Journal of Computational Intelligence in Bioinformatics.
ISSN: 0973-385X Volume 1 Number 2-3 (2008) pp. 141–146

http://www.ripublication.com/jcib.htm

In-silico prediction of the essentiality of Osmoprotectant Genes of *Pseudomonas aeruginosa*

Pramod Katara and Vinay Sharma

Department of Biosciences and Biotechnology,
Banasthali University,
Banasthali, India
E-mail: pmkatara@yahoo.co.in

Abstract

Genes that are very essential for any organism share similarities with the essential genes of other organisms because most of the basic protein requirements remain same for different organisms and essential genes are conserved more than the non essential counterparts in bacteria and other organism. The present study compares gene expression information with the sequence pattern. Genes that show considerable over and under-expression under osmotic stress and provide strength and suitable proteome to *Pseudomonas aeruginosa* to survive in such stress conditions are compared with the database of essential genes to observe their essentiality and also treated with MEME to analyze regulatory motif in their upstream region. It was found that 70 % of genes which show expression deviation under osmotic stress are essential and most of them are shared by *Mycobacterium tuberculosis.* Finding of regulatory motif shows that no motif was shared by all of the genes and the length of motif is vary very considerably.

Keyword: co-regulation, gene-expression, drug target, essential genes, Drug target, Information content

Background

Each cell has thousands of genes which are responsible for different cellular functions. Many of these are indispensable for the cell, which are collectively called as ‘essential genes’, rest may be called as ‘non essential genes’ which are not very important. The protein-coding essential genes encode ‘essential proteins’, which are vital for cell (1, 2). Such genes are considered as housekeeping genes of the cell,

which are frequently common for different organism. Bacteria are able to survive in very vast environmental condition, because of the presence of some genes which are responsible for their survival in such conditions. It is not essential that all housekeeping genes are expressed under each of these different environments. Their expression depends on the requirement of their product in such conditions. In absence of such essential genes cells are not able to survive. There is a DEG database which shows the collection of gene sequence which is essential for the life (10).

Pseudomonas aeruginosa is an opportunistic pathogen of human shows adaptive response against osmotic stress condition and is able to survive in osmotically stressful conditions (3). Differentially expressed genes in osmotic stress conditions are responsible for the prevalence and persistence of *P. aeruginosa* under these conditions (4). *P. aeruginosa* shows accumulation of osmoprotectants (5,6). It has been established that 66 genes of *P. aeruginosa* shows drastic changes in their expression, in response to the osmotic stress. These 66 genes are known as the steady-state osmotic stress regulon (7). Out of these 66 genes 40 are also associated with the virulence factor expression, encoding protein of a type III secretion system and two ancillary chaperones. Further, their role in cystic fibrosis has also been suggested (8, 9). The present paper attempts to sketch essentiality of these genes to *P. aeruginosa* and traces their presence in other organism. Paper also represents the finding of regulatory sites called transcription factor binding sites (TFBS) using upstream promoter regions of the genes.

Methodology

The genes that show over-expression and under down expression under osmotic stress (7) detected by cDNA microarray experiment have been short-listed. The corresponding sequences were downloaded from GenBank (http://www.ncbi.nlm.nih.gov/Genbank). The gene sequences were used to search for the presence of their orthologs in the DEG database (10) using Blast-n utility (www.ncbi.nlm.nih.gov/blast), we consider only those genes orthologs that shows similarity in form of E value below 1. Further for analysis of the regulatory sites of these genes first we found upstream region for all genes by using the service of RSAT server (http://rsat.ulb.ac.be/rsat/), and them we analyze this upstream region for regulatory motif finding by the help of MEME server (http://meme.sdsc.edu/meme/meme.html).

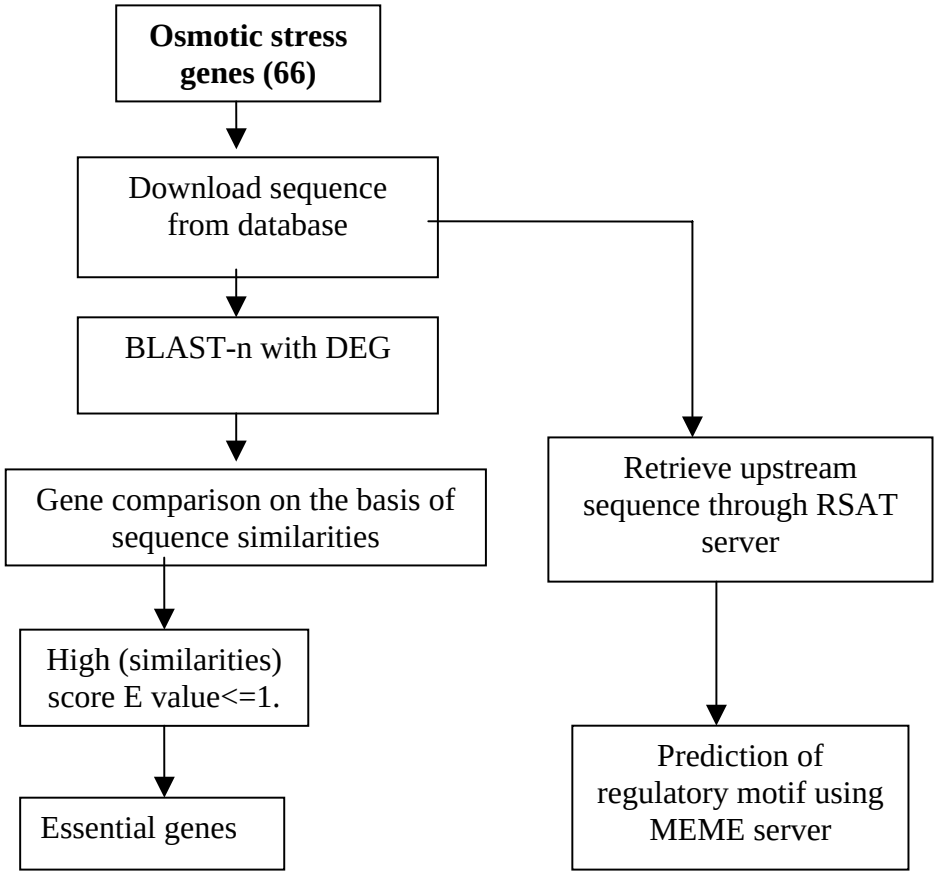

Figure 1: flow diagram of the methodology.

Result and Discussion

Genes that show active response against osmotic stress condition are compared with the DEG and found that most of the genes show very high similarities with the essential genes of many bacteria belonging to different *heterologous taxonomic groups*.

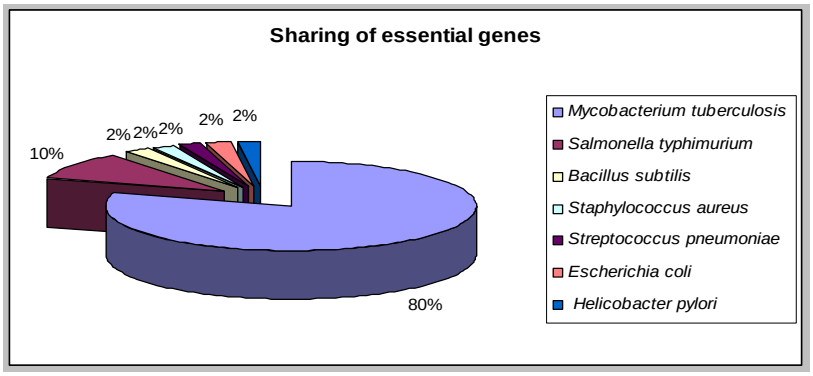

Figure 2: Sharing percentage of the essential genes by different bacterial species

By comparing the gene sequence we found that out of 66 Osmoprotectant genes, 52 shows significant sequence similarities with the genes collect as essential genes in DEG database. We observe that more than 80% of the genes show high levels of similarities with that of *Mycobacterium tuberculosis* which is well known for causing tuberclosis. The next largest group of genes shared sequence homologies with *Salmonella typhimurium* (10%). The rest of the genes shared similarities with genes of *Escherichia coli, Helicobacter pylori, Bacillus subtilis, Staphylococcus aureus* and *Streptococcus pneumoniae*.

A sub-set of 55 genes showing over-expression were found as essential genes for *Pseudomonas aeruginosa*. As described above, 80% of this set too shared similarities with *Mycobacterium tuberculosis*. Based on these results, it is indicated that *Mycobacterium tuberculosis* also follows similar metabolic pathway and genes to survive under conditions of osmotic stress. Further, we suggest that this set of 55 essential genes can also serves as potential targets against antibacterial drugs (10) for both *Pseudomonas aeruginosa* and *Mycobacterium tuberculosis.*

Finding of upstream region for all coexpressed genes shows that some of them are not shared any type of upstream region, means that they all are belongs to same type of operon system. By comparing upstream region we observe that all genes shared some type of conserved region (motif), which suppose to be work as a regulatory element called transcription factor binding sites, and responsible for their coexpression patterns.

Motif finding for up-regulated 55 genes shows that out of 55 genes 17 genes shows absence of considerable upstream region, thus provide indication for the presence of operon regulation process. For rest of the 38 genes we did not found any common motif which provides indication that these all are not belongs to same regulon. The best three motifs which we found are listed below (table 1). Here interestingly we found that these conserved motifs are ranging from 29-46 in width, the best motif which shared by maximum gene (10) is of 39 nucleotide in width.

Study of down regulated genes (11 genes) shows that the upstream regions are available just in three genes, by treating these genes we found 3 best conserved regulatory regions of variable length (table 2), here it is considerable that all three genes share all three types of conserved motifs with significant statistical value, an it is observable that T play a major contribution in these motifs.

The presence of same conserved regulatory motif in genes, indicate that this genes are more tightly coregulated in osmotic condition and information content shows that this motifs are relatively much conserved thus are informative for their respective genes and regulatory machinery.

Table 1: Best three conserved motifs located in upstream regions of over expressed genes.

Motif	width	Best possible Match	Presence in (total-38)	IC
1	39	CCGCTGCTCCTATCGGGATTGGCAGCCGTTTTCTCACTG	10	39.2
2	46	TGAGTTTTTTTGTAAATTAATGAATTGAAAGTGTTTTTATTTTTCA	4	60.9
3	29	TCTCATGATGTTTATCGGACAATCTAGAA	7	33.0

Table: 2 best three conserved motifs located in upstream regions of under expressed genes.

Motif	Width	Best possible Match	Presence in (total 3)	IC
1	8	TTTTTTTT	3	16.5
2	21	CGTTTTCTGATATCTCCCCGC	3	32.9
3	12	TTCCTATACT	3	18.6

Conclusion

Salmonella typhimurium Gene's that shows positive over-expression (55 genes), most of them (51 genes) are work as essential genes for it and also shared by the other bacteria's which indicate that their respective function is very important for organism to survive in osmotic stress conditions. Around 80% of essential over-expressed genes of *S. typhimurium* shows considerable sequence similarities with the gene sequence of *Mycobacterium tuberculosis* and provide indication that due to the presence of these genes it also may be able to survive in such osmotic stress conditions and associated with the virulence factor expression. Because of the essentiality, these all genes can be used as a target for drug molecules. Gene's shows considerable over-expression and under-expression are controlled in coregulated manner and most of them share common regulatory motifs which are responsible for their coexpression, coregulation and similar function. Gene for whom we did not found upstream region, probably may related to some type of operon system

Acknowledgement

I would like to thank Dr. Atul Grover for his valuable support.

References

[1] Kobayashi K., Ehrlich, S.D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P. *et al*. (2003) Essential *Bacillus subtilis* genes. *Proc. Natl Acad. Sci. USA*, **100**, 4678–4683.

[2] Itaya M. (1995). An estimation of minimal genome size required for life. *FEBS Lett.*, **362**, 257–260.

[3] Wood, J.M. (1999). Osmosensn.g by bacteria: signals and membrane based sensors. *Microbiol.Mol.Biol.Rev*.**63**:230-262.

[4] Govan, J.R, and V.Deretic (1996). Microbial pathogens in cystic fibrosis: mucoid Pseudomonas aeruginosa and burkholderia cepacia. *Microbiol. Rev*. **60**:539-574.

[5] Csonka, L. N. (1989).Physiological and genetic response of bacteria to osmatic stress. *Microbiol. Rev* **53**:121-147.

[6] Donaldson, S.H., and R.C. Boucher. (2003). Update on pathogenesis of cystis fibrosis lung disease. *Curr. Opin. Pulm*. Med**.9**:486-491.

[7] Aspedon, A., Palmer. K and Whiteley,.M (2006). Microarray analysis of the Osmotic Stress Response in *Pseudomonas aeruginosa. J. of Bacteriology*. **188**:2721-2725.

[8] Frank, D. W. (1997). The exozyme S regulon of p *aeruginosa. Mol. Microbiol.***26**:621-629.

[9] Hueck, C.J. (1998). Type III Protein secretion systems in bacterial pathogens of animal and plants. *Microbiol.Mol.Biol.Rev*, **62**: 379-433.

[10] Zhang, R., Hong and Zhang, C.T. (2004) DEG: a database of essential genes. *Nucleic Acids Res*.; **32**(Database issue): 271–272