2016

Genome sequences of nine gram-negative vaginal bacterial isolates

Grace E. Deitzler  
*Washington University School of Medicine in St. Louis*

Maria J. Ruiz  
*Washington University School of Medicine in St. Louis*

Wendy Lu  
*Washington University School of Medicine in St. Louis*

Cory Weimer  
*Washington University School of Medicine in St. Louis*

SoEun Park  
*Washington University School of Medicine in St. Louis*

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Recommended Citation  
Deitzler, Grace E.; Ruiz, Maria J.; Lu, Wendy; Weimer, Cory; Park, SoEun; Robinson, Lloyd S.; Hallsworth-Pepin, Kymberlie; Wollam, Aye; Mitreva, Makedonka; Lewis, Warren G.; and Lewis, Amanda L., "Genome sequences of nine gram-negative vaginal bacterial isolates." *Genome Announcements* 4, 5. e00889-16. (2016).  
https://digitalcommons.wustl.edu/open_access_pubs/5631
Authors
Grace E. Deitzler, Maria J. Ruiz, Wendy Lu, Cory Weimer, SoEun Park, Lloyd S. Robinson, Kymberlie Hallsworth-Pepin, Aye Wollam, Makedonka Mitreva, Warren G. Lewis, and Amanda L. Lewis

This open access publication is available at Digital Commons@Becker: https://digitalcommons.wustl.edu/open_access_pubs/5631
Reproductive and urinary tract infections are a major cause of morbidity and mortality for women worldwide (1, 2). Bacterial vaginosis (BV) is an imbalance of the vaginal microbiota that is associated with higher risks of sexually transmitted infections, urinary tract infections, and poor health outcomes among pregnant women (3–10). Women with BV have few lactic acid-producing bacteria (lactobacilli) and high levels of fastidious anaerobic bacteria. A variety of species within the Bacteroidetes and Fusobacteriae (among other taxa) have been isolated from women with BV, often from sites in the upper reproductive tract (e.g., placenta and amniotic fluid) (5, 11–14). Despite the widespread health complications associated with BV, its etiology is poorly characterized, and current treatment options are often met with recurrences (15).

Urinary tract infection (UTI) is another recurrent urogenital condition that is common among women and associated with poor pregnancy outcomes (1). Escherichia coli is the most common cause of UTI (16), and there are many dozens of available isolates and genomes of E. coli available for study. Citrobacter and Klebsiella spp. are less common etiologic agents of UTI. It is thought that the vagina can sometimes act as a reservoir for uropathogens; however, few vaginal isolates of uropathogenic bacterial species are available as fully sequenced deposited isolates. The lack of reference strains and corresponding reference genomes of urogenital bacteria hinders research progress aimed at understanding how bacteria cause infection in the genital and urinary tracts. Here, we present annotated genome sequences of nine Gram-negative vaginal isolates, which have been made available to the research community through BEI Resources.

Vaginal swabs were collected from nonpregnant and pregnant women according to Washington University institutional review board (IRB)-approved protocols 201108155 and 201103082. An aerobic vaginal swabs from reproductive-age pregnant and nonpregnant women were streaked onto agar medium and cultivated overnight. The resulting colonies were resuspended in normal saline and screened for contamination. The gene annotation process included gene prediction using GeneMark and Glimmer3 ab initio and evidence-based (BLAST) predictions. Coding sequences were identified using GeneMark and Glimmer3 and the GeneMark-Glimmer pipeline (17), with hash sizes of 31, 33, and 35 after downsizing the input data to 100X coverage. Postassembly, we set the minimum length for contigs to 200 bp, ran an internal core gene screen on the assembly (as defined by the Human Microbiome Project [HMP] [18]), removed adapters, trimmed low-quality regions, and screened for contamination. The gene annotation process included generating both ab initio and evidence-based (BLAST) predictions. Coding sequences were identified using GeneMark and Glimmer3 (19, 20). Loci were then defined by clustering predictions with the same reading frame. We evaluated predictions using the NR and Pfam databases (21) and resolved overlaps between adjacent coding genes. Intergenic regions not spanned by GeneMark and Glimmer3 were subjected to a BLAST search against NCBI’s nonredundant gene database.

### TABLE 1 Strain names and accession numbers

<table>
<thead>
<tr>
<th>Species</th>
<th>Strain name</th>
<th>BEI catalog no.</th>
<th>Nucleotide sequence accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citrobacter freundii</td>
<td>GED7749C</td>
<td>HMS-1280</td>
<td>LRPR000000000</td>
</tr>
<tr>
<td>Citrobacter koseri</td>
<td>GED7778C</td>
<td>HMS-1288</td>
<td>LRPS000000000</td>
</tr>
<tr>
<td>Fusobacterium sp.</td>
<td>CMW8396</td>
<td>HMS-1274</td>
<td>LRPX000000000</td>
</tr>
<tr>
<td>Fusobacterium nucleatum</td>
<td>MJR7757B</td>
<td>HMS-1289</td>
<td>LRPY000000000</td>
</tr>
<tr>
<td>Klebsiella pneumoniae</td>
<td>MJR8396D</td>
<td>HMS-1265</td>
<td>LRRQ000000000</td>
</tr>
<tr>
<td>Prevotella bivia</td>
<td>GED7880</td>
<td>HMS-1270</td>
<td>LTAG000000000</td>
</tr>
<tr>
<td>Prevotella bivia</td>
<td>GED7760C</td>
<td>HMS-1286</td>
<td>LRRQ000000000</td>
</tr>
<tr>
<td>Prevotella corporis</td>
<td>MJR7716</td>
<td>HMS-1294</td>
<td>LRGQ000000000</td>
</tr>
<tr>
<td>Proteus mirabilis</td>
<td>GED7834</td>
<td>HMS-1271</td>
<td>LSGS000000000</td>
</tr>
</tbody>
</table>
(NR) database, and predictions were generated based on protein alignments. tRNA genes and noncoding RNA genes were found using tRNAscan-SE, RNAmmer, and Rfam (22–24). The final gene set was annotated for metabolic pathway predictions using KEGG (25), subcellular localization using PSORTb (26), and functional domain associations using InterProScan (27).

**Accession number(s).** Nucleotide sequences have been deposited in GenBank under the accession numbers listed in Table 1. The sequences described in this paper are the first versions. We have also made the strains available to the research community by depositing them with the Biodefense and Emerging Infections (BEI) Research Resource Repository (see BEI numbers in Table 1).

**ACKNOWLEDGMENTS**

We acknowledge Justin C. Perry for technical assistance, the Women’s and Infant’s Health Specimen Consortium, and especially Michele Landeau. We also thank Jeff Peipert, Jenifer Allsworth, and Jennifer Bick of the Contraceptive CHOICE project for contributions of specimens from nonpregnant women.

**FUNDING INFORMATION**

This work, including the efforts of Makedonka Mitreva, was funded by HHS | National Institutes of Health (NIH) (U54HG004968). This work, including the efforts of Warren G. Lewis and Amanda L. Lewis, was funded by HHS | NIH | National Institute of Allergy and Infectious Diseases (NIAID) (NIH R01 AI114635). This work, including the efforts of Amanda L. Lewis, was funded by HHS | NIH | National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (P50 DK064540-11). This work, including the efforts of Grace E. Deitzler, was funded by HHS | NIH | National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (P20DK097798). This work, including the efforts of Amanda L. Lewis, was funded by Burroughs Wellcome Fund (BWF). This work, including the efforts of Amanda L. Lewis, was funded by March of Dimes Foundation (March of Dimes). This work was funded by a March of Dimes Basil O’Connor award (A.L.L.), NIH Specialized Centers of Research grant P50 DK064540-11 (Scott Hultgren, A.L.L.; project II), and the Human Microbiome Project grant U54HG004968. Summer stipends were provided by the Planning Center for Research in Benign Urology Summer Research Program (P20DK097798 to G.E.D.) and the Washington University in St. Louis Summer Research Program with the Center for Global Health & Infectious Disease (CGHID)/Institute for Public Health (3054–93396B to M.J.R.).

**REFERENCES**


