Genomic Identification of SinR Transcription Factor Binding Sites in Nitrogen Fixing Bacterium *Bradyrhizobium japonicum*[†]

Feroz Khan[1], Richa Sharma[1], Rakesh Kumar Shukla[2], Abha Meena[1], Ajit Kumar Shasany[2] and Ashok Sharma*[,1]

[1]Bioinformatics & In Silico Biology Division; [2]Genetic Resource Biotechnology Division, Central Institute of Medicinal & Aromatic Plants, Council of Scientific & Industrial Research, Lucknow-226015 (UP), India

Abstract: SinR is a transcription factor which controls expression of stress tolerance *sin* genes related to alternate development processes under stress condition. Identification of genome wide SinR-box motif and their regulated genes has not been worked out yet in *Bradyrhizobium japonicum*. For this, a weight matrix of 9 bp was developed from the known promoter sequences of *Bacillus subtilis*, which was then used for genome wide identification of co-regulated genes. The methodology first involves phylogenetic footprinting of SinR regulated genes and then construction of scoring matrix through 'Consensus' and confirmation through MEME & D-Matrix tools. Genomic prediction was done through 'Patser' program and confirmation through 'PossumSearch' program in Linux system. Results showed that all the 371 predicted genes belongs to 9 different functional classes, in which 221 found in operons with more than 80% Sin-box motif similarity. Similar approach can be used in other bacteria to explore hidden genomic regulatory network.

Keywords: SinR, Transcription factor, Sin-box, *Bradyrhizobium japonicum,* Stress tolerance.

INTRODUCTION

Genome scale experimental data is now increasing day by day with the rapid increase in the genomic sequences and the challenge is to identify functional elements at genome level. The most important functional element in any genome is the transcription factor (TF) and the sites within the DNA to which they bind called as transcription factor binding sites (TFBS). A more complete understanding of TFBS and their interactions provide a more comprehensive and quantitative mapping of the gene regulatory pathways. Also it gives deeper understanding of the potential functions of individual genes regulated by newly identified DNA-binding sites [1]. TFBS are short oligo-nucleotides (i.e. 5-15 base pair [bp] in length), mostly found in non-coding genomic sequences. Currently, degenerate sequences (i.e. DNA motif represented by IUPAC/IUB convention) are frequently used to depict the binding specificities of TF. But they do not contain precise information about the relative likelihood of observing the alternate nucleotides at various positions of a TFBS [2]. A common way of representing the degenerate sequence preferences of a DNA-binding protein is by a position specific scoring matrix (PSSM) [3, 4]. The elements of PSSM correspond to scores reflecting the likelihood of a particular nucleotide at a particular position [5]. Many studies have indicated that sequence-based prediction approaches can timely provide very useful information and insights for basic research and hence are widely welcome by science community [6, 7]. The present study is attempted to develop a novel method for genomic identification of SinR TFBS and co-regulated genes in nitrogen fixing bacterium *Bradyrhizobium japonicum* (Bradyrhizobiaceae family). However, the transcription factor, SinR is involved in the control of sporulation initiation in *B. subtilis* [8-13]. Known SinR regulatory binding sites are reported in *B. subtilis;* a best-characterized member of the Gram-positive bacteria and available at DBTBS database [14]. As reported, SinR controls sporulation through several independent genes, i.e., aprE, yqxM, yveK, epr, rok, spo0A, spoIIA, spoIIG and spoIIE, by repressing their transcription process [10]. Recent characterization of the *SinR/SinI* locus has been reported in nitrogen fixing bacterium *Sinorhizobium meliloti* [15], but the identification of SinR, their binding sites and co-regulated genes has yet not been worked out in recently sequenced bacterium *viz*., *B. japonicum* [16]. Genome wide detection of more SinR regulated stress related genes are expected in both the organisms as they belong to same eubacteria family. The genome of *B. japonicum* has a single circular chromosome of 9,105,828 bp in length with an average GC content of 64.1%, which is much larger then the genome size of the *B. subtilis i.e*.~ 4,214,630 bp [16]. In addition, it has also been found that many species characterized so far possess multiple *SinR* gene loci. Considering this, we designed our method to identify genome wide SinR TFBS in *B. japonicum*. For this we used our newly derived 9 bp long weight matrix to trace out *SinR-box* like DNA motifs in the non-coding upstream sequences. The methodology first involves identification of phylogenetic footprints of SinR and their regulated genes in *B. japonicum* and then construction of weight matrix through CONSENSUS [17], MEME [18] and D-Matrix programs. Genomic prediction of SinR TFBS was performed through PATSER [17] and later confirmed through POSSUMSEARCH [19, 20]. Genome wide prediction was considered highly accurate as further validated

*Address correspondence to this author at the Bioinformatics & In Silico Biology Division, Central Institute of Medicinal & Aromatic Plants, Council of Scientific & Industrial Research, P.O. CIMAP, Kukrail Picnic Spot Road, Lucknow-226015 (UP), India; Tel: +91 522 2717626; Fax: +91 522 2342666; E-mail: ashoksharma@cimap.res.in

[†]CIMAP Communication No. 2008-116-J.

through operon delineation. Out of total 371 predicted genes, 221 genes showed their localization in operons, as revealed by evolutionary gene pair conservation. Finally through this analysis we successfully identified the genes of *sin* operon similar to *B. subtilis* [21]. Thus, similar approach could be used in other recently sequenced bacterial genomes to pursue the study in detail and also to unfold the regulatory mechanism of stress tolerance in different families of bacteria.

RESULTS & DISCUSSION

Employing weight matrix based method for genome wide identification of regulatory binding site in *B. japonicum*, we found total 371 target genes having *SinR-box* like motif sequence conservation in their upstream sequences. These were predicted at lower cutoff weight matrix score '7.93' equivalent to 93.77% motif similarity (Fig. **1**). Results indicate that most of the predicted genes were related to alternate development pathways similar to *B. subtilis*; a soil bacterium which survive under adverse environmental conditions and evolved several adaptive mechanisms. For the successful completion at the end of exponential growth several unlinked genes are required during adaptive mechanisms such as production of extra-cellular enzymes & antibiotics, acquisition of competence and development of motility & sporulation [22, 23]. The regulation process of these alternate developmental pathways is not known, although many genes are involved. It was reported that this regulatory network involve sensing of changing nutritional conditions and this information then lead to turning on and off of appropriate genes. Several alternate developmental pathways genes e.g. spoO and other genes were reported to be involved in sensing environmental changes due to homology with effector's or sensor class of proteins of the two-component system [24]. Besides, cloning and characterization of *sinR* gene involved in controlling many late growth developmental processes have been reported and demonstrated the effect on competency [25]. Due to localization of conserved SinR binding sites, results support the view that an inactivation of the *sin* gene (loss of function) results in loss of competence and motility. While due to regulatory network of SinR factor, elevated levels of SinR (gain of function) repress sporulation and the production of extracellular proteases, since all these genes showed SinR binding sites in their upstream sequences. Our predictions also supported by the view in which *in vivo* repression of sporulation genes *viz*., *spoIIE*, *spoIIG* and *spoIIA*, by SinR regulatory factor has been reported, which showed SinR binding to *spolIA* promoter. Similarly, expression of the *aprE* gene, which encodes a major extracellular protease, subtilisin (also called an alkaline protease or extracellular serine protease), is also closely associated with the onset of sporulation, which also showed localization of *sin-box*. Similarly, other reports says that *spoO* genes that are also required for initiation of sporulation, *aprE* expression is controlled by several regulatory genes, such as *degU, degQ, degR, hpr, sen, abrB, pai* and *sin*. Target sites of hpr, degQ and degU regulatory genes in the aprE gene have been identified by using several *B. subtilis* strains that have a series of promoter upstream deletions in *aprE*. These results suggest that the target sites for these genes lie relatively far upstream from the *aprE* promoter. Recently, direct DNA binding of AbrB and SinR to the *aprE* promoter has been demonstrated by gel retardation assay and footprinting analysis [26], which also showed localization of *sin-box*.

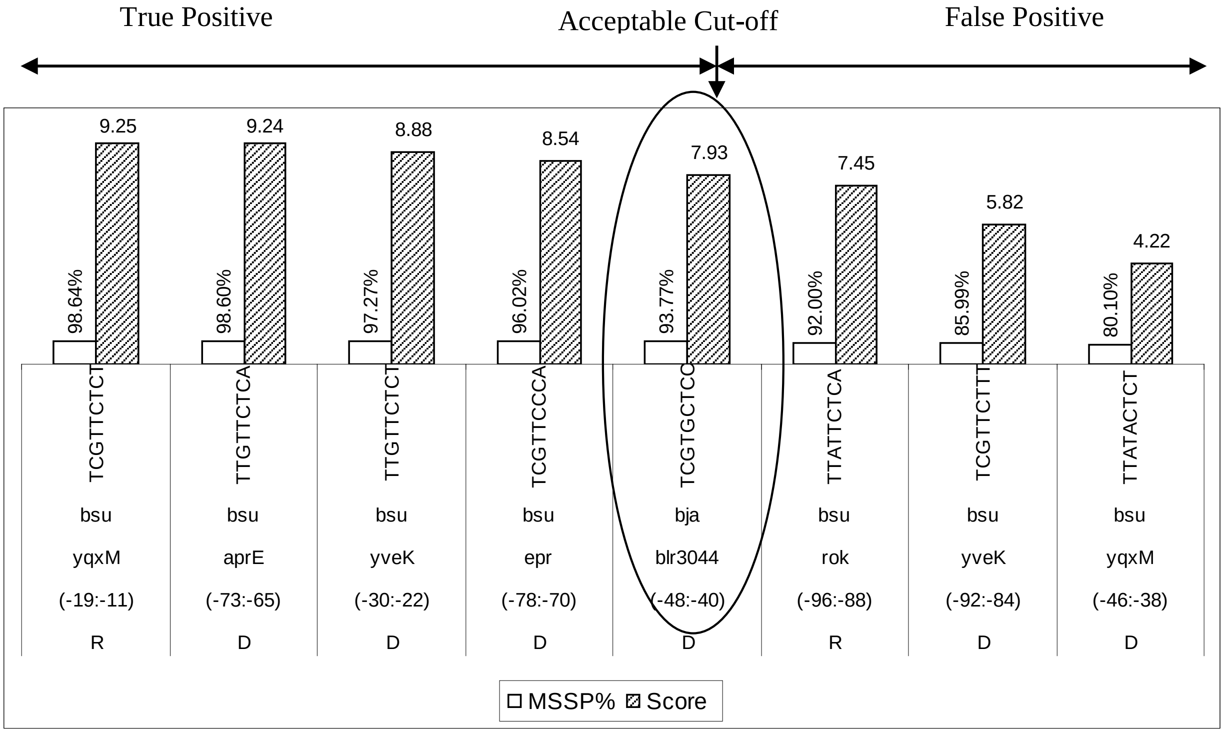

Fig. (1). Representation of correlation between score value and motif sequence similarity (MSSP) in both *B. subtilis* (bsu) and *B. japonicum* (bja) genes. Acceptable cut-off for genomic prediction of true positives in *B. japonicum* was identified as 7.93 score with 93.77% MSSP.

Distribution of Predicted Genes Based on Gene Function in *B. japonicum*

On the basis of gene function, all the predicted 371 SinR regulated genes of *B. japonicum* were classified into 9 different classes. Five genes classified under chemotaxis class, which showed high SinR motif similarity with aprE, epr and yveK genes of *B. subtilis* in the range 100 to 94.66%. Two genes classified under cytochrome class, which showed motif sequence similar to yveK gene of *B. subtilis* in the range 97.75 to 97.27%. Similarly 100 genes classified under enzyme class, which showed motif pattern similar to aprE, epr, yqxM and yveK genes of *B. subtilis* in the range 100 to 93.77%. It seems that different classes of enzymes may transcribe during stress condition and are regulated by SinR TF. Besides, 175 genes classified under hypothetical class, which showed motif patterns similar to aprE, epr and yveK genes of *B. subtilis* in the range 100 to 93.77%. On the other hand, 21 genes were categorized under other protein class, which showed motif pattern similar to aprE, epr and yveK genes of *B. subtilis* in the range 100 to 93.77%. These genes showed no relationship with any functional classes, thus indicates that apart from routine stress related genes few other genes may play a crucial role during stress metabolism. Moreover, 10 genes belong to receptors class, which showed motif pattern similar to aprE, epr and yveK genes of *B. subtilis* in the range 97.75 to 93.77%. Total 22 genes belong to regulatory class, which showed motif pattern similar to aprE, epr and yveK genes of *B. subtilis* in the range 100 to 93.77%. Similarly, total 5 genes belong to ribosomal class, which showed motif pattern similar to aprE and yveK genes of *B. subtilis* in the range 97.75 to 94.62%. Lastly, 31 genes belong to transporter class, which showed motif pattern similar to aprE, epr and yveK genes of *B. subtilis* in the range 100 to 93.77%. It seems that these predicted genes may play an important role in cell signaling and stress metabolism during environmental stress condition (Fig. **2**) (Suppl. Table **1**).

Gene Distribution According to SinR Binding Site Conservation

On the basis of SinR motif sequence similarity, 5 chemotaxis genes were clustered into 3 sub-groups with 96.02% SinR binding site conservation. Results indicate that mcpK gene comprises two binding sites signals, one each at promoter (i.e., -1 to -200 bp) and upstream (i.e., -201 to -400 bp) region, this may be due to the conservation of SinR binding sites in the non-coding genomic sequences. Similarly, all the 100 enzyme encoding genes divided into 9 sub-groups and after comparison 19 genes showed predicted motif similarity of 96.02% with known SinR binding sites, while 4 genes showed lowest conservation i.e., 97.27%, which again above the selectable cut-off. On the same way, all the 175 hypothetical genes divided in to 10 sub-groups and after comparison 36 hypothetical genes showed predicted motif similarity of 94.66%. Similarly, 10 receptor genes divided in to 4 sub-groups and after comparison higher gene frequency of 0.4 was showed by predicted motif with 96.02% & 94.66% similarity, while single gene frequency was showed by motifs with 97.75% and 93.77% conservation. On the same way, all the 22 regulatory genes were divided into 8 sub-groups and after comparison higher gene frequency of 0.32 was showed by predicted motif with 96.02% similarity, while single gene frequency was showed with motif conservation of 100%, 98.6%, 96.35% and 94.62%. Similarly, all the 31 transporter protein encoding genes divided in to 8 sub-groups and after comparison similar gene frequency of 0.26 was showed by predicted motifs with 96.02% & 94.66% sequence similarity, while single gene frequency was showed with motif conservation of 100% and 97.27%. These results indicate that predicted genes may be co-regulated by SinR and play an important role in stress metabolism of *B. japonicum* which are yet to be characterized (Suppl. Table **2**).

Gene Distribution According to Nucleotide Level Conservation

Position specific nucleotide level conservation was analyzed within predicted motif patterns in *B. japonicum*. Genes of chemotaxis class revealed high frequency of patterns similarity with known patterns of *B. subtilis* aprE and epr genes. Total 4 different DNA motif patterns similar to *B. subtilis* SinR binding sites were detected. The higher frequency of motif conservation was showed by pattern TCGTTCCCA.

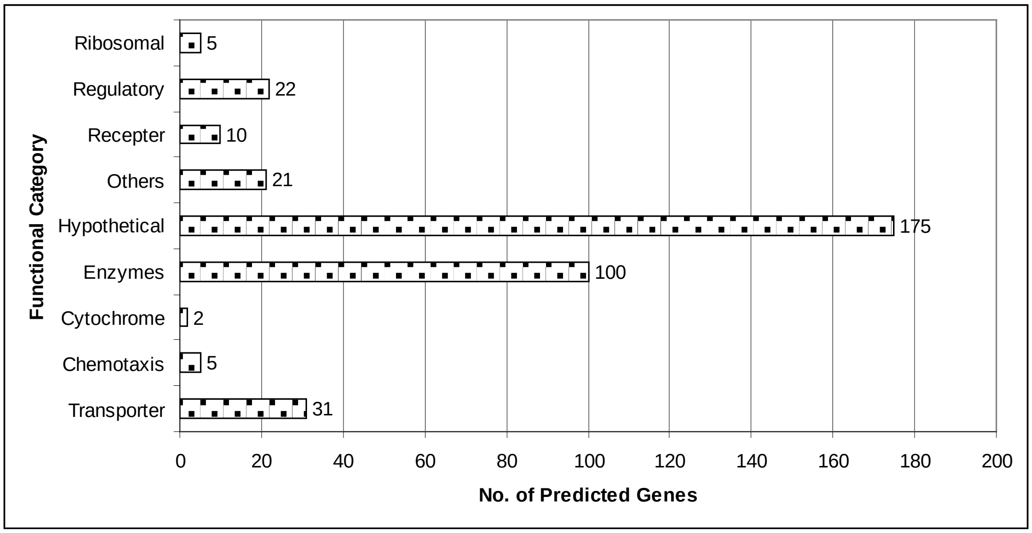

Fig. (2). Distribution of 371 predicted target genes of *B. japonicum* with *Sin-box* motifs in their upstream sequences detected through newly derived weight matrix.

Gene bll4327, encoding putative methyl accepting chemotaxis protein showed highly conserved DNA motif *i.e*. TCGTTCTCA at position -279 to -271 bp with 100% similarity. Besides, gene mcpK, which encode probable methyl accepting chemotaxis protein in *B. japonicum* showed two similar type DNA motifs *i.e.* TCGTTCCCA with 96.02% similarity and localized at position -352 to -344 bp. This motif was found similar to SinR binding site i.e. GTTCCCA of epr gene except first two nucleotides in *B. subtilis*, thus indicates conservation of multiple regulatory sites in *B. japonicum*. Cytochrome genes revealed single pattern distribution in each gene and showed similarity with SinR regulatory patterns of *B. subtilis* aprE and epr genes. Total of 2 different DNA motif patterns similar to *B. subtilis* were detected. No high frequency was detected in this class and single motif pattern distribution was found in each gene, which showed more than 97% similarity with yveK gene of *B. subtilis*. Gene bll0894, which encode putative cytochrome P450 protein in *B. japonicum* showed highly conserved *SinR-box* like DNA motif i.e. TCGTTCTCC at position -164 to -156 bp with 97.75% similarity. Enzymes encoding genes revealed high frequency of patterns similarity with aprE and epr genes of *B. subtilis*. Total 20 different DNA motif patterns were detected. The higher frequency of motif conservation was showed by two patterns TCGTGCTCA and TCGTTCTCC with their distribution in 12 genes each having >96% similarity. Similarly, hypothetical genes showed high frequency of patterns similarity with known regulatory patterns of *B. subtilis* aprE and epr genes. Total of 14 different DNA motif patterns similar to known SinR binding sites were detected in 175 hypothetical proteins. In other protein encoding genes category results indicates high frequency of patterns similarity. Total of 12 different DNA motif patterns similar to known SinR binding sites of *B. subtilis* were detected in the upstream sequences of 21 other protein encoding genes of *B. japonicum*. The higher frequency of motif conservation was shown by patterns TCGTGCTCT and TTGTTCTCC with their distribution in 4 genes each. While receptor genes revealed high frequency of patterns similarity. Total of 6 different DNA motif patterns similar to known SinR binding sites of *B. subtilis* were detected in the upstream sequences of 10 receptor protein encoding genes of *B. japonicum*. The higher frequency of motif conservation was showed by pattern TCGTGCTCA with their distribution in 3 genes. On the other hand, regulatory genes showed high frequency of patterns similarity with known regulatory patterns of *B. subtilis* aprE and epr genes. Total of 9 different DNA motif patterns similar to known SinR binding sites of *B. subtilis* were detected in the upstream sequences of 22 regulatory protein encoding genes of *B. japonicum*. The higher frequency of motif conservation was showed by pattern TCGTGCTCA with their distribution in 4 genes. Ribosomal genes revealed high frequency of patterns similarity. Total of 3 different DNA motif patterns similar to known SinR binding sites of *B. subtilis* were detected in the upstream sequences of 5 ribosomal protein encoding genes of *B. japonicum*. The higher frequency of motif conservation was showed by two patterns (i.e. TCGTTCTCC and TTGTGCTCA) with their distribution of 2 genes in each. Lastly transporter genes revealed high frequency of patterns similarity with known regulatory patterns of *B. subtilis* aprE and epr genes. Total of 11 different DNA motif patterns similar to known SinR binding sites of *B. subtilis* were detected in the upstream sequences of 31 transporter protein encoding genes of *B. japonicum*. The higher frequency of motif conservation was showed by pattern TCGTGCTCA with their distribution in 6 genes. Due to regulatory elements conservation, the predicted genes assumed to be similar in function too and expected to be regulated by SinR TF under stress condition, which are yet to be characterized (Suppl. Table **3**).

Detection of Co-Regulated Genes through Operon Delineation

Operon can be defined as co-transcribed genes to form polycistronic mRNA, which are present in prokaryotes. Most operons are under the control of a single transcriptional promoter located upstream of the first gene of the operon. More complex transcriptional regulations with multiple promoters in a single operon have been reported. It has been estimated that approximately 50% of genes in bacteria are located in operons. In order to characterize the regulatory network and considering above fact, we analyzed 371 predicted SinR regulated genes of *B. japonicum* to delineate operon structure. On the basis of genes conservation, we predicted only 221 genes revealing their presence in operons in the form of conserved gene pairs. Results of operon delineation support the different theories of transcriptional unit formation. Most of the predicted operons support the view that operons evolved to ensure that genes are co-regulated. This view is supported by the functionally related genes such as enzymes catalyzing subsequent steps within metabolic pathway or are members of a single protein complex etc. On the other hand, results of operons prediction also support the other view of 'Selfish operon'. In this view non-essential genes form operons via horizontal gene transfer to protect themselves from being removed from the genome. This view is based on the observation that numerous orthologous operons are conserved across bacterial and archaeal species. In the studied work, we have successfully showed the presence of predicted genes with their neighboring genes in the operons and divided the operons on the basis of gene functions. Identification of first regulatory gene in the operon is a difficult task and so far no bioinformatics tool has been developed, so we have considered all the 221 predicted genes as potential regulatory targets. Besides, highly conserved genes pairs of operons were scored with high confidence value, which define high probability for gene pair occurrence. The term 'gene pair' refers to two adjacent genes separated by 200 bp and thus considered in an operon.

Out of total 371 predicted genes, 221 genes showed their presence in operons. Four chemotaxis genes with gene pair's conservation upto 96 bacterial genomes, 2 cytochrome genes with gene pair's conservation upto 15 genomes, 75 enzyme coding genes with gene pair's conservation upto 225 bacterial genomes, 76 hypothetical genes with gene pairs conservation upto 118 bacterial genomes, 16 other protein encoding genes with gene pairs conservation upto 89 bacterial genomes. Similarly, 7 receptor genes showed their role in operons with gene pair's conservation upto 113 bacterial genomes, 8 regulatory genes with gene pair conservation upto 43 bacterial genomes, 5 ribosomal genes with gene pair conservation upto 367 bacterial genomes and 28 transporter genes with gene pair conservation upto 297 bacterial genomes (Suppl. Table **4**).

Table 1. Experimental Details of SinR Transcription Factor in *B. subtilis*

Transcription Factor	SinR
Domain	HTH_3 (Helix-turn-helix)
SubtiList	BG10754
UniProtKB/Swiss-Prot entry	P06533
Factor type	Xre
Motif (consensus seq.)	GTTCTCT
Length (bp)	7
Comment	Dual-function regulator which is essential for the late-growth processes of competence and motility and is also a repressor of others, e.g., sporulation and subtilisin synthesis. Might be a leucine zipper protein.
Reference	DBTBS database (Sierro *et al*., 2008)

MATERIALS & METHOD

Retrieval of Genomic Data

The genomic non-coding upstream sequences of both *B. subtilis* and *B. japonicus* were extracted through MicroBial Genome Database (MBGD) [27] and Regulatory Sequence Analysis Tool (RSAT) webserver [28], while the reported information of SinR regulatory protein and its binding sites were retrieved through *B. subtilis* known regulatory network database *viz*., DBTBS [14]. Note that no weight matrix of SinR-box is available in DBTBS database (Table **1**). The evidence of SinR potential ortholog along with conserved protein motifs and domains was detected with the help of BLASTp program [29] and MicroBial Genome Database (MBGD) [27]. Besides, similar conserved protein domains were also detected through CDD search program [30]. Detected SinR ortholog was further verified through COG database [31] (Suppl. Table **5**). An analysis of the predicted amino acid sequence of *B. japonicum* SinR revealed a potential leucine zipper protein dimerization motif which is flanked by two helix-turn-helix motifs that could be involved in recognizing two different dyad symmetries similar to *B. subtilis*.

Selection of Reference Data Set

A set of reference data was selected on the basis of sequence characteristics and same kind of imperfections as observed with real sequences. However, information of experimentally known *SinR-box* motif for SinR transcription factor and related real sequences were retrieved from DBTBS database and referred as reference data set (or control data set), which comprises conserved upstream sequences of *B. subtilis* five genes namely *aprE* encoding serine alkaline protease (subtilisin E); *epr* encoding extracellular serine protease; *yqxM* encoding hypothetical protein; *yveK* encoding hypothetical protein and *rok* encoding repressor protein [32-37]. To define acceptable threshold or cut-off during weight matrix based genomic predictions, we included the upstream sequence of *B. japonicum* gene *viz*., *blr3044* encoding extra-cellular protease. Gene *blr3044* was considered as an ortholog of *B. subtilis* aprE gene, keeping in mind that SinR binding site would be there, since functional sites are evolutionary conserved in nature (Table **2**). Promoter segments for the respective genes have been retrieved on the basis of non-coding sequences without overlapping open reading frame (ORF) upstream sequences because in bacteria, due to the organization of genes into operon, it is preferable to prevent overlap with upstream ORFs and also to prevent including too many coding sequences when the genes are located in operon, e.g. in *Escherichia coli*, 25% of the genes have an upstream neighbor closer than 50 bp. When this neighbor is on the same strand as the gene, it might indicate that they belong to the same operon. For upstream sequences, the reference position is the ORF start, *i.e.* the first nucleotide of the start codon *i.e.* in the 5’ end flank of the gene, e.g. from -1 to -400 bp (Suppl. Table **6**). Both reference data set and genomic upstream sequences were retrieved through RSAT nucleotide database parsed from GenBank (NCBI) [38].

Construction of *SinR-box* Weight Matrix

There are several approaches to derive the parameters of a weight matrix characterizing the sequences specificity of a DNA binding protein. Here we used the method that starts with a reference set of known *B. subtilis* SinR binding sites, as natural promoter sequences compiled from the biological literature & DBTBS database and *in silico* predicted binding site of aprE ortholog in *B. japonicum*. The sequences of the reference data set were typically longer than the actual recognition sequences and therefore we first aligned all the reference data set sequences before weight matrix construction. Finally the 9 bp conserved pattern was converted into a table of base frequencies from which the position specific weights were calculated according to CONSENSUS algorithm [17, 28]. The newly derived matrix was further verified through MEME [18, 39] and D-Matrix (http://203.190.147.116/dmatrix) programs.

A total of six genes upstream sequences were used for weight matrix construction and considered as reference data (Tables **3** & **4**). Out of those six, five known promoters with known *SinR-box* were from *B. subtilis* genes namely, *aprE*, *epr*, *yqxM*, *yveK* & *rok* and one upstream sequence with *SinR-box* was from *B. japonicum* gene namely, *blr3044*. Following their alignment a matrix was constructed from the relative frequencies of A, T, C or G at each position of the 9 bp *SinR-box* motif sequence. This matrix was used to determine an information-based measure of potential binding sites

Table 2. Experimental Details of Known SinR Binding Sites Used as Reference Data for Weight Matrix Construction

Organism	Gene	Gene ID	Synonym	Start: End	SinR Binding Site	Length	Strand	Function	Reference
B. subtilis	aprE	939313	sprE	-250:-350	GTTCTCA	101	R	Serine alkaline protease (subtilisin E)	Ferrari *et al*., 1988
B. subtilis	epr	937332	ipa-15r	-350:-450	GTTCCCA	101	D	Extracellular serine protease	Sloma *et al*., 1988; Crutz & Steinmetz, 1992
B. subtilis	yqxM	938532	yqhD	-30:-145	ATTCTTT GTTCTTT AGAGAAC	116	R	Hypothetical protein	Yoshida *et al*., 2003 ; Kearns *et al*., 2005
B. subtilis	yveK	938582	epsA	-100:-225	GTTCTTT GTTATTT GTTCATT AGAGAAC GTTCTCT	126	R	Hypothetical protein	Yoshida *et al*., 2003 ; Kearns *et al*., 2005
B. subtilis	rok	938793	ykuW	-70:-170	ATTCTCA	101	D	Repressor protein	Hoa *et al*., 2002
B. japonicum	blr3044	1053287	aprE	-300:-400	GTGCTCC	101	D	Probable extracellular protease	Predicted aprE ortholog with predicted SinR binding site.

Table 3. Construction of *Sin-box* Weight Matrix through RSAT Tool

Gene	Length (bp)	Strand	Start Position	Alignment of Known SinR Binding Site (width = 9 bp)
aprE	101	R	29	T T G T T C T C A
epr	101	D	24	T C G T T C C C A
yqxM	116	D	98	T C G T T C T C T
yveK	126	R	97	T T G T T C T C T
rok	101	R	6	T T A T T C T C A
blr3044	101	D	54	T C G T G C T C C

Position	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
A	0	0	1	0	0	0	0	0	3
C	0	3	0	0	0	6	1	6	1
G	0	0	5	0	1	0	0	0	0
T	6	3	0	6	5	0	5	0	2
SUM	6	6	6	6	6	6	6	6	6
Consensus	**T**	**C/T**	**G**	**T**	**T**	**C**	**T**	**C**	**A**

according to the standard method [40]. A 9 bp motif region was moved over the entire genome on both strands and the score (S_i) at each nucleotide position (having base i) was calculated according to Equation:

$$S_i = (1/9)\ \Sigma_j\ [2 + \log2(F_{ij})]$$

where F_{ij} is the frequency matrix for base *i* at position *j*.

This score, which ranges from −17.513 (the score of the worse match) to 9.616 (the score of the good or exact match with known *B. subtilis SinR-box* consensus sequence), is a measure of the information content of a potential binding site measured against the reference data set (Table **4**). The lowest reference score for *B. subtilis* genome, that of gene *yqxM*,

Table 4. Reconstruction of *Sin-box* Weight Matrix through MEME & D-Matrix Tools

Gene	Length (bp)	Strand	Start Position	ln (P)	Alignment of Known SinR Binding Site (width = 9 bp)
aprE	101	+	29	-11.20	T T G T T C T C A
epr	101	+	24	-10.38	T C G T T C C C A
yqxM	116	-	98	-11.20	T C G T T C T C T
yveK	126	+	97	-10.72	T T G T T C T C T
rok	101	-	6	-9.40	T T A T T C T C A
blr3044	101	+	54	-9.96	T C G T G C T C C

Position	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
A	0	0	1	0	0	0	0	0	3
C	0	3	0	0	0	6	1	6	1
G	0	0	5	0	0	0	0	0	0
T	6	3	0	6	6	0	5	0	2
Sum	6	6	6	6	6	6	6	6	6
Consensus	**T**	**C/T**	**G**	**T**	**T**	**C**	**T**	**C**	**A**

was 4.22 (80.1%), while lowest reference score for *B. japonicum* genome, that of gene blr3044, was 7.93 (93.77%) and thus, a threshold of 7.92 (93%) was used to define a 'Genome wide good & acceptable hits'. A scan of the entire *B. japonicum* genome produced about 371 motif hits on both strands [refers as Direct (D) and Reverse strand (R)]. These motif were again filtered to retain only those that were in between −400 to -1 bp on both strands from an annotated translational start site and having weight score ≥ 7.93 (93.77%) using CONSENSUS & PATSER algorithms [28].

Matrix Based Identification of SinR Binding Sites in the Reference Data Set

The use of weight matrix based DNA binding motif prediction is justified by the statistical mechanical theory [1]. Moreover, weight matrices were shown to be quite accurate binding affinity predictors for several transcription factors of bacterial origin. Therefore, newly derived weight matrix was then used to search out the known *SinR-box* like motif within reference data set sequences by using PATSER algorithm at RSAT webserver. For setting acceptable cut-off range in scanning of *B. japonicum* genomic non-coding sequences, we used score value of blr3044 as standard threshold. For better understanding all the matching scores values were transformed into the form of percentage and referred as matrix sequence similarity percentage (Table **5**).

Genome Wide SinR Binding Site Identification

To identify regulated genes, the known reference data set was first analyzed and validated with high-scoring matches

Table 5. Predicted Known SinR Binding Sites in the Upstream Sequences of Reference Data Set Genes through Constructed Weight Matrix

Gene	Strand	Start	End	Predicted *Sin-box*[1]	Score	MSSP[2] (%)	ln(P)	Sn= TP/n
yqxM	R	-19	-11	acaaTCGTTCTCTttaa	9.25	98.64	-11.65	7/8 =0.875 or 87.5%
aprE	D	-73	-65	tccaTTGTTCTCAcgga	9.24	98.60	-11.20	
yveK	D	-30	-22	gattTTGTTCTCTaaag	8.88	97.27	-10.72	
epr	D	-78	-70	cacTCGTTCCCAaaca	8.54	96.02	-10.38	
blr3044	D	-48	-40	gcgcTCGTGCTCCaggc	7.93	93.77	-9.95	
rok	R	-96	-88	tcctTTATTCTCAaggg	7.45	92.00	-9.40	
yveK	D	-92	-84	ttttTCGTTCTTTataa	5.82	85.99	-7.66	
yqxM	D	-46	-38	ttgaTTATACTCTattt	4.22	80.10	-6.43	

Note: [1]Predicted *Sin-box* motifs similar to known *B. subtilis* binding sites (in shaded & upper case).
[2]MSSP (Motif Score Similarity Percentage) = Similarity of predicted motif score with known *Sin-box* consensus sequence score.
Sn= Sensitivity, TP= no. of true positive predicted sites, n= no. of total known sites.

to the known *SinR-box* by using frequency matrix based pattern matching tool 'PATSER' implemented at RSAT. We considered one nucleotide variation in conserved pattern and set the search parameter at '1' substitution level instead of '0', because at 'zero' substitution level only limited genes were resulted while at '1' mismatch the probability of mutation by one nucleotide is expected. In this method, a position specific frequency matrix representing a consensus sequence was converted to a positional weight matrix or PSSM, which was later used to score the motif sequence according to the standard scoring system [4]. Finally we observed significant conservation in non-coding upstream sequences and then used this conservation to improve the predictions. After setting the standard cut-off or threshold, in *B. japonicum*, genes having *SinR-box* like conserved consensus pattern in their upstream sequences were predicted through the PATSER algorithm at RSAT webserver. This computational search was done at pattern width 9 bp and at one substitution level parameter because the length of known *SinR-box* motif was long enough i.e. 9 bp, therefore at one mismatch the probability of mutation by one nucleotide was considered. It is preferential to consider one nucleotide variation in conserved region of genes upstream because at zero substitution level only limited genes were matched with the matrix. Here, matching score was calculated as per PATSER algorithm using prior nucleotide frequency as A:T 0.2 and G:C 0.3 and at a threshold of 7.92. The constructed matrix was searched in the whole genome and identified the genes having similarity $\geq$ 90% so that the spurious and false positive predictions may be avoided. Here, score 9.62 is taken for 100% conservation of SinR binding sites. Genomic predictions through RSAT were further verified through PossuMsearch program.

Calculations of Statistical Parameters

To measure the prediction accuracy of newly derived *SinR-box* matrix, predicted motifs were analyzed in term of MSSP (Motif Score Similarity Percentage). MSSP means similarity of predicted motif score with known *SinR-box* consensus sequence (TTGTTCTCA) score in terms of percentage. For RSAT program based predictions, maximum score of 9.616 revealed best or exact similarity while minimum score of -17.513 for poor or worst similarity. Percentage was calculated out of 'range of scores' i.e. 27.129. Beside, score based predictions were further evaluated by probability value (P-value). A P-value for matrix based scoring was computed from the score distribution obtained with the weight matrix applied to 1000 randomized sequences with the same length and AT content as the original sequence. It is widely accepted measure of the significance. The MSSP was calculated as:

For known SinR TFBS,

Maximum score (S_{max}) = 9.616

Minimum score (S_{min}) = - 17.513

Range of scores (R_{known}) = Maximum score (S_{max}) – Minimum score (S_{min})

= 9.616 - (-17.513) = 27.129 = 27.13

Range of predicted motif score (Observed score) = Predicted score– Minimum score

$R_{pred} = S_{pred} - S_{min}$

Motif Score Similarity % = (Range of predicted motif score/Range of score) x 100

(MSSP) $P_{\%} = (R_{pred} / R_{known}) \times 100$

Calculation of Sensitivity

At weight score cutoff 4.22 (80.1% MSSP) – in case of *B. subtilis*

No. of true positive sites (TP) = 7

No. of total predicted sites (n) = 8

Sensitivity = TP/n =7/8 = 0.875 or 87.5%

To reduce the false positives, we increased the cutoff score (or threshold) value as follows:

At weight score cutoff 7.93 (93.77% MSSP) – in case of *B. japonicum*

No. of true positive sites (TP) = 5

No. of total predicted sites (n) = 5

Sensitivity = TP/n =5/5 = 1 or 100%

We also compared the predicted results with known SinR binding sites of *B.subtilis* and calculated the sensitivity as follows:

At weight score cutoff 4.22 (80.1% MSSP) – in case of *B. subtilis*

No. of true positive predicted sites (TP) = 8

No. of total known sites of *B. subtilis* (n) = 12

Sensitivity = TP/n =8/12 = 0.666 or 66.6%

Thus, sensitivity of predictions through newly constructed weight matrix is good enough and is in acceptable range.

Operon Delineation

Most of the genes in bacteria are organized into operons. These are transcribed into a single mRNA molecule. The co-transcribed genes often play related roles in the function of the organism, sometimes binding to one another or acting as part of the same metabolic pathway. In such an organization, the co-transcribed genes are co-regulated at the transcriptional level. Identifying the genes that are grouped together into operons may enhance our knowledge of gene regulation and function. Computational algorithms have been developed to locate the operons. The previously developed methods are based on finding the signals that occur on the boundaries of operons: transcription promoters on the 5' end and terminators on the 3' end. Such approaches can only be effective for organisms whose promoters and terminators are well known i.e., *E. coli*. An alternative method to predict operons is based on finding gene clusters where gene order and orientation is conserved in two or more genomes. This approach does not rely on experimental data, but instead uses the genome sequence and gene locations [41].

Identification of gene pairs has imparted the strong support in operon delineation. The presence of predicted genes in operon is based on the comparison of complete bacterial genomes. The analysis reveals a large number of conserved gene clusters sets of genes having the same order in two or more different genomes. Furthermore, we have analyzed some of the predicted genes having binding site for SinR transcription factor and paired with other neighboring genes in *B. japonicum* genome. Therefore, it can be assumed that the SinR regulatory protein have some role in operon regulation. Here we used a computational approach that finds such conserved gene clusters and assigns to each one a probability that the cluster is an operon. The predicted genes with conserved *SinR-box* like motif were analyzed for operon organization using web based 'Operon Finder' tool [41]. The tool follow a method to detect and analyze conserved gene pairs that are located close on the same DNA strand in two or more bacterial genomes. The gene pairing was estimated by confidence value (C-value) which is an estimation of the lower boundary of the probability that the two corresponding genes are located in the same operon, while 'n' is a number of other genomes that have the same pair of genes located in the same direction.

CONCLUSION

As an initial step toward understanding the molecular mechanism of stress tolerance through SinR master regulator in *B. japonicum*, we described in this paper the prediction of genes having SinR binding sites, distribution of different class of genes on the basis of motif weight matrix score (in %) and motif patterns on the basis of nucleotide level conservation. After that we identified the presence of predicted genes in operons and their role in stress tolerance, so that to unfold the unexplored regulatory gene network. Through constructed weight matrix of SinR binding site or *sin-box*, we first identified the known DNA motifs within the reference data set of *B. subtilis* and set the acceptable cut-off for genome wide search in *B. japonicum*. Our newly derived weight matrix successfully predicted the known SinR binding sites in genes of *B. subtilis* namely, *yqxM, aprE, yveK, epr* and *rok* with 98.64%, 98.60%, 97.27%, 96.02% and 92% motif similarity and showed weight scores 9.25, 9.24, 8.88, 8.54 and 7.45 respectively. While genes *viz*., *yvek* and *yqxM* showed low motif similarity i.e. 85.99% and 80.10% (score 5.82 and 4.22 respectively). While the identified orthologous gene of extra-cellular protease (aprE) in *B. japonicum* namely, *blr3044* showed 93.77% motif similarity with known sites and 7.93 motif score. To analyze the genome of *B. japonicum*, we therefore set the acceptable cut-off of 93.77% MSSP (or 7.93 score) for scanning of whole genome upstream sequences, so that to avoid false positives. We found total 371 genes with *SinR-box* like sequences in their upstream sequences (-1 to -400 bp) in the range of 100 to 93.77% (or score 9.62 to 7.93) as true positives and therefore hypothesized as potential SinR regulated genes in *B. japonicum* genome. Results showed that most of the predicted genes of *B. japonicum* were already reported in other system such as *B. subtilis*, *Sinorhizobium meliloti* and *Clostridium acetobutylicum*, thus it seems that our predictions are nearer to experimental results. To identify the regulatory network, we further analyzed their presence in operons. Out of total 371 genes, 221 showed their presence in operons. However, 11 genes were found false positives, thus rejected. We also identified the ortholog of *bsu24610* (*sinR*) gene of *B. subtilis* in *B. japonicum* as *bll1669*, encode for hypothetical protein, but after sequence as well as structural comparison both showed similar protein domain i.e. cd00093 and motif i.e. HTH-3. Finally we deduced that apart from reported genes, number of other genes may also be required in alternate development pathways during adaptive mechanisms such as production of extra-cellular enzymes & antibiotics, acquisition of competence and development of motility & sporulation regulated under adverse environmental conditions. In the present communication, we are reporting several high affinity genes of *B. japonicum* for SinR binding. The study was first modeled in *B. subtilis* and then complete genome of *B. japonicum* was traced for the identification of co-regulated genes. Through this method we successfully identified the genes of alternate development pathways regulated under different environmental adverse conditions. The same approach could be utilized in other unexplored family members and later this information could be utilized to work out the regulatory network more clearly.

ACKNOWLEDGEMENT

We acknowledge Department of Biotechnology, New Delhi for financial support at Central Institute of Medicinal & Aromatic Plants (Council of Scientific & Industrial Research), Lucknow (UP) India.

SUPPLEMENTARY MATERIAL

Supplementary material can be viewed at
www.bentham.org/open/tobioij

REFERENCES

[1] O.G. Berg, and P.H. Von Hippel, "Selection of DNA binding sites by regulatory proteins. Statistical mechanical theory and application to operators and promoters", *J. Mol. Biol.*, Vol. 193, pp. 723-750, February 1987.

[2] S. Henikoff, G.W. Haughn, J.M. Calvo, and J.C. Wallace, "A large family of bacterial activator proteins", *Proc. Natl. Acad. Sci. USA*, Vol. 85, pp. 6602-6606, September 1988.

[3] G.Z. Hertz, G.W. Hartzell, and G.D. Stormo, "Identification of consensus patterns in unaligned DNA sequences known to be functionally related", *Comput. Appl. Biosci*., Vol. 6(2), pp. 81-92, April 1990.

[4] G.Z. Hertz, and G.D. Stormo, "Identifying DNA and protein patterns with statistically significant alignments of multiple sequences", *Bioinformatics*, Vol. 15(7-8), pp. 563-77, July-August 1999.

[5] G.D. Stormo, "DNA binding sites: representation and discovery", *Bioinformatics*, Vol. 16(1), pp. 16-23, January 2000.

[6] H. Li, V. Rhodius, C. Gross, and E. D. Siggia, "Identification of the binding sites of regulatory proteins in bacterial genomes", *Proc. Natl. Acad. Sci. USA*, Vol. 99(18), pp. 11772-77, September 2002.

[7] F. Khan, S. Agarwal, and B.N. Mishra, "Genome wide identification of DNA binding motifs of NodD-factor in *Sinorhizobium meliloti* and *Mesorhizobium loti*", *J. Bioinform. Comput. Biol.*, Vol. 3(4), pp. 773-801, August 2005.

[8] N.K. Gaur, J. Oppenheim, and I. Smith, "The *Bacillus subtilis* sin gene, a regulator of alternate developmental processes, codes for a DNA-binding protein", *J. Bacteriol.*, Vol. 173(2), pp. 678-86, January 1991.

[9] U. Bai, M.M. Innes, and I. Smith, "SinI modulates the activity of SinR, a developmental switch protein of *Bacillus subtilis*, by protein-protein interaction", *Genes Dev.*, Vol. 7(1), pp. 139-48, January 1993.

[10] M.M. Ines, L. Doukhan, and I. Smith, "The *Bacillus subtilis* SinR protein is a repressor of the key sporulation gene *spo0A*", *J. Bacteriol.*, Vol. 177(16), pp. 4619-27, August 1995.

[11] M.A. Cervin, R.J. Lewis, J.A. Brannigan, and G.B. Spiegelman, "The *Bacillus subtilis* regulator SinR inhibits spoIIG promoter transcription *in vitro* without displacing RNA polymerase", *Nucleic Acids Res*., Vol. 26(16), pp. 3806-12, August 1998.

[12] N. Gottig, M.E. Pedrido, M. Méndez, E. Lombardía, A. Rovetto, V. Philippe, L. Orsaria, and R. Grau, "The *Bacillus subtilis* SinR and RapA developmental regulators are responsible for inhibition of spore development by alcohol", *J. Bacteriol.*, Vol. 187(8), pp. 2662-72, April 2005.

[13] F. Chu, D.B. Kearns, A. McLoon, Y. Chai, R. Kolter, and R. Losick, “A novel regulatory protein governing biofilm formation in *Bacillus subtilis*”, *Mol. Microbiol.*, Vol. 68(5), pp. 1117-27, June 2008.

[14] N. Sierro, Y. Makita, M.J.L. de Hoon, and K. Nakai, “DBTBS: a database of transcriptional regulation in *Bacillus subtilis* containing upstream intergenic conservation information”, *Nucleic Acids Res.*, Vol. 36(Database issue), pp. D93-D96, January 2008.

[15] M.M. Marketon, M.R. Gronquist, A. Eberhard, and J.E. González, “Characterization of the *Sinorhizobium meliloti* SinR/sinI locus and the production of novel N-acyl homoserine lactones”, *J. Bacteriol.*, Vol. 184(20), pp. 5686-95, October 2002.

[16] T. Kaneko, Y. Nakamura, S. Sato, K. Minamisawa, T. Uchiumi, S. Sasamoto, A. Watanabe, K. Idesawa, M. Iriguchi, K. Kawashima, M. Kohara, M. Matsumoto, S. Shimpo, H. Tsuruoka, T.Wada, M. Yamada, and S. Tabata, “Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110”, *DNA Res.*, Vol. 31, 9(6), pp. 189-97, December 2002.

[17] M. Thomas-Chollier, O. Sand, J.V. Turatsinze, R. Janky, M. Defrance, E. Vervisch, S. Brohee, and J. van Helden, **“**RSAT: Regulatory Sequence Analysis Tools”, *Nucleic Acids Res.*, Vol. 36(Suppl 2), pp. W119-W127, July 2008.

[18] T.L. Bailey, N. Williams, C. Misleh, and W.W. Li, “MEME: discovering and analyzing DNA and protein sequence motifs”, *Nucleic Acids Res.*, Vol. 34(Web Server issue), pp. W369-73, July 2006.

[19] M. Beckstette, D. Strothmann, R. Homann, R. Giegerich, and S. Kurtz, “Fast index based algorithms for matching position specific scoring matrices”, *BMC Bioinformatics*, Vol. 7, p. 389, August 2006.

[20] C. Pizzi, P. Rastas, and E. Ukkonen, “Fast search algorithms for position specific scoring matrices”, in *Bioinformatics Research and Development*, Vol. 4414, Springer Berlin: Heidelberg publisher, May 2007, pp. 239-250.

[21] N.K. Gaur, K. Cabane, and I. Smith, “Structure and expression of the *Bacillus subtilis* sin operon”, *J. Bacteriol.*, Vol. 170(3), pp. 1046-53, March 1988.

[22] L. Liu, M.M. Nakano, O.H. Lee, and P. Zuber, “Plasmid-amplified coms enhances genetic competence and suppresses SinR in *Bacillus subtilis*”, *J. Bacteriol.*, Vol. 178(17), pp. 5144-52, September 1996.

[23] J. Olmos, R. de Anda, E. Ferrari, F. Bolivar, and F. Valle, “Effects of the SinR and degU32 (Hy) mutations on the regulation of the aprE gene in *Bacillus subtilis*”, *Mol. Gen. Genet.*, Vol. 253(5), pp. 562-67, February 1997.

[24] P. Kodgire, M. Dixit, and K.K. Rao, “ScoC and SinR negatively regulate *epr* by corepression in *Bacillus subtilis*”, *J. Bacteriol*., Vol. 188(17), pp. 6425-28, September 2006.

[25] M. Ogura, K. Shimane, K. Asai, N. Ogasawara, and T. Tanaka, “Binding of response regulator DegU to the aprE promoter is inhibited by RapG, which is counteracted by extracellular PhrG in *Bacillus subtilis*”, *Mol. Microbiol.*, Vol. 49(6), pp. 1685-97, September 2003.

[26] M.H. Rashid, and J. Sekiguchi, “*flaD* (*SinR*) mutations affect SigD-dependent functions at multiple points in *Bacillus subtilis*”, *J. Bacteriol.*, Vol. 178(22), pp. 6640-43, November 1996.

[27] I. Uchiyama, “MBGD: microbial genome database for comparative analysis”, *Nucleic Acids Res.*, Vol. 31(1), pp. 58-62, January 2003.

[28] J. Van Helden, B. Andre, and J. Collado-Vides, “A web site for the computational analysis of yeast regulatory sequences”, *Yeast*, Vol. 16(2), pp. 177-187, January 2000.

[29] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman, “Gapped BLAST and PSI-BLAST: a new generation of protein database search programs”, *Nucleic Acids Res*., Vol. 25(17), pp. 3389-402, September 1997.

[30] A. Marchler-Bauer, A.R. Panchenko, B.A. Shoemaker, P.A. Thiessen, L.Y. Geer, and S.H. Bryant, “CDD: a database of conserved domain alignments with links to domain three-dimensional structure”, *Nucleic Acids Res.*, Vol. 30(1), pp. 281-3, January 2002.

[31] R.L. Tatusov, N.D. Fedorova, J.D. Jackson, A.R. Jacobs, B. Kiryutin, E.V. Koonin, D.M. Krylov, R. Mazumder, S.L. Mekhedov, A.N. Nikolskaya, B.S. Rao, S. Smirnov, A.V. Sverdlov, S. Vasudevan, Y.I. Wolf, J.J. Yin, and D.A. Natale, “The COG database: an updated version includes eukaryotes”, *BMC Bioinformatics*, Vol. 4, p. 41, September 2003.

[32] E. Ferrari, D.J. Henner, M. Perego, and J.A. Hoch, “Transcription of *Bacillus subtilis* subtilisin and expression of subtilisin in sporulation mutants”, *J. Bacteriol.*, Vol. 170(1), pp. 289-295, January 1988.

[33] A. Sloma, A. Ally, D. Ally, and J. Pero, “Gene encoding a minor extracellular protease in *Bacillus subtilis*”, *J. Bacteriol.*, Vol. 170(12), pp. 5557-5563, December 1988.

[34] A.M. Crutz, and M. Steinmetz, “Transcription of the *Bacillus subtilis* sacX and sacY genes, encoding regulators of sucrose metabolism, is both inducible by sucrose and controlled by the DegS-DegU signalling system”, *J. Bacteriol*., Vol. 174(19), pp. 6087-95, October 1992.

[35] K. Yoshida, H. Yamaguchi, M. Kinehara, Y.H. Ohki, Y. Nakaura, and Y. Fujita, “Identification of additional TnrA-regulated genes of *Bacillus subtilis* associated with a TnrA box”, *Mol. Microbiol*., Vol. 49(1), pp. 157-65, July 2003.

[36] D.B. Kearns, F. Chu, S.S. Branda, R. Kolter, and R. Losick, “A master regulator for biofilm formation by *Bacillus subtilis*”, *Mol. Microbiol.*, Vol. 55(3), pp. 739-49, February 2005.

[37] T.T. Hoa, P. Tortosa, M. Albano, and D. Dubnau, “Rok (YkuW) regulates genetic competence in *Bacillus subtilis* by directly repressing comK”, *Mol. Microbiol.*, Vol. 43(1), pp. 15-26, January 2002.

[38] D.A. Benson, K.M. Ilene, D.J. Lipman, J. Ostell, and D.L. Wheeler, “GenBank”, *Nucleic Acids Res.*, Vol. 34(Database issue): pp. D16-D20, January 2006.

[39] C.E. Lawrence, and A.A. Reilly, “An expectation maximization (EM) algorithm for the identification and characterization of common sites in unaligned biopolymer sequences”, *Proteins Struct. Funct. Bioinform.*, Vol. 7(1), pp. 41-51, February 1990.

[40] T.D. Schneider, G.D. Stormo, L.Gold, and A. Ehrenfeucht, “Information content of binding sites on nucleotide sequences”, *J. Mol. Biol.*, Vol. 188(3), pp. 415-31, April 1986.

[41] M.D. Ermolaeva, O. White and S. L. Salzberg, “Prediction of operons in microbial genomes”, *Nucleic Acid Res.*, Vol. 29(5), pp. 1216-21, March 2001.

Received: December 27, 2008 Revised: March 14, 2009 Accepted: March 19, 2009