experiments: BF. Performed the experiments: BF. Analyzed the data: BF LL YC KF. Contributed reagents/materials/analysis tools: RB JV XW. Wrote the paper: BF LL YC RB JV.

## References

1. Chen XH, Koumoutsis A, Scholz R, Eisenreich A, Schneider K, Heinemeyer I, et al. Comparative analysis of the complete genome sequence of the plant growth-promoting bacterium *Bacillus amyloliquefaciens* FZB42. *Nat Biotechnol.* 2007; 25(9):1007–14. PMID: [17704766](#).
2. Chen XH, Scholz R, Borriss M, Junge H, Mogel G, Kunz S, et al. Difficidin and bacilysin produced by plant-associated *Bacillus amyloliquefaciens* are efficient in controlling fire blight disease. *J Biotechnol.* 2009; 140(1–2):38–44. PMID: [19061923](#). doi: [10.1016/j.jbiotec.2008.10.015](#)
3. Kalyon B, Helaly SE, Scholz R, Nachtigall J, Vater J, Borriss R, et al. Plantazolicin A and B: structure elucidation of ribosomally synthesized thiazole/oxazole peptides from *Bacillus amyloliquefaciens* FZB42. *Organic Letters.* 2011; 13(12):2996–9. doi: [10.1021/ol200809m](#) PMID: [21568297](#)
4. Scholz R, Molohon KJ, Nachtigall J, Vater J, Markley AL, Sussmuth RD, et al. Plantazolicin, a novel microcin B17/streptolysin S-like natural product from *Bacillus amyloliquefaciens* FZB42. *J Bacteriol.* 2011; 193(1):215–24. PMID: [20971906](#). doi: [10.1128/JB.00784-10](#)
5. Koumoutsis A, Chen XH, Henne A, Liesegang H, Hitzeroth G, Franke P, et al. Structural and functional characterization of gene clusters directing nonribosomal synthesis of bioactive cyclic lipopeptides in *Bacillus amyloliquefaciens* strain FZB42. *J Bacteriol.* 2004; 186(4):1084–96. PMID: [14762003](#).
6. Scholz R, Vater J, Budiharjo A, Wang Z, He Y, Dietel K, et al. Amylocyclicin, a novel circular bacteriocin produced by *Bacillus amyloliquefaciens* FZB42. *Journal of Bacteriology.* 2014; 196(10):1842–52. doi: [10.1128/JB.01474-14](#) PMID: [24610713](#)
7. Koumoutsis A, Chen X-H, Vater J, Borriss R. DegU and YczE Positively Regulate the Synthesis of Bacilomycin D by *Bacillus amyloliquefaciens* Strain FZB42. *Applied and Environmental Microbiology.* 2007; 73(21):6953–64. PMID: [17827323](#)
8. Schneider K, Chen XH, Vater J, Franke P, Nicholson G, Borriss R, et al. Macrolactin is the polyketide biosynthesis product of the pks2 cluster of *Bacillus amyloliquefaciens* FZB42. *J Nat Prod.* 2007; 70(9):1417–23. PMID: [17844999](#).
9. Idris EE, Iglesias DJ, Talon M, Borriss R. Tryptophan-dependent production of indole-3-acetic acid (IAA) affects level of plant growth promotion by *Bacillus amyloliquefaciens* FZB42. *Mol Plant Microbe Interact.* 2007; 20(6):619–26. PMID: [17555270](#).
10. Fan B, Chen XH, Budiharjo A, Bleiss W, Vater J, Borriss R. Efficient colonization of plant roots by the plant growth promoting bacterium *Bacillus amyloliquefaciens* FZB42, engineered to express green fluorescent protein. *J Biotechnol.* 2011; 151(4):303–11. doi: [10.1016/j.jbiotec.2010.12.022](#) PMID: [21237217](#)
11. Mariappan A, Makarewicz O, Chen XH, Borriss R. Two-component response regulator DegU controls the expression of bacilysin in plant-growth-promoting bacterium *Bacillus amyloliquefaciens* FZB42. *J Mol Microbiol Biotechnol.* 2012; 22(2):114–25. PMID: [22677979](#). doi: [10.1159/000338804](#)
12. Sharma CM, Hoffmann S, Darfeuille F, Reigner J, Findeiss S, Sittka A, et al. The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature.* 2010; 464(7286):250–5. PMID: [20164839](#). doi: [10.1038/nature08756](#)
13. Thomason MK, Bischler T, Eisenbart SK, Forstner KU, Zhang A, Herbig A, et al. Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in *Escherichia coli*. *J Bacteriol.* 2015; 197(1):18–28. PMID: [25266388](#). doi: [10.1128/JB.02096-14](#)
14. Kroger C, Colgan A, Srikumar S, Handler K, Sivasankaran SK, Hammarlof DL, et al. An infection-relevant transcriptomic compendium for *Salmonella enterica* Serovar Typhimurium. *Cell Host and Microbe.* 2013; 14(6):683–95. doi: [10.1016/j.chom.2013.11.010](#) PMID: [24331466](#)
15. Albrecht M, Sharma CM, Dittrich MT, Muller T, Reinhardt R, Vogel J, et al. The transcriptional landscape of *Chlamydia pneumoniae*. *Genome Biol.* 2011; 12(10):R98. PMID: [21989159](#). doi: [10.1186/gb-2011-12-10-r98](#)

16. Schmidtke C, Findeiss S, Sharma CM, Kuhfuss J, Hoffmann S, Vogel J, et al. Genome-wide transcriptome analysis of the plant pathogen *Xanthomonas* identifies sRNAs with putative virulence functions. *Nucleic Acids Res.* 2012; 40(5):2020–31. PMID: [22080557](#). doi: [10.1093/nar/gkr904](#)
17. Imrov I, Sharma CM, Vogel J, Winkler WC. Identification of regulatory RNAs in *Bacillus subtilis*. *Nucleic Acids Research.* 2010. doi: [10.1093/nar/gkq454](#)
18. Cortes T, Schubert OT, Rose G, Arnvig KB, Comas I, Aebersold R, et al. Genome-wide mapping of transcriptional start sites defines an extensive leaderless transcriptome in *Mycobacterium tuberculosis*. *Cell Rep.* 2013; 5(4):1121–31. doi: [10.1016/j.celrep.2013.10.031](#) PMID: [24268774](#)
19. Li L, Huang D, Cheung MK, Nong W, Huang Q, Kwan HS. BSRD: a repository for bacterial small regulatory RNA. *Nucleic Acids Res.* 2013; 41(Database issue):D233–8. PMID: [23203879](#).
20. Raghavan R, Groisman EA, Ochman H. Genome-wide detection of novel regulatory RNAs in *E. coli*. *Genome Res.* 2011; 21(9):1487–97. PMID: [21665928](#). doi: [10.1101/gr.119370.110](#)
21. Kroger C, Dillon SC, Cameron AD, Papenfort K, Sivasankaran SK, Hokamp K, et al. The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proc Natl Acad Sci U S A.* 2012; 109(20):E1277–86. PMID: [22538806](#). doi: [10.1073/pnas.1201061109](#)
22. Vogel J, Luisi BF. Hfq and its constellation of RNA. *Nature Reviews Microbiology.* 2011; 9(8):578–89. doi: [10.1038/nrmicro2615](#) PMID: [21760622](#)
23. Christiansen JK, Larsen MH, Ingmer H, Sogaard-Andersen L, Kallipolitis BH. The RNA-binding protein Hfq of *Listeria monocytogenes*: role in stress tolerance and virulence. *J Bacteriol.* 2004; 186(11):3355–62. PMID: [15150220](#).
24. Bohn C, Rigoulay C, Boulou P. No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*. *BMC Microbiology.* 2007; 7(10):10.
25. Dambach M, Imrov I, Winkler WC. Association of RNAs with *Bacillus subtilis* Hfq. *PLoS ONE.* 2013; 8(2):e55156. doi: [10.1371/journal.pone.0055156](#) PMID: [23457461](#)
26. Hammerle H, Amman F, Vecerek B, Stulke J, Hofacker I, Blasi U. Impact of Hfq on the *Bacillus subtilis* transcriptome. *PLoS ONE.* 2014; 9(6):e98661. doi: [10.1371/journal.pone.0098661](#) PMID: [24932523](#)
27. Rasmussen S, Nielsen HB, Jarmer H. The transcriptionally active regions in the genome of *Bacillus subtilis*. *Mol Microbiol.* 2009; 73(6):1043–57. PMID: [19682248](#). doi: [10.1111/j.1365-2958.2009.06830.x](#)
28. Wiegand S, Dietrich S, Hertel R, Bongaerts J, Evers S, Volland S, et al. RNA-Seq of *Bacillus licheniformis*: active regulatory RNA features expressed within a productive fermentation. *BMC Genomics.* 2013; 14(1):667.
29. Brantl S, Bruckner R. Small regulatory RNAs from low-GC Gram-positive bacteria. *RNA Biology.* 2014; 11(5):443–56. PMID: [24576839](#)
30. Gimpel M, Heidrich N, Mader U, Krugel H, Brantl S. A dual-function sRNA from *B. subtilis*: SR1 acts as a peptide encoding mRNA on the gapA operon. *Mol Microbiol.* 2010; 76(4):990–1009. doi: [10.1111/j.1365-2958.2010.07158.x](#) PMID: [20444087](#)
31. Silvaggi JM, Perkins JB, Losick R. Genes for small, noncoding RNAs under sporulation control in *Bacillus subtilis*. *J Bacteriol.* 2006; 188(2):532–41. PMID: [16385044](#)
32. Swiercz JP, Hindra, Bobek J, Bobek J, Haiser HJ, Di Berardo C, et al. Small non-coding RNAs in *Streptomyces coelicolor*. *Nucleic Acids Res.* 2008; 36(22):7240–51. PMID: [19008244](#). doi: [10.1093/nar/gkn898](#)
33. Landt SG, Abeliuk E, McGrath PT, Lesley JA, McAdams HH, Shapiro L. Small non-coding RNAs in *Caulobacter crescentus*. *Mol Microbiol.* 2008; 68(3):600–14. PMID: [18373523](#). doi: [10.1111/j.1365-2958.2008.06172.x](#)
34. Perez N, Trevino J, Liu Z, Ho SC, Babitzke P, Sumby P. A genome-wide analysis of small regulatory RNAs in the human pathogen group A *Streptococcus*. *PLoS One.* 2009; 4(11):e7668. PMID: [19888332](#). doi: [10.1371/journal.pone.0007668](#)
35. Preis H, Eckart RA, Gudipati RK, Heidrich N, Brantl S. CodY activates transcription of a small RNA in *Bacillus subtilis*. *J Bacteriol.* 2009; 191(17):5446–57. PMID: [19542274](#). doi: [10.1128/JB.00602-09](#)
36. Axmann IM, Kensche P, Vogel J, Kohl S, Herzel H, Hess WR. Identification of cyanobacterial non-coding RNAs by comparative genome analysis. *Genome Biol.* 2005; 6(9):R73. PMID: [16168080](#).
37. Arnvig KB, Young DB. Identification of small RNAs in *Mycobacterium tuberculosis*. *Mol Microbiol.* 2009; 73(3):397–408. PMID: [19555452](#). doi: [10.1111/j.1365-2958.2009.06777.x](#)
38. Saito S, Kakeshita H, Nakamura K. Novel small RNA-encoding genes in the intergenic regions of *Bacillus subtilis*. *Gene.* 2009; 428(1–2):2–8. PMID: [18948176](#). doi: [10.1016/j.gene.2008.09.024](#)
39. Heidrich N, Chinali A, Gerth U, Brantl S. The small untranslated RNA SR1 from the *Bacillus subtilis* genome is involved in the regulation of arginine catabolism. *Mol Microbiol.* 2006; 62(2):520–36. PMID: [17020585](#).

40. Silvaggi JM, Perkins JB, Losick R. Small untranslated RNA antitoxin in *Bacillus subtilis*. *J Bacteriol*. 2005; 187(19):6641–50. PMID: [16166525](#)
41. Mars RA, Nicolas P, Ciccolini M, Reilman E, Reder A, Schaffer M, et al. Small Regulatory RNA-Induced Growth Rate Heterogeneity of *Bacillus subtilis*. *PLoS Genetics*. 2015; 11(3):e1005046. doi: [10.1371/journal.pgen.1005046](#) PMID: [25790031](#)
42. Heidrich N, Moll I, Brantl S. In vitro analysis of the interaction between the small RNA SR1 and its primary target *ahrC* mRNA. *Nucleic Acids Res*. 2007; 35(13):4331–46. PMID: [17576690](#).
43. Fan B, Carvalhais LC, Becker A, Fedoseyenko D, von Wiren N, Borriss R. Transcriptomic profiling of *Bacillus amyloliquefaciens* FZB42 in response to maize root exudates. *BMC Microbiol*. 2012; 12:116. doi: [10.1186/1471-2180-12-116](#) PMID: [22720735](#)
44. Chao Y, Papenfort K, Reinhardt R, Sharma CM, Vogel J. An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. *Embo J*. 2012; 31(20):4005–19. PMID: [22922465](#). doi: [10.1038/emboj.2012.229](#)
45. Forstner KU, Vogel J, Sharma CM. READemption—a tool for the computational analysis of deep-sequencing-based transcriptome data. *Bioinformatics*. 2014; 30(23):3421–3. PMID: [25123900](#). doi: [10.1093/bioinformatics/btu533](#)
46. Nicol JW, Helt GA, Blanchard SG Jr, Raja A, Loraine AE. The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics*. 2009; 25(20):2730–1. doi: [10.1093/bioinformatics/btp472](#) PMID: [19654113](#)
47. Dugar G, Herbig A, Forstner KU, Heidrich N, Reinhardt R, Nieselt K, et al. High-resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates. *PLoS Genet*. 2013; 9(5):e1003495. doi: [10.1371/journal.pgen.1003495](#) PMID: [23696746](#)
48. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*. 2005; 102(43):15545–50. doi: [10.1073/pnas.0506580102](#)
49. Luo W, Brouwer C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics*. 2013; 29(14):1830–1. PMID: [23740750](#). doi: [10.1093/bioinformatics/btt285](#)
50. Kim D, Hong JS, Qiu Y, Nagarajan H, Seo JH, Cho BK, et al. Comparative analysis of regulatory elements between *Escherichia coli* and *Klebsiella pneumoniae* by genome-wide transcription start site profiling. *PLoS Genet*. 2012; 8(8):e1002867. PMID: [22912590](#). doi: [10.1371/journal.pgen.1002867](#)
51. Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, Riley M, et al. The complete genome sequence of *Escherichia coli* K-12. *Science*. 1997; 277(5331):1453–62. PMID: [9278503](#).
52. Kozak M. Comparison of initiation of protein synthesis in prokaryotes, eucaryotes, and organelles. *Microbiol Rev*. 1983; 47(1):1–45. PMID: [6343825](#).
53. Nicolas P, Mader U, Dervyn E, Rochat T, Leduc A, Pigeonneau N, et al. Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science*. 2012; 335(6072):1103–6. PMID: [22383849](#). doi: [10.1126/science.1206848](#)
54. Margolis PS, Driks A, Losick R. Sporulation gene *spoIIIB* from *Bacillus subtilis*. *Journal of Bacteriology*. 1993; 175(2):528–40. PMID: [8419299](#).
55. Siervo N, Makita Y, de Hoon M, Nakai K. DBTBS: a database of transcriptional regulation in *Bacillus subtilis* containing upstream intergenic conservation information. *Nucleic Acids Research*. 2008; 36(suppl\_1):D93–6.
56. Rudner DZ, Losick R. Morphological coupling in development: lessons from prokaryotes. *Dev Cell*. 2001; 1(6):733–42. PMID: [11740935](#).
57. Hilbert DW, Piggot PJ. Compartmentalization of gene expression during *Bacillus subtilis* spore formation. *Microbiol Mol Biol Rev*. 2004; 68(2):234–62. PMID: [15187183](#).
58. Ryter A. Etude morphologique de la sporulation de *Bacillus subtilis*. *Ann Inst Pasteur (Paris)*. 1965; 108:40–60.
59. Eijlander RT, de Jong A, Krawczyk AO, Holsappel S, Kuipers OP. SporeWeb: an interactive journey through the complete sporulation cycle of *Bacillus subtilis*. *Nucleic Acids Res*. 2014; 42(Database issue):D685–91. PMID: [24170806](#). doi: [10.1093/nar/gkt1007](#)
60. Winkler WC, Breaker RR. Regulation Of Bacterial Gene Expression By Riboswitches. *Annu Rev Microbiol*. 2005; 59(1):487–517. doi: [10.1146/annurev.micro.59.030804.121336](#) PMID: [16153177](#).
61. Waters LS, Storz G. Regulatory RNAs in Bacteria. 2009; 136(4):615–28. S0092-8674(09)00125-1 doi: [10.1016/j.cell.2009.01.043](#) PMID: [19239884](#)
62. Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR. Rfam: an RNA family database. *Nucleic Acids Res*. 2003; 31(1):439–41. PMID: [12520045](#)

63. Burge SW, Daub J, Eberhardt R, Tate J, Barquist L, Nawrocki EP, et al. Rfam 11.0: 10 years of RNA families. *Nucleic Acids Res.* 2013; 41(Database issue):D226–32. doi: [10.1093/nar/gks1005](https://doi.org/10.1093/nar/gks1005) PMID: [23125362](https://pubmed.ncbi.nlm.nih.gov/23125362/)
64. Barrick JE, Corbino KA, Winkler WC, Nahvi A, Mandal M, Collins J, et al. New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc Natl Acad Sci U S A.* 2004; 101(17):6421–6. PMID: [15096624](https://pubmed.ncbi.nlm.nih.gov/15096624/)
65. Weinberg Z, Wang JX, Bogue J, Yang J, Corbino K, Moy RH, et al. Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol.* 2010; 11(3):R31. PMID: [20230605](https://pubmed.ncbi.nlm.nih.gov/20230605/). doi: [10.1186/gb-2010-11-3-r31](https://doi.org/10.1186/gb-2010-11-3-r31)
66. Bonner ER, D'Elia JN, Billips BK, Switzer RL. Molecular recognition of pyr mRNA by the *Bacillus subtilis* attenuation regulatory protein PyrR. *Nucleic Acids Res.* 2001; 29(23):4851–65. PMID: [11726695](https://pubmed.ncbi.nlm.nih.gov/11726695/).
67. Georg J, Hess WR. cis-antisense RNA, another level of gene regulation in bacteria. *Microbiology and Molecular Biology Reviews.* 2011; 75(2):286–300. doi: [10.1128/MMBR.00032-10](https://doi.org/10.1128/MMBR.00032-10) PMID: [21646430](https://pubmed.ncbi.nlm.nih.gov/21646430/)
68. Mao F, Dam P, Chou J, Olman V, Xu Y. DOOR: a database for prokaryotic operons. *Nucleic Acids Research.* 2009; 37(Database issue):D459–63. doi: [10.1093/nar/gkn757](https://doi.org/10.1093/nar/gkn757) PMID: [18988623](https://pubmed.ncbi.nlm.nih.gov/18988623/)
69. Wurtzel O, Sesto N, Mellin JR, Karunker I, Edelheit S, Becavin C, et al. Comparative transcriptomics of pathogenic and non-pathogenic *Listeria* species. *Molecular Systems Biology.* 2012; 8(583):583.
70. Sesto N, Wurtzel O, Archambaud C, Sorek R, Cossart P. The excludon: a new concept in bacterial anti-sense RNA-mediated gene regulation. *Nature Reviews Microbiology.* 2013; 11(2):75–82. doi: [10.1038/nrmicro2934](https://doi.org/10.1038/nrmicro2934) PMID: [23268228](https://pubmed.ncbi.nlm.nih.gov/23268228/)
71. Kuwana R, Yamamura S, Ikejiri H, Kobayashi K, Ogasawara N, Asai K, et al. *Bacillus subtilis* spoVIF (yjcC) gene, involved in coat assembly and spore resistance. *Microbiology.* 2003; 149(Pt 10):3011–21. PMID: [14523132](https://pubmed.ncbi.nlm.nih.gov/14523132/)
72. Zhao X, Wang Y, Shang Q, Li Y, Hao H, Zhang Y, et al. Collagen-like proteins (ClpA, ClpB, ClpC, and ClpD) are required for biofilm formation and adhesion to plant roots by *Bacillus amyloliquefaciens* FZB42. *PLoS ONE.* 2015; 10(2):e0117414. doi: [10.1371/journal.pone.0117414](https://doi.org/10.1371/journal.pone.0117414) PMID: [25658640](https://pubmed.ncbi.nlm.nih.gov/25658640/)
73. Taylor CB. Nature of the factor in soil-extract responsible for bacterial growth-stimulation. *Nature.* 1951; 168(4264):115–6. PMID: [14852978](https://pubmed.ncbi.nlm.nih.gov/14852978/)
74. Somers E, Vanderleyden J, Srinivasan M. Rhizosphere bacterial signalling: a love parade beneath our feet. *Crit Rev Microbiol.* 2004; 30(4):205–40. PMID: [15646398](https://pubmed.ncbi.nlm.nih.gov/15646398/).
75. Badri DV, Vivanco JM. Regulation and function of root exudates. *Plant Cell Environ.* 2009; 32(6):666–81. PMID: [19143988](https://pubmed.ncbi.nlm.nih.gov/19143988/). doi: [10.1111/j.1365-3040.2008.01926.x](https://doi.org/10.1111/j.1365-3040.2008.01926.x)
76. Nelson DL, Kornberg A. Biochemical studies of bacterial sporulation and germination. 18. Free amino acids in spores. *Journal of Biological Chemistry.* 1970; 245(5):1128–36. PMID: [4984698](https://pubmed.ncbi.nlm.nih.gov/4984698/).
77. Turner BL, Papházy M, Hargarth P, McKelvie I. Inositol phosphates in the environment. *Philos Trans R Soc Lond B Biol Sci.* 2002; 357(1420):449–69. PMID: [12028785](https://pubmed.ncbi.nlm.nih.gov/12028785/)
78. Beisel CL, Storz G. The base-pairing RNA spot 42 participates in a multioutput feedforward loop to help enact catabolite repression in *Escherichia coli*. *Mol Cell.* 2011; 41(3):286–97. PMID: [21292161](https://pubmed.ncbi.nlm.nih.gov/21292161/). doi: [10.1016/j.molcel.2010.12.027](https://doi.org/10.1016/j.molcel.2010.12.027)
79. Thomason MK, Fontaine F, De Lay N, Storz G. A small RNA that regulates motility and biofilm formation in response to changes in nutrient availability in *Escherichia coli*. *Mol Microbiol.* 2012; 84(1):17–35. PMID: [22289118](https://pubmed.ncbi.nlm.nih.gov/22289118/). doi: [10.1111/j.1365-2958.2012.07965.x](https://doi.org/10.1111/j.1365-2958.2012.07965.x)
80. Romeo T, Vakulskas CA, Babitzke P. Post-transcriptional regulation on a global scale: form and function of Csr/Rsm systems. *Environmental Microbiology.* 2013; 15(2):313–24. PMID: [22672726](https://pubmed.ncbi.nlm.nih.gov/22672726/). doi: [10.1111/j.1462-2920.2012.02794.x](https://doi.org/10.1111/j.1462-2920.2012.02794.x)
81. Adamson DN, Lim HN. Rapid and robust signaling in the CsrA cascade via RNA-protein interactions and feedback regulation. *Proceedings of the National Academy of Sciences of the United States of America.* 2013; 110(32):13120–5. PMID: [23878244](https://pubmed.ncbi.nlm.nih.gov/23878244/). doi: [10.1073/pnas.1308476110](https://doi.org/10.1073/pnas.1308476110)
82. Edwards AN, Patterson-Fortin LM, Vakulskas CA, Mercante JW, Potrykus K, Vinella D, et al. Circuitry linking the Csr and stringent response global regulatory systems. *Molecular Microbiology.* 2011; 80(6):1561–80. PMID: [21488981](https://pubmed.ncbi.nlm.nih.gov/21488981/). doi: [10.1111/j.1365-2958.2011.07663.x](https://doi.org/10.1111/j.1365-2958.2011.07663.x)
83. Vakulskas CA, Potts AH, Babitzke P, Ahmer BM, Romeo T. Regulation of Bacterial Virulence by Csr (Rsm) Systems. *Microbiology and Molecular Biology Reviews.* 2015; 79(2):193–224. PMID: [25833324](https://pubmed.ncbi.nlm.nih.gov/25833324/). doi: [10.1128/MMBR.00052-14](https://doi.org/10.1128/MMBR.00052-14)
84. Boysen A, Moller-Jensen J, Kallipolitis B, Valentin-Hansen P, Overgaard M. Translational regulation of gene expression by an anaerobically induced small non-coding RNA in *Escherichia coli*. *J Biol Chem.* 2010; 285(14):10690–702. PMID: [20075074](https://pubmed.ncbi.nlm.nih.gov/20075074/). doi: [10.1074/jbc.M109.089755](https://doi.org/10.1074/jbc.M109.089755)



85. Durand S, Storz G. Reprogramming of anaerobic metabolism by the FnrS small RNA. *Mol Microbiol.* 2010; 75(5):1215–31. PMID: [20070527](#). doi: [10.1111/j.1365-2958.2010.07044.x](#)
86. Mika F, Hengge R. Small RNAs in the control of RpoS, CsgD, and biofilm architecture of *Escherichia coli*. *RNA Biology.* 2014; 11(5):494–507. PMID: [25028968](#).
87. Shao Y, Bassler BL. Quorum-sensing non-coding small RNAs use unique pairing regions to differentially control mRNA targets. *Mol Microbiol.* 2012; 83(3):599–611. PMID: [22229925](#). doi: [10.1111/j.1365-2958.2011.07959.x](#)
88. Sonnleitner E, Gonzalez N, Sorger-Domenigg T, Heeb S, Richter AS, Backofen R, et al. The small RNA PhrS stimulates synthesis of the *Pseudomonas aeruginosa* quinolone signal. *Mol Microbiol.* 2011; 80(4):868–85. PMID: [21375594](#). doi: [10.1111/j.1365-2958.2011.07620.x](#)
89. Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, Balestrino D, et al. The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature.* 2009; 459(7249):950–6. PMID: [19448609](#). doi: [10.1038/nature08080](#)
90. Loh E, Dussurget O, Gripenland J, Vaitkevicius K, Tiensuu T, Mandin P, et al. A trans-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*. *Cell.* 2009; 139(4):770–9. PMID: [19914169](#). doi: [10.1016/j.cell.2009.08.046](#)
91. Moszer I, Jones LM, Moreira S, Fabry C, Danchin A. SubtiList: the reference database for the *Bacillus subtilis* genome. *Nucleic Acids Res.* 2002; 30(1):62–5. PMID: [11752255](#).