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 time{Consortium:2012bb, Schloissnig:2012hx}, changes in gut community structure during disease states has been identified{Blumberg:2012en}. Pregnant women, a population at a higher risk to develop UTIs{Law:2012jr}, undergo shifts in the ecology of their gut microbiome, including expansion of proteobacteria{Koren:2012ji}. Changes in the periutheral microbiota have also been found, as the prevalence of *E. coli* increases in the days preceding the onset of a UTI{Czaja:2009jx}. This is concordant with the rectal-perineal-urethral hypothesis, which states that UPEC originate in the gut.

*Experimental methods*

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 previously{Consortium:2012bb}. 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 level (*E. coli* is found in the Enterobacteriaceae family){Caporaso:2010bi}. Statistical analysis will include application of the Shannon diversity index to identify intra- and inter-host differences in community structure, Mann-Whitney non-parametric test of means to identify changes in the representation of Enterobacteriaceae during the UTI, and prinicipal component analysis (PCA) to measure the tendency of the samples to cluster to their sample groups.

Antibiotic treatment has been shown to dramatically effect the composition of the gut microbiota, and some genera may take months to reappear {Dethlefsen:2008kn}. 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 studies{Koren:2012ji}.

*Expected results*

The above methods will enable discrimination between the sample groups (during UTI, after antibiotic treatment, and after recovery) and gut microbiomes collected from the HMP. 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 indicates that gut *E. coli* carriage is higher during UTI than after antibiotic treatment and recovery.

*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.