With these factors in mind, ***I hypothesize that there is a bloom of uropathogenic E. coli in the gut that coincides with the onset of a UTI.***

**Sub-aim1: Compare the community structure of the gut microbiota during UTI and after recovery.**

*Rationale:* While the human microbiome is generally stable through time1,2, changes in gut community structure during disease states has been identified3. Changes in the periutheral microbiota have also been found, as the prevalence of *E. coli* increases in the days preceding the onset of a UTI4. This is concordant with the rectal-perineal-urethral hypothesis, which states that the gut contains a reservoir of UPEC, which escape the gut and colonize the periurethral area before ascending into the bladder5. While a bloom of growth in the gut would increase transmission to the periurethral area, no study has investigated changes in the relative abundance of UPEC during the onset of a UTI episode. However, at least one population susceptible to UTIs, pregnant women6, undergo shifts in the ecology of their gut microbiome during the course of their pregnancy, including expansion of proteobacteria7. Thus, ***I hypothesize that there is a greater representation of E. coli in the gut microbiota of a patient at the onset of a UTI than at times when the patient is healthy.*** To test this hypothesis, I will measure the relative abundance of Enterobacteriaceae, the Family taxon that contains *E. coli*, in fecal samples collected from patients during UTI and after their recovery. If overgrowth of the gut reservoir of UPEC coincides with the onset of a UTI, then the abundance of Enterobaceriaceae will be higher in the patient during UTI than when the patient is healthy.

*Experimental methods:* In collaboration with Case Western University and their clinical facilities, women aged 18-41 years who present with symptoms of an uncomplicated, lower UTI will be open to enroll in a new cohort if they pass the entry requirements, which reduce the chance of enrolling patients with anatomical or functional abnormalities. Patients will supply three fecal samples, one at the time of entry, before the administration of antibiotics, one after completing the antibiotic regimen at 14 days after enrollment, and one at 28 days after enrollment; these samples will be labeled as UTI, Treated, and Recovered, respectively (Figure 2b). Fifty patients will be enrolled with the expectation that 30 will submit all the samples, be free from recurrent UTIs during the study, and have confirmed cases of UTI caused by UPEC. A cohort of 30 patients will give the study the power to detect an effect size of 1.40 (Figure 2a).

Genomic DNA from the fecal samples collected from cohort will be extracted and used as template 16S rRNA gene using 454 pyrosequencing. Sequence data from this will then be used to estimate the community structure of the gut microbiota, as has been done previously1. Binning of the amplicons into phylotypes using the QIIME software package will allow for the representation of those phylotypes in the gut microbiota to be measured at the family level8. Statistical analysis will include application of the Shannon diversity index to identify intra- and inter-host differences in community structure, principle component analysis (PCA) to measure the tendency of the samples to cluster to their sample groups, and Mann-Whitney non-parametric t-tests to identify changes in the representation of Enterobacteriaceae during the UTI and after recovery.

Antibiotic treatment has been shown to dramatically effect the composition of the gut microbiota. While the majority of taxons regenerate after 28 days

9-13, some genera may take months to reappear14. This will reduce the overall species richness of the gut microbiota, and may artificially reduce the divergence between the UTI and Recovered groups. As an additional control against this error, data on the gut community structure of healthy adults available from the Human Microbiome Project (HMP) will be included in these analyses in order to identify differences between healthy microbiomes and microbiomes that may be altered after antibiotic treatment, as has been done in other studies7.

*Expected results:* Analysis using the Shannon diversity index will measure the differences between the sample groups and will likely show that the samples from the different groups will show more inter-group diversity than intra-group diversity, indicating that the groups have shared features of their community structure that differ between the groups. This analysis will be supported by PCA which will show that communities from the Recovered group will cluster with the data from the HMP, while the UTI group and the Treated group will each cluster separately. Finally, the Mann-Whitney t-test will show that the UTI group will have a statistically significant higher representation of Enterobacteriaceae than the Recovered group, which is an indication that gut *E. coli* carriage is higher during UTI than after recovery. Taken together, these data will show that the gut microbiota during UTI is in an altered state distinct from a healthy state, and that the abundance of Enterobacteriaceae is higher in this altered state than in the healthy state.

*Anticipated challenges:* Recurrence of a UTI episode within 6-12 months after an initial UTI caused by UPEC occurs in approximately 25-30% of cases. The chance of recurrence is greatest within the first 3 months following the initial UTI, and is often caused by the same UPEC strain as the one that caused the initial UTI. The chance of another UTI episode caused by the same strain drops after 6 months. This indicates that there may be a short-lived reservoir of UPEC in the patient that survives antibiotic treatment and immune system clearance. This reservoir may be located in the gut, which would complicate analysis of the gut carriage of UPEC during an acute UTI episode. As a result, any patients who self-reported another UTI episode within the 6 months following entry into the cohort will be excluded from the final analyses.

It is possible that there are members of the Enterobacteriaceae that could be have increased representation in the gut at the onset of a UTI episode due to host factors, such as immunodeficiency. If there are large discrepancies between the estimated abundance of Enterobacteriaceae and *E. coli* specific markers identified in sub-aim2, then the relative representation of *E. coli* in the gut microbiome will be re-estimated using qPCR targeting *E. coli* specific housekeeping genes and comparing those results to a control region conserved in the family Enterobacteriaceae. This will identify the proportion of *E. coli* in the fecal sample relative to the abudance of the Enterobacteriaceae family.