Characterization of the urogenital microbiome in patients with urinary tract infections

Fitri Nadifah1*, Wayan Tunas Artama2, Budi Setiadi Daryono3, Endah Retnaningrum4

1 Program of Doctoral, Faculty of Biology, Universitas Gadjah Mada. Jl. Teknika Selatan, Yogyakarta 55281, Indonesia
2 Department of Biochemistry, Faculty of Veterinary Medicine, Universitas Gadjah Mada. Jl. Fauna No.2 Karangmalang, Yogyakarta 55281, Indonesia
3 Laboratory of Genetics, Faculty of Biology, Universitas Gadjah Mada. Jl. Teknika Selatan, Yogyakarta 55281, Indonesia
4 Laboratory of Microbiology, Faculty of Biology, Universitas Gadjah Mada. Jl. Teknika Selatan, Yogyakarta 55281, Indonesia
* Corresponding author: fitri.nadifah@mail.ugm.ac.id

ABSTRACT Standard microbiological culture techniques can only identify a fraction of the urogenital microbiome. Meanwhile, identifying and characterizing infectious microorganisms are very important for the success of diagnosis and treatments, especially for Urinary Tract Infection (UTI) patients. This study aimed to characterize the urogenital microbiome of UTI patients using 16S rRNA gene sequencing. We sequenced two pooled DNA samples from voided urine of UTI patients (21 females and 13 males). To determine the structure and composition of taxa in the samples, 16S rRNA gene sequencing was performed using the Illumina Mi-Seq paired-end platform. The most abundant genera were Burkholderia-Caballeronia-Paraburkholderia (71%) followed by Prevotella (33%), Escherichia-Shigella (24%), Klebsiella (23%) and Sneathia (10%). The female microbiome was dominated by Prevotella bivia (28%), Escherichia coli (24%), Sneathia sanguinegens (7%) and Klebsiella pneumoniae (4%). On the other hand, the male microbiome was dominated by K. pneumoniae (23%) and E. coli (2%). K. pneumoniae and E. coli were the most abundant species found in both microbiomes. The 16S rRNA gene sequencing used in this study successfully uncovered the composition of the urogenital microbiome, which might not have been possible with conventional culture methods.

KEYWORDS urogenital microbiome; UTIs; 16S rRNA sequencing; bacteria

1. Introduction

A urinary tract infection (UTI) has become one of the most common infectious diseases in the world (Price et al. 2018). Its nomenclature and concept have begun since the discovery of microbiota that existed in the urine of a healthy person, making the old paradigm that urine is sterile disproven (Dixon et al. 2020). Generally, UTI is described as an acute urinary health problem because of the presence or predominance of uropathogenic microbes (Brubaker and Wolfe 2017). The pathogenesis of this disease starts from uropathogenic contamination in a periurethral space, which then makes colonies in the urethra and finally ascends to the bladder (Meštrović et al. 2020). However, there are still limitations to standards for UTI diagnosis (Price et al. 2018).

The Human Microbiome Projects which were conducted in 2007, have provided a new insight into the human microbial communities, especially in the gut, nasal cavity, mouth, skin and vagina, that correlate with health and diseases. A staggering number of researches has emerged and focused on revealing microbiome (the genetic materials of microbiota) roles in other body niches, including the urinary tract (Lee et al. 2020). Investigations on female bladder microbiota have uncovered Lactobacillus as the most common urotyp. Other urotypes found were Gardnerella, Streptococcus and Escherichia (Fok et al. 2019; Komesu et al. 2018; Price et al. 2019). Furthermore, many investigations have been conducted to find the correlation between urobiome (microbial communities in the bladder and urinary tract) and clinical conditions, including UTI and some forms of urinary incontinence (Karstens et al. 2016; Wolfe and Brubaker 2019).

Several studies on the human urobiome have been conducted using different types of urinary samples, such as catheterized and voided urine. The use of different urinary samples, as well as the collecting method, can affect the results obtained. Catheterized specimens of urine are better at giving information about the relatedness of urogenital microbiome than...
catheterized urine (Pohl et al. 2020).

A urogenital tract (UT) includes reproductive organs and the formation and excretion of urine. Reproductive organs, as the parts of UT may contribute to urine microbial load. Vagina, cervix, periurethral skin, penis, pubic skin surfaces, and perineal area, are ideal sites for pathogen habitation. This is called the urogenital microbiome since microbiome composition in urine can come from genital sites and organs (MacIntyre et al. 2017; Neugent et al. 2020).

A culture-independent method for identifying and characterizing microbiota has been conducted to increase an appreciation of a microbial community profile in the UT (MacIntyre et al. 2017). Determination of a bacterial community profile composing the human microbiome is primarily facilitated by a 16S rRNA gene sequencing assay. The 16S rRNA gene sequencing data are used to conduct metagenomic analysis, including compiling taxonomic classifications by comparing the reads to reference databases (Deurenberg et al. 2017; Hiergeist and Gessner 2017). The results of 16S rRNA sequencing can provide information about the presence of bacterial species undetectable by conventional culture. Furthermore, it can conclusively identify a huge number of sequence reads of a common uropathogen, as well as a new bacterial species associated with infection (Moustafa et al. 2018). This study aimed to characterize the urogenital microbiome from voided urine samples of UTI patients using 16S rRNA gene sequencing.

2. Materials and Methods

2.1. Ethical approval

This research has received approval from the Ethics Committee of the Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada Yogyakarta (Number KE/FK/1031/EC/2019).

2.2. Specimens collection

Sequenced samples were pooled of isolated DNA from voided urine specimens from 21 females and 13 males of UTI patients collected from Regional Public Hospitals in Yogyakarta, Indonesia. The specimen is the first catch urine with the possibility that the bacteria identified can come from the urinary tract or genitalia (Pohl et al. 2020). The diagnosis was made by the treating physician based on an assessment of symptoms and laboratory results, including the leukocyte esterase and bacterial counts in the urine. The female subjects ages ranged from 0-90 years old, while the male ranged from 23-80 years old. The specimens were collected into a sterile container and kept at -4°C then processed for DNA extraction within 24 h after collection.

2.3. Genomic DNA Extraction

Thirty mL of the urine specimen was pelleted by centrifugation (2000 rpm, 5 min). Twenty-five mL of the supernatant was decanted. The remaining 5 mL of the pellet was re-suspended and pelleted again by centrifugation for 10 min at 16000 × g (4 °C). In the next step, up to 100 μL of the pellet and some supernatants were processed. Isolation of the genome DNA from the urine pellets used the protocol of QIAamp DNA Mini Kit (QIAGEN). The DNA was eluted in 100 μL of AE buffer in the final step.

2.4. DNA Amplification

V3-V4 distinct regions of the 16S rRNA gene were amplified using specific primers 314F and 805R. The amplification process was performed with Phusion® High-Fidelity PCR Master Mix (New England Biolabs). The DNA was mixed with the same volume of 1× loading buffer (containing SYB green) and the products were visualised on 2% agarose gel. The samples with a bright main band between 400 bp-450 bp were chosen for further experiments. PCR products were mixed at equal density ratios. The mixed PCR products were purified with Qiagen Gel Extraction Kit (Qiagen, Germany). The libraries generated with NEBNext® UltraTM DNA Library Prep Kit for Illumina and quantified via Qubit and Q-PCR would be analysed by using Illumina platform.

2.5. 16S rRNA gene sequencing

Amplicon sequencing was performed using Illumina paired-end platform to generate 250 bp paired-end raw reads (Raw PE). The paired-end reads were assigned to the samples based on their unique barcodes and truncated by cutting off the barcode and primer sequences. The paired-end reads were merged using FLASH (V1.2.7), an analysis tool that was designed to merge paired-end reads when at least some of the reads overlapped the read generated from the opposite end of the same DNA fragment, and the splicing sequences were called raw tags (Magoč and Salzberg 2011). The quality of the raw tags was filtered under specific filtering conditions to obtain high-quality clean tags according to the Qiime (V1.7.0) quality control process (Caporaso et al. 2010). The tags were compared with the reference database using UCHIME algorithm to detect chimera sequences and then the chimera sequences were removed. The Effective Tags were finally obtained and could be used for subsequent analysis. The sequences analysis was performed by Uparse software (Uparse v7.0.1001) using all the effective tags (Edgar 2013).

2.6. Operational taxonomic units (OTUs) analysis

Sequences with ≥97% similarity were defined as the same OTUs. From each OTU, a representative sequence was screened for further annotation. For each representative sequence, Mothur software was performed against the SSUrRNA database of SILVA Database for species annotation at each taxonomic level (Threshold:0.8 1) (kingdom, phylum, class, order, family, genus, species) (Wang et al. 2007). The phylogenetic relationship of all OTUs representative sequences was compared with multiple sequences rapidly by using the MUSCLE (Version 3.8.31)
OTUs abundance information was normalized using a standard sequence number corresponding to the sample with the least sequences. Subsequent analysis of alpha diversity was performed based on this output normalized data. Alpha diversity was applied to analyze the richness and diversity of microbial community in each sample. The alpha diversity was applied to analyse complexity of biodiversity for a sample through 5 indices (Chao1, Shannon, Simpson, ACE, Good-coverage) which was calculated with QIIME (Version 1.7.0) and displayed with R software (Version 2.15.3).

3. Results and Discussion

3.1. Alpha diversity analysis

Amplicon was sequenced on Illumina paired-end platform, producing clean tags and effective tags after a series of treatments. Clean Tags represented tags after filtering, while effective Tags represented tags after filtering a chimera and could finally be used for subsequent analysis. Sequencing successfully produced 127,234 and 124,440 effective tags for female and male samples, respectively.

Sequencing of 16S rRNA gene has successfully identified all of the species that composed microbiome. The Good’s coverage was 100%, indicating a complete sampling of species. Species abundance, richness, evenness, and combinations thereof were measured with alpha diversity indices. Species richness, including unique species present, was rated with Chao1 index and the abundance-based coverage estimator (ACE). Meanwhile, the diversity of a community, considering evenness, richness, and abundance, was described by Shannon and Simpson indices (Hill 1973).

The Shannon diversity index in the female sample (3.780) was 3 times more than the male sample (1.367). Moreover, the Simpson’s index is also higher in the female sample than in the male (0.845 versus 0.446). This result shows that the female sample has a higher diversity of unseen species but also a high probability that two randomly selected sequences are of the same species.

To assign the species richness in the sample, Chao1 and ACE indices were calculated. Chao1 and ACE are indicators of species richness (the total number of species in a sample) sensitive to rare OTUs (singletons and doubletons). The result shows that the female sample has a higher value of both Chao1 (female: 200.000; male: 137.929) and ACE (female: 201.491; male: 138.668) indices (Table 1).

3.2. Composition of microbial community analysis

The histogram of relative abundance was constructed to visually refer to the taxa with a higher relative abundance and their proportion in different classification levels. Proteobacteria was the dominant phylum in both samples. Moreover, the abundance of this phylum is 99% in the microbial community of the male microbiome. In the female microbiome, apart from Proteobacteria (41%), the dominant phyla were Bacteroidetes (38%), Fusobacteria (12%) and Firmicutes (5%) (Figure 1).

The phylogeny of the top 100 most abundant genera of bacteria was further studied by phylogenetic analysis, including relative abundance (Figure 2). The most abundant genera were Burkholderia-Caballeronia-Caballeronia-Paraburkholderia (71%), followed by Prevotella (33%), Escherichia-Shigella (24%), Klebsiella (23%) and Sneathia (10%).

In order to quickly and intuitively display the microbial community composition and abundance information in the samples, a heatmap of taxonomic annotation corresponding to OTUs was constructed (Figure 3). The female microbiome was dominated by P. bivia (28%), E. coli (24%), S. sanguinegens (7%), and K. pneumoniae (4%). On the other hand, the male microbiome was dominated by K. pneumoniae (23%) and E. coli (2%).

To determine the predilection of the dominant microbiota, the nature of these species was investigated. Based on references, voided urine specimens were composed of pathogenic (P. bivia, S. sanguinegens, S. amnii, P. amnii, G. vaginalis group) and opportunistic (E. coli group, Enterobacteriaceae group, Citrobacter werkmanii, Citrobacter portucalensis, B. cepacia) species. Most of them are the causal agents of urogenital disruption, such as UTI and bacterial vaginosis or BV (Table 2).

### 3.3. Discussion

We collected 34 samples of voided urine from UTI patients and pooled the isolated DNA. The community profile of the urine samples was identified by 16S rRNA gene sequencing. These sequencing results from uncultivated bacteria described the composition of the microbiota inhabiting the urinary tract and genitalia. Since the microbiota in the samples might come from the lower urinary tract or post-urethral niches, the results described the urogenital microbiome (MacIntyre et al. 2017; Wolfe and Brubaker 2019). This term should be distinguished from the urinary microbiome (urobiome), which describes the microbial community in the urine (Brubaker and Wolfe 2017; Wolfe and Brubaker 2019).

Although a number of factors (age, hormonal status

<table>
<thead>
<tr>
<th>Sample</th>
<th>No. of clean tags</th>
<th>No. of effective tags</th>
<th>Observed species</th>
<th>Chao1</th>
<th>Shannon</th>
<th>Simpson</th>
<th>ACE</th>
<th>Good’s coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>150,824</td>
<td>127,234</td>
<td>197</td>
<td>200.000</td>
<td>3.780</td>
<td>0.845</td>
<td>201.491</td>
<td>1.000</td>
</tr>
<tr>
<td>Male</td>
<td>151,763</td>
<td>124,444</td>
<td>134</td>
<td>137.929</td>
<td>1.367</td>
<td>0.446</td>
<td>138.668</td>
<td>1.000</td>
</tr>
</tbody>
</table>
et al. could also cause perturbation to its composition (Karsens et al. 2016; Markowski et al. 2019; Qin et al. 2021). All the patients diagnosed with UTIs were recruited in this study. In this study, we would like first to identify the general composition of the microbiota by gender, allowing us to determine the structure of the urogenital microbiome, especially the bacteria that dominate the urine of UTI patients. In addition, we also recorded the patient’s medical history. Most of the UTI patients in this study experienced a variety of symptoms (fever, abdominal pain, cephalgia, and urinary retention) and complications (diabetes mellitus, benign prostatic hyperplasia or BPH, renal colic, and hydronephrosis).

We found that species that composed urogenital microbiome were mainly known as a pathogen of the genitalia organs. This may be due to the interconnectedness of urinary and genital microbiota (Thomas-White et al. 2018b; Perez-Carrasco et al. 2021). The crossstalk of these two may leads to the pathogenesis of UTI (Shreiner et al. 2015; Meštrović et al. 2020). A recent study shows a highly similar strains between vaginal and bladder microbiota, including E. coli, Streptococcus anginosus, L. iners, and L. crispatus (Thomas-White et al. 2018a).

Based on the 16S rRNA gene sequencing results, the most abundant phylum in both samples was Proteobacteria. This phylum significantly dominated the male urogenital microbiome, and with a lower proportion in female. This domination was followed by Bacteroidetes, Fusobacteria, Firmicutes and Actinobacteria (Figure 1).

Several investigations reported that sex is an important factor determining the microbiome composition. Sex hormones probably give direct impact. However, environment and stressors (such as infection or antibiotics) factors could also cause perturbation to its composition (Kim et al. 2020; Valeri and Endres 2021). This result is comparable with previous studies on the urinary microbiome, reporting Proteobacteria as the dominant phylum in the specimens of patients with urogenital disruption, including UTI (Christine et al. 2018; Moustafa et al. 2018; Thomas-White et al. 2018a). Moreover, this phylum is also detected in patients with bladder cancer and recipients of kidney transplantations, confirming its high tolerance as well as stability to immunosuppressive drugs (Bi et al. 2019; Bučević Popović et al. 2018; Colas et al. 2020).

After knowing the relative abundance of phyla and genus from both samples, we observed the microbial composition of both samples through Operational Taxonomic Units (OTUs) clustering. Sequences analysis was performed by Uparse software using all the effective tags. Sequences with ≥97% similarity were assigned to the same OTUs. Clustering analysis generated 488 and 321 OTUs from the female and male samples, respectively. Heatmap representation showed that K. pneumoniae and E. coli were the most abundant species in both samples (Figure 3.).

Gardnerella and Prevotella had significant roles as the virulence factors in BV as they formed biofilms (Castro et al. 2019; Kunze and Larsen 2019; Meštrović et al. 2020; Morrill et al. 2020). B. cepacia, a Gram-negative bacillus and aerobic opportunistic pathogen may play a role as a causative agent in community-acquired UTIs and nosocomial UTI. An infection of this pathogen is more common in adult male patients (Anggi et al. 2019; Du et al. 2021; Tubuh et al. 2019). Enterobacteriaceae group was found in both female and male microbiome but with a greater proportion in male. As a member of this family, K. pneumoniae is the most dominant species in the male microbiome. Other members of this family are the most common cause of UTI, including Citrobacter freundii, Enterobacter cloacae, E. coli, Enterobacter sp./Leclercia sp., Escherichia hermannii (Leski et al. 2016).

Sequencing has also detected specific opportunistic pathogens of the female reproductive tract: S. sanguinogens and S. amnii. Recent studies have reported a sig-

**FIGURE 1** A summary of the taxa relative abundance at the phylum level composing the female and male urogenital microbiome. Y-axis represented the relative abundance. “Others” represented a total relative abundance of the rest.
FIGURE 2 The top 100 genera composing the female and male urogenital microbiome with a phylogenetic tree showing the relationship among these genera. The colors of the branches represented corresponding phyla. The relative abundance of each genus was shown by the outer bar plots.

FIGURE 3 The heatmap of relative abundance indicating the most abundant bacterial species in the two samples. Red and blue respectively indicates high and low relative abundance.

significant association between *S. sanguinegens* and spontaneous abortion (Fettweis et al. 2019; Seo et al. 2017). Meanwhile, *S. amnii* has been known as an opportunistic pathogen that causes infection during pregnancy and in the post-partum period. The existence of *Sneathia* in the sample may indicate pregnancy in patients. These pathogens produce a cytotoxic exotoxin that can attack cells of the fetal membrane (Duployez et al. 2020; Gentile et al. 2020). As these pathogens were then detected in the voided urine, an invasion mechanism may have occurred from the genital to the urinary tract.

*B. cepacia* detected in the male specimen was a common pathogen in a male urogenital tract and a causal agent of catheter-associated UTI (Tubuh et al. 2019). In addition, this species also causes nosocomial UTI from the use of a contaminated anesthetic gel (Du et al. 2021). The existence of *B. cepacia* as the dominant taxon in the male microbiome brought out the allegation that UTI experienced by the patients were obtained during hospital treatments. Further studies are needed to find out the potential sources of nosocomial UTIs.

As the most common uropathogen, *E. coli* was de-
Nadifah et al.

*Indonesian Journal of Biotechnology* 27(3), 2022, 142-150

TABLE 2 Pathogenicity and clinical roles of 10 significate OTUs

<table>
<thead>
<tr>
<th>OTUs</th>
<th>Pathogenicity</th>
<th>Clinical roles</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Prevotella bivia</em></td>
<td>Pathogenic</td>
<td>BV</td>
<td>Alauzet et al. (2019); Veloo et al. (2018)</td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td>opportunistic</td>
<td>BV, UTI</td>
<td>Brannon et al. (2020); Chagneau et al. (2020)</td>
</tr>
<tr>
<td><em>Sneathia sanguinegens</em></td>
<td>Pathogenic</td>
<td>Preterm-birth, spontaneous abortion</td>
<td>Fettweis et al. (2019); Seo et al. (2017)</td>
</tr>
<tr>
<td><em>Enterobacteriaceae</em></td>
<td>opportunistic</td>
<td>UTI</td>
<td>Tayh et al. (2019)</td>
</tr>
<tr>
<td><em>Gardnerella</em></td>
<td>Pathogenic</td>
<td>Urethritis, amnionitis, BV, post-partum bacteremia</td>
<td>Duployez et al. (2020); Fettweis et al. (2019); Gentile et al. (2020)</td>
</tr>
<tr>
<td><em>Prevotella amnii</em></td>
<td>Pathogenic</td>
<td>BV</td>
<td>Alauzet et al. (2019); Veloo et al. (2018)</td>
</tr>
<tr>
<td><em>Citrobacter</em></td>
<td>opportunistic</td>
<td>UTI</td>
<td>Christine et al. (2018); Liu et al. (2020)</td>
</tr>
<tr>
<td><em>Citrobacter</em></td>
<td>opportunistic</td>
<td>UTI</td>
<td>Christine et al. (2018); Liu et al. (2020)</td>
</tr>
<tr>
<td><em>Gardnerella</em></td>
<td>Pathogenic</td>
<td>BV</td>
<td>Castro et al. (2019); Kunze and Larsen (2019)</td>
</tr>
<tr>
<td><em>Burkholderia</em></td>
<td>opportunistic</td>
<td>UTI</td>
<td>Anggi et al. (2019); Du et al. (2021)</td>
</tr>
</tbody>
</table>

Pathogens, especially anaerobes, are undetectable and difficult to culture. E. coli and other aerobe pathogens are easily detected with the conventional culture method, making them well known as the most common cause of UTI (Dune et al. 2017; Yildrim et al. 2020). On the other hand, many pathogens, especially anaerobes, are undetectable and difficult to culture (Moustafa et al. 2018). A new approach of bacterial identification based on 16S rRNA gene sequencing successfully identifies microbiota composition, even with some fastidious bacteria. Therefore, urology practices should consider using this technique for the prevention, diagnosis and appropriate treatments of UTI (Dixon et al. 2020; Mouraviev and McDonald 2018).

Naturally, body niches are made up of two types of ecological community: driven by the dominant taxa and without any dominant taxa. Although they have different compositions, both have specific mechanisms to keep community’s functions and stability. However, an imbalance may occur in urinary microbiota because of some temporal dynamics of the host (immune system, hormones, sex activities). Alterations in the microbial communities may occur in urinary disorders, including UTI (Greenbaum et al. 2019; Kim and Park 2018; Magistro and Stief 2019).

Understanding the urogenital microbiome is not only about the composition of the microbiota but also about the ecological aspects. Metagenomic sequencing technology has been beneficial in the studies of the human microbiome. The results of this study indicated the existence of a biodiversity of bacteria that composed the urogenital microbiome of UTI patients. Using this technology, the structure of the bacterial community in healthy people and patients with urogenital disorders can be identified. Studies on the urogenital microbiome should be further developed to prevent and properly treat patients with infectious diseases of the urogenital system. Further research should be conducted using negative controls to determine background contamination as well as sequencing of a mock community for better accuracy of the entire pipeline. On the other hand, replications should also be used for more reproducible research.

3.4. Data availability

The 16S rRNA gene amplicon data set is available at NCBI with SRA accession number PRJNA686823.

4. Conclusions

16S rRNA gene sequencing used in this study successfully uncovered the composition of the urogenital microbiome in UTI patients which might not have been possible with conventional culture methods. Most dominant species are pathogenic or opportunistic and have been known as the causal agents of urogenital disruption, including UTI. *K. pneumoniae* and *E. coli* were the most abundant species in both female and male microbiomes.
Acknowledgments

We thank the Indonesian Ministry of Education and Culture for providing funds for this work through the scheme of Doctoral Dissertation Research Grant No. 3064/UN1.DITLIT/DIT-LIT/PT/2020.

Authors’ contributions

WTA, BSD designed the study. FN, ER carried out the laboratory work. FN, BSD, ER analyzed the data. FN, WTA, BSD, ER wrote the manuscript. All authors read and approved the final version of the manuscript.

Competing interests

There were no competing interests during the research and in the preparation of this manuscript.

References


Meštrović T, Matijašić M, Perić M, Čipčić Paljetak H, Barešić A, Verbanac D. 2020. The role of gut, vagi-

Nadifah et al. Indonesian Journal of Biotechnology 27(3), 2022, 142-150