



Detecting the form of selection in the outer membrane protein C of *Enterobacter aerogenes* strains and *Salmonella* species

Abinash Padhi^{a,*}, Bindhu Verghese^a, Subhendu K. Otta^b

^aDepartment of Biological Science, University of Tulsa, 600 S. College Ave., Tulsa, OK-74104, USA

^bBaylor College of Medicine, Department of Pediatrics-Hematology/Oncology, 6621 Fannin Street, MC 3-3320, Houston, TX-77030, USA

Received 5 December 2006; accepted 13 December 2006

KEYWORDS

Positive selection;
Purifying selection;
Enteric bacteria;
Enterobacter aerogenes;
Salmonella

Summary

The types of selective pressure operating on the outer membrane protein C (*ompC*) of *Enterobacter aerogenes* strains, the causative agent for nosocomial infections, and *Salmonella* sp., the hazardous pathogen are investigated using the maximum likelihood-based codon substitution models. Although the rate of amino acid replacement to the silent substitution (ω) across the entire codon sites of *ompC* of *E. aerogenes* ($\omega = 0.3194$) and *Salmonella* sp. ($\omega = 0.2047$) indicate that the gene is subjected to purifying selection (i.e. $\omega < 1$), approximately 3.7% of *ompC* codon sites in *E. aerogenes* ($\omega = 21.52$) are under the influence of positive Darwinian selection (i.e. $\omega > 1$). Such contrast in the intensity of selective pressures in both pathogens could be associated with the differential response to the adverse environmental changes. In *E. aerogenes*, majority of the positively selected sites are located in the hypervariable cell-surface-exposed domains whereas the trans-membrane domains are functionally highly constrained.

© 2007 Elsevier GmbH. All rights reserved.

Introduction

Enterobacter aerogenes and *Salmonellae* are gram-negative pathogenic bacteria belonging to

the family Enterobacteriaceae. Since 1990s, *E. aerogenes* has emerged as the third leading cause for respiratory tract nosocomial infections (Schaberg et al., 1991; Jarvis and Martone, 1992). These pathogens exhibit high resistance to commonly used antibiotics and are usually predominant in hospitals causing complicated secondary infections (Bornet et al., 2000; Bosi et al., 1999). These

*Corresponding author. Tel.: +1 918 631 3062;
fax: +1 918 631 2762.

E-mail address: abinash-padhi@utulsa.edu (A. Padhi).

resistant forms have been reported from many countries (Allerberguer et al., 1996; Arpin et al., 1996; Georgiou et al., 1995; Jalaluddin et al., 1998). The increased antibiotic resistance is due to functional alternation of the outer-membrane permeability associated with decrease in porin (outer-membrane proteins) function (Bornet et al., 2004; Thiolas et al., 2004). Another bacterium of interest in the present study is *Salmonellae*, which is also a hazardous pathogen infecting both humans and animals causing diseases ranging from gastroenteritis to thyroid fever depending on their serotype.

The outer membrane of these pathogens contain pore forming proteins (porins) that allow the passive diffusion of small solute molecules (Lutenberg and van Alphen, 1983; Nikaido and Vaara, 1985). Among these porins, outer membrane protein C (*ompC*) is the major surface antigen which is expressed throughout the infection period (Muthukkaruppan et al., 1992; Sujatha et al., 2001). *OmpC* is an important virulent factor constantly exposed to the host immune system and harsh environmental conditions. Therefore, it is more susceptible to differential selection pressure during the course of evolution. Thiolas et al (2004) reported high sequence divergence among *ompC* from different strains of *E. aerogenes*. These lines of evidences indicate that *ompC* of this bacterium might have been subjected to differential selective pressure. *Salmonellae ompC* also display heterogeneous epitopes on the cell surface (Arockiasamy and Krishnaswamy, 2000), therefore, it is possible that this gene in *Salmonellae* might also be under differential selective pressure.

Here we report the types of selective pressure operating on the *ompC* genes of *E. aerogenes* strains and *Salmonella* sp. Two types of selective pressure shape the evolution (Kimura, 1983). While purifying selection favors the conservation of existing phenotypes, positive Darwinian selection leads to functional divergence of protein coding genes. One of the most widely used methods to detect positive selection from the DNA sequence is by comparing the rate of nonsynonymous nucleotide substitutions per nonsynonymous site (d_N) with that of synonymous substitutions per synonymous site (d_S) (Hughes and Nei, 1989). When d_N/d_S (here after referred as ω) > 1 , positive selection is said to be operating, whereas $\omega < 1$ indicates the gene is under the influence of purifying selection.

Maximum likelihood-based codon substitution models which account for variable ω ratios among sites and detect codon sites that are subjected to positive selection (Yang et al., 2000) have been widely used in detecting positive selection in a

number of membrane-associated proteins (e.g. Smith et al., 1995; Fares et al., 2001; Jiggins et al., 2002; Urwin et al., 2002; Andrews and Gojobori, 2004; Fitzpatrick and McInerney, 2005; Chen et al., 2006). Here we performed codon substitution analyses to determine types of selection pressure operating on the *ompC* of *E. aerogenes* and *Salmonella*.

Materials and methods

Phylogenetic analyses

For the analysis, six published *ompC* type (*Omp36*) coding region sequences representing different strains of *E. aerogenes* (AF335467, AF373860, AF336095, AF336096, AF336097, AF336098) and eight *ompC* sequences representing three *Salmonella* species were retrieved from GenBank [AE008801 (*S. typhimurium*), AE014613 (*S. enterica* subsp. *enterica* serovar Typhi Ty2), AF039309 (*S. typhimurium*), Y15844 (*S. minnesota*), AY081183 (*S. enterica* subsp. *enterica* serovar Gallinarum), AY081184 (*S. enterica* subsp. *enterica* serovar Dublin), AY081185 (*S. enterica* subsp. *enterica* serovar Pullorum) and AY341077 (*S. enterica* subsp. *enterica* serovar Gallinarum)]. Sequences were aligned using MacClade 4.03 (Maddison and Maddison, 2001).

The total length of *E. aerogenes* and *Salmonella* sp. nucleotide sequences were 1128 (376 codon sites) and 1134 base pairs (378 codon sites), respectively. The general-time reversible (GTR) model with proportion of invariable sites (I) was the appropriate nucleotide substitution model selected by the Akaike Information Criterion (AIC) (Huelsenbeck and Crandall, 1997) implemented in MODELTEST ver. 3.5 (Posada and Crandall, 1998). Unrooted maximum likelihood (ML), maximum parsimony (MP) and Bayesian inference (BI) phylogenies were inferred. MP analysis was conducted using heuristic search option, implementing stepwise addition with 100 random addition replicates and TBR branch swapping using PAUP* ver. 4.0b10 (Swofford, 2002). PHYML ver. 2.4.4 (Guindon and Gascuel, 2003) was used for ML analyses and MrBayes ver. 3.04 (Huelsenbeck and Ronquist, 2001) was used for BI. Nodal supports for the MP and ML trees were estimated using 1000 non-parametric bootstrap replicates. MrBayes was used to conduct a Bayesian approach to phylogenetic inference by running 2×10^6 generations (10,000 burn-in) with four Metropolis coupled MCMC to optimize efforts to find peaks in tree-space.

Parameters were set to $nst = 6$ and $rates = inv$ and one tree was sampled in every 100. Convergence of the tree was checked using Tracer ver. 1.3.1 (Rambout and Drummond, 2003), resulting trees were used to generate a majority consensus tree with posterior probability values.

Tests for selection

We performed two different types of analyses; first, we estimated the pairwise d_N and d_S among all the *ompC* sequences within each species using a maximum-likelihood (ML) approach described by Goldman and Yang (1994) implemented in CODEML program of PAML package (Yang, 1997). Second, to account for among site variations and to test for positive selection on different codon sites, we estimated parameters under seven different codon substitution models (Yang et al., 2000) and their performances were evaluated using likelihood ratio tests (LRTs). Tests for positive selection were carried out using ML approach implemented in PAML ver. 3.15 (Yang, 1997). In order to account for uncertainty regarding the true topology, we repeated the tests for positive selection using inferred ML, BI and MP trees. These trees were incorporated in the PAML program and parameters under each tree were estimated and compared. The seven codon substitution models implemented in CODEML of the PAML package are; M0: one-ratio, M1a: nearly neutral, M2a: positive selection, M3: discrete, M7: beta, M8: beta+ $\omega > 1$: continuous (Yang et al., 2000), and M8a: beta+ $\omega = 1$ (Swanson et al., 2003). The M1a model estimates a single parameter: p_0 , the frequency of conserved sites with $\omega_0 = 0$ and the remaining sites with frequency p_1 ($p_1 = 1 - p_0$) assuming $\omega_1 = 1$. The M2a model adds a class of positively selected sites with frequency p_2 (where $p_2 = 1 - p_1 - p_0$), with ratio ω_2 estimated from the data. Thus, while M1a estimates a single parameter (p_0), M2a estimates three parameters (p_0 , p_1 , and ω_2). In the M7 model, ω follows a beta distribution such that $0 \leq \omega \leq 1$ and the two parameters (p and q) of the beta distribution are estimated from the data. In the M8 model, a proportion p_0 of sites have ω drawn from the beta distribution. The remaining sites with proportion p_1 are positively selected and have $\omega_1 > 1$. Thus, M7 estimates two parameters (p and q), while M8 estimates four parameters (p , q , p_0 , and ω_1). The LRTs between nested models were conducted by comparing twice the difference in log-likelihood values ($2 \ln \Delta l$) against a χ^2 -distribution, with degrees of freedom (df) equal to the difference in the number of parameters between

models (Yang et al., 2000). Five LRTs were conducted. First comparison was made between M0, a model that fits a single ω for all sites with M2a, which allows three site classes ($0 < \omega < 1$, $\omega = 1$ or $\omega > 1$). The second comparison was between M0 and M3. The third comparison was between M1a, which allows for two site classes ($0 < \omega < 1$, $\omega = 1$) with M2a. The fourth comparison was between a model of beta distributed selective pressures, which allows for 10 site classes, each with $\omega < 1$ (M7) and M8, which has 11 site classes, one of them allowed for $\omega > 1$. The last comparison was between M8 and M8a, in which an additional parameter is constrained to have $\omega = 1$ (Swanson et al., 2003). The comparison of M7 and M8 is the most stringent test of positive selection (Anisimova et al., 2001). In all LRTs, good evidence for positive selection is found if the LRT indicates that models which allow for selection (i.e. M2a and M8) are significantly better than their respective null models (M1a, M7 and M8a) (Yang et al., 2000).

Three-dimensional structure of *E. aerogenes ompC*

We used Geno3D ver. 2 (<http://geno3d-pbi-l.ibcp.fr>), an automated web server for protein molecular modeling (Combet et al., 2002) to predict the structure of *E. aerogenes ompC*. The osmoporin (OmpK36) crystal structure of *Klebsiella pneumoniae* (Protein Data Bank number = 1OSM; Dutzler et al., 1999) was used as reference point to predict the *ompC* structure of *E. aerogenes*. The structure was displayed and positively selected sites were mapped using RasMol V2.7.2.1.1 (<http://www.openrasmol.org/software/rasmol/>).

Results and discussion

Phylogenetic relationship among six strains of *E. aerogenes* and among *Salmonella* species are shown in Fig. 1a and b, respectively. Distribution of pairwise estimates of ω values for *E. aerogenes* and *Salmonella* sp. are shown in Fig. 2a and b, respectively. From the pairwise comparison, it is unequivocal that the entire coding regions of *ompC* in both species are under the influence of purifying selection ($\omega < 1$). However, estimates of ω for *E. aerogenes* are relatively larger than that of *Salmonella* sp. Although pairwise comparison indicated the evidence of purifying selection operating on the entire coding region, from the pairwise comparison it is difficult to infer whether any individual codon sites are subjected to positive

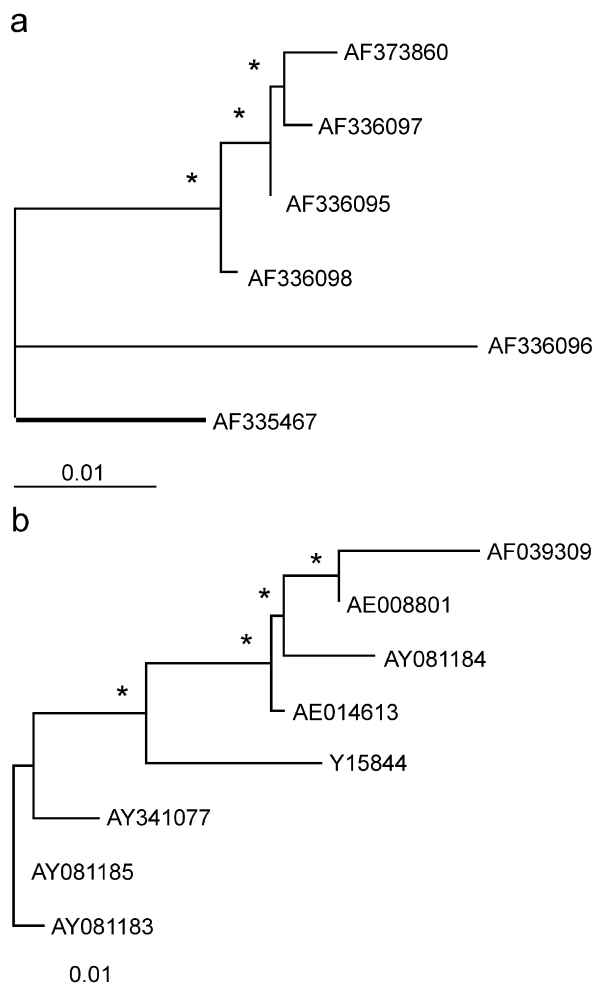


Figure 1. Unrooted ML phylogenies (a) *Enterobacter aerogenes* and (b) *Salmonella* sp. inferred from the *ompC* genes of respective species. Nodal supports (ML/BI/MP) ≥ 80 are indicated by asterisks.

Darwinian selection. If majority of codon sites are under the influence of purifying selection and few codons are subjected to positive Darwinian selection, the overall ω ratio is expected to be less than 1, and the codon sites that are under the influence of positive selection are underestimated (Yang et al., 2000). Considering this fact, we performed ML-based codon substitution analyses that accounted for individual codon sites. These methods offer a number of advantages over the pairwise comparisons of d_N and d_S among taxa that average over all sites and lineages (Akashi, 1999; Crandall et al., 1999).

To account for uncertainty regarding the true topology, we repeated the tests of positive selection using all the three trees resulting from ML, MP and BI analyses. The results were nevertheless very similar. Parameter estimates and log likelihood

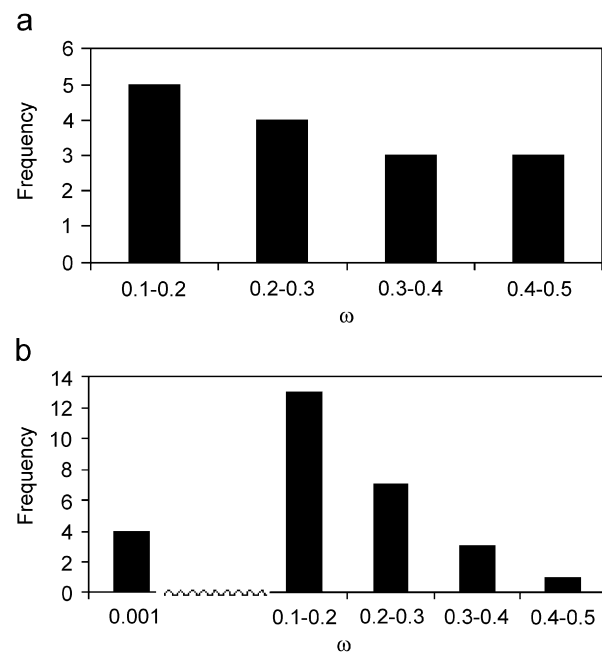


Figure 2. Distribution of pairwise estimates of ω -ratios among *ompC* sequences estimated by the maximum likelihood approach (Goldman and Yang, 1994) implemented in PAML (Yang, 1997): (a) *Enterobacter aerogenes* and (b) *Salmonella*.

values under models of variable ω ratios among codon sites of *ompC* gene and their likelihood ratio statistic (LRTs) are shown in Table 1a and b, respectively. Under the simplest model, which allows for only a single ω ratio across all sites (M0), the ML estimate of ω for *E. aerogenes* and *Salmonella* sp. are 0.3194 and 0.2047, respectively. These estimates are statistically indistinguishable from $\omega = 1$, the expected value under a completely neutral model of sequence evolution. In both species, a model allowing for variation among sites (M3) provides a significantly better fit to the data, showing that there is variation among sites in the strength of selection. However, M0 is a highly unrealistic model and the M0–M3 comparison thus provides a test for variation in ω among sites rather than variation in strength of selection among sites (Nielsen and Yang, 1998; Yang et al., 2000). A more stringent model (M8) revealed that roughly 3.7% codon sites with ω value 21.52 were positively selected for *E. aerogenes*. In contrast, LRTs (Table 1b M1a–M2a, M7–M8, and M8–M8a $p > 0.05$) for *Salmonella* sp. failed to support the evidence of positive Darwinian selection on any codon sites. Recently, Chen et al (2006) reported significantly higher ω (999.00) in *ompC* gene among *Escherichia coli* strains. In the present study, although *E. aerogenes* revealed the evidence of adaptive

Table 1a. Parameter estimates and log-likelihood values under models of variable ω -ratios among sites

Model	Free parameters	Parameter estimates	Likelihood scores	Positively selected sites*
<i>Enterobacter aerogenes</i>				
M0: One-ratio	1	$\omega = \mathbf{0.3194}$	−1915.505285	None
M1a: Nearly neutral	1	$\omega_0 = 0, \omega_1 = 1, (p_0 = 0.82, p_1 = 0.18)$	−1893.397228	Not allowed
M2a: Positive selection	3	$\omega_0 = 0, \omega_1 = 1, \omega_2 = \mathbf{22.53}; (p_0 = 0.82, p_1 = 0.15, p_2 = 0.03)$	−1878.626396	14, <u>133</u> , <u>228</u> , <u>231</u> , <u>232</u> , 233, 234, <u>236</u> , <u>237</u> , <u>238</u> , <u>239</u> , <u>240</u> , 280, <u>355</u>
M3: discrete [†]	5	$\omega_0 = 0.07, \omega_1 = \mathbf{8.69}, \omega_2 = \mathbf{96.47}, (p_0 = 0.945, p_1 = 0.05, p_2 = 0.005)$	−1876.821387	14, <u>133</u> , <u>228</u> , <u>231</u> , <u>232</u> , 233, <u>234</u> , <u>236</u> , ω , <u>238</u> , <u>239</u> , <u>240</u> , 280, <u>355</u>
M7: β	2	$p = 0.005, q = 0.02$	−1893.570174	Not allowed
M8: $\beta + \omega_s > 1$	4	$p_0 = 0.97, p_1 = 0.03, p = 0.013, q = 0.072, \omega = \mathbf{21.52}$	−1878.79428	14, <u>133</u> , <u>228</u> , <u>231</u> , <u>232</u> , 233, <u>234</u> , <u>236</u> , <u>237</u> , <u>238</u> , <u>239</u> , <u>240</u> , 280, <u>355</u>
M8a: $\beta + \omega_s = 1$	3	$p_0 = 0.82, p_1 = 0.17, p = 0.005, q = 1.72, \omega = 1$	−1893.39726	Not allowed
<i>Salmonella</i>				
M0: One-ratio	1	$\omega = \mathbf{0.2047}$	−1865.640191	
M1a: Nearly neutral	1	$\omega_0 = 0, \omega_1 = 1, (p_0 = 0.81, p_1 = 0.19)$	−1859.778848	
M2a: Positive selection	3	$\omega_0 = 0, \omega_1 = 1, \omega_2 = 1.83; (p_0 = 0.88, p_1 = 0, p_2 = 0.12)$	−1859.093214	
M3: discrete	5	$\omega_0 = 0, \omega_1 = 0, \omega_2 = 1.83, (p_0 = 0.88, p_1 = 0, p_2 = 0.12)$	−1859.093214	
M7: β	2	$p = 0.005, q = 0.02$	−1859.819392	
M8: $\beta + \omega_s > 1$	4	$p_0 = 0.88, p_1 = 0.12, p = 0.005, q = 2.82, \omega = 1.83$	−1859.093214	
M8a: $\beta + \omega_s = 1$	3	$p_0 = 0.81, p_1 = 0.19, p = 0.005, q = 1.58, \omega = 1$	−1859.778859	

Positively selected sites with posterior probability ≥ 90 are underlined, 0.8–0.9 in bold, 0.7–0.8 in italics, and 0.5–0.7 in plain text. M0: one-ratio ω value is the average for all codon sites, whereas M2a, M3 and M8 models ω values are the estimated values for the positively selected codon sites under respective models.

*Sites with a posterior probability $> 50\%$ of having $\omega > 1$.

[†]Model detected posterior probabilities based on naive empirical Bayes (NEB) analysis.

Table 1b. Likelihood ratio statistics among different models given in Table 1a

Comparison	2 Δl	df	<i>p</i>
<i>Enterobacter aerogenes</i>			
M0 vs. M2a	73.76	2	0.00
M0 vs. M3	77.37	4	0.00
M1a vs. M2a	29.54	2	0.00
M7 vs. M8	29.55	2	0.00
M8 vs. M8a	29.21	1	0.00
<i>Salmonella</i>			
M0 vs. M2a	13.09	2	0.00
M0 vs. M3	13.09	4	0.01
M1a vs. M2a	1.37	2	0.50
M7 vs. M8	1.45	2	0.48
M8 vs. M8a	1.37	1	0.24

evolution, the discrepancy in ω values between *E. aerogenes* and *E. coli* strains could be attributed to adaptive behavior of the strains. However, a detail investigation would reveal the causative factors for such discrepancy. Surprisingly, *Salmonella* sp. indicated the evidence of strong purifying selection. Although the reasons for such contrast are not known, it could be possible that cell-surface exposed domains of *E. coli* and *E. aerogenes* might be more adaptive to the adverse environmental conditions and caused rapid functional diversification, whereas *Salmonella ompC* is not influenced by the extrinsic environments, therefore functionally more conserved. Another possible reason could be its pathogenic nature by which it prefers to maintain the virulence nature without bringing substantial change to its genetic structure.

There is a large degree of overlap in the positively selected amino acid sites identified from models M2a, and M8 in *E. aerogenes* (Table 1a). While LRTs indicated strong evidence of positive selection (M1a vs. M2a; M7 vs. M8 in Table 1b) in *E. aerogenes*, the test failed to support the evidence of positive Darwinian selection in *Salmonella* sp. Although maximum likelihood-based codon substitution analysis, has several advantages over the pairwise comparison methods (but see Suzuki and Nei, 2001), tests for positive selection based on $\omega > 1$ are extremely stringent and most likely to fail to identify adaptive evolution when selection is weak (Anisimova et al., 2001).

In *E. aerogenes*, Thiolas et al (2004) reported the pattern of amino acid substitutions in a group of enteric bacteria and found significant proportion of amino acid replacement on the cell-surface exposed domains (L1–L8, except L3) in

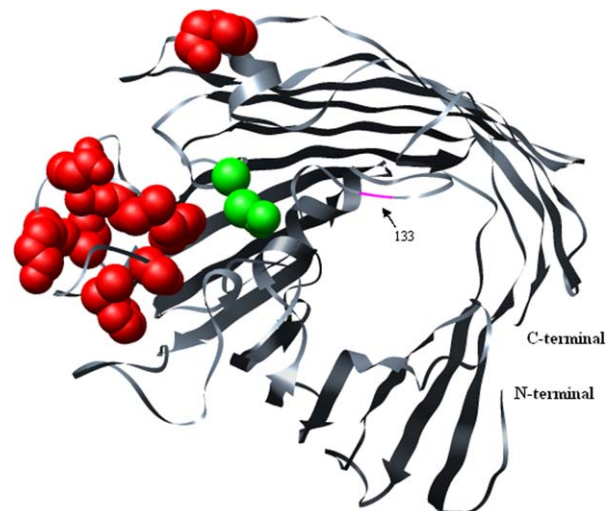


Figure 3. The predicted three-dimensional structure of *Enterobacter aerogenes ompC* gene reconstructed using Geno3D ver. 2 (Combet et al., 2002). Sites shown are those sites predicted to be under positive selection with posterior probability $> 50\%$ (see M2a in Table 1a). The structure was displayed using RasMol V2.7.2.1.1 (<http://www.openrasmol.org/software/rasmol/>). Positively selected sites at the hypervariable cell-surface exposed domains L5 and L3 are coded by dark grey and black (red and magenta, online), respectively. Trans-membrane domain residues is coded light grey (green, online).

ompC (*omp36*), whereas the trans-membrane domains ($\beta 1$ – $\beta 16$) were highly conserved. To test for positive selection we used all the sequences of *E. aerogenes* used in Thiolas et al (2004) study. Our study revealed that 10 codon sites ($\sim 71\%$ of the total positively selected sites in M8) in the hypervariable L5 domain of *E. aerogenes* were subjected to positive Darwinian selection, whereas only one codon site each in L3 and trans-membrane domain were positively selected. Trans-membrane domains of proteins are likely to be highly conserved (Jiggins et al., 2002; Fitzpatrick and McInerney, 2005), therefore it is obvious to find very few positively selected sites in those domains. The three-dimensional structure of *ompC* of *E. aerogenes* and the positively selected sites are shown in Fig. 3.

Despite the fact that both *E. aerogenes* and *Salmonella* are the members of enteric bacteria, *ompC* gene of both species are influenced by differential selection pressure. The rapid adaptive evolution of *E. aerogenes* could be the likely explanation for high resistance to commonly used antibiotics. This information on *ompC* could be important in analyzing its potential role in diagnostic assay, in antibiotic resistance and as immunogens for vaccination.

Acknowledgement

We thank the University of Tulsa for providing facilities to carry out this work.

References

- Akashi H. Within and between-species DNA sequence variation and the "footprint" of natural selection. *Gene* 1999;238:39–51.
- Allerberguer F, Koeuth T, Lass-Flörl C, Dierich MP, Putensen C, Schmützhard E, et al. Epidemiology of infections due to multiresistant *Enterobacter aerogenes* in a university hospital. *Eur J Clin Microbiol Infect Dis* 1996;15:517–21.
- Andrews TD, Gojobori T. Strong positive selection and recombination drive the antigenic variation of the PilE protein of the human pathogen *Neisseria meningitidis*. *Genetics* 2004;166:25–32.
- Anisimova M, Bielawski JP, Yang Z. Accuracy and power of the likelihood ratio tests in detecting adaptive molecular evolution. *Mol Biol Evol* 2001;19:1585–92.
- Arockiasamy A, Krishnaswamy S. Purification of integral outer-membrane protein OmpC, a cell surface antigen from *Salmonella typhi* for structure–function studies: a method applicable to enterobacterial major outer membrane protein. *Anal Biochem* 2000;283:64–70.
- Arpin C, Coze C, Rogues AM, Gachie JP, Bebear C, Quentin C. Epidemiological study of an outbreak due to multidrug-resistant *Enterobacter aerogenes* in a medical intensive care unit. *J Clin Microbiol* 1996;34:2163–9.
- Bornet C, Davin-Regli A, Pages JM, Bollet C. Imipenem resistance of *Enterobacter aerogenes* mediated by outer membrane permeability. *J Clin Microbiol* 2000;38:1048–52.
- Bornet C, Saint N, Fetnaci L, Dupont M, Davin-Regli A, Bollet C, et al. Omp35, a new *Enterobacter aerogenes* porin involved in selective susceptibility to cephalosporins. *Anat Agents Chemother* 2004;48:2153–8.
- Bosi C, Davin-Regli A, Bornet C, Mallea M, Pages JM, Bollet C. Most *Enterobacter aerogenes* strains in France belong to a prevalent clone. *J Clin Microbiol* 1999;37:2165–9.
- Chen SL, Hung CS, Xu J, Reigstad CS, Magrini V, Sabo A, et al. Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: a comparative genomics approach. *Proc Natl Acad Sci USA* 2006;103:5977–82.
- Combet C, Jambon M, Deleage G, Geourjon C. Geno3D: automatic comparative molecular modelling of protein. *Bioinformatics* 2002;18:213–4.
- Crandall KA, Kelsey CR, Imamichi H, Lane HC, Salzman NP. Parallel evolution of drug resistance in HIV: failure of nonsynonymous/synonymous rate ratio to detect selection. *Mol Biol Evol* 1999;16:372–82.
- Dutzler R, Rummel G, Alberti S, Hernandez-Alles S, Phale P, Rosenbusch J, et al. Crystal structure and functional characterization of OmpK36, the osmoporin of *Klebsiella pneumoniae*. *Structure Fold Des* 1999;7:425–34.
- Fares MA, Moya A, Escarmis C, Baranowski E, Domingo E, Barrio E. Evidence for positive selection in the capsid protein-coding region of the foot-and-mouth disease virus (FMDV) subjected to experimental passage regimens. *Mol Biol Evol* 2001;18:10–21.
- Fitzpatrick DA, McInerney JO. Evidence of positive Darwinian selection in Omp85, a highly conserved bacterial outer membrane protein essential for cell viability. *J Mol Evol* 2005;60:268–73.
- Georgioudis PR, Hamill RJ, Wright CR, Versalovic J, Koeuth T, Watson DA, et al. Molecular epidemiology of infections due to *Enterobacter aerogenes*: identification of hospital outbreak-associated strains by molecular techniques. *Clin Infect Dis* 1995;20:84–94.
- Goldman N, Yang ZA. Codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol Biol Evol* 1994;11:725–36.
- Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 2003;52:696–704.
- Huelsenbeck JP, Crandall KA. Phylogeny estimation and hypothesis testing using maximum likelihood. *Ann Rev Ecol Syst* 1997;28:437–66.
- Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 2001;17:754–5.
- Hughes AL, Nei M. Nucleotide substitution at major histocompatibility complex class II loci: Evidence for overdominant selection. *Proc Natl Acad Sci USA* 1989;86:958–62.
- Jalaluddin S, Devaster JM, Scheen R, Gerard M, Butzler JP. Molecular epidemiological study of nosocomial *Enterobacter aerogenes* isolates in a Belgian hospital. *J Clin Microbiol* 1998;36:1846–52.
- Jarvis WR, Martone WJ. Predominant pathogens in hospital infections. *J Antimicrob Chemother* 1992;29(Suppl A):19–24.
- Jiggins FM, Hurst GD, Yang Z. Host-symbiont conflicts: positive selection on an outer membrane protein of parasitic but not mutualistic Rickettsiaceae. *Mol Biol Evol* 2002;19:1341–9.
- Kimura M. The neutral theory of molecular evolution. New York: Cambridge University Press; 1983.
- Luttenberg B, van Alphen L. Molecular architecture and function of the outer membrane of *Escherichia coli* and other gram negative bacteria. *Biochem Biophys Acta* 1983;737:51–115.
- Maddison WP, Maddison D. MacClade version 4.03. Sunderland, MA: Sinauer; 2001.
- Muthukkaruppan VR, Nandakumar KS, Palanivel V. Monoclonal antibodies against *Salmonella* porins: generation and characterization. *Immunol Lett* 1992;33:201–6.
- Nielsen R, Yang Z. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* 1998;148:929–36.

- Nikaido H, Vaara M. Molecular basis of bacterial outer membrane permeability. *Microbiol Rev* 1985;49:1–32.
- Posada D, Crandall KA. Modeltest: testing the model of DNA substitution. *Bioinformatics* 1998;14:817–8.
- Rambout A, Drummond AJ. 2003. Tracer. Version 1.3.1. Available from <<http://evolve.zoo.ox.ac.uk/>>.
- Schaberg DR, Culver DH, Gaynes RP. Major trends in the microbial etiology of nosocomial infection. *Am J Med* 1991;91(Suppl. 3B):72S–5S.
- Smith NH, Maynard-Smith J, Spratt BG. Sequence evolution of the *porB* gene of *Neisseria gonorrhoeae* and *Neisseria meningitidis*: evidence of positive Darwinian selection. *Mol Biol Evol* 1995;12:363–70.
- Sujatha S, Arockiasamy A, Krishnaswamy S, Usha R. Molecular modeling of epitope presentation using the membrane protein *OmpC*. *Ind J Biochem Biophys* 2001;38:294–7.
- Suzuki Y, Nei M. Reliabilities of parsimony-based and likelihood based methods for detecting positive selection at single amino acid sites. *Mol Biol Evol* 2001;18:2179–85.
- Swanson WJ, Nielsen R, Yang Q. Pervasive adaptive evolution in mammalian fertilization proteins. *Mol Biol Evol* 2003;20:18–20.
- Swofford DL. PAUP* phylogenetic analysis using parsimony (* and other methods), version 4. Sunderland, MA: Sinauer Associates; 2002.
- Thiolas A, Bornet C, Davin-Regli A, Pages JM, Bollet C. Resistance to imipenem, cefeme, and ceftiofime associated with mutation in *Omp36* osmoporin of *Enterobacter aerogenes*. *Biochem Biophys Res Comm* 2004;317:581–6.
- Urwin R, Holmes EC, Fox AJ, Derrick JP, Maiden MC. Phylogenetic evidence for frequent positive selection and recombination in the meningococcal surface antigen *PorB*. *Mol Biol Evol* 2002;19:1686–94.
- Yang Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* 1997;13:555–6.
- Yang Z, Nielsen R, Goldman N, Pederson AM. Codon-substitution models for heterogeneous selection pressures at amino acid sites. *Genetics* 2000;155:431–49.