

Simultaneous profiling of sexually transmitted bacterial pathogens, microbiome, and concordant host response in cervical samples using whole transcriptome sequencing analysis

Catherine M. O'Connell^{1,*,#}, Hayden Brochu^{2,#}, Jenna Girardi¹, Erin Harrell², Aiden Jones², Toni Darville¹, Arlene C. Seña³ and Xinxia Peng^{2,4,*}

¹ Department of Pediatrics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA.

² Department of Molecular Biomedical Sciences, North Carolina State University, Raleigh, North Carolina, USA.

³ Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA.

⁴ Bioinformatics Research Center, North Carolina State University, Raleigh, North Carolina, USA.

These authors contributed equally to this work.

* Corresponding Authors:

Catherine O'Connell, PhD, University of North Carolina at Chapel Hill, 8340B MBRB, CB# 7509, 111 Mason Farm Road, Chapel Hill, NC 27599-7509; phone (919) 962-3422; fax:(919) 966-7277; E-mail: catherine.oconnell@unc.edu;

Xinxia Peng, PhD, North Carolina State University, Raleigh, NC 27695; phone (919) 515-4481; fax: (919)-513-6464; E-mail: xpeng5@ncsu.edu

ABSTRACT Pelvic inflammatory disease (PID) is a female upper genital tract inflammatory disorder that arises after sexually transmitted bacterial infections (STI). Factors modulating risk for reproductive sequelae include co-infection, microbiota, host genetics and physiology. In a pilot study of cervical samples obtained from women at high risk for STIs, we examined the potential for unbiased characterization of host, pathogen and microbiome interactions using whole transcriptome sequencing analysis of ribosomal RNA-depleted total RNAs (Total RNA-Seq). Only samples from women with STI infection contained pathogen-specific sequences (3 to 38% transcriptome coverage). Simultaneously, we identified and quantified their active microbial communities. After integration with host-derived reads from the same data, we detected clustering of host transcriptional profiles that reflected microbiome differences and STI infection. Together, our study suggests that total RNA profiling will advance understanding of the interplay of pathogen, host and microbiota during natural infection and may reveal novel, outcome-relevant biomarkers.

doi: xxx

Received originally: 24.09.2018;

In revised form: 03.12.2018,

Accepted 05.12.2018,

Published 24.01.2019.

Keywords: RNA-seq, chlamydia, gonorrhoea, natural infection, microbiome.

Abbreviations:

CT – *Chlamydia trichomatis*,

MG – *Mycoplasma genitalium*,

NG – *Neisseria gonorrhoeae*,

PID – pelvic inflammatory disease,

rRNA – ribosomal RNA,

STI – sexually transmitted infection,

TV – *Trichomonas vaginalis*.

INTRODUCTION

Pelvic inflammatory disease (PID) continues to pose great risk to the reproductive health of women worldwide [1, 2]. In addition to the cost of treating acute illness, its potentially devastating long-term sequelae can persist throughout a woman's reproductive years. PID occurs when sexually transmitted infections (STIs) ascend from the cervix to the uterus and oviducts, resulting in endometritis and salpingitis. Infection with *Chlamydia trachomatis* (CT), *Neisseria gonorrhoeae* (NG), or both, may lead to PID and associated morbidities including infertility, ectopic pregnancy, and chronic pelvic pain. *Mycoplasma genitalium* (MG) has also emerged as a PID-causing pathogen [3-5]. Currently, it

is impossible to predict an individual's risk for developing disease with infection so biomarkers that identify women at elevated risk for reproductive morbidities are needed. Transcriptional profiling that compared systemic responses of women with chlamydial PID to women with local cervical infection, identified specific, disease-associated inflammatory pathways [6] and revealed how NG co-infection depressed developing immunity. We also defined a blood-borne, 21 gene biomarker panel that diagnoses endometritis in asymptotically-infected women with high chlamydial burden [7]. However, our work, and that of others, has exposed knowledge gaps regarding the impact of pathogen burden [8, 9], co-infection [8, 10-12], vaginal

TABLE 1. Demographic, clinical and microbiological characteristics of pilot samples (N=10).

Patient Identifier	Age	Bacterial STI Pathogen (genome equivalents/swab)			TV Diagnosis*
		CT	NG	MG	
1	19	1.08E+05	-	-	positive
2	27	-	-	-	negative
3	28	-	-	-	negative
4	23	-	2.27E+06	-	positive
5	23	-	-	-	negative
6	22	-	-	3.66E+03	positive
7	25	-	-	-	negative
8	21	-	-	2.50E+02	negative
9	21	-	-	-	negative
10	20	-	-	-	negative

* Assessed via wet mount.

- not detected by NAAT and/or qPCR using pathogen-specific primers.

microbiota [13, 14] and oral contraceptives [8] on infection outcome and their potential to serve as additional sources of outcome-relevant biomarkers.

In this pilot study, we examined the potential for unbiased characterization of host, pathogen and microbiome interactions at a molecular level and identification of novel, outcome-relevant biomarkers in a single, easily obtained, clinical specimen using whole transcriptome sequencing analysis of ribosomal RNA (rRNA) depleted total RNAs (Total RNA-Seq).

RESULTS

Baseline characteristics of participants

A total of ten cervical specimens were available for this pilot study (Table 1) from young women with a median age of 22.5 years (range, 19-28 years). Four participants were tested positively for STI pathogens (40%). Two women had NAAT-confirmed CT or NG infection with *Trichomonas vaginalis* (TV) coinfection (Table 1). One other participant was diagnosed with TV, and we subsequently detected MG coinfection via qPCR. MG genomic DNA was also detected in one other sample.

Simultaneous profiling of STI pathogens, microbiome, and host response using Total RNA-Seq

An average of 26 million +/- 2.8 million Total RNA-Seq short read pairs were obtained for each sample (Table 2) and assigned as human, bacterial or TV origin after removal of reads corresponding to residual rRNA sequences (0.1- 2.5%) that persisted despite rRNA depletion. Distribution of raw counts for each classification varied across samples. We exploited the genome coverage derived from aligning short reads to a large, non-redundant collection of more than 13,000 bacterial genomes, to identify and quan-

tify active microbial communities present in the cervical samples. We noted an inverse correlation between total numbers of bacterial- and human-derived reads, indicating a possible tradeoff between coverage of the human transcriptome and the microbial transcriptomes represented in individual libraries (Supplementary Figure 1A). Improved resolution of human genes was not detected once samples exceeded 10 million human-derived reads, suggesting that human transcriptome coverage was saturated for many of these samples (Supplementary Figure 1B). Sequencing depths obtained for cervico-vaginal microbiota did not appear to correlate with the number of species detected, their diversity or abundance (Supplementary Figure 1C-D) indicating that sample composition influenced measurement of human and microbiome transcriptomes more strongly than sequencing depth for these cervical samples.

Samples obtained from individuals diagnosed with TV via wet mount contained high percentage reads for TV transcripts (4 to 40%) but all remaining samples contained some TV-assigned reads, ranging from <0.01 to 0.22% of the total. Only samples from women with CT, NG or MG infection contained pathogen-specific reads (Figure 1A, Supplemental Tables 1-4), with sequences that mapped to common commensals of the human female genital tract such as *M. hominis* and *Ureaplasma urealyticum* [15]. Assigning pathogen-specific reads to annotated genes, a total of 2,285 reads were aligned to 329 protein encoding genes, tRNAs or non-coding RNAs for CT (38.1% transcriptome coverage) (Supplemental Table 1) and 483 reads to 217 GC targets (10.9% coverage) (Supplemental Table 2). For patient samples containing MG, we mapped 60 and 36 reads to 20 and 19 targets respectively (~3.7% coverage/sample). We observed limited overlap, with only *rnpB*, a noncoding RNA and *tufA*, encoding the elongation factor Ef-Tu, com-

TABLE 2. Summary of read distribution from cervical RNA libraries after sequencing.

Patient Identifier	Total	rRNA	rRNA (%)	Human	Human (%)	Bacterial	Bacterial (%)	TV	TV (%)
1	24,473,792	332,825	1.36	8,443,636	34.5	4,386,302	17.92	7,804,744	31.89
2	22,111,682	319,791	1.45	4,114,718	18.61	15,215,299	68.81	28,842	0.13
3	31,615,660	752,169	2.38	7,628,174	24.13	16,742,345	52.96	69,082	0.22
4	25,584,080	167,353	0.65	1,271,025	4.97	10,533,772	41.17	1,138,132	4.45
5	26,899,897	633,674	2.36	21,255,809	79.02	3,066,007	11.4	111	<0.01
6	23,333,959	580,004	2.49	11,471,258	49.16	1,090,474	4.67	9,272,575	39.74
7	24,492,681	35,422	0.14	455,956	1.86	13,542,434	55.29	254	<0.01
8	27,012,744	263,825	0.98	14,769,365	54.68	8,165,858	30.23	421	<0.01
9	25,347,522	324,514	1.28	4,319,634	17.04	18,043,328	71.18	7757	0.03
10	29,199,143	415,861	1.42	14,345,092	49.13	6,614,425	22.65	616	<0.01

mon between MG-containing samples (Supplemental Tables 3-4).

To explore interplay between host response, and STI pathogens, we selected the subset of human genes exhibiting the most variable expression (top 500 here) across eight cervical samples for further investigation (Supplementary Table 5). Two samples were excluded because the human-associated reads were insufficient (**Table 2**). Functional enrichment analysis of these 500 selected genes revealed multiple pathways were over-represented ($p < 0.05$) (Supplementary Table 6). Top pathways included HIPPO signaling and epithelial adherens junction signaling, important for epithelial cell health and proliferation. Other pathways included those that contain kalikrein-related peptidases and estrogen receptor signaling, reflecting specific pathways characteristic of cervical epithelium [17]. Interestingly, pathways previously identified as important for establishment of chlamydial infection (Actin Nucleation by ARP-WASP Complex) [18] and subsequent growth (p53 signaling) [19, 20] were also over-represented as were genes involved in granulocyte adhesion and diapedesis, suggesting innate inflammatory signaling activation in some samples.

Recent studies have linked vaginal dysbiosis and elevated bacterial diversity to increased vulnerability to STI such as CT [13, 14] and HIV [21-23]. Here we also observed differences in the microbial relative abundance profiles for each sample (**Figure 1B** and Supplemental Table 7). More interestingly, when eight of these samples were analyzed in the context of the human transcripts also present, we detected clustering that reflected microbiome differences (**Figure 1C**). Samples dominated by *Lactobacillus sp.* clustered away from those in which *Gardnerella* and *Prevotella*, frequently observed with dysbiosis, predominated. The final cluster included samples from women infected with CT or MG. The two samples lacking sufficient human reads,

including the specimen obtained from the NG-infected patient, were not assessed (**Table 2**).

DISCUSSION

Understanding the mechanisms that promote development of genital tract disease and damage after STI infection has proven challenging. CT is an obligate intracellular pathogen with a complex developmental cycle [24] while NG [25] and MG [15] are highly fastidious in their nutritional requirements. Animal models are few [26-28], and large human studies are required to investigate the contribution of host genetic diversity to outcome severity [29]. Polymicrobial communities of anaerobes have been associated with bacterial vaginosis and elevated risk for a variety of reproductive morbidities [30]. In contrast, *Lactobacillus crispatus*-dominated community state types play a protective role against chlamydial infection [13, 14, 31]. It is also possible that uncharacterized PID pathogens hide within the cervico-vaginal microbiome [32]. Our pilot study suggests that total RNA profiling can serve as a valuable tool to investigate the interplay of pathogen, host and microbiota during natural infection.

The strengths of this approach are many. Vaginal and cervical swab samples are easily obtained in a clinical setting. The RNA/DNA preservation solution stabilized nucleic acids [33, 34] without need for immediate transfer to -80°C . Quantifying pathogen/co-pathogen abundance provides a route for determining primary drivers of inflammation and outcome. For example, the patient sample with very low levels of MG contained abundant TV transcripts suggesting that the latter was the predominating influence on her innate response. Information regarding gene expression during natural infection by bacterial STI pathogens is lacking [35, 36, 38, 39] but we were pleased to detect expression of between 3-30% of the relevant transcriptome without specific or targeted enrichment. These included path-

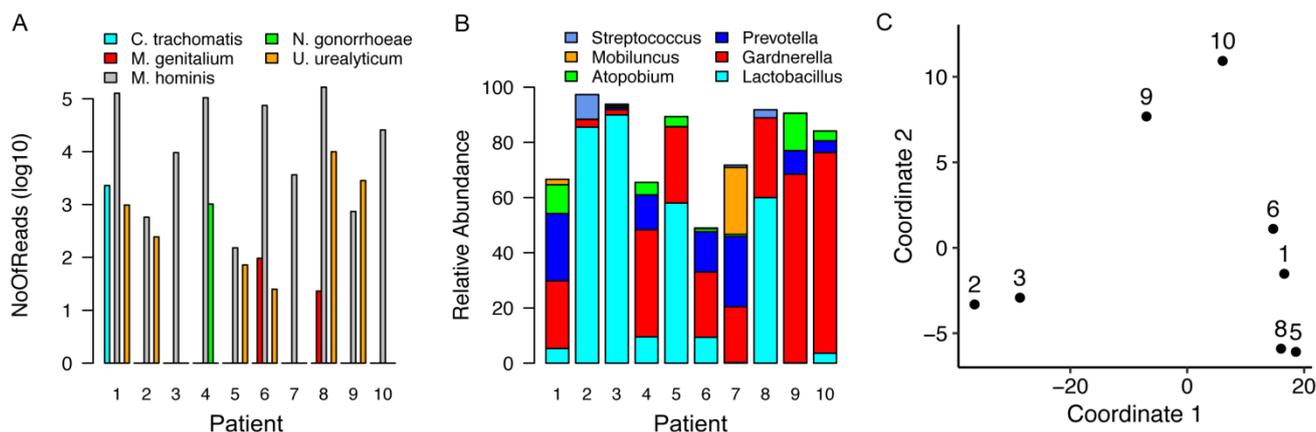


FIGURE 1: Characterization of cervical microbial communities using Total RNA-Seq. (A) CT, GC and genital mycoplasma-derived reads detected in cervical RNA libraries. **(B)** Relative abundances of dominant bacterial genera in each sample. **(C)** Clustering of samples based on human gene expression profiles measured in the same samples, visualized by multidimensional scaling. Patient 1 was infected with CT, patients 6 and 8 respectively were infected with MG and are present in the same cluster. Patients 4 (NG) and 7 are not graphed because a low percentage of reads in these samples aligned to the human genome.

ways critical for central metabolism, surface proteins and non-coding RNAs. Future studies may reveal the influence of host inflammatory or physiologic effects on their regulation and identify novel targets for antimicrobials or suppression of virulence effector expression.

Pathway analysis of the highly variable genes used in clustering host transcriptome profiles of the women in the study revealed significantly enriched pathways reflective of human cervical epithelium. These included pathways important for cell maintenance and proliferation (supplementary data) and pathways that seem of particular importance in the cervical milieu. For example, expression of a family of kalikrein-related peptidases in the epithelial cells of the lower genital tract of women is hormonally regulated [17] to control desquamation and remodeling of cervical mucus. Some members of this family may also contribute to innate defenses by processing/activating antimicrobial α and defensin-1 β proteins in mucus [17]. Chlamydial infection may impact epithelial cell health directly by down-regulating adherens junction molecule nectin-1 [37], disrupting epithelial tissue homeostasis [38] and inducing epithelial-mesenchyme transition [39] and we observed that these pathways, as well as pathways promoting chlamydial entry into cells [18] and subsequent growth [19, 20] were over-represented amongst the most dynamic human transcripts in our small sample population. The impact of MG infection on cervical epithelial responses has yet to be investigated *in vivo*, but *in vitro* chronic infection of human cervical epithelial cells elicits inflammatory cytokines [40, 41]. These very preliminary observations suggest that future studies exploiting Total RNA-Seq have the potential to provide unique insights into dynamic host-pathogen interactions at this infection portal.

Integrating transcriptional data from pathogen(s), microbiota and local host response revealed differential host-microbiome relationships between women with *Lactobacil-*

lus sp. versus those with *Gardnerella/Prevotella*. Furthermore, women with STI infections clustered separately, possibly reflecting local pathogen-specific inflammatory responses and paralleling systemic responses we have detected in blood [6]. Recent advances in analysis of high-dimensional data will assist interpretation of these observations in the context of carefully designed studies with sufficient power to reveal and validate key pathways and interactions that drive tissue damage and poor reproductive outcomes.

This pilot study also identified potential challenges analyzing clinical specimens. Infection severity is likely to influence representation of pathogen-derived transcripts. We did not have enough samples to compare the impact of lower or higher levels of STI infection on transcriptome coverage. Over-representation of TV transcripts as a result of high levels of co-infection might have also suppressed detection of pathogen sequences. It is possible that superior coverage will be obtained from mono-infected samples. Furthermore, in two of ten samples we did not recover sufficient host-derived material to be able to assess this compartment. Cytobrushes may improve collection of sufficient host material for extraction and analysis [42-44]. Depth of coverage for pathogen genes was low. Improving sensitivity may require a large number of samples for aggregation via bioinformatic analysis or specific, targeted enrichment of pathogen-derived transcripts [45-47]. High sequence conservation of genes encoding core microbial metabolic functions may introduce mis-assignment of pathogen- or microbiome-derived reads for individual genes. Our analysis is also likely biased toward currently available bacterial genomes that may not reflect the full diversity of genital tract microbiota. Future studies optimizing key technical procedures and with increased sample size can address or mitigate many of these challenges, to offer improved understanding of the interplay between STI

pathogens, local microbial communities and host responses in development of genital tract disease. Such information will accelerate identification of biomarkers that identify women with subclinical STI-induced reproductive morbidities for more intensive screening or for evaluation of novel therapeutics or vaccines.

MATERIALS AND METHODS

Study population

Symptomatic women between the ages of 18-29 years of age who presented for routine evaluation to the public STI clinic located in Durham, North Carolina, were recruited for the study. Sole exclusion criterion was use of antimicrobials in the four weeks prior to sampling. The Institutional Review Board for Human Subject Research at the University of North Carolina approved the study (#17-2175), and all participants provided written informed consent prior to sample collection. Vaginal and cervical specimens were collected from female participants during pelvic examinations in an ongoing pilot study intended to develop a new diagnostic/prognostic clinical tool for STIs. Participants were tested for TV by wet mount microscopy [48] and CT and NG using nucleic acid amplification tests, (Aptima Combo 2; Hologic, Marlborough, MA).

Sample collection and processing

Cervical swab specimens were collected and stored at -80°C in tubes containing zircon beads and preservative solution (RNA/DNA Shield, Zymo Research, Irvine, CA). Thawed samples were agitated vigorously using a FastPrep Beadbeater (MP Biomedicals, Santa Ana, CA) to release adherent material then DNA and total RNA were simultaneously extracted using a Quick DNA/RNA™ nucleic acid isolation kit (Zymo Research) with on-column DNase I treatment of the RNA.

Pathogen quantitation by qPCR

Cervical pathogen burden (CT and MG) for infected study participants was estimated via quantitative PCR [49, 50] using the genomic DNA extracted from swab eluates as template. Similarly, primers directed toward a conserved locus [51] (VT05_017777 F 5' GTCTCGGACAGTCATTCCTCA 3' and VT05_017777 R 5' CGGAGCAAGACCAACAGAA 3') were used to quantify NG load.

Library prep and sequencing

500 ng total RNA was depleted using Illumina Ribo-Zero Epidemiology kit. Libraries were prepared using KAPA RNA HyperPrep Kit with RiboErase (HMR) (12 cycles PCR amplification) and dual indexed with a KAPA Dual-Indexed Adapter Kits (1.5µM final concentration). Libraries were sequenced using a Nextseq High-Output Kit utilizing 76bp paired end reads.

Data processing

Raw reads were demultiplexed and quality trimmed using the standard Illumina bcl2fastq conversion software. Reads were aligned to a custom human rRNA index and then to the hg38 genome with GENCODE v25 primary assembly annotation using STAR v2.5.2b in quantmode [52]. The remaining unmapped reads were then aligned using Bowtie2 v2.3.4 [53] to

a pre-built index consisting of over 13,000 bacterial whole genome reference sequences [54]. SLIMM v0.2.2 was used to assign uniquely mapped reads using species-level taxonomic annotation [54]. HTSeq count v0.9.1 was used to generate gene counts for bacterial species of interest, namely CT, NG, and MG [55]. Finally, the leftover reads that did not map to the collection of bacterial genomes were aligned to the TV reference genome using Bowtie2 v2.3.4.

MDS and IPA pathway analysis

Multidimensional scaling (MDS) analysis was performed using human gene read counts. Counts were normalized and log transformed using DESeq2 [56] and the top 500 variable genes were selected based on their variance across the eight samples included in the analysis. Two samples (4 and 7) were excluded due to insufficient human expression. Functional enrichment analysis of these 500 genes was conducted using Ingenuity Pathway Analysis (QIAGEN Inc., <https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis>), which reported pathways detected and assessed whether they were significantly over-represented using right-tailed Fisher Exact Tests.

All sequence data arising from this study has been deposited into NCBI's GEO under accession GSE120192.

ACKNOWLEDGEMENTS

This work was supported by the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health, through Grant Award Number UL1TR002489, via pilot funding to COC.

We thank the women who agreed to participate in this study; Adriane Osborne and Anjeleigh Partridge at the Durham County Department of Public Health STD Clinic for their assistance with recruitment and sample collection.

SUPPLEMENTAL MATERIAL

All supplemental data for this article are available online at www.microbialcell.com.

CONFLICT OF INTEREST

None.

COPYRIGHT

© 2019 O'Connell *et al.* This is an open-access article released under the terms of the Creative Commons Attribution (CC BY) license, which allows the unrestricted use, distribution, and reproduction in any medium, provided the original author and source are acknowledged.

Please cite this article as: Catherine M. O'Connell, Hayden Brochu, Jenna Girardi, Erin Harrell, Aiden Jones, Toni Darville, Arlene C. Seña and Xinxia Peng (2019). Simultaneous profiling of sexually transmitted bacterial pathogens, microbiome, and concordant host response in cervical samples using whole transcriptome sequencing analysis. *Microbial Cell: in press*.

REFERENCES

- Workowski KA, Bolan GA, Centers for Disease Control and Prevention (2015). Sexually transmitted diseases treatment guidelines, 2015. **MMWR Recomm Rep** 64(RR-03): 1-137. doi: 10.1093/cid/civ771
- Rein DB, Kassler WJ, Irwin KL, Rabiee L (2000). Direct medical cost of pelvic inflammatory disease and its sequelae: decreasing, but still substantial. **Obstet Gynecol** 95(3): 397-402. doi: 10.1097/00006250-200003000-00016
- Bjartling C, Osser S, Persson K (2012). *Mycoplasma genitalium* in cervicitis and pelvic inflammatory disease among women at a gynecologic outpatient service. **Am J Obstet Gynecol** 206(6): 476 e471-478. doi: 10.1016/j.ajog.2012.02.036
- McGowin CL, Anderson-Smits C (2011). *Mycoplasma genitalium*: an emerging cause of sexually transmitted disease in women. **PLoS Pathog** 7(5): e1001324. doi: 10.1371/journal.ppat.1001324
- Wiesenfeld HC, Manhart LE (2017). *Mycoplasma genitalium* in women: current knowledge and research priorities for this recently emerged pathogen. **J Infect Dis** 216(suppl_2): S389-S395. doi: 10.1093/infdis/jix198
- Zheng X, O'Connell CM, Zhong W, Nagarajan UM, Tripathy M, Lee D, Russell AN, Wiesenfeld H, Hillier S, Darville T (2018). Discovery of blood transcriptional endotypes in women with pelvic inflammatory disease. **J Immunol** 200(8): 2941-2956. doi: 10.4049/jimmunol.1701658
- Zheng X, O'Connell CM, Zhong W, Poston TB, Wiesenfeld HC, Hillier SL, Trent M, Gaydos C, Tseng G, Taylor BD, Darville T (2018). Gene expression signatures can aid diagnosis of sexually transmitted infection-induced endometritis in women. **Front Cell Infect Microbiol** 8: 307. doi: 10.3389/fcimb.2018.00307
- Russell AN, Zheng X, O'Connell CM, Taylor BD, Wiesenfeld HC, Hillier SL, Zhong W, Darville T (2016). Analysis of factors driving incident and ascending infection and the role of serum antibody in *Chlamydia trachomatis* genital tract infection. **J Infect Dis** 213(4): 523-531. doi: 10.1093/infdis/jiv438
- Schillinger JA, Katz BP, Markowitz LE, Braslins PG, Shrier LA, Madico G, Van Der Pol B, Orr DP, Rice PA, Batteiger BE (2016). Genotype-specific concordance of *Chlamydia trachomatis* genital infection within heterosexual partnerships. **Sex Transm Dis** 43(12): 741-749. doi: 10.1097/OLQ.0000000000000525
- Navarro C, Jolly A, Nair R, Chen Y (2002). Risk factors for genital chlamydial infection. **Can J Infect Dis** 13(3): 195-207. doi: 10.1155/2002/954837
- Hwang LY, Ma Y, Moscicki AB (2014). Biological and behavioral risks for incident *Chlamydia trachomatis* infection in a prospective cohort. **Obstet Gynecol** 124(5): 954-960. doi: 10.1097/AOG.0000000000000429
- Aghaizu A, Reid F, Kerry S, Hay PE, Mallinson H, Jensen JS, Kerry S, Kerry S, Oakeshott P (2014). Frequency and risk factors for incident and redetected *Chlamydia trachomatis* infection in sexually active, young, multi-ethnic women: a community based cohort study. **Sex Transm Infect** 90(7): 524-528. doi: 10.1136/sextrans-2014-051607
- van Houdt R, Ma B, Bruisten SM, Speksnijder A, Ravel J, de Vries HJC (2018). *Lactobacillus iners*-dominated vaginal microbiota is associated with increased susceptibility to *Chlamydia trachomatis* infection in Dutch women: a case-control study. **Sex Transm Infect** 94(2): 117-123. doi: 10.1136/sextrans-2017-053133
- van der Veer C, Bruisten SM, van der Helm JJ, de Vries HJ, van Houdt R (2017). The cervicovaginal microbiota in women notified for *Chlamydia trachomatis* Infection: a case-control study at the sexually transmitted infection outpatient clinic in Amsterdam, The Netherlands. **Clin Infect Dis** 64(1): 24-31. doi: 10.1093/cid/ciw586
- Citti C, Blanchard A (2013). Mycoplasmas and their host: emerging and re-emerging minimal pathogens. **Trends Microbiol** 21(4): 196-203. doi: 10.1016/j.tim.2013.01.003.
- Kramer A, Green J, Pollard J, Jr., Tugendreich S (2014). Causal analysis approaches in Ingenuity Pathway Analysis. **Bioinformatics** 30(4): 523-530. doi: 10.1093/bioinformatics/btt703
- Shaw JL, Petraki C, Watson C, Bocking A, Diamandis EP (2008). Role of tissue kallikrein-related peptidases in cervical mucus remodeling and host defense. **Biol Chem** 389(12): 1513-1522. doi: 10.1515/BC.2008.171
- Jiwani S, Ohr RJ, Fischer ER, Hackstadt T, Alvarado S, Romero A, Jewett TJ (2012). *Chlamydia trachomatis* Tarp cooperates with the Arp2/3 complex to increase the rate of actin polymerization. **Biochem Biophys Res Commun** 420(4): 816-821. doi: 10.1016/j.bbrc.2012.03.080
- Siegl C, Prusty BK, Karunakaran K, Wischhusen J, Rudel T (2014). Tumor suppressor p53 alters host cell metabolism to limit *Chlamydia trachomatis* infection. **Cell Rep** 9(3): 918-929. doi: 10.1016/j.celrep.2014.10.004
- Gonzalez E, Rother M, Kerr MC, Al-Zeer MA, Abu-Lubad M, Kessler M, Brinkmann V, Loewer A, Meyer TF (2014). *Chlamydia* infection depends on a functional MDM2-p53 axis. **Nature Commun** 5: 5201. doi: 10.1038/ncomms6201
- Borgdorff H, Tsvitsivadze E, Verhelst R, Marzorati M, Jurriaans S, Ndayisaba GF, Schuren FH, van de Wijgert JH (2014). Lactobacillus-dominated cervicovaginal microbiota associated with reduced HIV/STI prevalence and genital HIV viral load in African women. **ISME J** 8(9): 1781-1793. doi: 10.1038/ismej.2014.26
- Gosmann C, Anahtar MN, Handley SA, Farcasanu M, Abu-Ali G, Bowman BA, Padavattan N, Desai C, Droit L, Moodley A, Dong M, Chen Y, Ismail N, Ndung'u T, Ghebremichael MS, Wesemann DR, Mitchell C, Dong KL, Huttenhower C, Walker BD, Virgin HW, Kwon DS (2017). Lactobacillus-deficient cervicovaginal bacterial communities are associated with increased HIV acquisition in young South African women. **Immunity** 46(1): 29-37. doi: 10.1016/j.immuni.2016.12.013
- Anahtar MN, Byrne EH, Doherty KE, Bowman BA, Yamamoto HS, Soumillon M, Padavattan N, Ismail N, Moodley A, Sabatini ME, Ghebremichael MS, Nusbaum C, Huttenhower C, Virgin HW, Ndung'u T, Dong KL, Walker BD, Fichorova RN, Kwon DS (2015). Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract. **Immunity** 42(5): 965-976. doi: 10.1016/j.immuni.2015.04.019
- O'Connell CM, Ferone ME (2016). *Chlamydia trachomatis* genital infections. **Microb Cell** 3(9): 390-403. doi: 10.15698/mic2016.09.525
- Thayer JD, Martin JE, Jr. (1964). A selective medium for the cultivation of *N. gonorrhoeae* and *N. meningitidis*. **Public Health Rep** 79: 49-57. doi: 10.2307/4592051
- De Clercq E, Kalmar I, Vanrompay D (2013). Animal models for studying female genital tract infection with *Chlamydia trachomatis*. **Infect Immun** 81(9): 3060-3067. doi: 10.1128/IAI.00357-13
- Rice PA, Shafer WM, Ram S, Jerse AE (2017). *Neisseria gonorrhoeae*: Drug resistance, mouse models, and vaccine development. **Ann Rev Microbiol** 71: 665-686. doi: 10.1146/annurev-micro-090816-093530
- McGowin CL, Spagnuolo RA, Pyles RB (2010). *Mycoplasma genitalium* rapidly disseminates to the upper reproductive tracts and knees of female mice following vaginal inoculation. **Infect Immun** 78(2): 726-736. doi: 10.1128/IAI.00840-09
- Cohen CR, Koochesfahani KM, Meier AS, Shen C, Karunakaran K, Ondondo B, Kinyari T, Mugo NR, Nguti R, Brunham RC (2005). Immu-

- noepidemiologic profile of *Chlamydia trachomatis* infection: importance of heat-shock protein 60 and interferon- γ . **J Infect Dis** 192(4): 591-599. doi: 10.1086/432070
30. Onderdonk AB, Delaney ML, Fichorova RN (2016). The human microbiome during bacterial vaginosis. **Clin Microbiol Rev** 29(2): 223-238. doi: 10.1128/CMR.00075-15
31. Parolin C, Foschi C, Laghi L, Zhu C, Banzola N, Gaspari V, D'Antuono A, Giordani B, Severgnini M, Consolandi C, Salvo M, Cevenini R, Vitali B, Marangoni A (2018). Insights into vaginal bacterial communities and metabolic profiles of *Chlamydia trachomatis* infection: positioning between eubiosis and dysbiosis. **Front Microbiol** 9: 600. doi: 10.3389/fmicb.2018.00600
32. Haggerty CL, Totten PA, Tang G, Astete SG, Ferris MJ, Norori J, Bass DC, Martin DH, Taylor BD, Ness RB (2016). Identification of novel microbes associated with pelvic inflammatory disease and infertility. **Sex Transm Infect** 92(6): 441-446. doi: 10.1136/sextrans-2015-052285
33. Wu H, de Gannes MK, Luchetti G, Pilsner JR (2015). Rapid method for the isolation of mammalian sperm DNA. **Biotechniques** 58(6): 293-300. doi: 10.2144/000114280
34. Hugerth LW, Seifert M, Pennhag AAL, Du J, Hamsten MC, Schuppe-Koistinen I, Engstrand L (2018). A comprehensive automated pipeline for human microbiome sampling, 16S rRNA gene sequencing and bioinformatics processing. doi: 10.1101/286526
35. Nudel K, McClure R, Moreau M, Briars E, Abrams AJ, Tjaden B, Su XH, Trees D, Rice PA, Massari P, Genco CA (2018). Transcriptome analysis of *Neisseria gonorrhoeae* during natural infection reveals differential expression of antibiotic resistance determinants between men and women. **mSphere** 3(3): e00312-18. doi: 10.1128/mSphereDirect.00312-18
36. McClure R, Nudel K, Massari P, Tjaden B, Su X, Rice PA, Genco CA (2015). The gonococcal transcriptome during infection of the lower genital tract in women. **PLoS One** 10(8): e0133982. doi: 10.1371/journal.pone.0133982
37. Sun J, Kintner J, Schoborg RV (2008). The host adherens junction molecule nectin-1 is downregulated in *Chlamydia trachomatis*-infected genital epithelial cells. **Microbiol** 154(Pt 5): 1290-1299. doi: 10.1099/mic.0.2007/015164-0
38. Kessler M, Zielecki J, Thieck O, Mollenkopf HJ, Fotopoulou C, Meyer TF (2012). *Chlamydia trachomatis* disturbs epithelial tissue homeostasis in fallopian tubes via paracrine Wnt signaling. **Am J Pathol** 180(1): 186-198. doi: 10.1016/j.ajpath.2011.09.015
39. Igietseme JU, Omosun Y, Stuchlik O, Reed MS, Partin J, He Q, Joseph K, Ellerson D, Bollweg B, George Z, Eko FO, Banda C, Liu H, Yang G, Shieh WJ, Pohl J, Karem K, Black CM (2015). Role of epithelial-mesenchyme transition in *Chlamydia* pathogenesis. **PLoS One** 10(12): e0145198. doi: 10.1371/journal.pone.0145198
40. McGowin CL, Radtke AL, Abraham K, Martin DH, Herbst-Kralovetz M (2013). *Mycoplasma genitalium* infection activates cellular host defense and inflammation pathways in a 3-dimensional human endocervical epithelial cell model. **J Infect Dis** 207(12): 1857-1868. doi: 10.1093/infdis/jit101
41. McGowin CL, Annan RS, Quayle AJ, Greene SJ, Ma L, Mancuso MM, Adegboye D, Martin DH (2012). Persistent *Mycoplasma genitalium* infection of human endocervical epithelial cells elicits chronic inflammatory cytokine secretion. **Infect Immun** 80(11): 3842-3849. doi: 10.1128/IAI.00819-12
42. Mitra A, MacIntyre DA, Mahajan V, Lee YS, Smith A, Marchesi JR, Lyons D, Bennett PR, Kyrgiou M (2017). Comparison of vaginal microbiota sampling techniques: cytobrush versus swab. **Sci Rep** 7(1): 9802. doi: 10.1038/s41598-017-09844-4
43. Chalvardjian A, De Marchi WG, Bell V, Nishikawa R (1991). Improved endocervical sampling with the Cytobrush. **CMAJ** 144(3): 313-317. PMID: 1989710
44. Virtanen S, Kalliala I, Nieminen P, Salonen A (2017). Comparative analysis of vaginal microbiota sampling using 16S rRNA gene analysis. **PLoS One** 12(7): e0181477. doi: 10.1371/journal.pone.0181477
45. Humphrys MS, Creasy T, Sun Y, Shetty AC, Chibucos MC, Drabek EF, Fraser CM, Farooq U, Sengamalay N, Ott S, Shou H, Bavoi PM, Mahurkar A, Myers GS (2013). Simultaneous transcriptional profiling of bacteria and their host cells. **PLoS One** 8(12): e80597. doi: 10.1371/journal.pone.0080597
46. Chung M, Teigen L, Liu H, Libro S, Shetty A, Kumar N, Zhao X, Bromley RE, Tallon LJ, Sadzewicz L, Fraser CM, Rasko DA, Filler SG, Foster JM, Michalski ML, Bruno VM, Dunning Hotopp JC (2018). Targeted enrichment outperforms other enrichment techniques and enables more multi-species RNA-Seq analyses. **Sci Rep** 8(1): 13377. doi: 10.1038/s41598-018-31420-7
47. Kumar N, Lin M, Zhao X, Ott S, Santana-Cruz I, Daugherty S, Rikihisa Y, Sadzewicz L, Tallon LJ, Fraser CM, Dunning Hotopp JC (2016). Efficient enrichment of bacterial mRNA from host-bacteria total RNA samples. **Sci Rep** 6: 34850. doi: 10.1038/srep34850
48. Heine P, McGregor JA (1993). *Trichomonas vaginalis*: a reemerging pathogen. **Clin Obstet Gynecol** 36(1): 137-144. doi: 10.1097/00003081-199303000-00019
49. Russell AN, Zheng X, O'Connell CM, Wiesenfeld HC, Hillier SL, Taylor BD, Picard MD, Flechtner JB, Zhong W, Frazer LC, Darville T (2016). Identification of *Chlamydia trachomatis* antigens recognized by T Cells from highly exposed women who limit or resist genital tract infection. **J Infect Dis** 214(12): 1884-1892. doi: 10.1093/infdis/jiw485
50. Taylor BD, Zheng X, O'Connell CM, Wiesenfeld HC, Hillier SL, Darville T (2018). Risk factors for *Mycoplasma genitalium* endometritis and incident infection: a secondary data analysis of the T cell Response Against *Chlamydia* (TRAC) Study. **Sex Transm Infect** 94(6): 414-420. doi: 10.1136/sextrans-2017-053376
51. Abrams AJ, Trees DL, Nicholas RA (2015). Complete Genome Sequences of Three *Neisseria gonorrhoeae* laboratory reference strains, determined using PacBio single-molecule real-time technology. **Genome Announc** 3(5): e01052-15. doi: 10.1128/genomeA.01052-15.
52. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. **Bioinformatics** 29(1): 15-21. doi: 10.1093/bioinformatics/bts635
53. Langmead B, Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. **Nat Methods** 9(4): 357-359. doi: 10.1038/nmeth.1923
54. Dadi TH, Renard BY, Wieler LH, Semmler T, Reinert K (2017). SLIMM: species level identification of microorganisms from metagenomes. **PeerJ** 5: e3138. doi: 10.7717/peerj.3138
55. Anders S, Pyl PT, Huber W (2015). HTSeq--a Python framework to work with high-throughput sequencing data. **Bioinformatics** 31(2): 166-169. doi: 10.1093/bioinformatics/btu638
56. Love MI, Huber W, Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. **Genome Biol** 15(12): 550. doi: 10.1186/s13059-014-0550-8