Sequence analysis of a salt tolerant metagenomic clone

Raj Kishor Kapardar · Ravi Ranjan · Munish Puri · Rakesh Sharma

Received: 29 April 2010 / Accepted: 17 May 2010

Abstract Metagenome represent an unlimited resource for discovery of novel genes. Here we report, sequence analysis of a salt tolerant metagenomic clone (6B4) from a pond water metagenomic library. Clone 6B4 had an insert of 2254 bp with G+C composition of 64.06%. DNA sequence from 6B4 showed homology to DNA sequences from proteobacteria indicating origin of 6B4 metagenomic insert from a yet uncharacterized proteobacteria. Two encoded proteins from clone 6B4 showed match with ATP-dependent Clp protease adaptor protein (ClpS) and phasin, while two truncated encoded proteins showed match with poly-3-hydroxybutyrate synthase and permease. Clp complex is known to play a role in stress tolerance. Expression of ClpS from metagenomic clone is proposed to be responsible for salt tolerance of the metagenomic clone 6B4.

Keywords Metagenomic · Metagenome · Salt tolerance · ClpS

Introduction

The phenomenon of bacterial salt stress tolerance is quite complex. It has been observed that organisms which thrive in saline environmental conditions possess specific mechanisms to maintain their osmotic balance. The salt stress tolerance genes have been involved in the recovery of turgour loss due to osmotic shock by uptake and/or de novo synthesis of organic (amino acids, sugars, polyols) and inorganic (K^+) compatible solutes and inducing general or specific stress tolerance genes [1]. Bacterial genes coding for enzymes catalyzing the synthesis of different classes of compatible solutes were successfully used for generating salt tolerant transgenic plants, for example, carbohydrate; trehalose [2], polyols; mannitol [3], amino acid derivatives; glycine betaine [4] and ectoine [5].

The microbial diversity present in the environment has been severely underestimated. This occurred profoundly since the determination of bacterial diversity relied on identifying bacteria by culturing from the environmental samples, without realizing the degree of bias introduced by the chosen culturing conditions [6]. Metagenomics has changed the way microbiologists approached different problems; it has also accelerated the rate of gene discovery. Metagenomics has vast potential which can be used in the field of biotechnology. The functional screens have identified new enzymes and antibiotics and other reagents from metagenomic libraries of diverse environments [7].

To exploit the gene pool of pond water microbial communities, a metagenomic library was prepared earlier and novel lipolytic genes were identified [8]. We envisaged that some of the clones in our metagenomic library would also possess salt tolerance genes, which may find application in

R. K. Kapardar^{1,2} · R. Ranjan¹ · M. Puri^{2,3} · R. Sharma¹ (✉)

¹Institute of Genomics and Integrative Biology, Council of Scientific and Industrial Research (CSIR) Mall Road, Delhi - 110 007, India.

²Fermentation and Protein Biotechnology Laboratory, Department of Biotechnology, Punjab University, Patiala - 147 002, India.

³Present affiliation: Centre for Biotechnology and Interdisciplinary Sciences (Biodeakin), Institute of Technology Research and Innovation (ITRI), Deakin University, Victoria, Australia.

E-mail: rsharma@igib.in

production of salt tolerant recombinant bacterial strains and transgenic plants [2, 3, 4]. To identify salt tolerant clones, we enriched the pond water metagenomic library for clones with an ability to grow in 750 mM NaCl containing medium. The enrichment resulted in isolation of three salt tolerant clones; two of the clones were sequenced and characterized earlier [1]. The third clone showed slower growth even in normal LB medium supplemented with ampicillin and hence was not characterized initially. In an attempt to identify salt tolerance genes from the third clone and to understand why the clone behaved differently from the other two clones in normal LB medium, we sequenced the third clone. Sequence analysis and gene expected to be involved in salt tolerance of this clone is presented in this report.

Clone 6B4 was found to grow at 750 mM NaCl in comparison to control which failed to grow under similar conditions [1]. The clone was further tested for its salt tolerance in LB medium supplemented with 750 mM KCl and ampicillin. The clone showed maximum growth at 28 h (Fig. 1) but the control did not show significant growth under same conditions.

The DNA sequence of both strands of p6B4 was obtained by sequencing transposon-flanking regions of the plasmids mutated by template generation system and primer walking [1]. DNA sequence of 2254 bp having G+C composition

of 64.06% was obtained. The sequence obtained from the clone 6B4 was deposited in GenBank under accession number HM117648. Nucleotide BLAST analysis resulted in match in short stretches 227/264 (85% identity) with *Erythrobacter litoralis* HTCC2594, 239/287 (83% identity) with *Sphingobium japonicum* UT26S, 59/70 (84% identity) with *Bordetella avium* 197N and 30/30 (100% identity) with *Polaromonas* sp. JS666. Previously characterized salt tolerant clones from the library had a G+C composition of 59.58 and 63.97% [1]. Clone 6S2 showed significant identity with *Rhodospirillum rubrum* and *Polaromonas* but clone 5B2 did not show any nucleotide identity with known sequences in the databases [1]. Even though the sequences from three clones had similar G+C composition, nucleotide BLAST analysis indicated at the possibility that metagenomic inserts in three clones originated from unrelated proteobacteria.

Four Open Reading Frames (ORFs) were identified by ORF finder. Two ORFs showed match with ClpS and phasin, while two truncated ORFs showed match with poly-3-hydroxybutyrate synthase and permease (Table 1). Organization of these ORFs is shown in Fig. 2. One of the encoded proteins showed match with 81% identity to ATP-dependent Clp protease adaptor protein, ClpS from *Erythrobacter* sp. NAP1. It had conserved domains of ClpS family (COG2127, pfam02617, PRK00033, PRK13019). Another encoded protein showed match with 55% identity to phasin from *Erythrobacter* sp. SD-21 and it had conserved domains of Phasin2 family (pfam09361, COG5490). Whereas one of the truncated encoded proteins showed match with 63% identity to poly-3-hydroxybutyrate synthase from *Erythrobacter* sp. SD-21 and it had truncated conserved domains for PhaC Poly(3-hydroxyalkanoate) synthetase (COG3243), PHA_synth_I, poly(R)-hydroxyalkanoic acid synthase, class I (TIGR01838), and PHA_synth_II, poly(R)-hydroxyalkanoic acid synthase, class II (TIGR1839). The other truncated encoded protein showed match with 78% identity to predicted permease (YjgP/YjgQ) from *Novosphingobium aromaticivorans* DSM 12444. It had truncated domains for predicted permease YjgP/YjgQ family (pfam03739) and predicted permeases (COG0795) from permease YjgP/YjgQ family.

The two truncated genes are not expected to form functional proteins and only ClpS and Phasin genes are expected to produce functional protein. These two genes may be involved in salt tolerance of the clone 6B4. Phasin proteins are important for polyhydroxy alkanoates storage and metabolism [9]. Polyhydroxy alkanoates are produced by the bacteria under environmental stress conditions [10] but are not reported to play any role in salt stress tolerance.

ClpS is a monomeric adaptor protein that binds to the N-terminal domain of ClpA subunits and alters substrate

Fig. 1 Growth curves of *Escherichia coli* strains, (□) *E. coli*/p6B4, (■) control *E. coli*/pUC19 in LB medium with 750 mM KCl. Experiments were performed in triplicates, values are averages and standard deviations are shown as error bars.

Table 1 ORF Blast results of Open Reading Frames identified from clone 6B4

ORF	BLAST RESULTS					
	Protein, Organism (Acc No.)	Size aa	Identity	e-value	Family	CONSERVED DOMAINS
1	Poly-3-hydroxybutyrate synthase <i>Erythrobacter</i> sp. SD-21 (ZP_01864750)	626	61/96 (63%),	8e-30	TIGR01838 COG3243 TIGR01839	PHA_synth_I (Class I subfamily of poly(R)-hydroxyalkanoate synthases) PhaC Poly(3-hydroxyalkanoate) synthetase [Lipid metabolism] PHA_synth_II (Class II subfamily of poly(R)-hydroxyalkanoate synthases)
2	Hypothetical protein ED21_23143 <i>Erythrobacter</i> sp. SD-21 (ZP_01864749)	267	87/157 (55%)	4e-40	pfam09361 COG5490	Phasin protein Uncharacterized conserved protein (Function unknown)
3	ATP-dependent Clp protease adaptor protein ClpS <i>Erythrobacter</i> sp. NAP1 (ZP_01040407)	140	100/122 (81%),	3e-52	COG2127 pfam02617 PRK00033 PRK13019	Uncharacterized conserved protein (Function unknown) ClpS, In the bacterial cytosol, ATP-dependent protein ClpS, ATP-dependent Clp protease adaptor protein ClpS ClpS, ATP-dependent Clp protease adaptor
4	Permease YjgP/YjgQ <i>Novosphingobium aromaticivorans</i> DSM 12444 (ABD24944)	414	84/107 (78%),	5e-41	pfam03739 COG0795	Predicted permease YjgP/YjgQ family Predicted permeases (General function prediction only)

Fig. 2 Schematic representation of the genetic organization of the salt tolerant clone p6B4 (2254 bp), with respect to pUC19 *lacZ* promoter and *EcoRI* and *HindIII* sites of the vector. A bold arrow with a pointed head represents a complete gene and with a round head represents a truncated gene. Figure is not to scale; size of the encoded proteins is indicated in the brackets.

recognition by ClpAP [11]. The ClpAP protease is composed of ClpA, a hexameric AAA+ ATPase/protein unfoldase, and ClpP, a 14-subunit protease [11]. ClpS directly interacts with the destabilizing N-terminal residues and transfers them to the ClpAP protease complex for degradation [12]. Clp family has been shown to act as molecular chaperones. These proteins can undo protein aggregates resulting from heat shock or other stresses and also play important role in degradation of polypeptide chains in an ATP-dependent manner [13].

ClpS is usually encoded immediately upstream of ClpA in the genomes of proteobacteria. ClpS homologs exist in bacteria. A group of ClpS is found in the genomic context of ClpA, which is universally distributed in proteobacteria and another group that is not linked to a ClpA is found broadly

in actinobacteria, cyanobacteria and plant chloroplasts [14]. ClpS is distantly related to a domain of the eukaryotic E3 ubiquitin ligase, N-recognin. E3 ubiquitin ligase has been shown to be important for abiotic stress responses [15].

Salt stress and other stresses result in aggregation of cytoplasmic proteins. We propose that the multicopy expression of ClpS in metagenomic clone might be involved in increased recognition of aggregated protein for refolding or degradation by ClpAP complex and allow stress tolerance. Unwanted degradation of essential proteins due to expression of ClpS under normal growth conditions may be responsible for the slower growth of clone 6B4 in LB medium. Overexpression of the *clpS* gene alone under an inducible heterologous promoter will help in testing this hypothesis in near future.

References

1. Kapardar RK, Ranjan R, Grover A, Puri M and Sharma R (2010) Identification and characterization of genes conferring salt tolerance to *Escherichia coli* from pond water metagenome. *Biores. Technol* 101(11):3917–3924
2. Pilon-Smits EA, Terry N and Sears T (1998) Trehalose producing transgenic tobacco plants show improved growth performance under drought stress. *J Plant Physiol* 152: 525–532
3. Thomas JC, Sepahi M, Arendall B and Bohnert HJ (1995) Enhancement of seed germination in high salinity by engineering mannitol expression in *Arabidopsis thaliana*. *Plant Cell Environ* 18:801–806
4. Waditee R, Bhuiyan MN, Rai V, Aoki K, Tanaka Y, Hibino T, Suzuki S, Takano J, Jagendorf AT, Takabe T and Takabe T (2005) Genes for direct methylation of glycine provide high levels of glycinebetaine and abiotic stress tolerance in *Synechococcus* and *Arabidopsis*. *Proc Natl Acad Sci USA* 102:1318–1323
5. Nakayama H, Yoshida K, Ono H, Murooka Y and Shinmyo A (2000) Ectoine, the compatible solute of *Halomonas elongata*, confers hyperosmotic tolerance in cultured tobacco cells. *Plant Physiol* 122:1239–1247
6. Sharma R, Ranjan R, Kapardar RK and Grover A (2005) Unculturable bacterial diversity: an untapped resource. *Curr Sci* 89:72–77
7. Singh J, Behal A, Singla N, Joshi A, Birbian N, Singh S, Bali V and Batra N (2009) Metagenomics: Concept, methodology, ecological inference and recent advances. *Biotechnol J* 4: 480–494
8. Ranjan R, Grover A, Kapardar RK and Sharma R (2005) Isolation of novel lipolytic genes from uncultured bacteria of pond water. *Biochem Biophys Res Commun* 335: 57–65.
9. York GM, Stubbe J and Sinskey AJ (2001) New insight into the role of the PhaP Phasin of *Ralstonia eutropha* in promoting synthesis of polyhydroxybutyrate. *J Bacteriol* 183(7): 2394–2397
10. Singh M, Patel SKS and Kalia VC (2009) *Bacillus subtilis* as potential producer for polyhydroxyalkanoates. *Microbial Cell Factories* 8:38
11. Wang KH, Sauer RT and Baker TA (2007) ClpS modulates but is not essential for bacterial N-end rule degradation. *Genes and Dev* 21:403–408
12. Schmidt R, Zahn R, Bukau B and Mogk A (2009) ClpS is the recognition component for *Escherichia coli* substrates of the N-end rule degradation pathway. *Mol Microbiol* 72: 506–517
13. Feng H and Gierasch LM (1998) Molecular chaperones: Clamps for the Clps? *Curr Biol.* 8:R464–R467
14. Lupas AN and Koretke KK (2003) Bioinformatic analysis of ClpS, a protein module involved in prokaryotic and eukaryotic protein degradation. *J Struct Biol* 141:77–83
15. Cho SK, Ryu MY, Song C, Kwak JM and Kim WT (2008) Arabidopsis PUB22 and PUB23 are homologous U-box E3 ubiquitin ligases that play combinatorial roles in response to drought stress. *Plant Cell* 20:1899–1914