Microbiological Research 164 (2009) 282–289

Available online at www.sciencedirect.com
ScienceDirect

Microbiological Research

www.elsevier.de/micres

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]

[a]*Department of Biological Science, University of Tulsa, 600 S. College Ave., Tulsa, OK-74104, USA*
[b]*Baylor 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 transmembrane domains are functionally highly constrained.

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).

0944-5013/$ - see front matter © 2007 Elsevier GmbH. All rights reserved.
doi:10.1016/j.micres.2006.12.002

resistant forms have been reported from many countries (Allerberguer et al., 1996; Arpin et al., 1996; Georghiou 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 *Salmolella*.

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 Akiake 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.