Using culture independent approaches to gain insights into human urinary tract microbiome in healthy male and female individuals

Sarah Hussein El-Alawi

Follow this and additional works at: https://fount.aucegypt.edu/etds

Recommended Citation

APA Citation

MLA Citation

This Master's Thesis is brought to you for free and open access by the Student Research at AUC Knowledge Fountain. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of AUC Knowledge Fountain. For more information, please contact thesisadmin@aucegypt.edu.
Using Culture Independent Approaches to Gain Insights Into Human Urinary Tract Microbiome in Healthy Male and Female individuals

A Thesis Submitted to
Biotechnology graduate program

In partial fulfilment of the requirements for the degree of Master of Science

By
Sarah Hussein El-Alawi

Under the supervision of
Dr. Hamza El-Dorry
January / 2012
Using Culture Independent Approaches to Gain Insights into Human Urinary Tract Microbiome in Healthy Male and Female individuals

A Thesis Submitted by

Sarah Hussein El-Alawi

To the Biotechnology Graduate Program

January / 2012

In partial fulfilment of the requirements for
The degree of Master of Science

Has been approved by

Thesis Committee Supervisor/Chair

Affiliation

Thesis Committee Reader/Examiner

Affiliation

Thesis Committee Reader/Examiner

Affiliation

Thesis Committee Reader/External Examiner

Affiliation

____________________________  ____________________  ____________________

Director                     Date                     Dean                     Date
DEDICATION

This work is dedicated to my parents and my brother who supported me at all time of stress and happiness. Without their encouragement, I couldn’t be able accomplish my work.
ACKNOWLEDGEMENTS

It is honourable to mention all people who contributed with their guidance and assistance that was provided during my thesis project. First of all, I would like to express my deepest gratitude to Prof. Dr. Hamza El-Dorry, the dissertation advisor for his guidance, expertise and patience that was continuously provided and assisted in broadening my graduate experience.

A very special thanks to Associate Research Professor, Dr. Ari Jose' Scattone Ferreira and Mr. Mustafa Adel for their suggestions and efforts in the computational analyses part.

I would like to thank Dr. Ahmed A. Sayed and Mr. Amged Ouf for their efforts in the sequencing part and their continuous guidance in the Laboratory. Thanks also go to my colleagues in Biotechnology master program, Rehab Abdallah, Nelly A. ElGhany and Laila Ziko.
ABSTRACT

Here we present comprehensive insights into the diverse collection of bacterial communities that indigenously inhabit the urinary tract of healthy adult male and female individuals using cultivation independent techniques. Several studies address the importance of uncultured microbial flora that inhabits the human urinary tract in health and diseases. Thus, a deep and accurate understanding of the factors influencing this ecological niche and its microbiome is important to prevent possible infectious diseases.

Genomic DNA was extracted from midstream urine samples collected from 4 adult, healthy and culture-negative male and female individuals at ages ranging from 18 to 40. Following PCR amplifications using 63F and 1387R primers for amplification of 16S rRNA genes from bacteria, 16S rRNA gene libraries were constructed using the amplified DNA fragments. From each library, we sequenced approximately 60-70 clones’ inserts from both ends. The generated 16S rRNA gene sequences were compared against RDP II database. Separate phylogenetic analyses of the two hypervariable regions, V1-V3 and V6-V8, revealed complex bacterial profiles with predominant taxa of Prevotella, Corynebacterium, Lactobacillus, Weeksella and Gardnerella. Different bacterial genera that have been identified in urinary tract overlap the microbial communities that were formerly reported in superficial skin, gut and vaginal microbiome. Disease causing bacteria were characterized as a part of the normal flora. Inter-individual variation within the same gender was observed, and no single characteristic bacterial community was evident. However, significant variations of microbial signatures between male and female individuals were observed. 16S rRNA gene sequences assigned to bacterial genera of Streptococcus, Staphylococcus, Escherichia/Shigella and others were characterized only in urines of male subjects whereas the genera of Atopobium, Megashpera and Oligella were identified only in urines of female subjects. Moreover, the richness and evenness of bacterial flora of the urinary tract was assessed by means of alpha diversity analysis. The experiments reported in this work represent the first study of the urinary tract microbiome in healthy Egyptian male and female
individuals, and should help in exploring the role of bacterial flora in health and diseases of both genders.
Table of Contents

DEDICATION ........................................................................................................... iii
ACKNOWLEDGEMENTS ........................................................................................ iv
ABSTRACT ................................................................................................................ v
GLOSSARY AND ABBREVIATIONS ........................................................................ x
CHAPTER 1: LITERATURE REVIEW ......................................................................... 1

1.1. Human Microbiome Project ............................................................................ 1
  1.1.1: Background .................................................................................................. 1
  1.1.2: Goals and Initiatives of Human Microbiome Project ................................. 1

1.2. Cultivation dependent and independent methods: ........................................... 1
  1.2.1: Limitations of the culture dependent and independent methods: .............. 2
  1.2.2: Characterization of the human microbiome using culture dependent and 
        independent methods: ................................................................................. 2
  1.2.3: Tools used for characterization of the human urogenital microbiome using 
        culture independent methods: ................................................................... 3

1.3. Factors affecting the microbiome structure and diversity: .............................. 4
  1.3.1: Female urogenital microbiome and probiotics: ....................................... 4
  1.3.2: Ethnic influences: ....................................................................................... 5
  1.3.3: Immunological influences: ......................................................................... 6

1.4. Female urinary system ..................................................................................... 6
  1.4.1: Anatomy of the female urinary system: .................................................... 6
  1.4.2: Antimicrobial defences of the urinary tract: ............................................. 7
  1.4.3: Female urinary system and microbiome: .................................................. 9
     1.4.3.1: Establishment of microbiota: ............................................................... 9
     1.4.3.2: The effect of age and hormonal status on the female microbiome: ...... 9
     1.4.3.3: The key members of the female urogenital tract microbiome: .......... 10
     1.4.3.4: Identification of *Lactobacillus* spp. and its role for the female 
            urogenital health: ..................................................................................... 10
     1.4.3.5: Other lactic acid producing bacteria and their role for a healthy 
            urogenital tract: ..................................................................................... 11

1.5. Male urinary system ....................................................................................... 11
  1.5.1: Anatomy of the male urinary system: ...................................................... 11
  1.5.2: Male urinary system and microbiome: .................................................... 11
     1.5.2.1: Antimicrobial defences of the urinary tract: .................................. 11
1.5.2.2: The effect of age and circumcision on the male microbiome: Error! Bookmark not defined.

1.5.2.3: The key members of the male urogenital tract microbiome: Error! Bookmark not defined.

CHAPTER 2: MATERIALS AND METHODS Error! Bookmark not defined.

2.1 Subjects and sample collection Error! Bookmark not defined.

2.2 DNA isolation Error! Bookmark not defined.

2.3 16S rRNA PCR amplification Error! Bookmark not defined.

2.4 Cloning, 16S rRNA gene library construction and Sequencing 15

2.5 Male and Female 16S rRNA urine sequence analyses 16

2.5.1: Alpha diversity analysis: Error! Bookmark not defined.

2.5.2: Phylogenetic analysis of 16S rRNA urine sequences: Error! Bookmark not defined.

CHAPTER 3: RESULTS 18

3.1 Processing of the urine samples, PCR amplification, library construction, and sequencing of the 16S rRNA gene 18

3.2 Computational and statistical analysis of 16S rDNA of human urine microbiome 18

3.2.1: Phylogenetic analysis of the female 16S rRNA sequences 19

3.2.2: Phylogenetic analysis of the male 16S rRNA sequences 21

3.2.3: Estimating the richness and diversity of urine microbial communities using alpha diversity 23

3.3 Inter-individual variation 24

CHAPTER 4: DISCUSSION 25

REFERENCES 31

FIGURES 41

TABLES 58

APPENDIX 64
LIST OF TABLES

Table 1 Summary of Female urine bacterial communities in 3’ and 5’ sequence datasets..... 58
Table 2 Summary of Male urine bacterial communities in 3’ and 5’ sequence datasets ....... 59
Table 3 Summary of the number of sequenced clones for each library, length of sequence obtained from each end and quality of reads................................................................. 61
Table 4 Alpha diversity measurements of V6-V8 sequence dataset of male and female 16S rRNA gene......................................................................................................................... 62
Table 5 Alpha diversity measurements of V1-V3 sequence dataset of male and female 16S rRNA gene......................................................................................................................... 63
LIST OF FIGURES

Figure 1 Distributions of different bacterial phyla ............................................................. 41
Figure 2 Gel electrophoresis of total genomics DNA from four male urine samples .......... 42
Figure 3 Gel electrophoresis of total genomics DNA from four female urine samples ...... 43
Figure 4 Gel electrophoresis of PCR amplification products of 16S rRNA gene from four
male urine samples........................................................................................................ 44
Figure 5 Gel electrophoresis of PCR amplification products of 16S rRNA gene from four
male urine samples........................................................................................................ 45
Figure 6 A phylogenetic tree showing the relationship between various phylotypes based on
analysis of 16S rRNA female urine sequences of V6-V8 dataset ..................................... 46
Figure 7 A phylogenetic tree showing the relationship between various phylotypes based on
analysis of 16S rRNA female urine sequences of V1-V3 dataset ..................................... 47
Figure 8 A phylogenetic tree showing the relationship between various phylotypes based on
analysis of 16S rRNA male urine sequences of V6-V8 dataset................................. 48
Figure 9 A phylogenetic tree showing the relationship between various phylotypes based on
analysis of 16S rRNA male urine sequences of V1-V3 dataset ..................................... 49
Figure 10 Comparison of major taxa taxonomically assigned to V1-V3 and V6-V8 reads of
human female 16S rRNA sequences urine sequences at the genus level.......................... 51
Figure 11 Comparison of major taxa taxonomically assigned to V1-V3 and V6-V8 reads of
human male 16S rRNA sequences urine sequences at the genus level............................ 53
Figure 12 Rarefaction curves plotted to estimate the phylotypes diversity and coverage level
of 16S rRNA urine sequences of 3’ reads from both male and female subjects ............... 54
Figure 13 Rarefaction curves plotted to estimate the phylotypes diversity and coverage level
of 16S rRNA urine sequences of 5’ reads from both male and female subjects ............... 55
Figure 14 The overall distribution of bacterial genera taxonomically assigned to 16S rRNA
urine sequences of V6-V8 dataset.................................................................................. 56
Figure 15 The overall distribution of bacterial genera taxonomically assigned to 16S rRNA
urine sequences of V1-V3 dataset.................................................................................. 57

Supplementary Figure 1 The approval letter by the institutional review board (IRB) of the
American university in Cairo......................................................................................... 64
Supplementary Figure 2 The Medical history.................................................................... 65
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALP</td>
<td>Anti-leukoprotease</td>
</tr>
<tr>
<td>AMPs</td>
<td>Anti microbial peptides</td>
</tr>
<tr>
<td>APCs</td>
<td>Antigen presenting cells</td>
</tr>
<tr>
<td>BV</td>
<td>Bacterial vaginosis</td>
</tr>
<tr>
<td>C-section</td>
<td>Caesarean section</td>
</tr>
<tr>
<td>Defb1</td>
<td>Defensin beta 1</td>
</tr>
<tr>
<td>DSGG</td>
<td>Disialosyl gal-globoside</td>
</tr>
<tr>
<td>GSLs</td>
<td>Glycosphingo lipids</td>
</tr>
<tr>
<td>HBD</td>
<td>Human beta defensin</td>
</tr>
<tr>
<td>HIMI</td>
<td>Human Intestinal Metagenome Initiative</td>
</tr>
<tr>
<td>HMP</td>
<td>Human Microbiome Project</td>
</tr>
<tr>
<td>HRT</td>
<td>Hormonal replacement therapy</td>
</tr>
<tr>
<td>HUSI-I</td>
<td>Human seminal inhibitor I</td>
</tr>
<tr>
<td>IDNS</td>
<td>SmartGene Integrated Database Network System</td>
</tr>
<tr>
<td>Ig</td>
<td>Immunoglobulin</td>
</tr>
<tr>
<td>IL12</td>
<td>Interleukin 12</td>
</tr>
<tr>
<td>INRA</td>
<td>French National Institute for Agricultural Research</td>
</tr>
<tr>
<td>IRB</td>
<td>Institutional Review Board</td>
</tr>
<tr>
<td>MALDI-TOF</td>
<td>Matrix Assisted Laser Desorption Ionization - Time of Flight</td>
</tr>
<tr>
<td>NCBI</td>
<td>U.S. National Center for Biotechnology Information</td>
</tr>
<tr>
<td>NIH</td>
<td>National Institute of Health</td>
</tr>
<tr>
<td>PLA2</td>
<td>Phospholipase A2</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal ribonucleic acid</td>
</tr>
<tr>
<td>SGG</td>
<td>Sialosyl gal-globoside</td>
</tr>
<tr>
<td>SLPI</td>
<td>Secretory leukocyte protease inhibitor</td>
</tr>
<tr>
<td>STDs</td>
<td>Sexually transmitted diseases</td>
</tr>
<tr>
<td>Th1</td>
<td>T-helper 1 cells</td>
</tr>
<tr>
<td>T-RFLPs</td>
<td>Terminal Restriction Fragment Length Polymorphisms</td>
</tr>
<tr>
<td>UTIs</td>
<td>Urinary tract infections</td>
</tr>
</tbody>
</table>
CHAPTER 1: LITERATURE REVIEW

1.1. Human Microbiome Project

1.1.1: Background

The Nobel laureate Joshua Lederberg was the first to define the concept of microbiome as “the ecological community of commensal, symbiotic, and pathogenic microorganisms that literally share our body space” [1]. Most of phylogenetic data and research focuses on bacteria, as they represent most of the microorganisms that inhabit our bodies. The number of resident microbes in the human body is estimated to be ten times that of the human somatic and germ cells, majority of which have not been identified or characterized yet [2].

Due to the fact that the majority of microbes inhabiting the different environments in the human body are uncharacterized, the National Institute of Health (NIH) established the Human Microbiome Project (HMP) in 2007. This project aims to identify the microbial communities that indigenously reside in different anatomical sites of the human body including the oral and nasal cavities, gastrointestinal tracts, genitourinary tract and the human skin [2] (Figure 1). One of the initiatives of the project is to establish the sequences of at least 900 bacterial genomes to later be utilized as a dataset of reference genomes. In addition to this objective, samples were collected from the main five body sites of 250 healthy volunteers: mouth, skin, vagina, the nasal cavity and the gastrointestinal tract. Using metagenomic approaches including 16S rRNA analysis and whole genome shotgun sequencing (WGS) an attempt was launched to characterize and describe the structure of the microbial communities that reside in these sites [2]. The main goal of the Human Microbiome Project is to determine and characterize the core microbiome at each site of the human body.

1.1.2: Goals of Human Microbiome Project

The first objective of the Human Microbiome Project aims to define the influences of age, lifestyle, gender, nutrition and medication on the diversity of the microbial community in the different sites of the human bodies (Figure 1). This work should help understanding the relationship between the resident microbial communities and their impact on human health [3]. The second objective is to create a reference dataset of the microbial genome sequences and characterize the human microbiota to determine their diversity and complexity. This
could be achieved by metagenomic analysis and also through sequencing of cultured microbial genomes. Sequencing techniques ranging from 16S rRNA and other deep metagenomic sequencing methods were applied. The vast information and studies collected are arranged precisely and organized by developing new computational tools [3]. In addition, HMP aims to design and implement new strategies for manipulation of microbiome in both healthy and diseased individuals.

1.2. Cultivation dependent and independent methods:

1.2.1: Limitations of the culture dependent and independent methods

Today, culture independent techniques are focused on discovering novel and more diverse microbes that reside in various sites of the human body. These works reveal slow growing, uncultivable, and rare bacteria. The most significant culture independent method to determine the community structure is the molecular approach based on sequencing of hypervariable regions from the small subunit of the ribosomal RNA gene. Although the cultivation dependent method is inexpensive and succeeds in identification of various bacteria, it requires specialized instruments for bacterial identification of anaerobic microorganisms, and suitable culture medium and specific growth condition for the rare and slow growing bacteria[4], [5], [6]. However culture independent approach (metagenomics) cannot substitute the culture-based methods since the phenotypic properties and microbial interactions with different human environments require isolation of viable microbes. However, molecular based methods succeeded in providing insight about microbial functional characteristics of slow and uncultured microbes [7], [8]. In addition, sequencing of a whole microbial genome using molecular based methods can be achieved [9].

1.2.2: Characterization of the human microbiome using culture dependent and independent methods

The differences between culture dependent and independent methods were evident in a study by Ian Kroes and colleges [10] where the bacterial diversity in the subgingival crevice was determined. The results obtained 56 bacterial isolates by cultivation methods. At the genus level, 37.5% of those isolates could not be identified, while 55% could not be identified at the species level. Sequencing of 16S rRNA from the 56 isolates showed 28 unique phylotypes. Sequences obtained by 16S rRNA amplification and cloning revealed
higher bacterial diversity than those obtained by cultivation. The uncharacterized sequences show less than 99% similarity to ARB database of the bacterial 16S rRNA sequences. From 2001 to 2007 about 215 novel bacterial species have been discovered and 29 are novel genera [4]. In a 5 year study conducted by Peter and colleges [11], they collected bacterial isolates from urine, faeces, cerebrospinal fluid, respiratory samples and genitourinary tract swabs. They were subjected to 16S rRNA sequencing and analyzed using NCBI Genbank database and IDNS (SmartGene Integrated Database Network System). Isolates giving sequence homology of less than 99% to the corresponding taxa were indicated as novel species and those with less than 95% of sequence homology were indicated as novel genera [4]. Out of a total of 1,663 isolates, 60 bacterial isolates were unknown 16S rRNA sequences giving less than 99% sequence homology on analysis and reported as novel species. After antibiotic therapy to the susceptible isolate and following the patient’s history and laboratory findings, 10 cases showed that isolates may play a pathogenic role and 8 out of these 10 were reported as infective agents. One of the novel bacterial species that was identified in the patients’ samples with urinary tract infections were Actinobaculum sp. and Gardnerella sp. During the study, the numbers of new 16S rRNA sequences in IDNS database increased by a factor of 4 while those in the NCBI Genebank database increased by a factor of 15.

1.2.3: Tools used for characterization of the human urogenital microbiome using culture independent methods

The techniques are based on analysis of the genetic material of the collected microbial community to describe the genomic composition, or sequencing of the 16S rRNA clones for bacterial phylogenetic analysis [12]. The most widely used molecular method is the sequencing of 16S rRNA clones for many reasons; 1) the approximate size of the 16S rRNA is 1.5 kb which is sufficient for sequence analysis, 2) the 16S rRNA gene encodes for a part of the small subunit of ribosome, which is highly conserved in nearly all prokaryotes required for accurate alignments, 3) the 16S rRNA contains nine hypervariable regions (V1-V9) which enables the identification of the diversity among different bacterial species due to sequence divergence along different evolutionary times and 4) pyrosequencing of 16 rRNA hypervariable regions provides in-depth taxonomic community profiling. Therefore, the sequencing of 16S rRNA has been applied to demonstrate taxonomical and phylogenetic relations among various species. The American biologist and physicist, Carl Woese, who made this technique accessible when he set a proposal to define the tree of life by dividing it
into three main domains: Archaea, Bacteria and Eucarya [13]. This tree of life relies on genetic diversity rather than on morphological similarities.

The hypervariable regions V2, V3 and V6 were the most appropriate targets for differentiation among various bacterial species while V4, V5, V7 and V8 were the least appropriate targets for species differentiation [14]. The three hypervariable regions succeeded in detection of 110 various bacterial species. PCR primers, or known as universal primers, bind to conserved regions of the 16S rRNA gene and amplify the hypervariable regions. The forward primer 63F and the reverse primer 1387R were used in a study aimed to evaluate the efficacy of both primers in identification of diverse bacterial species compared to the currently used PCR primers. They practically showed higher PCR specificity [15]. They were successfully able to amplify 16S rRNA genes of various organisms as a novel type of Proteobacteria (sulphate- and iron- reducing bacteria) and bacteria with high G+C rich sequences like organisms belonging to Micrococcus genera. However, the primer pair; 27F and 1392R failed to amplify those 16S rRNA genes [15].

1.3. Factors affecting the microbiome structure and diversity:

There are several unexplored and questionable issues concerns whether the host genetics, diet, lifestyle, physiological (e.g. pregnancy and menopause) and pathological factors (e.g. urinary tract infections, sexually transmitted diseases...etc) contribute to the microbiome structure and diversity of the genitourinary tract of both male and female individuals. In the following section, we present a literature review regarding these issues.

1.3.1: Female urogenital microbiome and probiotics:

In 1928, the concept of probiotics as beneficial group of bacteria such as lactobilli artificially administered to restore the unbalanced microflora [16] was first presented by Stanley Thomas. He performed in-vivo and in-vitro experiments, that resulted in eradication of Neisseria gonorrhoeae by the addition of a broth of cultured Lactobacillus acidophilus [17]. The urogenital infections, including urinary tract infections (UTIs), bacterial vaginosis (BV) and yeast vaginitis, represent a worldwide problem with over 300 million cases per year [16]. Recurrent infections, bacterial and yeast resistance, and side effects have been the major obstacles for the usage of antimicrobial agents in treating the urogenital infections [18], [19]. During urogenital infection, the flora is predominated by the infecting pathogen unlike healthy conditions where the urogenital flora is predominated by the indigenous organisms.
The incidence of BV was found to be decreased by four fold in case of vaginal and rectal colonization by *Lactobilli* [20]. The strain of *lactobacillus* and the route of administration influence the effectiveness of probiotics. *Lactobacillus crispatus* is one of the prevalent strains of *Lactobilli* in the vagina [20], so it is considered to be an optimal choice as a probiotic agent rather than other strains such as *L. fermentum* and *L. rhamnosus*. Both strains presented poor results as a probiotic agent in the treatment of urogenital infections [21]. In a study by G. Reid and colleges, they demonstrated that direct instillation of *L. rhamnosus* GR-1 and *L. fermentum* B-54 into the vagina reduce the recurrences of urogenital infections [22]. In 2001, it was firstly reported that probiotics could be administered orally with a dose of more than $10^8$ viable organisms per day [23]. In another study by G. Reid and colleges, they showed the first clinical evidence that the administration of *Lactobacillus rhamnosus* GR-1 and *Lactobacillus fermentum* RC-14 strains suspended in skimmed milk could be recovered from the rectum and finally reach the vagina [24]

### 1.3.2: Ethnic influences:

It has been reported that differences in ethnic groups have an impact on the urogenital microbiome. For instance, the vaginal microbiome of four different ethnic groups (white, asian, black, and hispanic) have revealed higher prevalence of *Lactobacillus* species in Asian and white women compared to black and hispanic women. The findings were supported by the median pH values of black and hispanic ethnic groups who revealed higher values in comparison to the other two ethnicities [25]. There is no well supported explanation for the differences seen among the ethnic groups. It might be due to genetic differences between the hosts or immunological differences or a potential outcome of the vaginal secretions and composition [25]. The data of this study was highly similar to those of Zouh and colleges (2010) where they characterized the vaginal microbiome of three different ethnic groups; Japanese, white and black healthy women [26].

In another study by Zouh and colleges (2007), the phylogenetic analysis and profiles of Terminal restriction fragment length polymorphism (T-RFLPs) of 16S rRNA genes has proved the significant differences in vaginal bacterial communities of 144 black and caucasian women from North America [27]. Black women showed higher incidence of vaginal bacterial communities not dominated by *Lactobilli* such as *Atopobium* species and *Clostridiales* order compared to Caucasian women. Caucasian women were significantly dominated by diverse species of *Lactobacillus* as *L.crispatus, L. jensenii* and *L. gasseri*. The
differences in ethnic group from the same geographical origin was found to be highly correlated in determining what strain of *Lactobacillus* to be dominant in the vaginal milieu [28]. *L. crispatus* was the dominant species in the healthy urogenital microflora in most populations [29], [30]. Those differences in vaginal microflora between various ethnic groups may account for their susceptibility to BV and sexually transmitted diseases (STDs) [27].

**1.3.3: Immunological influences:**

The defect of the host immunological response was found to contribute to the host susceptibility to urinary tract infections. Pirrka and colleges [31] studied the effect of immunological abnormalities on the susceptibility of women to recurrent urinary tract infections (UTIs). In this study, they collected urine, blood and vaginal samples from 22 subjects prone to recurrent UTIs in their disease free period and compared them to control subjects. They observed the enhanced production of IL-12 by antigen presenting cells (APCs) which are myeloid dendritic cells and peripheral monocytes. However, this was not translated into T-cell type-1 activation which is known to be responsible for cell mediated immune response by defence against uropathogens and antibody production [32]. Those findings were supported by the improper induction of CDC25, which is the T cell activation marker, compared to control subjects. This suggests that defective translation of APCs into T cell activation and differentiation into Th1 memory cells might contribute to increased susceptibility to recurrent UTI.

**1.4. Female urinary system**

**1.4.1: Anatomy of the female urinary system:**

The urinary system consists of two kidneys, two ureters, a bladder and a urethra. Several reasons account for discussing the anatomy of the male and female urinary systems separately. First, there are differences in the anatomy and physiology of both systems. Second, the urethra is a part of the reproductive system of male which differs in function from the female urethra. Finally, the anus is closer to the urethra in female than in male and the urethra is also close to the vaginal introitus. Both the anus and the vaginal introitus are heavily colonised with bacterial flora resulting in differences not only in the number and types of microbiota but also in susceptibility to infections between male and female urinary systems [33].
1.4.2: Antimicrobial defences of the urinary tract:

The antimicrobial defensive mechanisms of the urinary system play a major role in preventing the viable bacteria entering the lower urinary tract from colonization and then ascending to its upper part against urinary flow [34]. The flushing action of urine serves in preventing colonization of microorganisms in the urethra. Consequently, this creates a high pressure zone preventing the microbial ascent from the urethra to the upper urinary tract. Several studies demonstrated that there is a relationship between the urine flow rate and the number of bacterial population [35], [36, 37]. In a study by O’Grady and Cattell, it was observed that at the normal physiological state of urine flow rate, the bacterial population reaches the steady state of growth. However, at slower rate as overnight sleep or in cases of mild dehydration, there was a considerable increase in the number of the bacterial count [36]. Additional factors like urine pH, urine osmolality and urea concentration act as natural antibacterial activity. The urine pH of a healthy individuals ranges from 5.0 to 6.0. The acidity of urine is often inhibitory to bacterial growth but in some instances it can be bactericidal. However, the urea concentration shows greater influence as a defence mechanism against bacteria than high osmolality [38], [39]. Decreasing urine osmolality but preserving urea concentration as in dialysis will not affect the bacteriostatic activity of urine. However, when decreasing the urea concentration and maintaining the urine osmolality, this will result in reduction of the antibacterial activity [38]. Moreover, the osmotic shock can be avoided by some uropathogenic bacteria as in case of *E. coli* [40].

Shedding of the outermost urothelial cells with their adherent uropathogenic bacteria is considered to be one of the main antibacterial defence features of the urethra. The exfoliation rate depends to a great extent on the hormonal status of the female. Several urethral exfoliating cells are abundant in menstruating females and postmenopausal women on hormonal replacement therapy. On contrary, pre-menstruating girls and postmenopausal women not on HRT have fewer exfoliating cells [41]. Another defence mechanism is the mucous secretions of the paraurethral or Skene’s glands which were described in 1948 in a study by Huffman and colleges [42]. In another study by Hutch and colleges, they theorized that the paraurethral glands located at the lower end of the urethra produce mucous forming a layer at the epithelial lining of the urethra. Therefore, this will restrain bacteria from ascent into the upper urinary tract [43].
Several antimicrobial proteins are released into the epithelial lining of the urethra such as Tamm-Horsfall protein (THP), lactoferrin, and lipocalin. Tamm-Horsfall protein, Uromodulin, is a glycoprotein secreted by the kidney and is used in preventing Enterobacteria from adhering to the epithelia by interfering with their fimbrial structures and binding to adhesin receptors of Escherichia coli [44]. Lactoferrin was found to be abundant at extremely low levels in urine [45]. Its bactericidal activity mainly depends on iron sequestration, thus reducing the amount of free iron present in urine and reabsorbing it for later metabolism [34]. Free iron in urine acts as a fertile soil for Gram negative bacteria since sufficient concentration of iron is required for their nutritional need [45]. Moreover, antimicrobial peptides (AMPs), such as defensins and cathelicidin, are constitutively secreted showing bactericidal activity. In case of bacterial infection, AMPs will be inducibly secreted provoking the innate immune system through the activity of Toll like receptors, cytokines and chemokines. These receptors will attract more neutrophils and macrophages to the site of infection, resulting in phagocytosis of the microbes and release of more antimicrobial proteins and peptides [34]. However, the role of defensins is still not evident; they may act as an anti-infective agent as suggested by in-vitro studies [46] or serve other purposes. Strong evidence supports the role of beta defensins, anti-microbial agents, as demonstrated in a study by Morrison G. and colleges [47]. Using homologous recombination technology, they deleted the Defb1 gene, an analogous gene to HBD1 constitutively expressed in the collecting and distal tubules of the kidney [48], occurred to produce genetically engineered mice Defb1−/− mutants. The data presented 30% of the healthy mutants had Staphylococcus species in their urine compared to the wild type mutants.

The urinary secretion of immunoglobulins, e.g., IgA and IgG, is one of the antimicrobial defence mechanisms of the urinary tract [34]. Several studies have found correlation between the susceptibility to recurrent UTIs and the secretor status of the individual [49], [50]. Non secretors are individuals who secrete a little or none of their blood group antigens into the body fluids [51]. The blood group antigens are carbohydrate determinants located on the surface of erythrocytes, phagocytes, lymphocytes and urothelium onto which the microbial lectins attach [51], [52]. The blood group antigens are secreted into body fluids, thus serving as an attachment site of the fimbrial lectins of uropathogenic bacteria specifically Escherichia coli. Hence, this protects the individual from colonization by uropathogenic bacteria. The non secretors were observed to be more prone to recurrent UTIs when compared to the secretors since the composition of the carbohydrate residue is
altered in case of the secretors. In a study by Stapleton and colleges, the reason for increased adherence of *E. coli* to the urothelium of non secretors was explained. They radiolabelled an *E. coli* strain that had been isolated from a patient with UTI and bonded it to two extended globoseries GSLs, sialosyl gal-globoside (SGG) and disialosyl gal-globoside (DSGG). The two globoseries GSLs are selectively expressed on vaginal epithelial tissue in non secretors but not in secretors. It was concluded that sialylation of glycolipids occur in non secretors, resulting in higher affinity to the adherence of uropathogenic bacteria. However, the glycolipids are fucosylated in secretors and processed to blood group antigens [53].

1.4.3: Female urinary system and microbiome:

1.4.3.1: Establishment of microbiota:

Upon delivery, the healthy infant born from a sterile or bacteria free environment is exposed to a wide diversity of microbes after passage through the birth canal [28], [54]. This microbial community, some provided by the mother, determines the microbial inhabitants of the different anatomical sites of the healthy adult individual body. The delivery mode significantly influences the microbial community received by a newborn from its mother [55]. It was observed that vaginally delivered babies acquire microbiota dominated by *Lactobacillus, Prevotella, or Sneathia* spp., resembling most of their mothers’ vaginal microbiota. Moreover, C-section babies acquired microbiota dominated by *Staphylococcus, Corynebacterium, and Propionibacterium* spp., resembling those of the skin surface [55]. Therefore, the delivery mode affects the differentiation of microbiota in various habitats of the human body and consequently the disease predisposition.

1.4.3.2: The effect of age and hormonal status on the female microbiome:

The microbiome is influenced to a great extent by age and hormonal status of the female. Bollgren and colleges [56] observed heavy colonization of the periurethral region of healthy female infant by aerobic flora (*Escherichia coli, enterococci and staphylococci*) during the first weeks of life and this starts to diminish till the age of five. Healthy school children manifested considerably few Gram negative bacteria in their urine samples suggesting the maturation of a defensive mechanism against facultative anaerobe such as, *E. coli* and other bacteria like *Enterococci*, start to develop from the individual’s early years. The relationship between the glycogen content of mucosal cells and the hormonal influences creates a microenvironment that is more diverse and quite different from that of pre-
menarchal girls [57], [58]. The microbial community of the urethra of pre-menarchal and reproductive women is highly dominated by facultative anaerobe, whereas in post-menopausal women is roughly dominated by obligate anaerobes. Facultative gram negative bacilli are more frequent in the urethral microbiome of post-menopausal women, however absent in pre-menarchal and reproductive women [57], [59]. Contradictions aroused between studies that discussed the dominance of lactobacillus of post-menopausal women microbiome [59]. The diversity of microbial communities of the female urogenital flora is a direct consequence of the hormonal influences and age [56], [59].

1.4.3.3: The key members of the female urogenital tract microbiome:

The microbiotas were detected along the whole length of the female urethra; however they differ in their proportion densities. The prevalent microbiotas that indigenously colonize the female urogenital tract are Lactobacillus spp. Firmicutes phylum with Bacteriodetes, Actinobacteria, Fusobacteria, Proteobacteria and Corynebacterium spp.[60], [59], [54].

1.4.3.4: Identification of Lactobacillus spp. and its role for the female urogenital health:

Lactobacillus spp. was firstly discovered by Albert Doderlein in 1982 [61]. He observed that the in-vitro and in-vivo growth of pathogens was inhibited by a lactic acid producing organisms. In 1928 [17], it was identified as Lactobacillus acidophilus by Stanley Thomas and later in 1980 [62]; it was found that Lactobacillus acidophilus represents a group of organisms that were highly heterogeneous. The acidic vaginal milieu created by Lactobacillus spp. due to production of lactic acid, inhibits the growth of infective pathogens and bacteria predisposing to bacterial vaginosis [63]. Hydrogen peroxide (H2O2) [64], bacteriocins [65] and probiotics [22], [24] are also produced by Lactobillus which cause further reduction in acquisition of vaginal infections. From a total of 120 different species of Lactobacillus, only 20 species of Lactobacillus have been identified in the urogenital microflora. Only one or two of Lactobillus from the four dominant Lactobacillus species, L. crispatus, L. iners, L. jensenii and Lactobacillus gasseri, are commonly detected in the healthy female urogenital microflora [66]. However, other species rarely inhabit the vaginal microflora and tends to be novel phylotypes [67], [29]. This may correlate with the host or competitive factors that support the growth of one species over another or pre-emptive colonization by a particular species [30].
1.4.3.5: Other lactic acid producing bacteria and their role for a healthy urogenital tract

A significant proportion, 7-33%, of females reflected healthy urogenital state despite the absence or the occurrence of low numbers of \textit{Lactobacillus} spp. Several studies demonstrated that the production of lactic acid and the maintenance of the vaginal acidity by other bacteria such as \textit{Atopobium vaginae}, \textit{Megasphaera}, and \textit{Leptotrichia} species were highly correlated with the inhospitality of the urogenital milieu to infectious pathogens [27], [67], [30]. In several investigations, urine and vaginal samples were collected from healthy females with no symptoms of urogenital infections or any other clinical problems. They were either highly dominated by \textit{Lactobacillus} spp. or had low numbers of \textit{Lactobacillus} spp. and exhibited rich microbial communities composing of \textit{Atopobium}, \textit{Gardnerella}, \textit{Prevotella}, \textit{Dialister} and \textit{Anaerococcus} [54]. Consequently, it was concluded that the prevalence of particular strains of bacteria such as \textit{Gardnerella vaginalis} does not represent an abnormal state. This issue was investigated in several studies in which it was noted that \textit{Gardnerella vaginalis} was not detected in 70% of healthy female subjects without BV [68]. Moreover, the concentrations of \textit{Gardnerella vaginalis} in healthy subjects with no evidence of infection were detected to be increased during menstruation and decreased after menses. The opposite was the levels of \textit{Lactobacillus} spp. which obviously decrease during menses except for \textit{L. iners} [69].

1.5. Male urinary system

1.5.1: Anatomy of the male urinary system:

The anatomy of the male urethra differs from the female urinary system since it constitutes a part of the male urinary system as well as its reproductive system. Therefore, it is responsible for excretion of urine and ejaculation of semen. The urethra consists of three main regions: the prostatic, the penile and the membranous urethra. The male urethra is five times longer than that of the female. This elucidates the cause of the frequent UTIs by ascending pathogens in the females compared to the males [33].

1.5.2: Male urinary system and microbiome:

1.5.2.1: Antimicrobial defences of the urinary tract:

The majority of the antimicrobial defense mechanisms in the female urinary system resemble those of the male, including urinary flow, epithelial desquamation and the release of antimicrobial peptides and proteins. The high zinc concentration in the prostatic tissue,
compared to other human tissue [70] and its secretion into prostatic fluid, represents one of the main antibacterial defense mechanisms of the male genitourinary tract [71]. Several compounds, such as lysozyme, lactoferrin [71], phospholipase A2 and secretory leukocyte protease inhibitor (SLPI), have been identified in the prostatic fluid and were observed to protect the male genitourinary tract from infectious agents. When lysozyme concentrations in healthy male individuals were compared to patients with chronic prostatitis, the concentrations were noticeably higher. Patients that were investigated with an initial remarkably reduced concentration of lysozyme, experienced recurrences after antibiotic treatment [72]. SLPI is a proteinase inhibitor secreted by the epithelial lining of the seminal fluid and also known as human seminal inhibitor I (HUSI-I) or anti-leukoprotease (ALP) [73]. The first NH2 terminal domain of SLPI showed anti bacterial activity against *Escherichia coli* and *Staphylococcus aureus* in vitro, but lesser activity when compared to lysozyme and defensins [75]. Phospholipase A2 (GIIA PLA2) shows potent antibacterial activity against Gram-positive bacteria [76], its concentration in the human seminal plasma is high, approximately 1000 times compared to the human blood plasma [77]. The positive charge on PLA2 attacks the peptidoglycan layer of the bacterial cell wall, facilitating penetration of the phosphobilipid layer [76].

In an immunohistochemical study by Jeffrey Pudney and Deborah J. Anderson [78], it was demonstrated that the uroepithelial lining of the male penile urethra expresses the secretory component of the polymeric IgA, IgM and to lesser extent IgG transport molecules. The immunoglobulins represent the first line of immune defense against bacterial pathogens. The glands of Littre’ secrete a mucoidal layer containing the secretory IgA abundant in the epithelium lining the fossa navicularis and meatus. This shows similarity to the mucous secretions of the cervical gland of the female genital system which is rich in secretory immunoglobulins and acts as an immune barrier [78], [79]. The seminal fluid contains high levels of IgG and IgA and low levels of IgM [80].

**1.5.2.2: The effect of age and circumcision on the male microbiome:**

The relationship between age and the establishment of microbiome in healthy boys is correspondent to that previously mentioned in healthy girls [57]. Circumcision was found to be an additional factor that influences the establishment of microbiome and later the pattern of bacteria causing UTIs during the first years of life. Definite types of uropathogens such as *Klebsiella-enterobacter, Proteus mirabilis, and Pseudomonas aeruginosa* was observed to be
significantly higher in the urethra of uncircumcised boys. Moreover, there was a higher colony count for Gram negative uropathogens and *E. coli* isolated from the urethra of uncircumcised boys [74], [75]. This supports the idea that there is a direct correlation between periurethral bacterial colonization and pathogenesis with urinary tract infections.

1.5.2.3: The key members of the male urogenital tract microbiome:

The male urethra is approximately 20 cm in length, where various microbial communities are expected to exist in differently in various regions. It was found that the last 6 cm of the male urethral orifice is solely colonized by highly complex and variable strains of bacterial communities. Travelling further into the male urethra, the bacterial concentration and their complexity becomes less and even undetectable [76]. Several factors result in the differences in complexity and diversity of microbial strains between various individuals: sexual activity, age, gender, level of hygiene and circumcision. Several investigations observed that the indigenous microbiota of sexually active individuals were significantly more complex than that of non-sexually active individuals [77], [78]. Moreover, several organisms isolated from the vagina and cervixes of females were abundant in sexually active males. However, the conditions, specifically the pH, found in the penile urethra differ from those of the female vagina. Therefore, the isolation of bacterial species, such as *Lactobacillus* spp. that predominate the vaginal flora of healthy female, were not increased in sexually active men [79]. The most detectable microbiotas in the urethral orifice were *Staphylococcus*, *Streptococcus*, *Corynebacterium*, *Pseudomonas* and *Enterococcus* species [76], [80]. The microbiotas isolated from urine samples of older men might be less representative of the penile urethra due to the higher incidences of renal and prostatic disorders [84].

Two specific aims will be addressed in this thesis work. First, this work will focus on the application of cultivation independent approach, 16S rRNA analysis, in order to define the structure and diversity of the urinary tract microbiomes investigated in four healthy and culture negative male and female individuals. The results showed diverse bacterial profiles in both genders. Second, 16S rRNA gene libraries were constructed and 16S rRNA phylogenies of two different hypervariable regions, V1-V3 and V6-V8, were done. Comparing the results of the two regions demonstrated wider range of taxonomical classification of V1-V3 sequence dataset. Therefore, this thesis provides comparable taxonomical classification of bacterial genera diversity between male and female individuals.
CHAPTER 2: MATERIALS AND METHODS

2.1 Subjects and sample collection

Mid-stream urine samples, 100 ml, were collected from eight healthy adult volunteers, 4 male and 4 female individuals, in a sterile collection cup. All male subjects were circumcised. None of the female subjects reported taking vaginal douches, genital wipes and/or vaginal medications. They were all regularly menstruating (25-35 days menstrual cycles), at reproductive age and with no history of sexual experience. Urine samples were collected from female individuals who were not menstruating. All the subjects were provided with a written informed consent which was approved by the institutional review board (IRB) of the American university in Cairo (Appendix 1). The subjects’ age range between 18-40 years old. All enrolled subjects were asked to fill a written medical history form to make sure of the overall general health (Appendix 2). The urine samples of all subjects have been tested for any urinary tract infections. Urine was also tested for sexually transmitted pathogens as Chlamydia trachomatis and Neisseria gonorrhoeae [81]. None of the involved subjects in the study has reported taking an antibiotic for at least the last three months. All the previously mentioned tests in addition to the urine culture and sensitivity tests for the infected pathogens were carried out in El-Borg clinical laboratory. El-Borg is the first medical laboratory in Egypt accredited by the Swedish Board for Accreditation and Conformity Assessment (SWEDAC).

2.2 DNA isolation

Urine samples were centrifuged at 4,000 xg for 15 minutes at 4°C. The supernatant was discarded and the cell pellet was stored at -80°C until DNA isolation [82]. DNA was extracted from the cell pellet by using QIAmp DNA blood mini kit (QIAGEN, Valencia, CA), according to the manufacturer’s instructions. Genomic DNA was resuspended in 30 ml of buffer AE and stored at -20°C until 16S rRNA PCR amplification and sequencing analyses.

2.3 PCR amplification 16S rRNA gene

All PCR reactions were performed using thermocycler (Applied Biosystems GeneAmp PCR 9700). The two universal primers used for PCR amplification of 1,300 bp of the consensus 16S rRNA gene were forward primer 63F (5’-CAGGCC TAA CAC ATG
CAA GTC-3’) and reverse primer 1387R (5’-GGG CGG WGT GTA CAA GGC-3’) [15]. The PCR approach was performed in a total volume of 25µL containing 50 ng of the template DNA (3µL), 100 mM of dNTP, 20 pmol of each primer, 10X DreamTaq reaction buffer and 5U / µL of DreamTaq DNA polymerase (Fermentas, Burlington, CA). The reaction conditions were initial DNA denaturation at 94°C for 10 min followed by 25 cycles of denaturation at 94°C for 40 sec, annealing at 56°C for 40 sec and elongation at 72°C for 1 min, and a final extension at 72°C for 7 min. To confirm the presence of amplicon, it was electrophoresed on 1% agarose gel, stained with ethidium bromide and visualized using short wavelength ultraviolet light.

2.4 Cloning, 16S rRNA gene library construction, and Sequencing

PCR products were purified using QIAquick PCR Purification Kit (QIAGEN, Valencia, CA) according to the manufacturer’s instructions. The purified DNA was cloned into pGEM®-T Easy Vector (Promega, USA) using a vector/insert ratio of 1:3 and performed according to the manufacturer’s instructions. The ligation mixtures were transformed using high efficiency competent E.coli TOP 10 cells using Micropulser electroporation device (Bio-Rad, Hercules, CA) according to the manufacturer’s instructions and plated on Luria-Bertani agar plates containing 100 µg/ml ampicillin followed by incubation overnight at 37°C. Approximately 100 white colonies were selected from each library and inoculated into 96-well microtitre plate containing 150 µL of Luria-Bertani broth supplemented with 100 µg/ml ampicillin followed by incubation overnight at 37°C. 15% glycerol stock was prepared for further use and stored at -80°C. Colony PCR was performed to confirm the 1.3 kb insert size. Bacterial cultures were reinoculated into 96-well block containing 1300µL of Luria-Bertani broth (supplemented with 100 µg/ml ampicillin) followed by incubation overnight at 37°C with shaking. Plasmid DNAs were isolated using R.E.A.L. Prep 96 Plasmid Kit (QIAGEN, Valencia, CA) and stored at -20°C. The isolated plasmids with their cloned inserts were sequenced from both ends with universal primers equivalent to the plasmid vector sequences M13F 5’- TGTAAAACGACGGCCAGT- 3’ and M13R 5’- CAGGAAACAGCTATGACC- 3’. Cycle sequencing was performed using Applied Biosystems Big Dye terminator chemistry and the sequences were resolved on the 96-capillary Applied Biosystems 3730XL DNA Analyzer.
2.5 Male and Female 16S rRNA urine sequence analyses

2.5.1: Alpha diversity analysis:

*Mothur* v.1.18.1 pipeline’s alpha diversity analyses package [83] was used for estimating the richness, evenness and overall diversity of the bacterial communities of male and female urinary tract. Plotting rarefaction curves and calculating the three non-parametric estimators of alpha diversity (Chao1, Shannon and Simpson indices) were executed to measure the overall coverage of the tested urine samples.

2.5.2: Phylogenetic analysis of 16S rRNA urine sequences:

Sequences were processed by trimming the plasmid vector bases with cross match and discarding low quality sequences < 300 bp in length and/or had average PHRED quality score of Q= < 20 using “Codoncode Aligner” (Codoncode Corporation, Dedham, MA). The 3’ and 5’ reads could not be assembled and therefore the phylogenetic approach was based upon separate analysis for the 3’ and 5’ reads of both male and female samples. BLASTN of the 16S rRNA sequences generated from libraries of both male and female urine samples were conducted against RDP II database [83]. The database is already installed in our local server maintaining the sequences with the best two blast hits and E-value threshold 10-5. On each of the four sequence sets “5’ males, 5’ females, 3’ males, 3’ females” chimera detection was performed, that can arise during PCR.

The infernal 1.0 software [84] was used to create a single multiple sequence alignments (MSA) of all 16S rRNA sequences of the male and female urine samples. It is a fast sensitive tool that builds a consensus RNA secondary structure profiles called covariance models where MSA of each final set together with reference reads already identified in previous blastn search on RDP database was done. Therefore, two alignments were created; the 3’ sequence alignment (including reads ending with the reverse complement of the 1387R primer sequence) and the 5’ sequence alignment (including reads ending with the 63F primer sequence). For purposes of phylogenetic deduction, further refinement to remove uncertainty in MSA was accomplished using ZORRO masking program [85] which assigns confidence scores to each alignment column ranging from 0-1, where 1 represents the most reliable match and 0 represents the ambiguous match. This program was applied in order to minimize errors resulting from manual trimming of MSA columns for the removal of poorly aligned columns and highly divergent alignments. For taxonomy assignments, a maximum likelihood
phylogenetic tree of the refined alignment was generated using phyML [86]. Several taxonomic assignments that have bootstrap values of < 70% at the species level were not considered and removed as most investigators agree that the phylogenetic clade is strongly supported at bootstrap value of > 70% [87].
CHAPTER 3: RESULTS

3.1 Processing of the urine samples, PCR amplification, library construction, and sequencing of the 16S rRNA gene

Multiple reports have investigated that either urine or urethral swabs can be used similarly in the detection of urethral microbiomes of males [81]. In this study, urine samples were used as a representative for sampling in males since urethral swabs cause discomfort particularly in case of multiple sampling.

After sample collection and DNA extraction, agarose gel electrophoreses of extracted DNAs show high molecular weight with good quality, figures (2 and 3). Using universal primers, 63F and 1387R, the entire 16S rRNA gene was amplified. Figures 4 and 5 showed that the amplified DNA fragments for the four males and females have the expected size of around 1,300 bp. Amplified DNAs were ligated into pGEM®-T Easy Vector (Promega, USA) and used to transform competent E. coli cells. For each individual transformant sample, library of 96 random individual colonies was grown in 96-well plates. Recombinant plasmids from each of the eight plates-representing four males and four females-were sequenced from both ends of the inserts (here called 3’ and 5’ ends) using the Sanger chain termination approach and analyzed on ABI 3730XL DNA analyzer. Tables (1 and 2) show a summary of the sequenced samples. The average length of sequenced DNA was 478 bp with at least 20 of PHRED quality score. The number of reads for each individual (forward plus reverse) varied from a minimum of 76 (male 3) to a maximum 157 (female 3). It is important to note that our inserts have size of around 1300 bp, and the maximum read length obtained from the 5’ and the 3’ of the inserts did not permit overlapping of the sequences from both ends of the recombinant plasmids as shown in Table 3. Therefore, the phylogenetic analysis was based upon separate analysis for the 3’ and 5’ reads of both male and female samples. Table (3) summarizes the number of sequenced clones for each library, length of sequence obtained from each end and quality of reads.

3.2 Computational and statistical analysis

To identify microbial communities that coexist in the urinary tract of male and female individuals, the 16S rRNA sequences were classified using the RDP II database [88] keeping
the E-value cutoff $10^{-5}$. For taxonomy assignments, the phylogenetic tree was constructed at bootstrap values of $>70\%$ at the species level.

### 3.2.1: Phylogenetic analysis of the female 16S rRNA sequences

The core microbiome of the female urinary tract was better characterized by sorting the 16S rRNA sequences at the genus level. The phylogenetic analysis of the two different regions of the 16S rRNA amplicon revealed diverse bacterial genera (Figure 6 and 7). Out of a total of 15 bacterial genera identified in all female urine specimens, 4 different genera were detected in both 3’ and 5’ reads. The 3’ reads of V6-V8 regions of the urine 16S rRNA sequences were approximately 50% dominated by bacterial phylotypes that were phylogenetically related to *Prevotella*, *Corynebacterium*, *Gardnerella* and *Weeksella* spp. The 16S rRNA sequences belonging to *Lactobacillus*, *Unclassified Flavobacteriaceae family*, *Sneathia*, *Atopobium* and *Oligella* spp were less abundant while the remaining genera represent approximately less than 1%. Likewise, The 5’ reads of V1-V3 regions of the urine 16S rRNA sequences were approximately 50% consisting of the same bacterial phylotypes that were previously identified in the V6-V8 regions. *Facklamia* spp. was the only bacterial genus being identified in the V1-V3 reads but not in the V6-V8 reads. Most reads for the female subject1 detected bacterial phylotypes belonging to *Gardnerella* and uncultured taxa. Coincidence was revealed in case of *Atopobium vaginae* and *Megasphaera* spp, each one being present only when the other one being found.

At the order level, the results of phylogenetic analysis data of the female urine 16S rRNA sequences identified 9 different orders in both sequence datasets, V1-V3 and V6-V8. The most abundant orders of the 16S rRNA sequences were *Actinomycetales*, *Bifidobacteriales*, *Flavobacteriales* and *Bacteriodales*. *Actinomycetales* and *Flavobacteriales* were present in 3 / 4 urine samples. In contrast, *Bifidobacteriales* was only detected in 1 / 4 urine samples however it accounted for significant proportion of sequences. All bacterial genera and orders being identified in the female urine flora are summarized in Table 1.

*Actinomycetales* comprised approximately 40 sequences and contains 3 genera, mainly represented by the genus *Corynebacterium* and to a lesser extent, the genera *Arthrobacter* (*A. albus* and *A. cumminsii*) and *Zimmermannella* (*Pseudoclavibacter timone* and uncultured taxa). *Corynebacterium* spp contain diverse species (*C. coyleae*, *C.
of *Candida tuscaniense*, *C. riegelii* and uncultured taxa) and has been identified as a member of the superficial skin flora [89]. *Flavobacteriales* comprised approximately 40% of sequences and it contains 3 genera, (*Planobacterium*, *Weeksellaria* and uncultured taxa), represented at various relative abundances. Unclassified *Flavobacteriaceae* family was represented by a single clade containing high numbers of clones assigned to uncultured bacteria (at > 90% 16S rRNA sequences similarity). Closest RDP matches of all *Weeksellaria* sequences, represented as one of the dominant genera of the female urinary microbiome, were assigned to *W. virosa* spp. Numerous clones from female subjects 2 and 4 of the V1-V3 reads were taxonomically related (at different percent similarity, mostly greater than 70%) to 16S rRNA sequences of *Weeksellaria virosa*. While numerous clones representing the female subjects 1 and 3 showed 100% similarity to the 16S rRNA sequences of V6-V8 reads of *Weeksellaria virosa*. Closest RDP matches of all *Bacteriodales* and *Bifidobacteriodales* sequences were assigned to the genera *Prevotella* and *Gardnerella* respectively. The genus *Prevotella* contains diverse species (*P. dientes*, *P. buccalis*, *P. timonensis*, *P. amnii* and mainly uncultured taxa) while the genus *Gardnerella* was mostly assigned as *G. vaginalis* and both comprised approximately half of the female urinary 16S rRNA sequences. *Lactobacillus* and *Facklamia* spp. (only detected in female urine 16S rRNA sequences of V6-V8 reads) were the main representative genera of the order *Lactobacillales*.

Other bacterial orders: In addition to bacterial phylotypes that were previously discussed, 4 orders representing 6 bacterial genera were abundant at low proportions. The bacterial genera represent approximately less than 1% of the female urine microbiome. For instance, the *Coriobacteriales* (genus *Atopobium*), the *Selenomondales* (including the following genera *Megasphera* and *Dialister* spp.) and the *Fusobacteriales* (genus *Sneathia*) were known to be directly correlated with the female urogenital flora [25], [54], [30]. It should be noted that those bacterial genera were entirely abundant in female subject 3. The abundance of diverse bacterial communities in female subject 3 might be attributed to the high number of clones recovered allowing examination of additional bacterial phylotypes in the female urine flora. The most dominant bacterial phylotype of the *Burkholderiales* order, the genus *Oligella* (mostly assigned to *O. urethralis*), coexists in female urine subjects 2 and 4. *Oligella urethralis* was previously classified as *Moraxella urethralis* [90] which is a small aerobic and Gram negative bacterium usually isolated from human urine.
At the phylum level, a total of 5 phyla have been identified. The most frequently detected sequences corresponding to 4 bacterial phyla were *Firmicutes, Actinobacteria, Fusobacteria* and *Bacteroidetes* while *Proteobacteria* phylum was less frequently observed.

### 3.2.2: Phylogenetic analysis of the male 16S rRNA sequences

The 16S rRNA sequences of the male urine samples were similarly sorted at the genus level to characterize the core microbiome of the male urinary tract. Likewise, the phylogenetic analysis of the two different regions of the 16S rRNA amplicons revealed various bacterial genera (Figure 8 and 9). Out of 39 bacterial genera being identified in all male urine samples, 19 different genera were solely detected in both reads. The 3’ reads of V6-V8 regions of the urine 16S rRNA sequences were approximately 50% dominated by bacterial phylotypes that were phylogenetically related to *Prevotella, Escherichia/Shigella, lactobacillus* and *Corynebacterium*. 16S rRNA sequences belonging to *Bacteroides, Fecalibacterium, Acinetobacter, Anaerococcus*, unclassified *Lachnospiraceae* and *Flavobacteriaceae* families were less abundant while the other genera represents considerable low proportions of the whole bacterial community. The 5’ reads of V1-V3 regions of the urine 16S rRNA sequences were approximately 50% consisting of the same bacterial phylotypes that were previously identified in the V6-V8 regions. Other microbial communities were only observed in the V1-V3 reads; *Peptoniphilus, Ruminococcus, Coprococcus, Sphingomonas, Alistipes, Empedobacter, Sphingomonas, pseudoxanthomonas, Kocuria* and *Cupriavidus spp.* However, *Arthrobacter, Negativicoccus, Salmonella, Weeksella* and *Veilonella spp* were only identified in the V6-V8 reads. Unlike the 16S rRNA sequences of the female urine specimens, uniqueness of a single bacterial genus to a specific male subject was not observed. Distinct differences between kinds of bacterial communities of the urinary tract of the female and male individuals were significantly observed as shown in Figures 10 and 11.

At the order level, the two sequence datasets revealed a total of different 17 orders being identified. The most abundant orders of the 16S rRNA sequences were *Actinomycetales, Lactobacillales, Enterobacteriales, Bacteroidales* and *Clostridiales*. *Sphingomonadales, xanthomonadales* and *Burkholderiales* were only identified in the V6-V8 reads while virtually no order was specifically identified in the V6-V8 reads. All bacterial genera and orders being identified in the male urine flora were summarized in Table 2.
Actinomycetales comprised approximately 50 sequences and contains 3 genera, mainly represented by the genus Corynebacterium and to a lesser extent, the genera Arthrobacter (A. albus and A. cumminsii) and Kocuria. Corynebacterium spp contains diverse species (C. riegelii and mostly uncultured taxa). Flavobacteriales comprised nearly 25 sequences and contains 3 genera, (Flavobacterium and the genus Weeksella that was recovered from the 16S rRNA sequences of V6-V8 reads of male subjects 3 and 4). The Flavobacteriales order was represented at various relative abundance rates. Unclassified Flavobacteriaceae family was represented by a single clade containing high numbers of clones mainly assigned to uncultured bacteria (at > 90% 16S rRNA sequences similarity).

Similarly, the closest RDP matches of all Weeksella sequences were assigned to W. virosa spp. Unlike the female urine 16S rRNA sequence similarity, the Bacteroidales order is represented by 5 different genera, Prevotella, Bacteroides, Porphyromonas, and the genera Alistipes and Barnesiella that were detected only in V1-V3 reads, constituting approximately 25% of 16S rRNA sequences of the phylogenetic tree. The genus Prevotella contains diverse species (P. melaninogenica which was recently classified as Bacteroides melaninogenicus [91], B. buccalis and mostly uncultured taxa). Bacteroides spp. is a host member of the vaginal and the gastrointestinal tract flora [92]. Enterobacteriales contains 3 genera, mainly represented by the genus Escherichia/Shigella. However, the remaining genera (Enterobacter aerogenes and Salmonella) accounted for a considerable small proportion of the Enterobacteriales order where they have been identified as members of the gastrointestinal flora. The genus Escherichia/Shigella accounted for significant proportions of male urine sample 1 indicating a possible urinary tract infection. The Clostridiales order comprises nearly 30 sequences of 16S rRNA sequences of V1-V3 dataset and contains more diverse bacterial phylotypes than those of V6-V8 sequence dataset. Unclassified Lachnospiraceae family was represented by several clades containing high numbers of clones assigned to uncultured bacteria (at > 97% 16S rRNA sequences similarity). Closest RDP matches of all Caulobacteriales and Pseudomonadales sequences were assigned to the genera Acinetobacter and Brevundimonas respectively. High number of clones of V1-V3 and V6-V8 datasets were assigned to Acinetobacter spp which is an environmentally widespread aerobic Gram-negative bacterium predominately found in moist regions of the skin [33].

In addition to the previously mentioned bacterial phylotypes, other 9 orders representing 12 bacterial genera were represented by approximately less than 1% of the female urine microbiome. For instance, bacterial communities of the Selenomondales order
(including the following genera *Phascolarctobacterium, Veillonella* and *Negativicoccus succinicivorans* spp.) have been identified before as indigenous members of the microbiota of the human gut [93], female urogenital tract [60], [30] and the skin [94]. Closest RDP matches of all *Bacillales* sequences were assigned to the genus *Streptococcus* (mainly represented by the species of *S. agalactiae*).

At the phylum level, a total of 5 phyla have been identified. The most frequently detected sequences corresponding to 4 bacterial phyla were *Firmicutes, Actinobacteria, Proteobacteria* and *Bacteroidetes* while *Fusobacteria* phylum was less frequently observed.

### 3.2.3: Estimating the richness and diversity of urine microbial communities using alpha diversity

*Mothur* v.1.18.1 pipeline’s alpha diversity analyses package [83] was performed for estimating the richness, evenness and overall diversity of microbial communities for all urine host samples under study. Rarefaction curves, Chao1, Shannon and Simpson indices estimations were measured to determine the different depths of sampling and the overall richness and diversity of the host communities. Rarefaction curves have been plotted for both the 3’ and 5’ reads of the male and female samples at a 1% genetic difference level. The number of OTUs range from 2 to 31 and 12 to 53 OTUs of 3’ and 5’ sequence reads respectively. The results of rarefaction curves are shown in Figures 12 and 13.

Chao1 is one of the three non parametric estimators extrapolated from data to explore the true species richness by using an equation to identify the undiscovered species in the sample. Additionally, Simpson and Shannon indices were calculated to estimate the richness of the sample by measuring the species diversity and the relative abundance of each species taking into account the whole community (evenness). Both Shannon and Simpson indices showed high species diversity in all samples. The least diverse sample was observed only in the female subject1 as it was the only specimen that reached asymptote phase and its Shannon and Simpson estimations showed complete evenness. Interestingly, the 5’ reads were more diverse when compared to the 3’ reads. However, both require sufficient coverage and the communities need more sampling in order to reach completion. The overall results of the three non parametric estimators are listed in Tables 4 and 5.
3.3 Inter-individual variation

The female and male urine samples revealed remarkable differences in the distribution of various taxa at the genus level as shown in figures 14 and 15. Species of Gardnerella and Prevotella were the dominant bacterial population recovered from different 3/4 female urine specimens. In case of the male urine 16S rRNA sequences, species of Prevotella, Corynebacterium and Escherichia/Shigella were the dominant bacterial population recovered from all male urine specimens. Although Lactobacillus spp. represents one of the most abundant bacterial genera, its reads have been identified in only 2/4 of the urine samples of both sequences reads. However, the bacterial species composition and evenness varied between different subjects. Species of Megasphaera, Dialister and Sneathia were detected once in subject 3, Arthrobacter and Facklamia spp. in female subject 4, and Zimmermannella spp. in female subject 2. Female subject 3 show the highest diversity and species richness in the 16S rRNA sequence dataset of the V6-V8 reads. In contrast, female subjects 1, 2 and 4 have nearly the same species diversity and evenness.

When examining V6-V8 reads of the 16S rRNA sequences of the male urine microbiome, reads assigned to Arthrobacter, Phascolarctobacterium, Negativicoccus and Succinivibrio spp. were detected once in sample 4. Species of Roseburia and Oscillibacter spp., Sneathia and Flavobacterium spp. were identified once in samples 1 and 3 respectively. When examining V6-V8 reads of the 16S rRNA sequences of the male urine microbiome, reads assigned to Sphingomonas, Pseudoxanthomonas, Cupriavidus, Barnesiella and kocuria were observed once in urine subject 4. Species of Alistipes, Ruminococcus and Coprococcus were detected once in sample 1, and Brevundimonomas spp in male subject 2.

When examining the 16S rRNA sequences of both the V1-V3 and V6-V8 reads, the results uncovered noteworthy differences in distribution of bacterial phyla in all urine specimens. Actinobacteria was present in all female urine samples. Fusobacteria and Bacteroidetes were observed in 3/4 urine samples. Firmicutes and Proteobacteria were detected in only two samples 3, 4 and 2, 4 respectively. Additionally, Fusobacteria was the only bacterial phylum that was detected once in male urine subject 3 while the remaining phyla were present in all male urine samples. This reveals that the inter-individual variation was even noticeable at the phylum level.
CHAPTER 4: DISCUSSION

In this work, mid-stream urine samples were collected from 8 healthy volunteers (n=8), 4 healthy male and female individuals, into sterile containers. Then, PCR amplification was performed using forward and reverse 16S universal primers yielding amplicons with the expected length of 1300 bp. Following library construction, M13 vector primers were used for Sanger sequencing of ends of the inserts. As previously mentioned in the results, the 16S rRNA gene was not fully sequenced. Thus, two different hypervariable regions, V1-V3 (represents the 5’ end) and V6-V8 (represent the 3’ end), of 16S rRNA sequences were sequenced because they were more suitable to distinguish between various bacteria at the genus level when compared to V4 and V5. Both V4 and V5 manifested more sequence conservation, therefore they were not suitable targets for species identification [14]. V1-V3 reads were noticeably more appropriate for enduringness of taxonomical classification while V6-V8 reads were more appropriate for studying the microbial diversity [95], [96]. At the phylum and the order level, similar bacterial groups were observed in V1-V3 and V6-V8 reads respectively. While considerable differences were detected at the genus and the species level.

The first glimpse on the results of the 16S rRNA sequences obtained from both human female and male urine samples showed that the female urine microbiome consists of major five bacterial phyla; *Firmicutes, Actinobacteria, Fusobacteria, Bacteroidetes* and *Proteobacteria*. The results of the female 16S rRNA urine sequences were comparable with previously identified 16S rRNA sequences of former studies describing the female urogenital microbiota [25], [54], [30]. *Prevotella, Corynebacterium, Gardnerella* and *Weeksella spp.* represents the most dominant bacterial genera. *Weeksella viroso* was first isolated from the human urine by Tatum and colleges (1974) [97] and was detected as dominant genus of the female genital tract by Holmes and colleges (1986) [98] with no clear pathogenicity. While *Prevotella, Escherichia/Shigella, lactobacillus* and *Corynebacterium* represent the most dominant bacterial genera of the male urine microbiome. Similar results on the male urine microbiome were previously observed in studies conducted by Nelson and colleges (2010), (2012) [79], [99] and Dong and colleges (2011) [81]. Although *Gardnerella spp.* has been assigned as one of the dominant genera of the female urogenital microbiota, the three studies on the male urine did not display similar results. However, there is no characteristic
identification of the types of microbial communities of either male or female urine core microbiota of all individuals.

Multiple bacterial species such as, *A. vaginae, Megasphaera* and *Sneathia spp.* were identified using culture independent methods, 16S rRNA sequencing of gene libraries, which was inapplicable using cultivation methods since they are slow growing, strict anaerobes and difficult to grow on selective media. Specific media as MRS and Rogosa media were applied to support the growth of *Lactobacillus* but not of *L.iners* [100]. The identification of uncultured bacterial species as members of the healthy female urine microbiome further assesses the limitations of cultivation dependent methods.

*Lactobacillus spp.* is the dominant bacteria of the female urogenital tract [54], [30], [25] which was similarly observed here but to a lesser extent. The majority of the clones of *Lactobacillus* recovered from the 16S female urine sequences were phylogenetically related to the following species, *L.crispatus* and *L.iners*. They represent two of the four most commonly found species in the human female vagina. The results of phylogenetic analyses were quite similar to the 16S rRNA phylogeny results discussed in several studies [54], [27], [25]. In 3/4 of the female urine samples, *Lactobacillus spp.* coexisted either at significantly low abundance rates or even undetectable. In both cases, it was noted that other lactic acid producing bacteria such as *Atopobium vaginae, Megasphaera*, and *Leptotrichia spp.* compensate for the low proportions of *lactobacillus spp.* They are well known for their ability to conserve the low pH and lactic acid production for a healthy vagina. Accordingly, the coexistence of bacterial species as *Prevotella disiens* and *Gardnerella vaginalis* as potential members of the female core microbiome does not represent an abnormal state as in case of bacterial vaginosis [21], [60]. Bacterial vaginosis [18] is defined as: a medical condition in which a considerable decrease in the number of lactic acid producing bacteria specifically *Lactobacillus spp.* and a concomitant increase in other facultative or anaerobic bacteria is observed resulting in a change in the composition of vaginal flora. It is noteworthy that the growth of *G. vaginalis* is positively influenced by the abundance of *Prevotella spp.* *Prevotella spp.* provides *Gardnerella spp.* with the key nutrients of amino acids and ammonia [25]. *Prevotella disiens* is well known for its key role in predisposition to female genital tract infections [60]. *Corynebacterium riegelii* was discovered as a novel species in female patients with UTIs [101]. *Gardnerella vaginalis* strains are one of the pathogenetic factors of BV [102].
The bacterial microbiota found in healthy men are quite similar in both urethral swabs and urine specimens [81]. Thus, both samples can be applied for taxonomical and phylogenetic analysis of the male core microbiome. Interestingly, the phylogenetic results were strongly correlated with studies [99], [81] that previously discussed the abundance of vaginal flora of healthy women and BV associated taxa in the urine specimens of healthy men. Lactobacillus, Prevotella, Gardnerella, Veillonella (detected in V6-V8 reads) and Sneathia spp. (detected in V1-V3 reads) were abundant in male urine sample 3 who previously reported sexual experience. Therefore, vaginal sexual exposure may result in broader exchange of urogenital microbiota. The Lactobacillales, genus Lactobacillus, coexists in 2/4 of the male individuals primarily in urine sample 3. Lactobacillus contains diverse species (L. crispatus and mostly represented as L. iners and uncultured taxa) and comprised approximately 20 sequences. A study by Rachel R. and Cindy A. proposed that the vaginal Lactobacillus plays a key role in reducing the incidence of sexually transmitted infections (STIs) following exposure to STI pathogens [103]. A study by Nelson and colleges [99] reported the abundance of bacterial genera, isolated from the vaginal flora, in urine and coronal sulcus specimens of participants with no history of sexual experience except for Sneathia spp. Based on the study observations, the similarity of both male and female genital tracts microbiome supports the assumption that a few microbial communities could survive during sexual transfer.

Multiple bacterial genera were previously identified as water-borne bacteria suggesting that they could be a tap water contaminant or a potential member of the bacterial community of the female and male urinary tract. For instance, the genus Planobacterium [104] which was observed in the 16S rRNA sequence phylogeny of female subject 3 and the genera Sphingomonas, Cupriavidus and Pseudoanthonomonas spp. which were observed in the 16S rRNA sequence phylogeny of male urine sample 4 (with > 97% sequence similarity). Likewise, the former observation was in concordance with the results of a study by Riemersma and colleges [80].

It is noteworthy that several abundant taxa detected in urine samples were previously classified as residents of the human gut microbiome [105] as Ruminococcus, Faecalibacterium, Coprococcus, Roseburia, Clostridiale and Oscillibacter while Peptoniphilus spp. as a member of indigenous microbiota of the skin respectively [33]. Barnesiella spp. was also known to inhabit the human distal gut flora [106]. Furthermore, it
has been detected in the conjunctival microbiota of neonates delivered vaginally [33]. *Alistipes spp.* has been detected in few studies done on paediatric patients suffering from irritable bowel syndrome [107]; however it has not been identified as a genuine member of the male urine microbiota. Here, it was observed as a potential member of the male urine flora. Most of the taxonomically assigned bacterial phylotypes of the *Clostridiales* order have been identified in the vaginal flora of healthy women. They are well known for their clinical diagnosis of BV as they metabolize carbohydrates and amino acids producing stinking compounds such as thiols, amines and organic acids [108]. The stinking odour is one of the main diagnostic features of BV. They include strict anaerobes such as: *Anaerococcus, Peptoniphilus* and unclassified *Lachnospiraceae* family that was recovered using cultivation independent methods. They represent considerable proportion of the indigenous microbiota of the male urogenital tract. Therefore, more efforts should be done to better characterize the roles of these microbial communities in the urinary flora.

Similarly, several abundant taxa detected in urine samples were previously classified as residents of the human skin microbiome. In a study by Helene Marchandin and colleagues (2010), the anaerobic and Gram negative coccus, *Negativococcus succinicivorans*, was firstly identified in clinical isolates from the human skin and soft tissues as a novel genus using 16S rRNA phylogeny with 99.9% identity. Several microbiotas were solely detected in urine like the species of *Lactobacillus, Veillonella* and *Gardnerella* which were similar to bacterial species detected in the mucosal surfaces [99].

In order to assess the diversity of 16S rRNA sequences of bacterial phylotypes that indigenously coexist in the urinary tract of both male and female individuals, alpha diversity analyses were performed using *mothur* pipeline. Based on a 0.01 phylogenetic distance, rarefaction curves were plotted for both V1-V3 and V6-V8 sequence datasets. Expectedly, neither 5’ nor 3’ reads showed good coverage as both did not reach the plateau phase indicating that more sequences needed to be sequenced in order to reach the predicted maximum richness. Those findings may be attributed to the fact that the numbers of analyzed clones are small. Interestingly, the female subject 1 solely reached the plateau phase in V6-V8 reads. Unlikely, the coverage of the same urine sample of its V1-V3 reads was incomplete and thus, more phylotypes remain to be identified. In addition, Chao1, Simpson and Shannon alpha diversity indices estimations were performed. Chao1 is a non-parametric estimate that measures the total number of phylotypes that are expected to exist and might not
be detected in each sample. For both sequence reads of male and female urine samples, Chao1 values were significantly exceeding the actual number of phylotypes present in each sample. Simpson and Shannon indices measure the species diversity which reflects the species richness and evenness. However, the sensitivity of Simpson index towards species evenness is higher than its richness while the sensitivity of Shannon index towards species richness is in turn higher than its evenness [109]. Simpson index values range from 0 (low diversity) to 1 (high diversity) while Shannon index values range from 1 (low diversity) to 5 (high diversity).

Based on the previous observations, our data implies a phylogenetic view of the urinary microbial communities in male and female individuals. First, our findings support that the bacterial communities of the urinary tract are complex and polymicrobial. Inter-individual variation was observed in all analyzed samples confirming that no single characteristic bacterial community was evident. Those results coincide with previous studies on male and female urogenital flora [79], [60]. Diverse bacterial genera identified in the urinary tract were highly similar to microbial communities that were formerly reported in superficial skin [105], [94], gut [33], [93], and vaginal microbiome [54], [26], [21]. In addition, the interpretations of phylogenetic analysis results of male 16S rRNA urine sequences suggest that there are evident differences in the composition of microbial communities of urine and penile coronal sulcus (CS). Coronal sulcus specimens are dominated by 16S rRNA sequences assigned to the family *Pseudomonadaceae spp.* [110] which was apparently rare in urine. Obligate and/or facultative anaerobes such as *Streptococcus*, *Lactobacillus*, *Gardnerella* and *Veillonella* were commonly detected in 16S urine sequences whereas they were rare in CS sequences [99]. Second, several BV associated taxa such as *Prevotella*, *Gardnerella*, *A.vaginae* and *Sneathia spp.* are suggested to be genuine members of both healthy male and female urinary microbiome. Some of them are lactic acid producing bacteria that compensate for the undetectable abundance of *Lactobacillus spp.* in female urine samples. Thus, additional future studies should be performed in order to clear out this issue as numerous cases of healthy women having normal urogenital microbiome could be misdiagnosed with BV. Finally, considerable differences between male and female urine microbiome were detected. Common bacterial taxa exist in high proportions in urine sequences of both genders as *Prevotella*, *Corynebacterium*, *Weeksella* and *Lactobacillus*. Few bacteria as *Gardnerella spp.* predominantly found in urine samples of female individuals, was apparently identified as minor representatives of male
urine samples. 16S rRNA sequences assigned to bacterial genera of *Streptococcus*, *Staphylococcus*, *Escherichia/Shigella*, *Acinetobacter* and others were abundant in male urine samples whereas sequences assigned to these genera were not detected in urines of female subjects. This was similarly observed in phylogenetic analysis of female urine sequences such as 16S sequences assigned to *O. urethralis*, *A. vaginae*, *Dialister* and other bacterial flora not identified in urines of male subjects.

It is noteworthy to mention the pitfalls of this thesis work which needs further investigation. First, the overall results of alpha diversity agree on implying that both male and female urine samples are highly diverse and more coverage is required to give better insight into the urine microbiome. This might be attributed to the small number of clones analyzed in each individual urine sample which was reflected on the results of the three non-parametric estimators of alpha diversity analysis. Therefore, repeated sampling of each individual will be necessary to validate the confidence of our results. Second, primer walking is required to ensure complete coverage of the whole 16S rRNA gene. This will broaden our perception on the diversity of the male and female urine microbiota and adds further level of complexity for data analysis at the bacterial species level. Finally, the inadequate number of male and female individuals under study needs further investigation in the future.

In conclusion, enhanced resolution obtained by 16S rRNA Sanger sequencing manifests the advantages of cultivation independent methods over independent ones in exploring the indigenous microbiome of urinary tract of healthy adult individuals. The study of phylogenetic profiles of 16S rRNA sequences of urine phylotypes shed the light on the occurrence of wide range of bacteria including many associated with disease causing bacteria as a part of the healthy male and female urine microbiome. Thus, additional future studies should uncover the relationship between the human urinary tract microbiota and their impact on human health upon study of various disease states particularly those resistant to traditional treatment methods.

REFERENCES

30


90 Rossau R, K.K., Falsen E, et al.: ‘Oligella, a new genus including Oligella urethralis comb. nov. (formerly Moraxella urethralis) and Oligella ureolytica sp. nov. (formerly CDC group IVe)’, Int J Syst. Bacteriol., 1987, 37, pp. 198-210


Figure 1 Distributions of different bacterial phyla

This figure illustrates the relative abundances of different bacterial phyla at various anatomical sites of the human body including the skin, gut, male and female urogenital tracts. This figure is modified from (Dethlefsen et al. 2007) [110].
Figure 2 Gel electrophoresis of total genomics DNA from four male urine samples

DNAs extracted from microbial community collected from urine of four healthy adult male individuals were examined on 1% gel electrophoresis. MW: 1kb DNA marker ladder (Invitrogen, Carlsbad, CA).
Figure 3 Gel electrophoresis of total genomics DNA from four female urine samples

DNAs extracted from microbial community collected from urine of four healthy adult female individuals were examined on 1% gel electrophoresis. MW: 1kb DNA marker ladder (Invitrogen, Carlsbad, CA).
DNA isolated from four urine males shown in Figure 2 were used to amplify the 16S rRNA gene. The PCR reactions were performed using the universal primers 63F and 1387 (details in Material and Methods). The products of the amplifications were analyzed on 1% gel electrophorasis. MW: 1kb DNA marker ladder (Invitrogen, Carlsbad, CA). A purified *E. coli* genomic DNA was used as PCR positive (+ ve) control and no added genomic DNA as PCR negative (-ve) control. All 16S rRNA amplicons had the expected length of 1.3 kb.
DNA isolated from four urine females shown in Figure 2 were used to amplify the 16S rRNA gene. The PCR reactions were performed using the universal primers 63F and 1387 (details in Material and Methods). The products of the amplifications were analyzed on 1% gel electrophorasis. MW: 1kb DNA marker ladder (Invitrogen, Carlsbad, CA). A purified *E. coli* genomic DNA was used as PCR positive (+ ve) control and no added genomic DNA as PCR negative (-ve) control. All 16S rRNA amplicons had the expected length of 1.3 kb.
Figure 6  A phylogenetic tree showing the relationship between various phylotypes based on analysis of 16S rRNA female urine sequences of V6-V8 dataset

The phylogenetic tree was constructed using a maximum - likelihood algorithm based on analysis of 16S rRNA gene sequences of V6-V8 dataset recovered from urine samples of four healthy female subjects. It also demonstrates the relationship between various bacterial phylotypes that were evolutionary compared to closely related organisms retrieved from RDP II database. For taxonomy assignments, the phylogenetic tree was constructed at bootstrap values of > 70% at the species level.
Figure 7 A phylogenetic tree showing the relationship between various phylotypes based on analysis of 16S rRNA female urine sequences of V1-V3 dataset

The phylogenetic tree was constructed using a maximum-likelihood algorithm based on analysis of 16S rRNA gene sequences of V1-V3 dataset recovered from urine samples of four healthy female subjects. It also demonstrates the relationship between various bacterial phylotypes that were evolutionary compared to closely related organisms retrieved from RDP II database. For taxonomy assignments, the phylogenetic tree was constructed at bootstrap values of > 70% at the species level.
Figure 8 A phylogenetic tree showing the relationship between various phylotypes based on analysis of 16S rRNA male urine sequences of V6-V8 dataset

The phylogenetic tree was constructed using a maximum-likelihood algorithm based on analysis of 16S rRNA gene sequences of V6-V8 dataset recovered from urine samples of four healthy male subjects. It also demonstrates the relationship between various bacterial phylotypes that were evolutionary compared to closely related organisms retrieved from RDP II database. For taxonomy assignments, the phylogenetic tree was constructed at bootstrap values of > 70% at the species level.
Figure 9 A phylogenetic tree showing the relationship between various phylotypes based on analysis of 16S rRNA male urine sequences of V1-V3 dataset

The phylogenetic tree was constructed using a maximum-likelihood algorithm based on analysis of 16S rRNA gene sequences of V1-V3 dataset recovered from urine samples of four healthy male subjects. It also demonstrates the relationship between various bacterial phylotypes that were evolutionary compared to closely related organisms retrieved from RDP II database. For taxonomy assignments, the phylogenetic tree was constructed at bootstrap values of > 70% at the species level.
**Figure 10 Comparison of major taxa taxonomically assigned to V1-V3 and V6-V8 reads of human female 16S rRNA sequences urine sequences at the genus level**

This illustration depicts the bacterial communities that indigenously inhabit the female urinary tract by sorting the 16S rRNA sequences at the genus level. Out of 17 bacterial genera being identified in all female urine specimens, 4 different genera were unique for both of the two reads. The 3’ reads representing V6-V8 regions of the urine 16S rRNA sequences were predominantly assigned to *Prevotella, Corynebacterium, Gardnerella* and *Weeksella spp.* While the 5’ reads representing V1-V3 regions of the urine 16S rRNA sequences were predominantly assigned to the same bacterial taxa. Other microbial communities were only observed in the V1-V3 reads; *Empedobacter, Facklamia* and *Flavobacterium spp.*
(A) Male Genera of V6-V8 reads

(B) Male Genera of V1-V3 reads
Figure 11 Comparison of major taxa taxonomically assigned to V1-V3 and V6-V8 reads of human male 16S rRNA sequences urine sequences at the genus level

This illustration depicts the bacterial communities that indigenously inhabit the male urinary tract by sorting the 16S rRNA sequences at the genus level. Out of 41 bacterial genera being identified, 19 different genera were unique for both of the two reads. The 3’ reads representing V6-V8 regions were predominantly assigned to *Prevotella, Escherichia/Shigella, lactobacillus* and *Corynebacterium*. Similarly, the taxonomical assignments of 16S rRNA sequences of 5’ reads to the major bacterial were dominantly observed. Other microbial communities were only observed in the V1-V3 reads as *Ruminicoccus, Coprococcus, Sphingomonas, Pseuadoxanthomonas, Kocuria* and *Cupriavidus spp*. However, *Negativicoccus, Oscillibacter, Brucella, Salmonella, Weeksella* and *Alistipes spp* were only identified in the V6-V8 reads.
Figure 12 Rarefaction curves plotted to estimate the phylotypes diversity and coverage level of 16S rRNA urine sequences of 3’ reads from both male and female subjects

This graph was plotted to assess the diversity of 16S rRNA sequences of bacterial phylotypes that indigenously inhabit the urinary tract both male and female individuals. Based on a 0.01 phylogenetic distance, rarefaction curve was plotted for 3’ reads representing V6-V8 sequence dataset. The graph also demonstrates that neither male nor female bacterial phylotypes showed good coverage as both expectedly failed to reach the plateau phase. The only exception was female subject 1.
Figure 13 Rarefaction curves plotted to estimate the phylotypes diversity and coverage level of 16S rRNA urine sequences of 5’ reads from both male and female subjects

This graph was plotted to assess the diversity of 16S rRNA sequences of bacterial phylotypes that indigenously inhabit the urinary tract both male and female individuals. Based on a 0.01 phylogenetic distance, rarefaction curve was plotted for 5’ reads representing V1-V3 sequence dataset. The graph also demonstrates that neither male nor female bacterial phylotypes, specifically female subject 1, showed good coverage as both expectedly failed to reach the plateau phase.
Figure 14 The overall distribution of bacterial genera taxonomically assigned to 16S rRNA urine sequences of V6-V8 dataset

This graph demonstrates the overall composition of 16S rRNA sequences of V6-V8 dataset being recovered from 16S libraries of male and female urine samples according to sequence similarity against RDP database. Nine bacterial taxa have been identified in both male and female subjects in different proportions. It also represents the relation between the number of sequence reads and their composition of bacterial communities in both subjects.
Figure 15 The overall distribution of bacterial genera taxonomically assigned to 16S rRNA urine sequences of V1-V3 dataset

This graph demonstrates the overall composition of 16S rRNA sequences of V1-V3 dataset being recovered from 16S libraries of male and female urine samples according to sequence similarity against RDP database. Eight bacterial taxa have been identified in both male and female subjects in different proportions. It also represents the relation between the number of sequence reads and their composition of bacterial communities in both subjects.
### Table 1 Summary of Female urine bacterial communities in 3’ and 5’ sequence datasets

<table>
<thead>
<tr>
<th>Order</th>
<th>Genus</th>
<th>Bootstrap value of phylogenetic clade</th>
<th>Total number of sequences of 3’ reads</th>
<th>Total number of sequences of 5’ reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bifidobacteriales</td>
<td>Gardnerella</td>
<td>99</td>
<td>97</td>
<td>100</td>
</tr>
<tr>
<td>Coriobacteriales</td>
<td>Atopobium</td>
<td>99</td>
<td>13</td>
<td>98</td>
</tr>
<tr>
<td>Actinomycetales</td>
<td>Zimmernallina</td>
<td>100</td>
<td>4</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>Arthrobacter</td>
<td>88</td>
<td>3</td>
<td>99</td>
</tr>
<tr>
<td>Corynebacterium</td>
<td></td>
<td>95</td>
<td>30</td>
<td>99</td>
</tr>
<tr>
<td>Burkholderiales</td>
<td>Oligella</td>
<td>99</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>Lactobacillales</td>
<td>Lactobacillus</td>
<td>83</td>
<td>18</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>Facklamia</td>
<td></td>
<td></td>
<td>98</td>
</tr>
<tr>
<td>Bacteriodales</td>
<td>Prevotella</td>
<td>100</td>
<td>71</td>
<td>92</td>
</tr>
<tr>
<td>Flavobacteriales</td>
<td>Weeksella</td>
<td>&lt; 92</td>
<td>55</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Planobacterium</td>
<td>100</td>
<td>3</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Unclassified taxa</td>
<td>96</td>
<td>10</td>
<td>&lt; 94</td>
</tr>
<tr>
<td>Fusobacteriales</td>
<td>Sneathia</td>
<td>100</td>
<td>13</td>
<td>100</td>
</tr>
<tr>
<td>Selenomondales</td>
<td>Dialister</td>
<td>93</td>
<td>6</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Megasphaera</td>
<td>98</td>
<td>5</td>
<td>89</td>
</tr>
</tbody>
</table>

The table represents all bacterial communities being identified in female urinary tract by comparing their 16S rRNA gene sequences to known organisms in the RDP database. This table demonstrates the dominant bacterial genera, *Prevotella, Weeksella, Corynebacterium* and *Gardnerella*, in both V1-V3 and V6-V8 sequence reads [dark orange colour]. It also verifies the results of phylogenetic analyses where higher taxonomical range and improved resolution of bootstrap value of phylogenetic clade were observed in V1-V3 dataset. Few bacterial genera showed lesser resolution in 5’ reads [light orange colour].
# Table 2 Summary of Male urine bacterial communities in 3’ and 5’ sequence datasets

<table>
<thead>
<tr>
<th>Order</th>
<th>Genus</th>
<th>Bootstrap value of phylogenetic clade</th>
<th>Total number of sequences of 3’ reads</th>
<th>Bootstrap value of phylogenetic clade</th>
<th>Total number of sequences of 5’ reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actinomycetales</td>
<td>Corynebacterium</td>
<td>100</td>
<td>40</td>
<td>99</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>Arthrobacter</td>
<td>92</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kocuria</td>
<td></td>
<td></td>
<td>96</td>
<td>3</td>
</tr>
<tr>
<td>Bifidobacteriales</td>
<td>Gardnerella</td>
<td>100</td>
<td>8</td>
<td>100</td>
<td>3</td>
</tr>
<tr>
<td>Lactobacillales</td>
<td>Streptococcus</td>
<td>98</td>
<td>5</td>
<td>100</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Lactobacillus</td>
<td>95</td>
<td>40</td>
<td>99</td>
<td>46</td>
</tr>
<tr>
<td>Fusobacteriales</td>
<td>Sneathia</td>
<td>100</td>
<td>3</td>
<td>100</td>
<td>3</td>
</tr>
<tr>
<td>Bacillales</td>
<td>Staphylococcus</td>
<td>98</td>
<td>6</td>
<td>95</td>
<td>6</td>
</tr>
<tr>
<td>Clostridiales</td>
<td>Anaerococcus</td>
<td>91</td>
<td>10</td>
<td>99</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Roseburia</td>
<td>95</td>
<td>9</td>
<td>90</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Unclassified Lachnospiraceae family</td>
<td>100</td>
<td>15</td>
<td>94</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Oscillibacter</td>
<td>100</td>
<td>3</td>
<td>99</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Unclassified Ruminococcaceae</td>
<td>99</td>
<td>3</td>
<td>99</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Ruminococcus</td>
<td></td>
<td></td>
<td>99</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Fecalibacterium</td>
<td>99</td>
<td>16</td>
<td>99</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Peptoniphilus</td>
<td></td>
<td></td>
<td>96</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Unclassified Clostridiales</td>
<td>100</td>
<td>3</td>
<td>99</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Coprocococcus</td>
<td></td>
<td></td>
<td>98</td>
<td>3</td>
</tr>
<tr>
<td>Selenomonadales</td>
<td>Phascolarctobacterium</td>
<td>97</td>
<td>3</td>
<td>99</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Veillonella</td>
<td>99</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Negativicoccus</td>
<td>99</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caulobacterales</td>
<td>Brevundimonas</td>
<td>97</td>
<td>7</td>
<td>98</td>
<td>11</td>
</tr>
<tr>
<td>Rhizobiales</td>
<td>Ochrobactrum</td>
<td>97</td>
<td>4</td>
<td>97</td>
<td>5</td>
</tr>
<tr>
<td>Pseudomonadales</td>
<td>Acinetobacter</td>
<td>99</td>
<td>16</td>
<td>99</td>
<td>36</td>
</tr>
<tr>
<td>Aeromonadales</td>
<td>Succinivibrio</td>
<td>100</td>
<td>4</td>
<td>97</td>
<td>14</td>
</tr>
<tr>
<td>Enterobacterales</td>
<td>Escherichia/Shigella</td>
<td>98</td>
<td>56</td>
<td>99</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Salmonella</td>
<td>82</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Enterobacter</td>
<td>100</td>
<td>3</td>
<td>100</td>
<td>6</td>
</tr>
<tr>
<td>Bacteroidales</td>
<td>Prevotella</td>
<td>98</td>
<td>62</td>
<td>95</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>Bacteroides</td>
<td>87</td>
<td>17</td>
<td>97</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Porphyromonas</td>
<td>100</td>
<td>6</td>
<td>100</td>
<td>8</td>
</tr>
<tr>
<td>Genus</td>
<td>Percentage</td>
<td>Number of Sequence Reads</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>------------</td>
<td>--------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barnesiella</td>
<td>99</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alistipes</td>
<td>99</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flavobacteriales</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weekellsa</td>
<td>75</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flavobacterium</td>
<td>78</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unclassified Flavobacteriaceae</td>
<td>99</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphingomonadales</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphingomonas</td>
<td>99</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xanthomonadales</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudoxanthomonas</td>
<td>97</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burkholderiales</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cupriavidus</td>
<td>99</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table represents all bacterial communities being identified in female urinary tract by comparing their 16S rRNA gene sequences to known organisms in the RDP database. This table demonstrates the dominant bacterial genera, *Prevotella, Eischerichia/Shigella, Corynebacterium and Lactobacillus*, in both V1-V3 and V6-V8 sequence reads [dark orange colour]. It also verifies the results of phylogenetic analyses where higher taxonomical range and improved resolution of bootstrap value of phylogenetic clade were observed in V1-V3 dataset. Few bacterial genera showed lesser resolution in 5’ reads [light orange colour].
Table 3 Summary of the number of sequenced clones for each library, length of sequence obtained from each end and quality of reads

<table>
<thead>
<tr>
<th></th>
<th>Female 1</th>
<th>Female 2</th>
<th>Female 3</th>
<th>Female 4</th>
<th>Male 1</th>
<th>Male 2</th>
<th>Male 3</th>
<th>Male 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of reads</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(using 63F primer)</td>
<td>50</td>
<td>67</td>
<td>77</td>
<td>43</td>
<td>52</td>
<td>49</td>
<td>45</td>
<td>72</td>
</tr>
<tr>
<td>Number of reads</td>
<td>63</td>
<td>63</td>
<td>80</td>
<td>36</td>
<td>48</td>
<td>33</td>
<td>31</td>
<td>27</td>
</tr>
<tr>
<td>(using 1387R primer)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of reads in</td>
<td>40</td>
<td>56</td>
<td>70</td>
<td>30</td>
<td>39</td>
<td>17</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>both forward and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>reverse sequences</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Length</td>
<td>530</td>
<td>497</td>
<td>524</td>
<td>433</td>
<td>470</td>
<td>478</td>
<td>429</td>
<td>468</td>
</tr>
<tr>
<td>of reads (bps)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum Length of</td>
<td>376</td>
<td>302</td>
<td>329</td>
<td>300</td>
<td>309</td>
<td>303</td>
<td>315</td>
<td>308</td>
</tr>
<tr>
<td>reads (bps)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum Length of</td>
<td>620</td>
<td>619</td>
<td>644</td>
<td>583</td>
<td>589</td>
<td>639</td>
<td>612</td>
<td>617</td>
</tr>
<tr>
<td>reads (bps)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum Quality of</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>reads</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table represents the number of sequenced clones for each library, length of sequence obtained from each end and quality of reads. The number of reads for each individual (forward plus reverse) varied from a minimum of 76 (male 3) to a maximum 157 (female 3). Sequences were processed by trimming the plasmid vector bases with cross match and discarding low quality sequences < 300 bp in length and/or had average PHRED quality score of Q= < 20. Therefore, the minimum length of reads is 302 (female 2) and varied up to a maximum of 644 (female 3). This explains why the reads obtained from the 5’ and the 3’ of the inserts did not permit overlapping of the sequences from both ends of the recombinant plasmids.
Table 4 Alpha diversity measurements of V6-V8 sequence dataset of male and female 16S rRNA gene

<table>
<thead>
<tr>
<th>Library</th>
<th>Number of Sequences</th>
<th>Number of OTUs</th>
<th>Chao1</th>
<th>Shannon index</th>
<th>Simpson index</th>
</tr>
</thead>
<tbody>
<tr>
<td>3’_plate1_female</td>
<td>58</td>
<td>2</td>
<td>2</td>
<td>0.08365</td>
<td>0.967213</td>
</tr>
<tr>
<td>3’_plate2_female</td>
<td>59</td>
<td>22</td>
<td>44.75</td>
<td>2.52451</td>
<td>0.111864</td>
</tr>
<tr>
<td>3’_plate3_female</td>
<td>76</td>
<td>31</td>
<td>69</td>
<td>2.897021</td>
<td>0.080253</td>
</tr>
<tr>
<td>3’_plate4_female</td>
<td>62</td>
<td>20</td>
<td>98</td>
<td>2.217736</td>
<td>0.191396</td>
</tr>
<tr>
<td>3’_plate1_male</td>
<td>55</td>
<td>30</td>
<td>205.5</td>
<td>2.613783</td>
<td>0.168831</td>
</tr>
<tr>
<td>3’_plate2_male</td>
<td>61</td>
<td>25</td>
<td>110.5</td>
<td>2.521391</td>
<td>0.136201</td>
</tr>
<tr>
<td>3’_plate3_male</td>
<td>43</td>
<td>24</td>
<td>87.33333</td>
<td>2.480538</td>
<td>0.170048</td>
</tr>
<tr>
<td>3’_plate4_male</td>
<td>45</td>
<td>24</td>
<td>45</td>
<td>2.927935</td>
<td>0.047343</td>
</tr>
</tbody>
</table>

All the table calculations were performed using mothur software [9] for 3’ reads of both male and female 16S rRNA sequences at a 1% divergence threshold. The number of observed OTUs range from 2 to 31. The results coincide with the plotted rarefaction curves where the Chao1 values of all urine samples, with exception of female subject1, were significantly exceeding the actual number of phylotypes present in each sample. Chao1, Shannon and Simpson indices were calculated for estimating the richness, evenness and overall diversity of microbial communities for urine samples under study.
Table 5 Alpha diversity measurements of V1-V3 sequence dataset of male and female 16S rRNA gene

<table>
<thead>
<tr>
<th>Library</th>
<th>Number of clones</th>
<th>Number of OTUs</th>
<th>Chao1</th>
<th>Shannon index</th>
<th>Simpson index</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’_plate1_female</td>
<td>44</td>
<td>12</td>
<td>34.5</td>
<td>1.136917</td>
<td>0.567177</td>
</tr>
<tr>
<td>5’_plate2_female</td>
<td>70</td>
<td>40</td>
<td>82</td>
<td>3.349299</td>
<td>0.041408</td>
</tr>
<tr>
<td>5’_plate3_female</td>
<td>79</td>
<td>43</td>
<td>88.11111</td>
<td>3.42104</td>
<td>0.037975</td>
</tr>
<tr>
<td>5’_plate4_female</td>
<td>59</td>
<td>41</td>
<td>115.3333</td>
<td>2.987554</td>
<td>0.068747</td>
</tr>
<tr>
<td>5’_plate1_male</td>
<td>43</td>
<td>24</td>
<td>58</td>
<td>2.869584</td>
<td>0.057082</td>
</tr>
<tr>
<td>5’_plate2_male</td>
<td>63</td>
<td>44</td>
<td>137.5</td>
<td>3.596609</td>
<td>0.020337</td>
</tr>
<tr>
<td>5’_plate3_male</td>
<td>46</td>
<td>23</td>
<td>53</td>
<td>2.662995</td>
<td>0.100177</td>
</tr>
<tr>
<td>5’_plate4_male</td>
<td>72</td>
<td>53</td>
<td>289.5</td>
<td>3.812269</td>
<td>0.013693</td>
</tr>
</tbody>
</table>

All the table calculations were performed using mothur software [9] for 5’ reads of both male and female 16S rRNA sequences at a 1% divergence threshold. The number of observed OTUs range from 12 to 53. The results coincide with the plotted rarefaction curves where the Chao1 values of all urine samples were significantly exceeding the actual number of phylotypes present in each sample. Unlike wise, the coverage of V1-V3 reads of female urine sample 1 was incomplete and thus, more phylotypes remain to be identified. Chao1, Shannon and Simpson indices were calculated for estimating the richness, evenness and overall diversity of microbial communities for urine samples under study.
Supplementary Figure 1 The approval letter by the institutional review board (IRB) of the American university in Cairo

All the subjects under the study were provided with a written informed consent which was approved by the institutional review board (IRB) of the American university in Cairo.
Supplementary Figure 2 The medical history

All enrolled subjects were asked to fill a written medical history form to make sure of the overall general health.