1. What is uropathogenic E. coli
   1. UPEC causes disease and is an excellent model organism for the study of opportunistic pathogens{Dobrindt:2010fe} as well as the study of mucosal infections by bacterial pathogens{Svanborg:2006it}.
   2. What is a UTI and how is it defined?
      1. Symptoms of UTI?
      2. Incidence of UTI
         1. Greater number and phenotypic variety are found among elderly patients with UTI{Vollmerhausen:2011fe}
      3. What is recurrent UTI?
      4. What bacteria are able to invade bladders and what bacteria cause disease
      5. Asymptomatic bacteria are also capable of existing in the bladder
   3. How is UPEC cleared from the host
      1. Natural clearance
      2. Antibiotic therapy
2. Population Structure of UPEC
   1. Description of Phenotypes of pathogenic E. coli.
      1. What are the clades
      2. Sub-categorization of intra-clade diversity is capable
         1. At least 9 groups of strains with clade B2{LeGall:2007bq}
      3. Pathovars are a phenotype that can be overlaid on the phylogentic tree.
   2. UPEC are predominantly found in clade B2, which is overrepresented in *E. coli* strains isolated from gut and bladder
      1. ExPEC fall predominantly into clade B2, to a lesser extent into clade D, and are generally absent from other clades{EscobarParamo:2004to}
      2. Majority of urine isolates of E. coli are from clade B2 {Zhang:2002wo, Moreno:2008eg, Moreno:2009jc}
      3. Interestingly, clade B2 has been shown to persist at a greater rate in the gut environment of human infants{Nowrouzian:2005uu}.
      4. Act as commensals. Eat stuff in the gut
   3. Impact of UTIs on strain richness in the gut and bladder
      1. During UTI population structure in the bladder
         1. The population dynamics of UPEC during the course of a UTI are complex and consist of a number of bottleneck events that occur both outside and within the host epithelium{Hannan:2012jh} which result in a drift to clonality in UPEC in the bladder{Schwartz:2011cy, Walters:2012hq}. These bottlenecks occur recursively during the course of the cyclical progression of UTI (Figure here). A stringent bottleneck occurs during the formation of intracellular bacterial colonies (IBCs), which is a critical step of UPEC pathogenesis that occurs during the acute phase of UTIs{Mulvey:1998wv, Anderson:2003kb, Justice:2004gx}. Although IBCs allow for significant clonal expansion of UPEC{Anderson:2003kb}, formation of the IBCs occurs at a very low rate, with only 50-700 IBCs persisting at 6h after inoculation of 107 UPEC {Schwartz:2011cy}. Formation of these IBCs requires known virulence factors, including the adhesin *fimH*{Wright:2007ha}. While the IBC bottleneck is important during the acute phase of UTI, the disappearance of IBCs at the end of the acute phase does not halt the continued loss in genetic diversity, suggesting a secondary bottleneck that occurs during the extracellular, chronic phase of UTI{Hannan:2012jh}. As with the IBC bottleneck, passage through the extracellular bottleneck is also mediated by known virulence factors{Hannan:2012jh}. These findings show that virulence factors have a significant effect on population structure, which, in turn, affects disease progression.
      2. Gut UPEC during UTI
         1. Transient strains and persistent strains. And blooms
         2. The gut populations of *E. coli* are surprisingly simple, with the majority of healthy women harboring fewer than four unique strains. The number of *E. coli* strains in the gut of women experiencing a UTI (~3) does not differ significantly from the number of *E. coli* strains in the guts of healthy women(2.5) as determined by PCR typing{Moreno:2008eg, Moreno:2009jc}. Interestingly, during acute UTI, The strains isolated from the urine are found to be the dominate strain in the rectal and fecal populations of *E. coli*{Gruneberg:1969wo, Yamamoto:1997wk, Moreno:2006ji, Moreno:2008eg}. Further, dominance of a B2 strain in the gut is correlated with both increased number of urovirulence factors in the dominant strain and reduced species richness in the gut habitat{Moreno:2008eg, Moreno:2009jc}. This suggests that more urovirulent strains are able to outcompete less urovirulent strains in the gut habitat, which may result in local extinction of those less virulent strains. This pattern mirrors the population dynamics that occur in the bladder during UTI.
   4. Recurrence of UPEC
      1. Recurrence occurs by the same strain approximately 50% of the time (LOOK UP CITATION)
      2. Perineum washing with antimicrobials did not prevent recurrence, indicating that another mechanism aside from anal to urogenital transmission is occurring{Cass:1985uc}
3. Virulence factors in UPEC
   1. Introduction to set of virulence factors
   2. UPEC genotypes are varied, but structured
      1. The pangenome of a species, defined as the collection of all genes found in at least one strain of the species, is comprised of the core genome, genes found in >95% of strains from that species, and accessory genes, which are genes that are found in at least one but less than 95% of strains for a species{Tettelin:2005jg, Medini:2008gi}. The composition of a bacterial pangenome has been shown to affect the evolution of virulence within a bacterial species (reviewed in Dobrindt *et al.* 2010). The *E. coli* pangenome is heavily biased towards accessory genes, as estimates of the total number of non-prophage, non-transposase genes in the *E. coli* reservoir is estimated to be over 10,000, almost five times as many as are expected to constitute the core genome shared by all *E. coli* strains{Rasko:2008bx, Touchon:2009kw}. The UPEC genomes that have been sequenced thus far, such as the model strains 536{Brzuszkiewicz:2006cu}, CFT0073{Welch:2002bj}, and UTI89{Chen:2006wz}, show similar patterns in pangenome composition. Additionally, like other *E. coli,* UPEC genomes contain a large number of accessory genes unique to specific strains, in part due to the prevalence of pathogenicity associated islands (PAIs) common to UPEC{Rasko:2008bx, Touchon:2009kw, Dobrindt:2010fe}. Despite the number of unique genes, members of the UPEC group have a greater genomic similarities and are more genetically distinct, as a group, than other pathovars{Rasko:2008bx}, which indicates that greater inter-group heterogeneity and less intra-group diversity. Taken together, these data indicate that, although the genomic content of the UPEC group is varied due to the presence and absence of accessory genes, they are generally similar in their total gene content. This is an indication that investigation of accessory genes is important in understanding phenotypic differences that exist between UPEC strains.
      2. Although the UPEC group has a high degree of genetic similarity, a definitive set of virulence factors has yet to be defined. Many UPEC genotypes are capable of causing disease in the bladder and there is no single set of urovirulence factors{Picard:1999uk, Norinder:2012fq, Touchon:2009kw}. Despite their variety, evidence suggests that the accumulation of virulence factors is non-random, as at least five virulence profiles can be delineated by analyzing the presence known virulence factors and clade membership of UPEC strains {Poey:2012be}. This is an indication that there is a pattern of co-occurrence of virulence factors, despite the variety of UPEC genotypes. Analysis of these virulence factors has shown that many factors co-occur and display low levels of intra-group diversity, indicating that structured, though frequent, horizontal gene transfer of virulence genes {Johnson:2001cl}. This pattern mirrors the homologous recombination in core genes, which has been shown to be high in *E. coli* {Touchon:2009kw}, which suggests that virulence factors move through *E. coli* populations through horizontal gene transfer and processes of recombination. These data are strong indicators that virulence gene networks, rather than single genes, define sets of virulent genotypes. As a result, single gene investigations may not capture a complete picture of UPEC pathogenicity{Picard:1999uk}.
   3. Virulence factors, phylogeny, and phenotype.
      1. While great variety exists in UPEC genotypes, single genes, or even small sets of virulence genes encoding complete virulence factors, are not sufficient to cause disease by themselves{Picard:1999uk, Marrs:2005ty}, which suggests that additional genetic factors are necessary for pathogenesis. Supporting this finding, genomic analysis of virulence factors presence and its correlation to fitness has shown that the genetic context of the virulence factors is important in determining its functional effect within UPEC strains{EscobarParamo:2004up, Nowrouzian:2005uu}. In many cases, non-pathogenic and pathogenic bacteria contain similar sets of virulence factors, as can be seen in commensal probiotic strain *E. coli* Nissle 1917 and the uropathogenic bacterium *E. coli* CFT073 {Grozdanov:2004bd}, however, despite their similar genomic architecture, these strains have widely different pathogenic potential. Additional evidence for the necessity of a proper genomic context for virulence gene pathogenicity comes phylogenetic analysis of the virulence genes. Virulence factors specific to pathogenic isolates are common in isolates from clades B2 and D and rare in other clades, indicating that they are ancestral to those clades (B2 and D){Boyd:1998ub}. Furthermore, genomic hybridization shows a correlation between the presence and absence of specific genomic content and the phylogenetic history of the core-genome of B2 isolates, indicating the co-evolution of the accessory and core genomes{LeGall:2007bq} possibly through a process of “fine-tuning” {EscobarParamo:2004to}. Taken together, these data indicate that the phenotypic effects of virulence genes is mediated by an interaction with the genomic milieu that has been fine-tuned by the evolutionary history of the strain, and that clades B2 and D may have the milieu most conducive to maximum virulence potential. Identification of the genetic factors, other than the accessory genes, that differ between clades B2 and D and other clades may reveal the context that enhances a virulent phenotype.
4. Swords or plowshares?
   * 1. Common extra-intestinal virulence genes have been found to affect the fitness of strains within the gut environment{LeGall:2007bq}.
     2. Virulence factors, such as hemolysin, type I fimbriae, and P fimbriae, are associated with persistence of E. coli in the gut{Nowrouzian:2003bs}.
     3. In healthy women, dominant E. coli clones had higher urovirulence scores than non-dominant clones{Moreno:2009jc}. In addition, persistent strains were more likely than transient strains to present uropathogenic phenotypes{Wold:1992tg}.
     4. Persistent strains in the gut environment were statistically more likely to belong to the uropathogenic subgroup of clade B2, indicating a potential link between fitness in the gut and pathogenicity in the bladder{Nowrouzian:2006bu}. Additionally, in patients with UTI, strains with greater numbers of urovirulence factors were associated with reduced species richness in the gut habitat and were more likely to belong to clade B2 {Moreno:2008eg}. This suggests that urovirulence factors may facilitate the clonal expansion of UPEC strains in the gut at the cost of competing *E. coli* strains, which results in a reduction of species diversity.
     5. Deletion of PAIs in CFT0073 reduces rate of intestinal colonization{Diard:2010fr}. Presence of PAIs in CFT0073 is linked to reduced growth rate in urine{Diard:2010fr}, indicating that there may be genetic factors that are maintained in the population despite the fitness cost of these factors in the bladder environment. This is an additional indication of selection pressure in habitats outside of the bladder.
5. Questions that remain in the host
   1. Why don’t more people have recurrent UTIs?
   2. Mounting evidence suggests that there are multiple ways to skin a cat, so single gene investigations may not capture a complete picture of UPEC pathogenicity{Picard:1999uk}
   3. What factors shape the abundance and distribution of virulence factors?
   4. Previous research into urovirulence genes in the gut have focused mainly on the presence or absence of the genes, rather than identification particular alleles.
   5. Something in addition to known virulence genes is enabling clade B2 strains to persist in the gut better than strains from other clades{Nowrouzian:2005uu}.
6. What studies of these questions will reveal
   1. Treatment plan – if you know that gut E. coli are part of the problem, then it must be addressed.
   2. Enable better understanding of how E. coli inhabit multiple environments, what it means to succeed in the gut and how to make sure that success isn’t abused.
   3. Better understanding of how E. coli have evolved, and why they are pathogenic in the first place.
      1. Accidental pathogen model (by-product hypothesis)
      2. Opportunistic pathogen model
      3. Source-sink model
7. Experimental plan

OVERALL QUESTION:  
What is happening with the gut population of UPEC, and how does that influence UTI susceptibility?

Hypothesis:

A bloom of UPEC growth coincides with the onset of UTI, which results in an increase in the relative abundance of urovirulence genes in the microbiome.

1. Measure the bloom (does it exist) – 454 pyrosequencing to identify community structure.
2. Measure relative abundance of virulence genes (does a bloom of E. coli coincide with a bloom of urovirulence genes?) – qPCR using BioMark
3. Sequence genomes of fecal and urine isolates (do these genomes differ?) – high througput sequencing using Illumina

~~PREDICTIONS:~~

~~If urovirulence genes offer a fitness advantage in the gut, then…:~~

~~…dominant strains in the bladder will be more fit in the gut environment than non-dominant strains.~~

~~…strains that are replaced during recurrent UTIs will be less virulent than strains that persist through recurrent infections.~~

~~Sub-aim 1: Assess population structure of UPEC in the host – determine population complexity using MLST – describe strain richness.~~

~~Sub-aim 2: Sequence representative genomes of MLST subtypes – allow for description of synteny and provide information on the genomic organization of strains. Perhaps identification of virulence profiles common to persistent versus replaced strains?~~

~~Sub-aim 3: Identify abundance and allelic distribution of virulence factors in patients suffering recurrent UTI – describe strain evenness. Do strains with more virulence factors dominate more frequently?~~

Questions I am trying to answer:

1. Are recurrent UTIs caused by the same strain (i.e., the same gene network?)
2. Wholesale shifts occur in the gut and the bladder at the same time. A minor member rising to prominence, or is it a factor of a secondary invasion?
   1. Alternatively, look to see if there is a concordance in change in virulence genes. Are they linked, or are they changing independently.
3. Is there a core UPEC genome? How different is the core UPEC genome compared to the core genome of E. coli?
   1. How variable is the UPEC accessory genome?
   2. Do virulence genes co-occur?
4. What is the competitive advantage of virulence genes? Do they offer fitness advantage in the gut?
   1. In patients with a secondary invasion of UPEC, could do subtraction between the genomes to identify regions that are different and ask if they are these virulence genes are responsible for the competitive advantage between isolates.

Interesting Side notes:

* 1. How many colicin producing strains are there in the gut?

What is the connection between UPEC population changes in the gut and susceptibility to UTI?

We are targeting recurrent UTI because people with an acute UTI are more likely to suffer another UTI.

Because UTIs diagnoses are so sketchy, specific rules will be made to ensure proper diagnosis{Foxman:2010bx}.