**­­Virulence factors and population dynamics in uropathogenic *E. coli.***

Urinary tract infections (UTIs) are a threat to human health, as approximately 50% of women suffer from a UTI at some time in their life. Additionally, UTIs are a burden on the healthcare system, as they result in 11 million physician visits annually, and a financial drain on society, as they cost to the United States of over $3.5 billion annually. More than 80%, of community acquired UTIs are caused by uropathogenic *Escherichia coli* (UPEC). During the acute phase of UTI, UPEC cells in the lumen of the bladder attach to and subsequently invade uroepithelial cells using extracellular type I pili fibers tipped with a FimH adhesin. After a successful invasion of the uroepithelial cells, UPEC cells form biofilm-like structures termed intracellular bacterial colonies (IBCs), which are necessary for persistence and pathogenicity in the bladder. In addition to factors governing formation of IBCs and type I pili, a number of other virulence factors have been characterized, including additional fimbrae structures, secreted toxins, iron acquisition elements, and immune evasion systems. Intriguingly, however, the presence of these virulence factors in an *E.* coli genome is not sufficient for uropathogenicity. A catalogue of factors necessary and sufficient for virulence in UPEC genome has yet to be compiled. This results, in part, from the technical difficulties that are caused by the stringent bottlenecks that exist during the acute phase of infection. During the invasion of uroepithelial cells, a stochastic population bottleneck severely reduces the genetic variability of the UPEC population. *In vivio* murine models routinely show a random loss of diversity of six orders of magnitude, where 108 lumenal bacteria result in approximately 1000 invasion events and the formation of 50-300 IBCs. Only UPEC that are able to form IBCs persist in the bladder while other bacteria that did not successfully form IBCs are expunged. The stringent and stochastic nature of these bottlenecks has rendered loss-of-function screens, such as transposon-mediated mutagenesis, ineffectual as the genomic coverage of deletions necessary to achieve statistical significant for a negative result becomes inhibitory. As such, although a number of virulence factors have been identified, the delineation of a core set of virulence genes remains elusive. The function and evolvability of these virulence factors are an important driver of the population dynamics of UPEC during UTI.

Currently, the accepted model of UPEC population dynamics during UTI is the source-sink model. Environmental factors, such as poor nutrient availability, that distinguish the bladder from other body habitats have led to the acceptance of the bladder as a “sink” habitat that must be replenished by a “source” habitat, which is postulated to be the gut. In this source-sink model, UPEC that invade the bladder quickly go extinct unless they have or quickly develop genetic or genomic changes that increase their fitness in the bladder environment. However, increased fitness in the sink habitat of the bladder is postulated to come at the cost of decreased fitness in the source habitat of the gut. This model predicts that, for the majority of UTIs, UPEC strains found in the bladder will be unstable, short-lived, and exist at low levels in other body habitats. Pilot experiments in our lab, however, have found that this prediction may not be biologically accurate in human UTIs. My research will be an examination of the biological relevancy of the source-sink explanatory model a description of inter- and intra-habitat allelic frequency of UPEC virulence factors. This research will aid understanding of population dynamics during UTIs and the evolution of virulence factors in UPEC populations.

**Specific Aim 1: Identification of virulence factors sufficient for uropathogenicity in *E. coli*.**

Prospective gain-of-function selections, such as functional genomic selections, are a viable alternative to ineffectual loss-of-function selections in UPEC UTI. In order to isolate virulence factors, I will construct a large-insert plasmid library from the prototypical UPEC strain *E. coli* UTI89 and transform the library into amodified *E. coli* strain containing the *fimH* gene from UTI89 that is capable of uroepithelial cell invasion, but not IBC formation or persistence in the bladder after 18hpi. The resultant transformants will then be inoculated into mice and, at 18hpi, the mouse bladders will be harvested. If any large-inserts derived from the UTI89 genome contain virulence genes sufficient for uropathogenicity, then IBCs will have formed, allowing for bacterial persistence in the bladder and kidneys. Individual plasmid clones will be isolated and sequenced to identify the gene regions that result in the infectious phenotype. Further functional characterization of genes found in the large inserts will reveal the identity of the virulence gene(s). Identification of the wild-type function of the gene products will be determined using GFP fusion proteins, knockout experiments, and time-course qRT-PCR.

**Specific Aim 2: Examination of population dynamics during UPEC UTI.**

The source-sink model of population dynamics predicts that for most UPEC strains containing pathoadaptive mutations that increase fitness in the bladder will exist at low frequency in the gut due to decreased fitness and stability in the source environment. I hypothesize, however, that UPEC strains that are dominant in the bladder are stable and dominant in the gut microbiota as well. This hypothesis predicts that UPEC strains isolated from patients suffering recurrent UTIs are stable across time points and found in high-titers in both the gut and bladder habitats. To test this prediction, I will isolate and identify UPEC strains in fecal and urine samples from patients with recurrent UTIs using single-nucleotide polymorphism biomarkers identified using high-throughput sequencing. This high-resolution analysis will enable comparisons of UPEC community and genomic structure across time points and body habitats through measurements of nucleotide variance at key loci and is more sensitive than 16S ribotyping or Multilocus Sequence Type identification.

**Specific Aim 3: Description of the allelic frequency of UPEC virulence factors in patients with Recurrent UTIs**

A number of genes have been shown to affect the virulence of UPEC during UTI, and ongoing research continues to identify new members of the virulence complement. Using the sequence data obtained from clinical isolates collected in Aim 2, I will assess the allelic frequency of known, and newly discovered, virulence factors in *E. coli* strains isolated from the gut and bladder habitats. This assessment will further test the predictions of source-sink population dynamics, as virulence factors that increase the fitness of UPEC in the bladder should exist at a low frequency in the gut microbiome. Allelic frequencies of these virulence factors will also be examined between patients and time points during recurrent UTI. These data can then be used to describe the stability of genetic and genomic mutations within UPEC virulence factors, an important consideration in population dynamics.

Taken together, these specific aims will result in a clearer understanding of population dynamics of UPEC virulence factors during UTI, an clinically-important infectious disease.