Sequence and expression divergence of an ancient duplication of the chaperonin groESEL operon in Vibrio species

Nityananda Chowdhury,† Joseph J. Kingston,† W. Brian Whitaker, Megan R. Carpenter, Analuisa Cohen† and E. Fidelma Boyd

Department of Biological Sciences, University of Delaware, Newark, DE 19716, USA

Heat-shock proteins are molecular chaperones essential for protein folding, degradation and trafficking. The human pathogen Vibrio vulnificus encodes a copy of the groESEL operon in both chromosomes and these genes share <80 % similarity with each other. Comparative genomic analysis was used to determine whether this duplication is prevalent among Vibronaceae specifically or Gammaproteobacteria in general. Among the Vibronaceae complete genome sequences in the database (31 species), seven Vibrio species contained a copy of groESEL in each chromosome, including the human pathogens Vibrio cholerae, Vibrio parahaemolyticus and V. vulnificus. Phylogenetic analysis of GroEL among the Gammaproteobacteria indicated that GroESEL-1 encoded in chromosome I was the ancestral copy and GroESEL-2 in chromosome II arose by an ancient gene duplication event. Interestingly, outside of the Vibronaceae within the Gammaproteobacteria, groESEL chromosomal duplications were rare among the 296 genomes examined; only five additional species contained two or more copies. Examination of the expression pattern of groEL from V. vulnificus cells grown under different conditions revealed differential expression between the copies. The data demonstrate that groEL-1 was more highly expressed during growth in exponential phase than groEL-2 and a similar pattern was also found in both V. cholerae and V. parahaemolyticus. Overall these data suggest that retention of both copies of groESEL in Vibrio species may confer an evolutionary advantage.

INTRODUCTION

Chaperones are proteins that are essential for the proper folding and function of other proteins (Ellis, 1993). The chaperonin heat-shock protein, also known as GroEL in Escherichia coli or Hsp60/Cpn60 in mitochondria, is a group I chaperonin that is widespread within Prokarya and Eukarya. GroEL forms a homopolymeric structure that is required for macromolecule assembly, proper protein folding as well as protein degradation and trafficking (Azem et al., 1994a, b; Braig et al., 1994; Ellis, 1993). Chaperones also play a role in buffering the deleterious effects of misfolding-related mutations (Rutherford, 2003). The function of GroEL is greatly enhanced by the presence of GroES (Hsp10) and ATP (Goloubinoff et al., 1997). The groEL and groES genes, encoding Hsp60 and Hsp10, respectively, in many bacterial species are present as an operon in the genome. In Escherichia coli, GroES–GroEL are the only chaperones essential for the bacterium to grow at all temperatures (Fayet et al., 1989). In addition to their ability to bind a wide range of misfolded protein substrates, GroEL is also highly antigenic (Brunham & Peeling, 1994; Kerner et al., 2005; Zügel & Kaufmann, 1999). Recent phylogenetic data suggest that the presence of multiple chaperonins within a species may play an important role in evolutionary and ecological change in certain bacterial groups (Garduño et al., 2011; Williams et al., 2010). Studies have demonstrated alternative functional roles for GroEL, such as invasion of host tissue by acting as an adhesion factor, modulation of the host immune system or biofilm formation (Clark & Tillier, 2010; Garduño et al., 2011; Hu et al., 2008; Lichtenwalner et al., 2004; Lund, 2009; Wuppermann et al., 2008; Zügel & Kaufmann, 1999).

The importance of GroEL is evidenced by its near universal presence among bacteria (Lund, 2009). A notable exception is within the Mollicutes, with many species of the genus Mycoplasma containing no groEL homologue (Clark & Tillier, 2010). In contrast to this finding, among other...
Gram-positive bacteria, the presence of multiple chaperonins has been documented to be encoded either by groEL alone or by both groES-EL (Lund, 2009). Actinobacteria, which include a number of important pathogens and commensals, were the first bacteria shown to encode two copies of groEL. Among the Actinobacteria that contain two copies of groEL, all contain only a single copy of groES (Lund, 2009; Rinke de Wit et al., 1992). The divergent copies in Mycobacterium were suggested to result from differential interaction with the host immune system (Hughes, 1993). However, more recent data suggest that in the mycobacteria, chaperones are also secreted as intercellular signalling molecules and can control the composition and functions at the cell wall (Henderson et al., 2010; Hu et al., 2008). Multiple highly divergent copies of groEL are also present in Chlamydiae species, including the human obligate intracellular pathogens Chlamydia pneumoniae and Chlamydia trachomatis, at least one copy of which is thought to be involved in pathogenesis (Brunham & Peeling, 1994; Karunakaran et al., 2003; Lichtenwalner et al., 2004; McNally & Fares, 2007; Wuppermann et al., 2008). In C. pneumoniae it was shown that the GroEL-1 protein is cell surface-associated and required for infection of HEp-2 cells (Wuppermann et al., 2008).

Within the Proteobacteria, the presence of multiple copies of groEL has been studied in the Alphaproteobacteria, particularly Bradyrhizobium, Mesorhizobium and Rhizobium species, where three to seven copies of groEL have been found (George et al., 2004; Gould et al., 2007a, b; Ivic et al., 1997; Rodríguez-Quiñones et al., 2005). Compared with Gram-positive bacteria, in Gram-negative bacteria the groEL copies are usually bicistronic transcriptional units containing both ORFs groES and groEL (Goyal et al., 2006).

Vibrio is a member of the family Vibrionaceae, subdivision Gammaproteobacteria, and is closely related to the Enterobacteriales. Vibrio vulnificus is abundant in warm coastal waters and occupies the full range of environmental niches, including associations with zooplankton, fish, and filter feeders such as oysters and molluscs (Blackwell & Oliver, 2008; Oliver et al., 1982; Wright et al., 1996). It is also a fish pathogen, a significant pathogen of eels and infection is associated with biotype 2 strains (Amaro & Biosca, 1996; Fouz et al., 2010). In addition, the bacterium is an opportunistic pathogen of humans, causing primary septicaemia and wound infections with mortality rates greater than 50% in susceptible individuals (Gulig et al., 2005; Jones & Oliver, 2009). Population genetic studies and genotyping data divide V. vulnificus isolates into at least two distinct but closely related lineages; lineage I consists of mainly clinical isolates and lineage II contains environmental isolates (Bisharat et al., 2005; Cohen et al., 2007; Warner & Oliver, 2008).

The first complete genome sequence of V. vulnificus was published by Chen et al. (2003) and, like other Vibrio species, V. vulnificus was shown to have two chromosomes of unequal size. Chromosome I, the larger of the two, is also the most conserved in terms of gene content and gene order in comparison with other sequenced Vibrio species. Chromosome II, the smaller of the two, usually contains more hypothetical proteins and proteins of unknown function compared with chromosome I and the gene content and order is much less conserved between species (Chen et al., 2003). The V. vulnificus genome sequence YJ016 contained 1143 more genes than V. cholerae N16961 and a total of 260 duplication events, which added 495 genes to the V. vulnificus gene count (Chen et al., 2003). The genome of V. vulnificus shows high plasticity with large genomic islands variably absent and present among strains (Quirke et al., 2006). To date, three complete genome sequences are available in the database, all lineage I clinical isolates, strains YJ016, CMCP6 and MO6-24/O. In all three complete V. vulnificus genomes the groES-EL operon is duplicated with a copy on each chromosome, and in 32 V. vulnificus genomes in progress, including lineage II strains, the two copies of groES-EL are present.

Multiple copies of groES and groEL can arise by gene duplication or by horizontal gene transfer events. In bacteria, depending on the species, both processes were shown to be involved in groES-EL duplications (Lund, 2009). The origin of each copy of groES-EL in V. vulnificus is not known nor is the prevalence of multiple copies among Vibrio species. In this study, we use bioinformatics and phylogenetic analyses to address these questions. In addition, we determined whether both copies of groES-EL differ in their gene expression under different growth conditions in V. vulnificus, Vibrio cholerae and Vibrio parahaemolyticus.

**METHODS**

**Sequence retrieval.** Searches using the BLAST program were used to identify homologues of groEL and groES among Vibrionaceae and Gammaproteobacteria species in the sequenced genome database. The groEL nucleotide sequences of Vibrio species and GroEL protein sequences of Gammaproteobacteria were retrieved from the web-based chaperonin sequence database (Hill et al., 2004) and the NCBI protein and gene databases (Table S1, available in the online Supplementary Material). Each genome sequence was interrogated for the presence of multiple copies of groEL and groES and their genome location and arrangement were examined. In addition, BLAST searches of the V. vulnificus GroEL protein were performed against all Proteobacteria genomes listed in the NCBI microbial genome database to determine the number of species within each subdivision that contained multiple copies (Altschul et al., 1997). The DNA sequences of the 16S rRNA gene from representative Gammaproteobacteria species were downloaded from GenBank (http://www.ncbi.nlm.nih.gov/gene) and the Ribosomal Database project (http://rdp.cme.msu.edu/index.jsp) (Table S2) (Maidak et al., 2001).

**Phylogenetic analysis.** Among Vibrio species, the copies of groES-EL present in chromosome I were termed groES-EL-I while those present in chromosome II were termed groES-EL-2. The nucleotide and protein sequences were aligned using the multiple sequence alignment program CLUSTAL W2 and the alignments were examined further by alignment explorer (MEGA5) before further phylogenetic analysis (Tamura et al., 2011; Thompson et al., 1994). A single
representative bacterial strain was included for bacterial species with multiple strains. A total of 168 GroEL protein sequences were aligned (Table S1). Two different tree-building methods were used, maximum-parsimony (MP) and neighbour-joining (NJ), as implemented in MEGAX for the analysis of groEL-1 and groEL-2 sequences among Vibrio species (Tamura et al., 2011). Bootstrap values for the MP and NJ trees were obtained after 1000 generations. The MEGAX tree viewer was used to visualize the trees and calculate confidence values (Tamura et al., 2011). The topologies of the MP and NJ trees were very similar, with differences in branch lengths and confidence values but not in branching pattern, and the NJ trees are displayed (Saitou & Nei, 1987). Rates of synonymous substitutions ($d_S$) and non-synonymous ($d_N$) substitutions per site were calculated for groEL among Vibrio species, and substitution ratio ($d_S/d_N$) estimated to understand the type of selection acting on the groEL genes (purifying, $d_S/d_N$<1; neutral, $d_S/d_N$=1; diversifying, $d_S/d_N$>1) (Nei & Gojobori, 1986; Nei & Jin, 1989).

**Bacterial strains, plasmids and growth conditions.** All bacterial strains used in this study are listed and described in Table 1. *V. vulnificus* strain YJ016 was used for expression analysis and was grown in Luria–Bertani (LB) broth containing 2% NaCl (LB 2% NaCl) (Fisher Scientific), pH 7.0, at 37°C with shaking (225 r.p.m.). *V. parahaemolyticus* RIMD 2210633 was grown under the following conditions: 4 h in LB 3% NaCl pH 7 (exponential phase), and 10 h in LB 3% NaCl pH 7 (stationary phase). *V. cholerae* N16961 was grown under the following conditions: 4 h in 1% NaCl pH 7 (exponential phase), and 10 h in LB 1% NaCl pH 7 (stationary phase). A solid medium was prepared by adding 1.5% agar (Fisher Scientific) to the broth when required. A 1% inoculum of *V. vulnificus* YJ016. Primers (Table S3) were designed to delete 1479 bp from *groEL-1* (1641 bp) or 1488 bp of *groEL-2* (1599 bp) to generate a truncated non-functional version of *groEL-1* (162 bp) or *groEL-2* (111 bp). Briefly, the AB and CD PCR products were fused together by a third PCR to generate truncated AD product deleting the internal part of the gene, which was then cloned into a suicide vector pDS132 (Philippe et al., 2004) and the resulting plasmid was named as pAVV3106NC for *groEL-1* and pAVVA1659NC for *groEL-2*. These plasmids were then transformed into an *E. coli* DAP auxotroph straingunakan conjugation with *V. vulnificus*. However, several attempts were made to isolate exconjugants from *V. vulnificus* strain YJ016 without success, indicating that these may be essential in *V. vulnificus*.

**RESULTS AND DISCUSSION**

**Sequence analysis and distribution of groES-EL among Vibrionaceae**

The *groEL* gene present in chromosome I (*groEL-1*) of *V. vulnificus* YJ016 encoded a protein of 546 aa with a predicted molecular mass of 57.4 kDa, while *groEL-2* present in chromosome II encoded a 532 aa protein with a predicted molecular mass of 55.8 kDa. The *V. vulnificus* YJ016 *groEL-1* protein shared only 79% amino acid identity with *V. vulnificus* GroEL-2 but shared 95% amino acid identity

**Table 1. Bacterial strains and plasmids used in this study**

<table>
<thead>
<tr>
<th>Strain/plasmid</th>
<th>Characteristic</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strain</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>V. vulnificus</em></td>
<td>YJ016, clinical isolate</td>
<td>Chen et al. (2003)</td>
</tr>
<tr>
<td><em>V. cholerae</em></td>
<td>N16961 O1 El Tor, clinical isolate</td>
<td>Heidelberg et al. (2000)</td>
</tr>
<tr>
<td><em>V. parahaemolyticus</em></td>
<td>RIMD2210633 O3:K6, clinical isolate</td>
<td>Makino et al. (2003)</td>
</tr>
<tr>
<td><strong>Escherichia coli</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH5α λ−pir</td>
<td>pir80dilacZM15 (lacZYA-argF)U169, recA1,hsdR17, deoRthi−1supE44 gyrA96 relA1</td>
<td></td>
</tr>
<tr>
<td>β2155 λpir</td>
<td>DAP auxotroph; thrh004 pro thi strA hsdS lacZ_M15 (F lacZ_M15 lacTRQI_36 proA_proB) dapA Ermr pirRP4 (KmR from SM10)</td>
<td></td>
</tr>
<tr>
<td><strong>Plasmid</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pDS132</td>
<td>Suicide plasmid, CM,R, SacB</td>
<td>This study</td>
</tr>
<tr>
<td>pAVV3106NC</td>
<td>ΔV3106 (AgroEL1) cloned into pDS132, CmR</td>
<td>This study</td>
</tr>
<tr>
<td>pAVVA1659NC</td>
<td>ΔVVA1659 (AgroEL2) cloned into pDS132, CmR</td>
<td>This study</td>
</tr>
</tbody>
</table>
with the GroEL homologues present in chromosome I of other Vibrio species analysed in this study. The groES genes in chromosomes I and II of V. vulnificus YJ016 encoded a 97 aa protein with a predicted mass of 10 kDa. The groEL and groES genes in chromosome I were separated by an intergenic spacer region of 34 bp whereas these genes in chromosome II had a 73 bp intergenic region. The nucleotide sequence of groEL-1 and groEL-2 showed an overall G+C content of 51.4 and 49.5 mol%, respectively. GroEL-1 and GroES-1 shared 86% identity (93% similarity) and 80% identity (91% similarity) with GroEL and GroES from E. coli. GroEL-2 and GroES-2 shared 78% identity (90% similarity) and 67% identity (84% similarity) with GroEL and GroES from E. coli. The level of identity among GroEL from other species was generally greater than 60% with the exception of those found among Chlamydiae, which were significantly lower (George et al., 2004; Karunakaran et al., 2003; Rao & Lund, 2010; Rodriguez-Quiones et al., 2005). Thus, the two copies of groES-EL in V. vulnificus share high identity with E. coli groES-EL, well within the range of values that suggests structural and functional conservation between the species.

Our BLAST searches using the V. vulnificus groES-EL-1 (chromosome I) genes as a probe against the Vibrionaceae genomes revealed that seven of 31 Vibrionaceae species (genera Vibrio, Aliivibrio, Grimontia and Photobacterium) examined from the genome database have a copy of groES-EL on each chromosome. The seven Vibrio species were Vibrio sp. EX25, Vibrio alginolyticus, V. cholerae, Vibrio furnissii, Vibrio mimicus, V. parahaemolyticus and V. vulnificus. Several representative strains among these seven species have been sequenced; for example, over 30 genome sequences of V. cholerae and V. vulnificus are available and all contain two copies of groES-EL. Similarly, multiple strains of V. mimicus, V. vulnificus, V. parahaemolyticus and V. furnissii all contained one copy of groES-EL in each chromosome. None of the copies was associated with transposable or integrase genes, which would have been indicative of horizontal transfer (Fig. 1).

We compared the genes flanking the groES-EL loci in chromosome I and II among seven Vibrio species to determine whether the gene arrangement differed among them (Fig. 1). The groES gene was preceded by the groEL gene in both chromosomes of the Vibrio species analysed. V. parahaemolyticus and V. alginolyticus groES-EL-1 had the most similar flanking gene homology and arrangement to that of V. vulnificus (Fig. 1). The orthologous genes flanking the groES-EL-1 loci in chromosome I of V. cholerae, V. mimicus and V. furnissii were essentially contiguous except for minor differences noted for V. furnissii. The location of groES-EL in the Vibrio Ex25 genome was most similar to that of V. vulnificus (Fig. 1). The genes flanking chromosome II groES-EL-2 loci in V. vulnificus did not show homology to the genes flanking the corresponding loci in any other Vibrio species examined (Fig. 1). The V. parahaemolyticus and V. alginolyticus groES-EL-2 copies shared homologous flanking genes with each other and with Vibrio sp. EX25. The groES-EL-2 loci of V. cholerae and V. mimicus shared common flanking genes with each other and to a lesser extent with V. furnissii. Thus, it appears that conservation of flanking genes is more extensive for groES-EL-1 genes in chromosome I than those located in chromosome II, which suggests that the copy on chromosome I may be the ancestral copy (Fig. 1).

**Nucleotide sequence polymorphism and phylogenetic analysis**

To further examine whether the groES-EL gene arrangement and genome context in chromosome I and II suggests horizontal gene transfer or a gene duplication event, 26 Vibrio sp. groEL-1 and 10 Vibrio sp. groEL-2 sequences were aligned using CLUSTAL W, and we calculated the nucleotide and amino acid variation among these sequences. The groEL-1 gene had 630 variable polymorphic sites [555 parsimonious-informative (Pi) sites] that resulted in 101 variable amino acid polymorphic sites (85 Pi sites) among 26 Vibrio sequences examined (Table 2). The groEL-2 gene had 596 polymorphic nucleotide sites (559 Pi sites) among 10 Vibrio sp. sequences examined that resulted in 121 polymorphic amino acid sites (115 Pi sites). Next, we estimated the number of synonymous (silent) substitutions per synonymous site (dS) and the number of non-synonymous (amino acid replacement) substitutions per non-synonymous site (dN) for groEL-1 and groEL-2 (Table 2). The data suggest that there is a selective constraint against amino acid replacements at the groEL-1 locus, given that the non-synonymous rate dN (0.04) is over 10-fold lower than the synonymous rate dS (0.5) (Table 2). For the groEL-2 locus, although there appears to be a selective constraint against amino acid replacements with a non-synonymous rate dN (0.08) fivefold lower than the synonymous rate dS (0.46), it was less rigorous than for groEL-1 (Table 2). Overall, the nucleotide diversity among 26 Vibrio groEL-1 sequences was 0.155 compared with 0.191 for 10 groEL-2 sequences examined, which suggests that groEL-2 is evolving faster (Table 2).

Next, we performed phylogenetic analysis of both copies of groEL from the Vibrio species. The phylogenetic relationships of the groEL-1 and groEL-2 sequences of all Vibrio species in the genome database were reconstructed using the NJ method based on evolutionary distances computed using the Tamura three-parameter method, which were in units of the number of base substitutions per site (Fig. 2). The bootstrap consensus tree inferred from 1000 replicates is shown to represent the evolutionary history of the taxa analysed (Felsenstein, 1985). On to the groEL-1 tree, we mapped the species that contained duplicates of groES-EL (filled squares) and their position suggests a phylogenetically clustered occurrence among closely related species and not a random distribution (Fig. 2a). An exception was the absence of a second copy in Vibrio sp. RC341, which is highly related to V. cholerae. In this species the ORFs (vcj003106 and vcj003107) homologous to vca0818 and
vca0822 that flank *V. cholerae* groES-EL-2 loci in chromosome II are adjacent to each other. The phylogeny of *groEL-2* from the seven species with two copies was broadly similar to that of *groEL-1* among the same species (Fig. 2). However, the mean distance within the taxa in the *groEL-2* lineage was 0.228, higher than the *groEL-1* lineage of 0.179. The mean distance between the lineages *groEL-1* and *groEL-2* is 0.368, which indicates that the mean distance of the taxa within the lineages is less than the distance between the lineages. Overall, the data suggest that *groEL-1* is probably the ancestral copy and that the two copies of *groES-EL* resulted from an ancient duplication event that occurred early in the divergence of the genus. The duplication may have occurred only in a few lineages or it may have occurred in all lineages with subsequent deletion events among different *Vibrio* species to give rise

---

**Table 2.** Nucleotide polymorphism in *groEL-1* and *groEL-2* sequences from *Vibrio* species

<table>
<thead>
<tr>
<th>Gene</th>
<th>No. of polymorphic:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nucleotides</td>
<td>V</td>
<td>Pi</td>
<td>S</td>
<td>V</td>
<td>Pi</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>groEL-1</em></td>
<td>630</td>
<td>555</td>
<td>75</td>
<td>101</td>
<td>85</td>
<td>16</td>
<td>0.50 + 0.01</td>
<td>0.04 + 0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>groEL-2</em></td>
<td>596</td>
<td>559</td>
<td>37</td>
<td>121</td>
<td>115</td>
<td>7</td>
<td>0.46 + 0.01</td>
<td>0.08 + 0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

V, variable; Pi, parsimonious-informative; S, singletons; *d_s*, synonymous substitutions per site; *d_N*, non-synonymous substitutions per site.
horizontal transfer events. The branch lengths of GroEL-2 in the evolution of the genus or less likely multiple ancient species was the ancestral copy and the GroEL-2 copy *Vibrio* by high bootstrap values suggested that the GroEL-1 of family main lineages, branching off from representatives of the Pasteurellaceae (Fig. 3). However, GroEL-2 clustered distantly from these lineages and GroEL-1 (Fig. 3). The GroEL-1 lineages showed an identical branching pattern to the same set of strains in the phylogeny of GroEL among representative Gammaproteobacteria species and compared this with a species tree for the same species as indicated that this protein has evolved at a much faster rate than GroEL-1 but among the *Vibrio* strains still showed phylogenetic relationships similar to GroEL-1 (Fig. 3). It appears the duplicated and the ancestral copies of groEL have evolved independently of each other. The most parsimonious explanation is that the entire operon was duplicated early in the evolution of *Vibrio* species and lost from several lineages, and in those species that retained the second copy, this copy is evolving faster.

Among the 296 Gammaproteobacteria genomes examined, an additional five species were found to contain multiple copies of groEL (Fig. S2). We identified only three additional species that encoded two chromosomal copies of groEL, *Allochromatium vinosum*, *Azotobacter vinelandi* and *Nitrosococcus halophilus*, and two species that encoded three copies of groEL, *Methyllococcus capsulatus* and *Psychromonas ingrahamii*. Two copies of groEL were also present in specific strains of *E. coli* and *Salmonella enterica*, although in these cases the second copy was present in a plasmid and was highly divergent from the chromosomal copy (Fig. 3). In *E. coli*, plasmids pAPEC-O1-R, pEH4H and pAR060302 and in *Salmonella enterica* serovar Newport SL254 plasmid pns254 encoded a groEL copy that clustered with each other but
were unrelated to the chromosomal copy of groEL present among other Enterobacteriales (Fig. 3). The plasmid copies branched with GroEL proteins from members of the Xanthomonadaceae, with Xylella fastidiosa, Xanthomonas campestris and Stenotrophomonas maltophilia indicating horizontal gene transfer mediated by plasmid transfer within E. coli and Salmonella enterica.

In Nitrosococcus halophilus, only the groEL gene was duplicated and the second copy was distantly related to the groES-EL copy and those of other Chromatiaceae (Fig. 3). Similarly, the two copies of GroEL from Azotobacter vinelandi clustered separately and divergently in the GroEL tree, suggesting acquisition by horizontal gene transfer (Fig. 3). In contrast, the two copies present in Allochromatium vinosum included both groEL and groES and both copies clustered together on the GroEL tree, suggesting a recent gene duplication event (Fig. 3). Within Methylcoccus capsulatus, three copies of groEL were present; two copies encompassed both groEL and groES and clustered together on the GroEL tree, suggesting a recent gene duplication event (Fig. 3). The third copy of GroEL was highly divergent from the others, indicating acquisition by horizontal gene transfer (Fig. 3). A similar pattern was seen for the three copies in Psychromonas ingrahamii 37: both a gene duplication and a horizontal
gene transfer event occurred (Fig. 3). These events appeared to be strain specific and not universally present within the species. Gene duplication and horizontal gene transfer of groEL homologues have also been demonstrated in other species, for example in Mycobacterium smegmatis (Rao & Lund, 2010).

**Expression analysis of groEL-1 and groEL-2 genes**

To determine whether the groES and groEL genes are bicistronic in *V. vulnificus*, similar to what is seen in *E. coli*, we extracted RNA and looked for the presence of each copy as a single mRNA transcript. RT-PCR using a forward primer designed within groES and a reverse primer designed within groEL was carried out on two cDNA samples isolated from exponential phase cells. PCR amplicons of the predicted sizes of 1492 and 1453 bp were obtained for groES-EL-1 and groES-EL-2 from chromosome I and chromosome II, respectively, indicating that the genes were co-transcribed (Fig. 4).

Given the significant sequence divergence between groEL-1 and groEL-2, next we determined whether there were differences in the expression pattern between these two genes in exponential phase and stationary phase cells. To examine this, we used qPCR to check whether groEL-1 and groEL-2 present in *V. vulnificus*, *V. cholerae* and *V. parahaemolyticus* were expressed. From *V. vulnificus* cells grown to exponential phase the expression of groEL-1 was approximately 27-fold ($P<0.0001$) higher than that of groEL-2 (Fig. 5). Similarly, for *V. cholerae* and *V. parahaemolyticus* we found that groEL-1 was ninefold ($P<0.005$) and 125-fold ($P<0.0001$) higher than that of groEL-2, respectively (Fig. 5). Our results were congruent with previous reports on *Chlamydiae* and *Rhizobium leguminosarum* where groEL-1 was found to be more highly expressed than groEL-2 or groEL-3 (Karunakaran et al., 2003; Rodríguez-Quinones et al., 2005).

Moreover, groEL-1 expression was significantly higher in exponential phase cells than in stationary phase cells both in *V. vulnificus* and in *V. cholerae* by about 12- and 8.5-fold ($P<0.005$), respectively, but it did not change significantly in *V. parahaemolyticus* (Fig. 5). In stationary phase cells, there were no significant changes in expression between groEL-1 and groEL-2 in all three *Vibrio* species examined ($P>0.05$).

**Deletion mutation in groEL-1 or groEL-2 of *V. vulnificus***

Repeated attempts were taken to construct in-frame deletion mutations in either groEL-1 alone leaving the groEL-2 gene untouched or vice versa in *V. vulnificus* VJ016. However, we were unable to mutate either of them. Note that we have used a similar splicing by overlap extension PCR and homologous recombination approach to successfully construct several mutant strains of other genes in *V. vulnificus* and in both *V. cholerae* and *V. parahaemolyticus* (Chowdhury et al., 2012; Kalburge et al., 2014; Lubin et al., 2012; Whitaker et al., 2014). Therefore, we believe that failure to construct a groEL-1 or groEL-2 mutant in *V. vulnificus* implies that both of the copies are important for multiple cellular functions in this species. This is the first analysis of multiple copies of groEL-ES in a species with two chromosomes. Thus, it is highly likely that the presence of a copy of groELES on each chromosome is important for cellular survival. In addition, the fact that all strains within a species retain a copy of groELES on each chromosome strongly suggests that the loss of one copy may have a significant fitness effect.

While most bacteria possess one copy of the groEL gene, 30% of sequenced bacterial genomes contain multiple groEL loci (Lund, 2009). In these species, there have been various reports regarding the essential requirement of the groEL genes but none examined a species with multiple chromosomes. For example, *Myxococcus xanthus* DK1622 contains two copies of groEL that have differing expression patterns and it was found that one copy is essential for growth and survival (Wang et al., 2013, 2014). Unlike *Vibrio* species where all strains contain two copies of groEL-ES, not all strains of *Myxococcus xanthus* contain both copies, indicating that two copies are not essential for

---

**Fig. 4.** Single transcript for groES–groEL genes in *V. vulnificus*. RT-PCR was performed with forward and reverse primer pair AB, VVgroES1-TF and VVgroEL1-TR, and primer pair DE, VVgroES2-TF and VVgroEL2-TR, to determine the presence of a single transcript that encompasses both groES and groEL. Two independent cDNA samples were examined along with a positive genomic DNA control and a negative no DNA template control.
suggests that the occurrence of duplicates of species and that the phylogenies of both loss from multiple lineages. This is substantiated by the result of an ancient gene duplication event with subsequent required for growth (Rodrı´guez-Quin˜ones et al., 2005). Additionally, Sinorhizobium melloti has five copies of groEL and it was found that the presence of either groEL-1′ or groEL-2 was necessary for growth (Bittner et al., 2007).

CONCLUSION

In this study, it was demonstrated that the presence of two copies of groES-EL among Vibrionaceae is probably the result of an ancient gene duplication event with subsequent loss from multiple lineages. This is substantiated by the fact that the duplication occurs in all representatives of a species and that the phylogenies of both groEL copies are congruent. In contrast, among Gammaproteobacteria duplications of groES-EL are rare and occur only in single strains of a species, where either gene duplication or horizontal transfer events have occurred. Among the nearly 300 Gammaproteobacteria species sequenced to date, only 12 species contain multiple copies of groEL, seven of which are in the genus Vibrio. The over-representation in the genus Vibrio suggests that the occurrence of duplicates of groES-EL is significant in the evolution of this group. Expression analysis results indicate that there are differences in expression patterns between the genes. The inability to construct a deletion mutation in either copy suggests both these genes are essential, which may be a reflection of the two-replicon genome structure of Vibrio.

ACKNOWLEDGEMENTS

This research was supported by a National Science Foundation CAREER award DEB-0844409 to E. F. B., and J. J. K. was supported by a BOYSCAST Indian government fellowship. M. R. C. was supported in part by a University of Delaware Graduate Fellowship award. W. B. W. was supported in part by a University of Delaware Dissertation Fellowship award. We thank Sai Siddharth Kalburge, J. B. Lubin, Nathan McDonald, Serge Ongagna-Yhombi and Abish Regmi for their review and comments on the manuscript.

REFERENCES

Vibrio cholerae constitute the sole route of sialic acid uptake in the human pathogen sialic acid-specific subfamily of TRAP transporters (SiaPQM) and structure constitute genome analysis of Vibrio vulnificus, a marine pathogen. Genome Res 13, 2577–2587.


Edited by: R. Lan