ABSTRACT

GENOMIC ANALYSIS OF PATHOGEN EVOLUTION: VIRULENCE GENE ACQUISITION AND GENETIC EROSION IN *ESCHERICHIA COLI*

By

Adam Michael Nelson

Escherichia coli is a gram-negative, rod-shaped bacterium that lives naturally and commensally in the intestinal tract of humans and other mammals. However, some types of *E. coli* have acquired genetic elements encoding virulence factors that contribute to disease. Different disease phenotypes are caused by strains carrying a variable array of virulence factors. Many of these virulence factors are largely acquired through horizontal transfer. Pathogenic *E. coli* are grouped into at least twelve classes, or pathotypes, based on the type of disease they cause. Observing the acquisition of virulence elements as a means to study evolution has been previously used to determine the genetic ancestry of *E. coli* and other pathogenic microorganisms. Conversely, gene loss can also be important in enhancing pathogenicity by removing genes that encode proteins that may hinder increased virulence, or are no longer functional because they are mutated or otherwise incomplete. The research presented here is intended to enhance the current understanding of the roles of genetic acquisition and genetic loss to virulence changes in pathogenic *E. coli* by applying three paths of inquiry.

To address correlations between virulence profiles and disease incidence, 392 *E. coli* isolates from 115 pediatric patients were screened for virulence gene content using fluorescently labeled PCR amplicons in a capillary based sequencing system. Virulence profiles were compared to a phylogenetic framework to determine virulence distribution, and correlations between presence of specific genes and the incidence of disease in

patients. The screening of this pediatric population lead to the discovery of a variant of *E. coli* (sequence type 29), which was found very frequently (19%) in all of the samples examined. This population of ST-29 isolates was further characterized by PCR to determine the frequency of a panel of attachment-related loci, and by RFLP to determine the capsular polysaccharide type.

To track genetic erosion events, PCR-based screening of individual components in a pathogenicity island, called ETT2, showed examples of gene loss that rids possibly non-functional genetic material from the genome. When a strain acquires new genetic material by horizontal transfer, the new material must provide some fitness benefit or it will be under selection to be removed. Type III secretion systems are complicated structures that require a large number of genes to encode the proteins necessary for proper assembly. If critical components of the type III assembly are missing, the assembly will not function properly, and is likely to be under selective pressure for deletion. A streamlined genome may result in a more efficient pathogen. Here, the loss of all or portions of ETT2 is shown in a variety of pathogenic isolates of *E. coli*. At least six different deletion variants were discovered in the 57 strains examined.

These results help advance our understanding of evolution in *E. coli* by both acquisition and loss of virulence elements and further demonstrate the dynamic and diverse nature of the *E. coli* genome.