# Identification of Vaginal Bacteriophages in Pregnant Women: A Maternal Microbiome Virome Study

**Dr. Sofia Almeida[1]***

[1]Hospital Clínico San Carlos, Department of Obstetrics and Microbiology, Madrid, Spain

Abstract

The presence and identity of bacteriophages (phages) in the vagina of pregnant women were determined. A convenience sample of 107 mid-trimester pregnant women in São Paulo, Brazil were analyzed for phages by metagenomic sequencing, utilizing previously unthawed vaginal aliquots from subjects whose vaginal bacterial composition and pregnancy outcome were already determined. Phages were detected in 96 (89.7%) of the samples. Six different phage families were identified: Siphoviridae in 69.2%, Myoviridae in 49.5%, Microviridae in 37.4%, Podoviridae in 20.6%, Herelleviridae in 10.3% and Inviridae in 1.9% of the women. Four phage families were present in 14 women (13.1%), 3 families in 20 women (18.7%), 2 families in 31 women (29.1%) and 1 family in 31 women (29.1%). The most common phage species detected were Bacillus phages in 48 (43.6%), Escherichia phages in 45 (40.9%), Staphylococcus phages in 40 (36.4%), Gokushovirus in 33 (30.0%) and Lactobacillus phages in 29 (26.4%) women. There were no associations between a particular phage family, the number of families present in the vagina or any particular phage species and either gestational age at delivery or the dominant bacterium present in the vagina. We conclude that in our pregnant population a number of phages are present in the vagina of most women and that their occurrence does not influence subsequent gestational age at delivery or numerical bacterial dominance in the vagina.

Keywords: bacteriophage, pregnancy, vaginal microbiome

INTRODUCTION

Bacteriophages (phages) are viruses that infect bacteria. They are by a large margin the most frequent components of the human virome. Most human-related phage studies have involved the gastro-intestinal tract, the oropharanx and the skin where multiple phage families have been identified.[1-5] It has been estimated that there may be up $10^{15}$ phages in various regions of the body. The consequences of phage infection are twofold. The phage can replicate within its bacterial host followed by cell lysis and the release of new progeny phages to infect other cells (lytic cycle). This will promote the selective presence and persistence of different bacterial species. Alternatively, following penetration into the cell the phage DNA becomes exposed, inserts into the bacterial genome and exists as a latent prophage (lysogenic cycle). The incorporated phage DNA, depending on its composition and location, may modify bacterial activity, such as increasing antibiotic resistance or remain without noticeable effect. In response to altered physiological events the phage DNA can separate from the bacterial DNA and initiate a lytic cycle.[2]

Variations in composition of bacteria in the vaginal microbiome between individual pregnant women and its association with pregnancy-related outcomes is an area of intense investigation.[6-9] However, the presence and activity of specific phages in the human vagina during pregnancy has not been studied. There have been only two investigations of phages in the human vagina; both limited to phages infecting Lactobacilli and involving non-pregnant women.[10,11]

In this communication we report the results of an exploratory study to determine the phage composition of the vagina in 107 pregnant women whose vaginal microbiota and pregnancy outcome were previously determined.

MATERIALS AND METHODS

**Subjects** From 560 second trimester pregnant women (20-22 weeks gestation) with singleton gestations in which we previously reported on the composition of their vaginal microbiome and pregnancy history[9] we selected a convenience sample of 107 women in whom previously unthawed aliquots that were maintained at -80$^{o}$C were available and that provided a range of pregnancy outcomes and vaginal microbiome composition. The subjects' mean age was 28.4 +/- 7.5 years, median parity was 1, median gravidity was 2 and mean body mass index was 26.8 +/- 5.5 kg/m$^2$. The majority (59.8%) of the subjects were White, 33.0% were of mixed race and 7.2% were Black. An early spontaneous preterm birth (SPTB), delivery at <32 weeks gestation, occurred in 12 women while 16 had a late SPTB, delivery $\geq$32 weeks and <37 weeks gestation. The bacteria that were dominant in the samples tested were *Lactobacillus crispatus* in 39*, Lactobacillus iners* in 34*, Gardnerella vaginalis* in 20, *Lactobacillus* jenseni in 10 women *Lactobacillus gasseri* in 4, *Lactobacillus vaginalis* in 1, Shuttleworthia sp in 1 and no dominant bacteria in 2 women. All subjects provided informed written consent. The study was approved by the Institutional Review Board at The University of São Paulo Medical School.

**Viral metagenomics**

The procedure used to perform deep sequencing is a combination of several previously described protocols that have been applied to viral metagenomics and/or virus discovery. [da Costa, A.C., Leal, E., Gill, D. et al. Discovery of Cucumis melo endornavirus by deep sequencing

of human stool samples in Brazil. Virus Genes 55, 332–338 (2019).] The vaginal samples were diluted with 500µl of Hanks' buffered salt solution, added to a 2ml impact-resistant tube containing lysing matrix D (MP Biomedicals, USA), and homogenized in a FastPrep-24 5G Homogenizer (MP biomedicals, USA). The homogenized samples were centrifuged at 12,000 × g for 10 min, and approximately 400µl of the supernatant was then percolated through a 0.22µm filter (Merck Millipore, Billerica, MA, USA) to remove eukaryotic and bacterial cell-sized particles. Approximately, 100µl, roughly equivalent to one-fourth of the volume of the tube of a cold PEG-it Virus Precipitation Solution (System Biosciences, CA, USA) was added to the obtained filtrate, and the contents of the tubes were gently mixed and then incubated at 4°C for 24 h. After the incubation period, the mixtures were centrifuged at 10,000 ×g for 30 min at 4°C. Following centrifugation, the supernatants (~ 350 µl) were discarded. The pellets rich in viral particles were treated with a combination of nuclease enzymes (TURBO DNase and RNase Cocktail Enzyme Mix—Thermo Fischer Scientific, CA, USA; Baseline-ZERO DNase-Epicentre, WI, USA; Benzonase-Darmstadt, Germany; and RQ1 RNaseFree DNase and RNase A Solution-Promega, WI, USA) in order to digest unprotected nucleic acids. The resulting mixtures were subsequently incubated at 37°C for 2h. After incubation, viral nucleic acids were extracted using the Maxwell 16 Viral Total Nucleic Acid Purification Kit (Promega, WI, USA) according to the manufacturer's protocol. The cDNA synthesis was performed with AMV Reverse transcription (Promega, WI, USA). A second strand of cDNA synthesis was performed using DNA Polymerase I Large (Klenow) Fragment (Promega, WI, USA). Subsequently, a Nextera XT Sample Preparation Kit (Illumina, CA, USA) was used to construct a DNA library, identified using dual barcodes. Library was then purified using ProNex® Size-Selective Purification System (Promega,

WI, USA). Following ProNex® purification, the quantity of each sample was normalized to ensure equally library representation in our pooled samples using the ProNex® NGS Library Quant Kit (Promega, WI, USA). For size range selection, Pippin Prep (Sage Science, Inc.) was used to select a 300 bp insert (range 200–400 bp), which excluded very short and long library fragments. Prior to cluster generation, libraries were quantified again by qPCR using the ProNex® NGS Library Quant Kit (Promega, WI, USA). The library was deep sequenced using the HiSeq 2500 Sequencer (Illumina, CA, USA) with 126 bp ends.

Bioinformatics analysis was performed according to a previously described protocol. Pipelines are already published and available for free access (Deng et al (Nucleic Acids Res. 2015 Apr 20;43(7):e46.) the source code of pipeline can be obtained at http://ensembleassembly.sourceforge.net, or https://github.com/xutaodeng/EnsembleAssembler. Briefly, raw reads obtained from Illumina sequencing were preprocessed before assembly as follows. Human host reads were subtracted by mapping the reads with human reference genome hg37, bacterial and fungal RefSeq genomes release using bowtie2. Reads that were identical from nucleotide positions 5–45 were considered clonal reads and only one random copy of clonal reads was retained. The other clonal sequences were replaced with sequence ‘A’ as a place holder; thus, the original order of the paired-end files was preserved. A paired-end sequence record was removed only if both ends were replaced. Low-quality sequences were trimmed using a Phred quality score 30 as the threshold. Adaptor and primer sequences were trimmed using the BLAST-based VecScreen at default parameters. Reads were considered duplicates if bases at position 5 to 55 from 5’ end were identical. One random copy of duplicates was kept. The filtered reads were

assembled using a de novo sequence assembler consists of SOAPdenovo2, ABySS, meta-Velvet, CAP3, MIRA and SPADES. The contigs, including viral sequences and others, sharing a percent nucleotide identity of 95% or less were assembled from the obtained sequence reads by de novo assembly. The resulting singlets and contigs were analyzed using BLASTx to search for similarity to viral proteins in GenBank. The contigs were compared to the GenBank non-redundant nucleotide and protein databases (BLASTn and BLASTx). After identification of the viruses, reference template sequences were used for mapping the full-length genomes with Geneious R9 software (Biomatters Ltd L2, 18 Shortland Street Auckland, 1010, New Zealand).

**Statistics** Differences in the number of women positive for a specific vaginal phage and the dominant bacterium in their microbiome or their pregnancy outcome were determined by Fisher's exact test. The association between the number of phages in a given sample and gestational age at delivery was determined by the Spearman rank correlation test. A p value $<0.05$ was considered significant.

**RESULTS**

In the vaginal samples from 107 women 6 different phage families were identified (Table 1). Member of the Siphoviridae were most prevalent, being present in 74 women (69.2%). The other families were Myoviridae in 53 women (49.5%), Microviridae in 40 women (37.4%), Podoviridae in 22 women (20.6%), Herelleviridae in 11 women (10.3%) and Inoviridae in 2 women (1.9%). In addition, six additional phages were present whose families could not be identified; no phages were present in 11 women (10.3%). The number of different phage families that were present in a given vaginal sample varied from 0 to 4. One or 2 phage families were each present in 31 women (29.1%), 3 families were found in 20 women (18.7%), while in 14 women (13.1%) 4 phage families were present in the vagina.

The association between the presence of different phage families in the vagina and gestational age at delivery is shown in Table 2. There were no significant differences between the presence of a specific phage family and pregnancy outcome. The 11 women in which no phages were detected in their vagina had a higher rate of preterm birth (45.5%) than did the 74 women positive for Siphoviridae (20.3%), or the 53 women positive for Myoviridae (20,7%), but these differences did not reach statistical significance ($p = 0.1197$, $p = 0.1262$, respectively).

The association between the number of different phage families present in the vagina and gestational age at delivery is shown in Table 3. No differences were detected. The lower rate of term births in women with no phages in their vagina (54.5%) compared to those with one or more phage families was not significant ($p \geq 0.3432$). There was no association between the number of different phages present in the vagina and gestational age at delivery (Spearman $r = - 0.034$, $p = 0. 717$).

The association between the presence of different phage families in the vagina and the dominant bacteria in the vaginal microbiome is shown in Table 4. Dominant bacteria is defined as the bacterium present at >50% of the total number of different bacteria identified in a given vaginal sample. There were no associations between the presence of any phage family, or the absence of all phage families, and differences in bacterial dominance.

The association between the number of phage families present in the vagina and the dominant bacterium in the vaginal microbiome is shown in Table 5. A similar distribution of dominant bacteria was present in vaginal samples in the absence of any phage families or when 1-4 different phage families were present.

The genera and species of phages identified in each of the phage families is provided in Supplemental Tables 1-3. The most prevalent phage species present in our population were Bacillus phages in 48 women, Escherichia phages in 45 women, Gokushovirus in 33 women, Staphylococcus phages in 32 women, Lactobacillus phages in 29 women and Mycobacteria in 22 women. The association between the most prevalent phage species and gestational age at delivery is shown in Table 6. The distributions of each of the phages analyzed were similar in women with early or late SPTB or term delivery.

Table 7 presents the data on the association between the 6 most prevalent phage species and dominant bacterium in the vaginal microbiome. Similar to the results with the phage families, there was no relationship between the presence of any phage species and the dominant bacterium in the vaginal microbiome.

## DISCUSSION

Ninety percent of second trimester pregnant women in this study from Brazil were positive for one or more different phage families in their vagina. The phage composition varied widely between individual women but there was no association between the presence of a specific phage family, or the number of different phage families present in the vagina, and SPTB. Similarly, there were no associations between the presence of a specific phage family, or the number of different phage families, and the dominant bacteria in the vaginal microbiome. Evaluating the most prevalent species in each of the phage families also failed to find an association with either pregnancy outcome or bacterial dominance in the vagina.

The phage families identified as being most prevalent in the vagina have also been reported to be the most prevalent families at other body sites. In the gastrointestinal tract Siphoviridae, Myoviridae, Microviridae and Podoviridae were most frequently identified.[1,12] Similarly, the Siphoviridae, Myoviridae and Podoviridae phage families have been shown to predominate in the oral cavity.[13,14] In the skin, most phage families remain uncharacterized.[1,15] The phage DNA sequences that were identified in samples of peripheral blood most commonly included Siphoviridae, Myoviridae and Microviridae.[16,17] At all sites, including our findings on phages in the vagina, many phage DNA sequences do not correspond to known sequences and, therefore, remain unidentified.[1] As has been pointed out, there are no consensus phage DNA sequences similar to 16S bacterial DNA sequences that that would easily facilitate identification of phages based on DNA homology.[18]

Similar to our observations on phage families in the vagina, investigations of phage families in the gastro-intestinal tract have reported a high degree of diversity among apparently healthy

individuals.[1-3] Our study also did not detect any association between the presence of a specific phage family or species and the dominant bacteria in the vagina. This absence of an association between phage families and bacterial composition is characteristic of all other body sites that have been evaluated.[1,,2,3,5,14] This variation in phage composition among healthy individuals may result from differences in phage exposure as a consequence of diet, sexual activity, contamination from the skin, fingers or gastrointestinal tract, as well as the rate of conversion of lysogenic to lytic phage in the vagina by changes in the local environment.

The reasons why a specific Lactobacillus species or other bacteria such as *G. vaginalis* are dominant in the vagina of individual women remain largely undetermined.[19] The results of the present study suggest that the phage composition of the vagina may not be a major determinant of this variation. The 20 women in our study in which *G. vaginalis* was numerically dominant, the 37 with *L. crispatus* dominance and the 34 in which *L. iners* was dominant all had indistinguishable vaginal phage populations.

A strength of our study was the use of parallel aliquots of vaginal samples in which microbiome analysis had already been performed and pregnancy outcomes were known[9]. Limitations of this investigation need to be acknowledged. We emphasize that our study was exploratory with the aim of making initial observations on the presence and variety of phages in the vagina during pregnancy. The limited number of women analyzed did not permit us to evaluate possible contributions of all individual phage species to microbiome composition or pregnancy outcome. In addition, our study was limited to women from Brazil who were between 20 and 22 weeks gestation. It remains to be determined if vaginal samples obtained from the same woman at different time points in gestation, or from women residing in different

countries, will yield similar or divergent results. Lastly, the protocol utilized to isolate phages from the vaginal samples was designed to recover intact extracellular phages. However, we cannot rigorously exclude the possibility that the integrity of some intact bacteria was disrupted resulting in the release of bacterial DNA that contained an incorporated phage genome(s). This could have led to our detection also of lysogenic phages. Treatment of our samples with deoxyribonuclease should have reduced this possible source of contamination.

In conclusion, phages appear to be present in the vagina of most second trimester pregnant women. In our exploratory investigation vaginal phage composition varied greatly among individual women and was independent of vaginal bacterial composition and pregnancy outcome. Future investigations on a much greater number of women at various pregnancy stages and from different countries are needed to probe more deeply into possible relationships between specific phages and bacteria in the vagina and their combined influence on pregnancy-related parameters.

ACKNOWLEDGEMENTS

This study was supported by a grant from the Bill and Melinda Gates Foundation, Brazilian Ministry of Health (DECIT) and the Brazilian National Research Council (CNPq – grant 401626/2013-0). Bacterial sequence data collection and analyses were performed by the IBEST Genomics Resources Core at the University of Idaho which is supported in part by NIH COBRE grant P30GM103324. Antonio Charlys da Costa is funded by FAPESP #2017/00021-9

References

1. Rascovan N, Duraisamy R, Desnues C. Metagenomics and the human virome in symptomatic individuals. Annu Rev Microbiol 2016;70:125-41.
2. Salmond GP, Fineran PC. A century of the phage: Past, present and future. Nat Rev Microbiol 2015;13:777-786.
3. Minot S, Sinha R, Chen J, Li H, Keilbaugh SA, Wu GD, et al. The human gut virome: Inter-individual variation and dynamic response to diet. Genome Res 2011;21:1616-1625.
4. Mills S, Shanahan F, Stanton C, Hill C, Coffey A, Ross RP. Movers and shakers: Influence of bacteriophages in shaping the mammalian gut microbiota. Gut Microbes 2013;4:4-16.
5. Norman JM, Handley SA, Baldridge MT, Droit L, Liu CY, Keller BC, et al. Disease-specific alterations in the enteric virome in inflammatory bowel disease. Cell 2015;160:447-460.
6. DiGiulio DB, Callahan BJ, McMurdie PJ, Costello EK, Lyell DJ, Robaczewsa A, et al. Temporal and spatial variation of the human microbiota during pregnancy. Proc Natl Acad Sci USA 2015;112:11060-11065.
7. MacIntyre DA, Chandiramoni M, Lee SY, KIndinger L, Smith A, Angelopoulos N, et al. The vaginal microbiome during pregnancy and the postpartum period in a European population. Sci Rep 2015;5:8988.
8. Elovitz MA, Gajer P, Riis V, Brown AG, Humphrys MS, Holm JB, et al. Cervicovaginal microbiota and local immune response modulate the risk of spontaneous preterm delivery. Nat Commun 2019;10:1305.

9. Witkin SS, Moron AF, Ridenhour BJ, Minis E, Hatanaka A, Mattar R, et al. Vaginal biomarkers that predict cervical length and dominant bacteria in the vaginal microbiome of pregnant women. mBio 2019;10: pii:e02242-19.
10. Kilic AO, Pavlova SI, Alpay S, Kilic SS, Tao L. Comprative study of vaginal *Lactobacillus* phages isolated from women in the United States and Turkey: Prevalence, morphology, host range, and DNA homology. Clin Diagnost Lab Immunol 2001;8:31-39.
11. Pavlova SI, Tao L. Induction of vaginal *Lactobacillus* phages by the cigarette smoke chemical benzo[a]pyrene diol epoxide. Mutat Res 2000;466:57-62.
12. Lim ES, Zhou Y, Zhao G, Bauer IK, Droit L, Ndao IM, et al. Early life dynamics of the human gut virome and bacterial microbiome in infants. Nat Med 2015;21:1228-34.
13. Abeles SR, Ly M, Santiago-Rodriguez TM, Pride DT. Effects of long term antibiotic therapy on human oral and fecal viromes. PLoS One 2015;10:e0134941.
14. Abeles SR, Robles-Sikisaka R, Ly M, Lum AG, Salzman J, Boehm TK, et al. Human oral viruses are personal, persistent and gender-consistent. ISME J 2014;8:1753-67.
15. Hannigan GD, Meisel JS, Tyldsley AS, Zheng Q, Hodkinson BP, SanMiguel AJ, et al. The human skin double-stranded DNA virome: topographical and temporal diversity, genetic enrichment, and dynamic associations with the host microbiome. mBio 2015;6:e01578-15.
16. Breitbart M, Rohwer F. Method for discovering novel DNA viruses in blood using viral particle selection and shotgun sequencing. Biotechniques 2005,39:729-36.

17. Li SK, Leung RK, Guo HX, Wei JF, Wang JH, Kwong KT, et al., Detection and identification of plasma bacterial and viral elements in HIV/AIDS patients in comparison to healthy adults. Clin Microbiol Infect 2012;18:1126-33.
18. Carroll-Portillo A, Lin HC. Bacteriophage and the innate immune system: Access and signaling. Microorganisms 2019;7:625.
19. Witkin SS, Linhares IM. Why do lactobacilli dominate the human vaginal microbiota? BJOG 2017;124:606-611.

Table 1. Phage families present in the vagina of pregnant women

| Phage Family | |
|---|---|
| Siphoviridae | 74 (69.2%) |
| Myoviridae | 53 (49.5%) |
| Microviridae | 40 (37.4%) |
| Podoviridae | 22 (20.6%) |
| Herelleviridae | 11 (10.3%) |
| Unclassified | 6 (5.6%) |
| Inoviridae | 2 (1.9%) |
| None | 11 (10.3%) |

Vaginal secretions from 107 mid-trimester pregnant women were analyzed for bacteriophages by metagenomics.

Table 2. Association between the presence of phage families in the vagina and gestational age at delivery

| Phage Family | No. Positive (%) | Gestational age at delivery (weeks) | | | |
|---|---|---|---|---|---|
| | | <32 | 32-36 | ≤36 | ≥37 |
| Siphoviridae | 74 (69.2%) | 8 (10.8%) | 7 (9.5%) | 15 (20.3%) | 59 (79.7%) |
| Myoviridae | 53 (49.5%) | 6 (11.3%) | 5 (9.4%) | 11 (20.7%) | 42 (79.2%) |
| Microviridae | 40 (37.4%) | 5 (12.5%) | 7 (17.5%) | 12 (30.0%) | 28 (70.0%) |
| Podoviridae | 22 (20.6%) | 2 (9.1%) | 4 (18.2%) | 6 (27.3%) | 16 (72.7%) |
| Herelleviridae | 11 (10.3%) | 2 (18.2%) | 0 | 2 (18.2%) | 9 (81.8%) |
| Unclassified | 6 (5.6%) | 0 | 0 | 0 | 6 (100%) |
| Inoviridae | 2 (1.9%) | 0 | 0 | 0 | 2 (100%) |
| None | 11 (10.3%) | 2 (18.2%) | 3 (27.3%) [a] | 5 (45.5%) [b] | 6 (54.5%) |

Vaginal secretions from 107 mid-trimester pregnant women were tested for phage families by metagenomics. Subsequent pregnancy outcomes were determined by chart review.
[a]$p = 0.1157$ vs. all others; [b]$p = 0.1336$ vs. all others

Table 3. Association between the number of different phage families present in the vagina and subsequent gestational age at delivery

| No. Phage Families In the vagina | No. women | Gestational age at delivery (weeks) | | | |
|---|---|---|---|---|---|
| | | <32 | 32-36 | ≤36 | ≥37 |
| | | N = 12 | N = 16 | N = 28 | N = 80 |
| 4 | 13 | 1 | 2 | 3 (23.1%) | 10 (76.9%) |
| 3 | 20 | 3 | 2 | 5 (25.0%) | 15 (75.0%) |
| 2 | 31 | 5 | 2 | 7 (22.6%) | 24 (77.4%) |
| 1 | 31 | 1 | 6 | 7 (22.6%) | 24 (77.4%) |
| 0 | 11 | 2 | 3 | 5 (45.5%)[a] | 6 (54.5%) |

[a]p = 0.1402 vs. all others

Table 4. Association between different phage families in the vagina of pregnant women and the dominant bacterium in the vagina

| Phage Family | No. Positive | Dominant Bacterium | | | | |
|---|---|---|---|---|---|---|
| | | G. vaginalis | L. crispatus | L. iners | L. jensenii | Other* |
| Siphoviridae | 74 | 14 (18.9%) | 23 (31.1%) | 25 (33.8%) | 8 (10.8%) | 4 (5.4%) |
| Myoviridae | 53 | 10 (18.9%) | 15 (28.3%) | 19 (35.8%) | 6 (11.3%) | 3 (5.7%) |
| Microviridae | 40 | 8 (20.0%) | 13 (32.5%) | 15 (37.5%) | 3 (7.5%) | 1 (2.5%) |
| Podoviridae | 22 | 4 (18.2%) | 10 (45.5%) | 6 (27.3%) | 2 (9.1%) | 0 |
| Herelleviridae | 11 | 2 (18.2%) | 4 (36.4%) | 6 (54.5%) | 1 (9.1%) | 0 |
| Unclassified | 6 | 2 (33.3%) | 1 (16.7%) | 2 (33.3%) | 0 | 1 (16.7%) |
| Inoviridae | 2 | 0 | 1 (50.0%) | 0 | 1 (50.0%) | 0 |
| None | 11 | 2 (18.2%) | 4 (36.4%) | 3 (27.3%) | 1 (9.1%) | 0 |

Bacteria present in the vagina was identified by gene amplification. Dominant bacterium is the organism present in >50% of all bacteria detected in a given vaginal sample.
*Other: L. gasseri (5), L. vaginalis (1), Shuttleworthia (1), no dominant bacterium (2)

Table 5. Association between the number of different phage families present in the vagina and dominant bacterium in the vaginal microbiome

| No. Phage Families In the vagina | No. women | Dominant bacterium | | | | |
|---|---|---|---|---|---|---|
| | | G. vaginalis | L. crispatus | L. iners | L. jensenii | Others |
| 4 | 13 | 3 (23.1%) | 3 (23.1%) | 6 (46.2%) | 1 (7.7%) | 0 |
| 3 | 20 | 5 (25.0%) | 7 (35.0%) | 5 (25.0%) | 3 (15.0%) | 0 |
| 2 | 31 | 4 (12.9%) | 14 (45.2%) | 7 (22.6%) | 1 (3.2%) | 5 (16.1%) |
| 1 | 31 | 6 (19.4%) | 9 (29.0%) | 14 (45.2%) | 1 (3.2%) | 1(3.2%) |
| 0 | 11 | 2 (18.2%) | 4 (36.4%) | 3 (27.3%) | 1 (9.1%) | 0 |

Dominant bacterium is the organism present in >50% of all bacteria detected in a vaginal sample. Others: L. gasseri (5), L. vaginalis (1), Shuttleworthia (1), no dominant bacterium (2)

Table 6. Bacteriophage species in the vagina of pregnant women and subsequent pregnancy outcome

| Phage species | No. positive | No. women positive (%) | | |
|---|---|---|---|---|
| | | ≤32 weeks | 33-36 weeks | ≥37 weeks |
| Bacillus | 48 | 5 (10.4%) | 6 (12.5%) | 37 (77.1%) |
| Escherichia coli | 45 | 5 (11.1%) | 9 (20.0%) | 31 (68.9%) |
| Gokushovirus | 33 | 3 (9.1%) | 0 | 30 (90.9%) |
| Staphylococcus | 32 | 5 (15.6%) | 3 (5.4%) | 24 (75.0%) |
| Lactobacillus | 29 | 6 (20.7%) | 0 | 23 (79.3% |
| Mycobacteria | 23 | 4 (17.4%) | 2 (8.7%) | 17 (73.9%) |

≤32, 33-36, ≥37 weeks refers to weeks gestation at delivery

Table 7. Bacteriophage species and dominant bacteria in the vagina of pregnant women

| Phage species | No. women | No. women positive (%) | | | |
|---|---|---|---|---|---|
| | | G. vaginalis | L. crispatus | L. iners | L. jensenii |
| Bacillus | 48 | 8 (16.7%) | 18 (37.5%) | 18 (37.5%) | 4 (8.3%) |
| Escherichia coli | 45 | 8 (17.8%) | 19 (42.2%) | 13 (28.9%) | 4 (40%) |
| Gokushovirus | 33 | 8 (24.2%) | 10 (30.3%) | 11 (33.3%) | 2 (20%) |
| Staphylococcus | 32 | 6 (18.8%) | 18 (56.3%) | 11 (34.4%) | 4 (40%) |
| Lactobacillus | 29 | 5 (17.2%) | 12 (41.4%) | 6 (20.7%) | 4 (40%) |
| Mycobacteria | 23 | 5 (21.7%) | 6 (26.1%) | 9 (39.6%) | 2 (20%) |

Supplement Table 1

Genera and species of Siphoviridae phage family present in the vagina

| Genera | Species | No. women |
|---|---|---|
| Andromedavirus | Bacillus virus Riggi | 1 |
| Biseptimavirus | Staphylococcus virus 108PVL,77, 13 | 15 |
| Bronvirus | Mycobacterium virus Faith1 | 2 |
| C2 virus | Lactococcus virus c2, bIL67 | 4 |
| C5 virus | Lactobacillus virus c5, LLKu | 19 |
| Cecivirus | Bacillus virus IEBH | 1 |
| Cequinquevirus | Lactobacillus virus c5, LLKu | 2 |
| Che8virus | Mycobacterium virus Che9D, Llij | 2 |
| Cheoctovirus | Mycobacterium virus Che9d, Pacc40, Gumbi | 3 |
| Cjw1 virus | Mycobacterium virus Kostya, Porky | 2 |
| Coetzeevirus | Lactobacillus virus phiJL1, ATCC8014 | 4 |
| EL25virus | Burkholderia virus phiE125 | 1 |
| Kostyavirus | Mycobacterium virus Kostya, Bask21, CJW1 | 4 |
| L5 virus | Mycobacterium virus Saintus, Kssjeb, Rebeuca | 3 |
| Lambdavirus | Escherichia coli HK97, HK022, Lambda | 11 |
| Omegavirus | Mycobacterium virus Courthouse, Littlee, Optimus | 5 |
| Phietavirus | Staphylococcus virus phiETA, CNPH82 | 16 |
| Phijl 1 virus | Lactobacillus virus ATCC 801 | 2 |
| Ravinvirus | Escherichia virus N15 | 1 |
| Skunavirus | Lactococcus virus 712, ASCC191, ASCC465,Bibb29,P008, Bil170, jj50, CB13 | 19 |
| Stanholtvirus | Burkholderia virus phi6442 | 1 |
| T5 virus | Escherichia coli T5 | 1 |
| Timquatrovirus | Mycobacterium virus Jaws, Pixie | 2 |
| Triavirus | Staphylococcus virus 47, 42e,IPLA35, phi12 | 9 |
| Unclassified | Streptococcus virus MS1, 9872, Geobacillus virus E3 Siphoviridae sp. | 25 |
| Webervirus | Klebsiella virus GML KpCol1 | 1 |

Supplement Table 2. Genera and species of Herelleviridae, Microviridae, Podoviridae and Inoviridae phage families present in the vagina

| Family | Genera | Species | No. women |
|---|---|---|---|
| Herelleviridae | Caeruleovirus | Bacillus virus BM15 | 1 |
| | Kayvirus | Staphylococcus virus K | 1 |
| | Okubovirus | Bacillus virus SPO1 | 2 |
| | Silviavirus | Staphyococcus virus SA11 | 3 |
| | Unclassified | Lactobacillus virus Lb-338-1, LP65 | 7 |
| Microviridae | Chlamydiamicrovirus | Chlamydia virus CPAR39 | 1 |
| | Phix174microvirus | Escherichia coli ph174 | 16 |
| | Unclassified | Gokushovirus WZ 2015a | 46 |
| Podoviridae | Lessievirus | Burkholderia virus DC1 | 1 |
| | Phi29virus | Bacillus virus B103, GA1, phi29 | 18 |
| | SP6 virus | Escherichia virus K1-5, K1E | 3 |
| | Unclassified | Actimomyces virus 1, Lactococcus virus KSY1, Alteromonas virus vB AspP-H4/4, Pseudomonas virus KNP, Mycoplasma virus P1 | 13 |
| Inoviridae | Lineavirus | Escherichia virus 122 | 2 |

Supplement Table 3. Genera and species of Myoviridae phage family present in the vagina

| Genera | Species | No. women |
|---|---|---|
| Bcepmuvirus | Burkholderia virus BcepMu | 2 |
| Biquartavirus | Aeromonas virus 31, 44RR2 | 4 |
| Hpunavirus | Aeromonas virus phiO18P | 2 |
| Muvirus | Escherichia virus Mu | 1 |
| P2virus | Escherichia virus P2 | 5 |
| Peduovirus | Burkholderia virus phiE122 | 1 |
| Punavirus | Escherichia virus P1 | 8 |
| Silviavirus | Staphylococcus virus SA11 | 3 |
| Svunavirus | Bacillus virus 1 | 1 |
| Tequatrovirus | Escherichia virus T4 | 1 |
| Unclassified | Aeromonas virus Aehi, Bacillus virus G, Escherichia virus RB43, Synechococcus virus S-PRM1, Lactobacillus virus LP65, Acinebacter virus 133, Myoviridae sp | 56 |
| Vi1virus | Escherichia virus CBA120 | 4 |