ABSTRACT

COMPARATIVE MOLECULAR EVOLUTIONARY ANALYSIS
OF VIRULENCE LOCI IN PATHOGENIC *ESCHERICHIA COLI*

By

David William Lacher

Escherichia coli is a diverse species of Gram-negative bacteria, some strains of which are pathogenic. The early stages of *E. coli* pathogenesis often involve bacterial attachment mediated by the expression of surface proteins. It has been hypothesized that pathogens alter their surface proteins in order to evade detection by their host's immune system. Therefore, it is likely that natural selection is acting to generate new allelic variants. It is the goal of this research to examine the allelic diversity of genes that encode a variety of surface structures in different classes of pathogenic *E. coli* (pathotypes). The specific aims are to: 1) develop a method to quickly and accurately subtype a highly polymorphic locus responsible for the hallmark phenotype of the attaching and effacing *E. coli*, 2) characterize enteropathogenic *E. coli* (EPEC) through multilocus sequence typing (MLST) and restriction fragment length polymorphism (RFLP) analyses, 3) assess the level of genetic polymorphism in a region of the operon encoding the type 1 fimbriae of *E. coli*, and 4) examine various genes encoding surface structures for the actions of positive selection and recombination.

To address specific aim 1, a new method to quickly and accurately type the *eae* locus was developed. This new technique addresses the limitations of existing typing schemes and was applied to a set of *E. coli* capable of the attaching and effacing phenotype. For specific aim 2, a system was designed to detect and identify the alleles of three EPEC virulence genes. The distribution of these virulence gene alleles was

assessed in a collection of strains representing a variety of EPEC serotypes and then compared to a phylogenetic framework generated from MLST analysis of conserved housekeeping loci. To address specific aim 3, a segment of DNA encompassing a regulatory region of the type 1 fimbrial operon was sequenced in various types of pathogenic *E. coli*. This region contains an invertible genetic element responsible for the phase variability of the type 1 fimbriae and has been shown to be inactive in some strains. For specific aim 4, a collection of allelic sequences was assembled for genes encoding five different surface proteins from several *E. coli* pathotypes. These genes were analyzed for evidence of positive selection and homologous recombination. The results of this work will give us a better understanding of how different types of pathogenic *E. coli* have evolved and may have important public health implications.