scientific reports

OPEN

Cervicovaginal *Gardnerella* sialidase-encoding gene in persistent human papillomavirus infection

Juliano Novak[1✉], Rafael Belleti[1], Gabriel Vitor da Silva Pinto[1],
Aline do Nascimento Bolpetti[1], Márcia Guimarães da Silva[1] & Camila Marconi[1,2]

Disturbed vaginal microbiota have a role in the persistence of high-oncogenic-risk human papillomavirus (hrHPV) and *Gardnerella* spp. is closely related with this condition. Such bacteria are the major source of cervicovaginal sialidases, important for microbiota alterations. The sialidase-encoding gene nanH3 is account for their sialidase activity. Thus, a subset of 212 women positive for hrHPV at the first visit were included in the analysis of the current study aiming to compare the loads of nanH3 in cervicovaginal fluid (CFV) of women with persistent hrHPV infection and with those cleared the infection after a year. Participants were assigned to two study groups named "persistence" (n = 124, 53.22%) or "clearance" (n = 88, 37.77%), according to the HPV status upon enrollment and follow-up. Absolute quantification of nanH3 gene was performed using quantitative real-time PCR (qPCR). Persistence and clearance group did not show statistical difference in the load of nanH3 gene (p = 0.19). When considering the subset of women with HPV16, differences in number of copies of nanh3 gene was observed between the persistent (7.39E+08 copies/μL) and clearance group (2.85E+07 copies/μL) (p = 0.007). Therefore, baseline loads of nanH3 gene is increased in women that persist with cervical HPV16 infection after 12 months.

The cervical infection by human papillomavirus (HPV) is the most frequent sexually transmitted infection (STI) worldwide[1,2] and the persistence of cervical infections by high-risk-HPV (hrHPV) for long periods of time cause virtually all precursor lesions and cervical cancers[3]. Despite that, the majority of the cases of cervical HPV infection are cleared within 2 years[4,5], being the immune responses important for clearance[6].

Factors associated with developing cervical lesion and cancer include smoking[7], hormonal contraceptive use[8,9], and parity[10]. In addition, the local cervical microenvironment, including the vaginal microbiota, may also influence the natural history of HPV infection[11]. Therefore, a *Lactobacillus*-deprived vaginal microbiota, as in bacterial vaginosis (BV), have been associated with persistent HPV infection and lesion progression[12–14].

Bacterial vaginosis is a polymicrobial dysbiosis, which occurs by the replacement of beneficial *Lactobacillus* and increased anaerobic and facultative anaerobic bacteria, of which, *Gardnerella* spp. that is present in nearly all cases[15,16]. Sialidase production is one of the most important virulence factors of *Gardnerella* spp.[17,18]. Among the deleterious effects of bacterial sialidases, it is the degradation of several protective factors of the vaginal mucosa and contribution to the exfoliation and detachment of vaginal epithelial cells[19] facilitating bacterial adhesion to the epithelium and biofilm formation[20–22], a condition already associated with persistence of BV[23].

At first, the putative sialidase gene, nanH1 (sialidase A gene), was identified in *Gardnerella* spp. and was thought to be the gene responsible for sialidase production[24]. However, the recent studies concluded that nanH2 and nanH3 account for the sialidase activity observed in cultured *G. vaginalis*[25]. Besides that, nanH2 is less prevalent in *Gardnerella* isolates and always detected when in the presence of nanH3[18,25].

Until very recently, *Gardnerella* was one single specie genus but now three other species in addition to *G. vaginalis* (*G. leopoldii*, *G. piotii* and *G. swidsinskii*) were described[26]. Interestingly, sialidase genes nanH2 and nanH3 were only detected in *G. piotii* and in a subset of *G.vaginalis* isolates[18].

Considering the importance of better understanding the relationship between the bacterial components of the vaginal microbiota and the outcome of hrHPV infection, the aim of this study was to compare the loads of

[1]Department of Pathology, Botucatu Medical School, UNESP, São Paulo State University, São Paulo, Brazil. [2]Department of Basic Pathology, Sector of Biologic Science, UFPR, Universidade Federal do Paraná, Curitiba, Brazil. ✉email: juliano.novak@unesp.br

the sialidase-encoding gene nanH3 of *Gardnerella* spp. in the cervicovaginal fluid of women between persistent hrHPV infection and those who cleared the infection after a 12-month period.

Results

Table 1 displays the sociodemographic, behavioral and clinical characteristics of the participants, according to the study groups. The groups showed differences in few variables such as age (p = 0.04), and the number of sex (p < 0.0001). When considering the number of hrHPV genotypes detected, persistent group had more mixed infection than clearance group (p = 0.002). In relation to the HPV genotypes, HPV52 differed between the groups when considering single or mixed infection (p = 0.01).

The most frequently genotype detected in the population (showed in Supplementary Table S1), was HPV16, followed by HPV31, HPV52, HPV51, HPV58 and HPV45, respectively. Of 233 women included in this study, 21 cleared their HPV genotype detected at baseline and a new genotype was detected at follow-up, therefore these women were not included in the analysis. Thus, clearance group included women without detection of HPV at follow-up (n = 88) and persistence group includes those women showing at least one genotype detected at baseline and follow-up (n = 124). Thus, clearance and persistence rates were 41.50% (n = 88) and 58.49% (n = 124), respectively. For HPV16, rate of clearance was 26.56% (n = 17) and persistence was 73.44% (n = 47).

Regarding the presence of nanH3 gene and the outcome of cervical hrHPV infection displayed on the Table 2, the groups did not show difference in relation to the positivity to the gene. When compared the absolute quantification of nanH3 gene, showed by median (minimum–maximum), in the presence of all hrHPV, no difference in the load was observed between the clearance (1.45E+08 copies/µL; 2.30E+04–6.15E+13) and persistence (9.16E+08 copies/µL; 3.02E+05–1.72E+13) groups (p = 0.19). For five most prevalent genotypes (HPV16, 31, 52, 51 and 58) clearance group showed 1.02E+08 copies/µL (2.30E+04–6.15E+13) and persistence 1.82E+09 copies/µL (3.67E+05–1.72E+13) (p = 0.02). When excluding HPV16, the 5 most prevalent genotypes (HPV31, 52, 51, 58 and 45) there was no difference, the clearance group showed 3.60E+08 copies/µL (2.30E+04–6.15E+13) and persistence 1.36E+09 (3.67E+05–1.72E+13) (p = 0.42). Loads of nanH3 gene differed between the clearance (2.85E+07 ng/µL; 1.90E+05–4.23E+07) and persistence group (7.39E+08 copies/µL; 4.15E+06–4.09E+10) in the presence of HPV16 (p = 0.007) (Table 3).

Analysis to test the association between vaginal microbiota and nanH3 gene were also performed. Of the 233 women, 127 (54.51%) had normal microbiota, 21 (9.01%) intermediate and 80 (34.33%) had BV, detected by Nugent scoring. Data of Nugent evaluation were not available for five participants. Among BV positive women, 52/80 (65%) were nanH3 positive, while 28/80 (35%) were nanH3 negative. Thus, sensitivity and specificity of the test was 65% (95% confidence interval, 54–75%) and 75% (66–82%), respectively. Figure 1A displays a bar graph illustrating the relationship of nanH3 gene status across normal and BV Nugent scoring, showing a significant difference between the categories (p < 0.0001). In relation to the loads of nanH3 gene expressed by median (minimum–maximum) in the cervicovaginal fluid shown in Fig. 1B, there was difference between normal microbiota (2.06E+08 copies/µL; 2.30E+04–1.37E+13) and BV (2.91E+09 copies/µL; 8.17E+03–6.15E+13) (p = 0.0041). Due the low representativeness, intermediate microbiota was not considered in this analysis.

Discussion

Recently our study group showed that women with persistent HPV16 and HPV18 infection have increased baseline loads of *G. vaginalis*[27]. Thus, it was hypothesized that gene encoding sialidase, frequently found in *Gardnerella* spp.[18] could have a role in the outcome of HPV infection, being a useful marker for hrHPV persistence. Surprisingly, loads of nanH3 gene differed between clearance and persistence group in the presence of 5 most prevalent genotypes, however such difference remains only in the presence of HPV16.

This study showed high rates of HPV persistence, 58.49% for all hrHPV infection and 73.44% when considered only the HPV16. Such rates were similar to the rates showed in another study conducted in Brazil, with 61.8% of persistence rate[28]. Persistent rates in a large retrospective cohort study in China was 42.7% within 24 months[29], while a multicentric study showed 54.1% of persistence rate over a year in American, Canadian and Brazilian women[30]. In relation to HPV16, the persistence rate was 52.1% after 12 months in Dutch cohort of young women[31].

In relation to sociodemographic and behavioral characteristics of the study population, few variables differed between the groups. The median of age was higher in clearance group, agreeing with a study that showed that the highest HR-HPV persistence occurred in the 22–27 years old group, whereas clearance increased in women aged 28–33 years[32]. However, other Brazilian study did not show differences of age when evaluating factors associated with clearance and persistence after 24 months[28]. The number of sex partners and the number of detected genotypes differed between the clearance and persistence group. Furthermore, one of five most prevalent HPV genotypes differed among the groups when considering their detection in single or mixed infection. In fact, a previous study showed that HPV genotypes differ in terms of the period for clearance, as women with mixed infections had a longer time to clear their infections compared to those with a single infection[30]. The increasing in number of sex partners could contribute to persistence since it increases the risk of acquiring new HPV genotypes.

Studies on vaginal microbiota have demonstrated the association between *Gardnerella* spp. and persistent hrHPV infection, as well as progression of the cervical lesion[33,34]. Of the results of association between *Gardnerella* spp. with persistence and progression of hrHPV, Usik et al. proposed that in addition to role of *Gardnerella* spp. in the viral persistence, it also contributes for the disruption on vaginal microbiota (i. e. increased bacterial diversity), which also leads to the persistence and progression of HPV infection[34].

Variable	Outcome of hrHPV infection		
	Clearance (88)	Persistence (124)	P value
Age[#]	29 (17–50)	26.5 (17–50)	**0.04**
Ethnicity*			0.52
Black	35 (39.77)	59 (47.58)	
White	51 (57.95)	62 (50.00)	
Other	2 (2.27)	3 (2.42)	
Living with partner[$]			0.89
No	41 (46.59)	60 (48.39)	
Yes	47 (53.41)	64 (51.61)	
Years at school[#]	11 (1–16)	11 (3–17)	0.14
Smoking[$]			0.66
No	57 (64.77)	85 (68.55)	
Yes	31 (35.23)	39 (31.45)	
Douching[$]			1.00
No	84 (95.45)	117 (94.35)	
Yes	4 (4.55)	7 (5.65)	
Hormonal contraceptive use[$]			0.14
No	33 (37.50)	34 (27.42)	
Yes	55 (62.50)	90 (72.58)	
Condom use[$]			0.88
No	63 (71.59)	87 (70.16)	
Yes	25 (28.41)	37 (29.84)	
Sex partners[#]	0 (1–15)	3 (1–100)	**<0.0001**
Sex partners (last year)[#]	1 (0–2)	1 (0–12)	0.26
Vaginal abnormal discharge[#]			0.25
No	27 (30.68)	48 (38.71)	
Yes	61 (69.32)	76 (61.29)	
Nugent*			0.59
Normal	50 (56.82)	68 (54.84)	
Intermediate	10 (11.36)	10 (8.06)	
BV	28 (31.82)	46 (37.10)	
Number of hrHPV genotypes[#]	1 (1–3)	1 (1–5)	**0.002**
HPV16[$]			1.00
Single	12 (13.64)	31 (25.00)	
Mixed	5 (5.68)	16 (12.90)	
HPV31[$]			0.43
Single	6 (6.82)	9 (7.26)	
Mixed	3 (6.41)	10 (8.06)	
HPV51[$]			0.12
Single	12 (13.64)	5 (4.03)	
Mixed	1 (1.14)	4 (3.23)	
HPV52[$]			**0.01**
Single	9 (10.23)	6 (4.84)	
Mixed	1 (1.14)	10 (8.06)	
HPV58[$]			1.00
Single	4 (4.55)	7 (5.65)	
Mixed	3 (3.41)	8 (6.45)	

Table 1. Sociodemographic, behavioral and clinical characteristics of the participants, according to the study groups. For all analysis P-value < 0.05 was considered as significant. *Chi-square test. [#]Mann–Whitney test, expressed by median (minimum–maximum). [$]Fisher's exact test. Significant values are in bold.

As recently demonstrated, not all newly described species of *Gardnerella* produce sialidase[18]. Therefore, studying the presence of sialidase encoding gene of *Gardnerella* is important for paving the way for further studies focusing in the role of each individual *Gardnerella* sp. for HPV infection and persistence.

	Detectable nanH3 gene		
	Yes	No	P value*
All hrHPV			0.77
Clearance	32 (36.36)	56 (63.64)	
Persistence	48 (38.71)	76 (61.29)	
5 most prevalent hrHPV genotypes[&]			0.60
Clearance	20 (37.74)	33 (62.26)	
Persistence	38 (42.70)	51 (57.30)	
5 most prevalent hrHPV genotypes, excluding HPV16[$]			0.70
Clearance	19 (38.78)	30 (61.22)	
Persistence	26 (44.07)	33 (55.93)	
HPV16			0.56
Clearance	5 (29.41)	12 (70.59)	
Persistence	19 (40.43)	28 (59.57)	

Table 2. Detection of nanH3 gene in the cervicovaginal fluid according to the study groups. For all analysis $P < 0.05$ was considered as significant. *Fisher's exact test. [&]Including HPV16, HPV31, HPV52, HPV51 and HPV58. [$]Including HPV31, HPV52, HPV51, HPV58 and HPV45.

	NanH3 loads (copies/µL)[#]		
	Clearance	Persistence	P value
All hrHPV	1.45E+08 (2.30E+04–6.15E+13)	9.16E+08 (3.67E+05–1.72E+13)	0.19
5 most prevalent hrHPV genotypes[&]	1.02E+08 (2.30E+04–4.34E+13)	1.82E+09 (3.67E+05–1.72E+13)	**0.02**
5 most prevalent hrHPV genotypes, excluding HPV16[$]	3.60E+08 (2.30E+04–6.15E+13)	1.36E+09 (3.67E+05–1.72E+13)	0.42
HPV16	2.85E+07 (1.90E+05–4.23E+07)	7.39E+08 (4.15E+06–4.09E+10)	**0.007**

Table 3. Load of nanH3 gene in the cervicovaginal fluid according to the study groups. For all analysis $P < 0.05$ was considered as significant. [&]Including HPV16, HPV31, HPV52, HPV51 and HPV58. [$]Including HPV31, HPV52, HPV51, HPV58 and HPV45. [#]Mann–Whitney Test, expressed by median (minimum–maximum). Significant values are in bold.
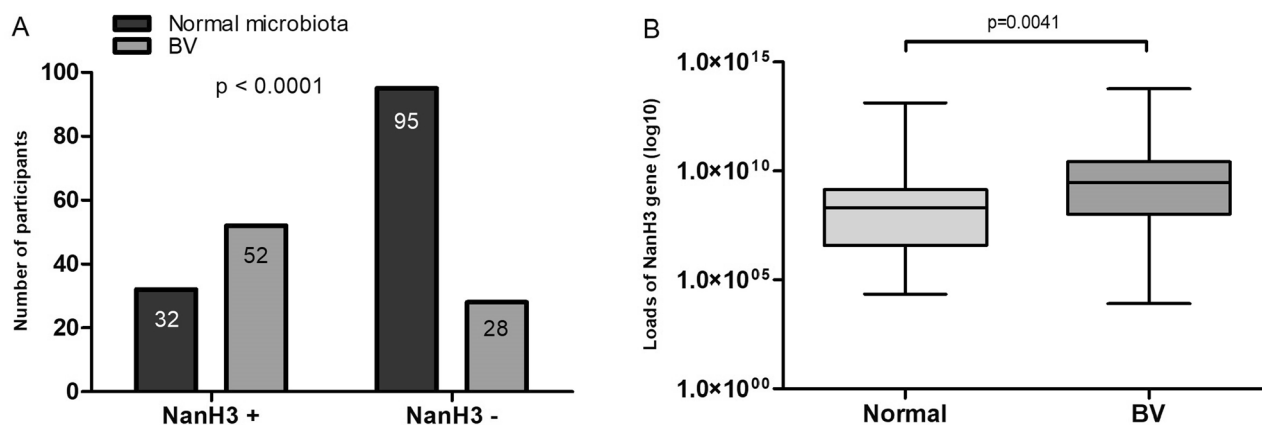

Figure 1. Characteristics of association between nanH3 gene and vaginal microbiota. (**A**) A bar graph illustrating the relationship of nanH3 gene status across the two Nugent score categories (p < 0.0001); (**B**) loads of nanh3 gene compared between normal microbiota and BV (p = 0.0041).

To the best of our knowledge, this is the first study to assess the relationship between hrHPV persistence infection and loads of nanH3 gene. It is known that sialidases hydrolyze the sialic acid of the host's epithelial cells, provides a carbon source to *Gardnerella* allowing their uptake and catabolism and expose binding sites, facilitating bacterial adhesion to the epithelium and further biofilms formation[20–22,35]. Sialidases can still cleave mucin oligosaccharides, that decreased the cervicovaginal mucus viscosity, compromising the physical and biochemical barrier against pathogens[36]. Besides that, sialidases compromise the immune barrier by hydrolyzing immunoglobulin A[35,37]. In the present study, results on the quantification of nanH3 did not differ when included

all hrHPV infection grouped, however loads of nanH3 were higher in the cervicovaginal fluid of women with HPV16 persistent.

Studies have been reporting the association of altered vaginal microbiota with HPV infection[12–14,33,34,38], despite the mechanisms by it can influence in HPV infection remain unknow. It is known that *Lactobacillus* spp. produce several microbicidal factors, including lactic acid[39], an important agent to acidify the vaginal environment and to control overgrowth of bacteria and preventing STIs[40,41]. The difference of nanh3 loads found in HPV16 infection is intriguing, are needed new approaches involving vaginal microbiota and HPV16, since this genotype is responsible for the most cases of high-grade intraepithelial lesions and invasive cervical carcinomas[42].

Human papillomaviruses modulate the host immune response by blocking immune-related gene expression and immune signaling pathways and different genotypes may act by different pathways[43]. In addition, cervical microbiome and cytokine profile showed notably different in all stages of the natural history of cervical cancer[44]. Thus, bacterial species may interact with immune signaling during HPV infection contributing to persistence and progression[44]. On the other hand, a consequence of HPV immune evasion is the imbalance in the vaginal microbiota due the reduced amino acid source sustaining the survival of *Lactobacillus* species[45]. Therefore, the association between BV and HPV seems to be bi-directional. In fact, a study showed that the stability of vaginal microbiome is necessary to maintenance of immune surveillance for HPV16 clearance[46]. In this sense, the differences in the load of nanH3 gene in the presence of HPV16 lead to hypothesize that the vaginal microenvironment would have a different modulation in the presence of different genotypes and the changes in the immune response might create a more permissive microenvironment for bacterial overgrowth.

Gardnerella spp. is present in virtually all cases of BV, the most frequent dysbiosis of the vaginal microbiota[47]. In this study, an association between the nanH3 gene and BV was observed. Robinson et al. suggested the potential use of nanH3 as a molecular diagnostic marker of BV, such PCR test showing 80.95% sensitivity and 78.26% specificity compared with the Nugent score for BV diagnosis[25]. Although the association, in this study values of sensitivity and specificity were lower than the mentioned study. It is worth mentioning that studies differ in terms of prevalence of BV and number of study participants.

Besides the association between nanH3 gene and BV, the load of the gene was significantly increased in BV compared to normal microbiota. As previously discussed, sialidases have deleterious effects in the vaginal microenvironment[20–22,35–37]. Recently, our study group showed that sialidase activity in molecular-BV, assessed by V3–V4 16S rRNA sequencing, is associated with changes in bacterial components of the local microbiome as well as increased sialidase levels[48]. Indeed, sialidase activity in cervicovaginal fluid already was demonstrated in the presence of microscopy-detected BV[49,50]. In fact, sialidase activity could promote the growth or colonization of *Prevotella*, *Bacteroides* as well as *G. vaginalis*, vaginal bacterial sialidase producers[51,52], being the production of sialidase an important step in the biofilm formation[24]. Such mechanism reinforces the role of *Gardnerella* spp. as the scaffold on vaginal mucosa for attachment of other bacterial species, which can start to form biofilm[53], a condition already associated with refractory treatment of BV and persistence of BV-associated bacteria after treatment[23,54].

Although the amount of sialidase-encoding nanH3 gene did not differ between clearance and persistence in the presence of all hrHPV infection, the loads are high in the persistent group and there is a difference among the groups when considering only HPV16 infection. A limitation of this analysis is due the low sample size when stratifying each genotype. Furthermore, such results contribute to better understanding the role of *G. vaginalis* sialidase-encoding gene in BV, since this condition seems have an important role in the persistence and progression of HPV infection. Thus, microbial products are promising tools for identifying women with increased risk for HPV persistence.

Methods

Data collection, sampling and laboratorial procedures.   Between 2012 to 2014, 1635 reproductive-aged HIV-negative women screened for HPV infection in Botucatu, SP, Brazil. A sociodemographic, behavioral, gynecological history and clinical data were obtained during a face-to-face interview using a structured questionnaire and were entered into a Microsoft Excel spreadsheet (Microsoft Corporation, Redmond, WA, USA). HPV positive women were then invited to a follow-up within 12 months. All participants were informed about the procedures of study and informed consent was obtained from all subjects as all reached the age of majority (18 years-old). All methods were carried out in accordance with relevant guidelines and regulations and this current study was reviewed and approved by the Ethics Committee of Botucatu Medical School (Approval number: 5.660.472) and the samples included in the present study belongs to the biorepository of principal study (Biorepository approved number: 3.140.843).

Same sampling procedures during physical examination were conducted by nurses/physicians at enrolment and at follow-up visit. After speculum insertion, vaginal content was obtained with swabs for smearing into microscopic glass slides for Nugent scoring, after gram staining[15]. Mid vaginal wall was utilized for *Trichomonas vaginalis* culture in Diamond's medium[55]. Samples obtained with endocervical brushes were stored at −20 ºC until *Chlamydia trachomatis* and *Neisseria gonorrhoeae* detection, as described previously[55,56] and for HPV detection and genotyping using the Linear Array HPV Genotyping kit (Roche Molecular Systems, Pleasanton, CA, USA). For cervicovaginal fluid sampling, 3 ml of sterile 0.9% NaCl was inserted into vaginal wall, homogenized with cervicovaginal secretion and recovered using a plastic sterile pipette. The cervicovaginal samples were stored at −20 ºC until genomic DNA extraction and *G. vaginalis* nanH3 gene absolute quantification.

Genomic DNA extraction of cervicovaginal fluid was performed using the UltraClean Soil DNA Isolation kit (MoBio Laboratories, Carlsbad, CA, USA) as recommended by the manufacturer. Extracted DNA was quantified using an Epoch spectrophotometer (Biotek, Winooski, VT, USA) and the quality of the extraction was confirmed by the 260/280 nm absorbance ratio.

A cloning step was performed to obtain the plasmid DNA with *Gardnerella* nanH3 gene, through polymerase chain reaction (PCR) products from clinical samples using the primers *Gardnerella* nanH3 F (5′-CAGTTCCAA TGGAAGTGTGC-3′) and *Gardnerella* nanH3 R (5′-AGCATCTGGGAATGCTCTTG-3′) with 51 ℃ of annealing temperature. Expected amplicon size was 322 bp for nanH3 gene, confirmed in 1.5% agaroses gel band[25]. Positive samples had the amplicon purified using Illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare UK Buckinghamshire, UK) and sequenced by Sanger method in ABI 3500 (Applied Biosystems, Foster City, CA, USA) and confirmed by blast (https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_2235500528). The amplified products were used for the ligation step using the CloneJET PCR Cloning Kit (Thermo Fisher Scientific, Carlsbad, CA, USA), cloned into *Escherichia coli* DH5-α (Invitrogen, Carlsbad, CA, USA). Plasmid DNA was extracted using GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific, Carlsbad, CA, USA).

The absolute quantification of nanH3 gene was performed by quantitative real-time PCR (qPCR) using 2× qPCRBIO SyGreen Blue Mix Hi-ROX (PCR Biosystems, London, UK), primers *Gardnerella* nanH3 F and *Gardnerella* nanH3 R under the following cycling conditions: 95 ℃ for 2 min, 25 cycles of denaturation at 95 ℃ for 5 s, annealing at 60 ℃ for 25 s, and an extension at 95 ℃ for 15 s followed by a dissociation stage: 95 ℃ for 15 s; 60 ℃ for 1 min and 95 ℃ for 15 min, according to the manufacturer's recommendations, performed in a StepOnePlus Real-Time System (Thermo Fisher Scientific, Waltham, MA, USA). Samples with a melting temperature value of 81 ℃ +/− 0.5 ℃ were considered positive for nanH3 gene. For constructing the standard curve, three plasmid dilutions (3.05E+05 copies/μL, 3.05E+07 copies/μL and 3.05E+09 copies/μL) were utilized for sample cycle threshold interpolation. Loads of nanH3 gene were expressed as the number of copies per volume (μL) of cervicovaginal fluid.

Selection of participants and group assignments. In the first visit, 544 (33.27%) women were positive for any HPV infection, of which 413 (25.26%) tested positive for hrHPV genotypes and were recruited for the follow-up visit after 12 months. Four hundred sixty-two (27.65%) women returned to the follow-up. For the present study, women positive for *Chlamydia trachomatis* (138, 8.44%), *Trichomonas vaginalis* (23, 1.41%), *Neisseria gonorrhoeae* (2, 0.12%) and only low-risk HPV (131, 8.01%) were not included in the analysis. Five women were excluded of this study due conflicting data in the spreadsheet database. From that, 245 (14.96%) women were hrHPV positive (HPV16, HPV18, HPV31, HPV33, HPV35, HPV39, HPV45, HPV51, HPV52, HPV56, HPV58, HPV59, HPV68, HPV73, HPV82), however, 12 samples were excluded due insufficient volume of cervicovaginal fluid for laboratorial analysis of the current study. Therefore, 233 cervicovaginal fluid samples were included for this study analysis. Participants were assigned at two groups according to the status of hrHPV infection at enrolment and at follow-up: 'clearance group' comprised participants who tested negative for any hrHPV at follow-up and in 'persistence group' were assigned participants positive for hrHPVs genotype at first visit and follow-up, matched with those detected at follow-up. For clearance group, 21 women who showed clearance of baseline genotype and had a new genotype detected at follow-up were not included.

Statistical analysis. For sociodemographic, behavioral and clinical variables comparisons between the clearance and persistence groups, chi-squared or Fisher's exact were utilized for categorical variables and Mann–Whitney tests was utilized for continuous variables. The loads and presence of nanH3 gene were compared between the study groups using, respectively, Mann–Whitney and chi-squared tests or Fisher's exact test. All analyses considered a P-value < 0.05 to be statistically significant and were performed using GraphPad Prism software (version 6.0, GraphPad, CA, USA).

Data availability
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

References
1. de Sanjosé, S. *et al.* Worldwide prevalence and genotype distribution of cervical human papillomavirus DNA in women with normal cytology: A meta-analysis. *Lancet Infect. Dis.* **7**, 453–459. https://doi.org/10.1016/S1473-3099(07)70158-5 (2007).
2. Bruni, L. *et al.* Cervical human papillomavirus prevalence in 5 continents: Meta-analysis of 1 million women with normal cytological findings. *J. Infect. Dis.* **202**, 1789–1799. https://doi.org/10.1086/657321 (2010).
3. Schiffman, M., Castle, P. E., Jeronimo, J., Rodriguez, A. C. & Wacholder, S. Human papillomavirus and cervical cancer. *Lancet* **370**, 890–907. https://doi.org/10.1016/S0140-6736(07)61416-0 (2007).
4. Jaisamrarn, U. *et al.* Natural history of progression of HPV infection to cervical lesion or clearance: Analysis of the control arm of the large, randomised PATRICIA study. *PLoS ONE* **8**, e79260. https://doi.org/10.1371/journal.pone.0079260 (2013).
5. Winer, R. L. *et al.* Early natural history of incident, type-specific human papillomavirus infections in newly sexually active young women. *Cancer Epidemiol. Biomark. Prev.* **20**, 699–707. https://doi.org/10.1158/1055-9965.EPI-10-1108 (2011).
6. Stanley, M. A. & Sterling, J. C. Host responses to infection with human papillomavirus. *Curr. Probl. Dermatol.* **45**, 58–74. https://doi.org/10.1159/000355964 (2014).
7. Eldridge, R. C. *et al.* Smoking and subsequent human papillomavirus infection: A mediation analysis. *Ann. Epidemiol.* **27**, 724-730.e721. https://doi.org/10.1016/j.annepidem.2017.10.004 (2017).
8. Smith, J. S. *et al.* Cervical cancer and use of hormonal contraceptives: A systematic review. *Lancet* **361**, 1159–1167. https://doi.org/10.1016/s0140-6736(03)12949-2 (2003).
9. Moreno, V. *et al.* Effect of oral contraceptives on risk of cervical cancer in women with human papillomavirus infection: The IARC multicentric case–control study. *Lancet* **359**, 1085–1092. https://doi.org/10.1016/S0140-6736(02)08150-3 (2002).
10. Muñoz, N. *et al.* Role of parity and human papillomavirus in cervical cancer: The IARC multicentric case–control study. *Lancet* **30**, 1093–1101. https://doi.org/10.1016/S0140-6736(02)08151-5 (2002).

11. Castle, P. E. & Giuliano, A. R. Chapter 4: Genital tract infections, cervical inflammation, and antioxidant nutrients—Assessing their roles as human papillomavirus cofactors. *J. Natl. Cancer Inst. Monogr.* https://doi.org/10.1093/oxfordjournals.jncimonographs.a003478 (2003).

12. Gillet, E. *et al.* Bacterial vaginosis is associated with uterine cervical human papillomavirus infection: A meta-analysis. *BMC Infect. Dis.* **11**, 10. https://doi.org/10.1186/1471-2334-11-10 (2011).

13. Guo, Y. L., You, K., Qiao, J., Zhao, Y. M. & Geng, L. Bacterial vaginosis is conducive to the persistence of HPV infection. *Int. J. STD AIDS* **23**, 581–584. https://doi.org/10.1258/ijsa.2012.011342 (2012).

14. King, C. C. *et al.* Bacterial vaginosis and the natural history of human papillomavirus. *Infect. Dis. Obstet. Gynecol.* **2011**, 319460. https://doi.org/10.1155/2011/319460 (2011).

15. Nugent, R. P., Krohn, M. A. & Hillier, S. L. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. *J. Clin. Microbiol.* **29**, 297–301. https://doi.org/10.1128/jcm.29.2.297-301.1991 (1991).

16. Marconi, C., Donders, G. G., Parada, C. M., Giraldo, P. C. & da Silva, M. G. Do *Atopobium vaginae*, *Megasphaera* sp. and *Leptotrichia* sp. change the local innate immune response and sialidase activity in bacterial vaginosis?. *Sex Transm. Infect.* **89**, 167–173. https://doi.org/10.1136/sextrans-2012-050616 (2013).

17. Schellenberg, J. J. *et al.* Gardnerella vaginalis subgroups defined by cpn60 sequencing and sialidase activity in isolates from Canada, Belgium and Kenya. *PLoS ONE* **11**, e0146510. https://doi.org/10.1371/journal.pone.0146510 (2016).

18. Kurukulasuriya, S. P., Patterson, M. H. & Hill, J. E. Slipped-strand mispairing in the gene encoding sialidase NanH3 in. *Infect. Immun.* https://doi.org/10.1128/IAI.00583-20 (2021).

19. Cauci, S. *et al.* Determination of immunoglobulin A against *Gardnerella vaginalis* hemolysin, sialidase, and prolidase activities in vaginal fluid: Implications for adverse pregnancy outcomes. *J. Clin. Microbiol.* **41**, 435–438. https://doi.org/10.1128/JCM.41.1.435-438.2003 (2003).

20. Swidsinski, A. *et al.* Adherent biofilms in bacterial vaginosis. *Obstet. Gynecol.* **106**, 1013–1023. https://doi.org/10.1097/01.AOG.0000183594.45524.d2 (2005).

21. Varki, A. Multiple changes in sialic acid biology during human evolution. *Glycoconj. J.* **26**, 231–245. https://doi.org/10.1007/s10719-008-9183-z (2009).

22. Varki, A. & Gagneux, P. Multifarious roles of sialic acids in immunity. *Ann. N. Y. Acad. Sci.* **1253**, 16–36. https://doi.org/10.1111/j.1749-6632.2012.06517.x (2012).

23. Verstraelen, H. & Swidsinski, A. The biofilm in bacterial vaginosis: Implications for epidemiology, diagnosis and treatment: 2018 update. *Curr. Opin. Infect. Dis.* **32**, 38–42. https://doi.org/10.1097/QCO.0000000000000516 (2019).

24. Santiago, G. L. *et al. Gardnerella vaginalis* comprises three distinct genotypes of which only two produce sialidase. *Am. J. Obstet. Gynecol.* **204**(450), e451-457. https://doi.org/10.1016/j.ajog.2010.12.061 (2011).

25. Robinson, L. S., Schwebke, J., Lewis, W. G. & Lewis, A. L. Identification and characterization of NanH2 and NanH3, enzymes responsible for sialidase activity in the vaginal bacterium. *J. Biol. Chem.* **294**, 5230–5245. https://doi.org/10.1074/jbc.RA118.006221 (2019).

26. Vaneechoutte, M. *et al.* Emended description of *Gardnerella vaginalis* and description of *Gardnerella leopoldii* sp. nov., *Gardnerella piotii* sp. nov. and *Gardnerella swidsinskii* sp. nov., with delineation of 13 genomic species within the genus *Gardnerella. Int. J. Syst. Evol. Microbiol.* **69**, 679–687. https://doi.org/10.1099/ijsem.0.003200 (2019).

27. Belleti, R. *et al.* Cervicovaginal loads of *Gardnerella* spp. are increased in immunocompetent women with persistent high-risk human papillomavirus infection. *J. Med. Microbiol.* https://doi.org/10.1099/jmm.0.001527 (2022).

28. Miranda, P. M. *et al.* Persistence or clearance of human papillomavirus infections in women in Ouro Preto, Brazil. *Biomed. Res. Int.* **2013**, 578276. https://doi.org/10.1155/2013/578276 (2013).

29. Li, M. *et al.* Incidence, persistence and clearance of cervical human papillomavirus among women in Guangdong, China 2007–2018: A retrospective cohort study. *J. Infect. Public Health* **14**, 42–49. https://doi.org/10.1016/j.jiph.2020.11.011 (2021).

30. Ramanakumar, A. V. *et al.* Incidence and duration of type-specific human papillomavirus infection in high-risk HPV-naïve women: Results from the control arm of a phase II HPV-16/18 vaccine trial. *BMJ Open* **6**, e011371. https://doi.org/10.1136/bmjopen-2016-011371 (2016).

31. van der Weele, P. *et al.* Correlation between viral load, multiplicity of infection, and persistence of HPV16 and HPV18 infection in a Dutch cohort of young women. *J. Clin. Virol.* **83**, 6–11. https://doi.org/10.1016/j.jcv.2016.07.020 (2016).

32. Sammarco, M. L., Del Riccio, I., Tamburro, M., Grasso, G. M. & Ripabelli, G. Type-specific persistence and associated risk factors of human papillomavirus infections in women living in central Italy. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **168**, 222–226. https://doi.org/10.1016/j.ejogrb.2013.01.012 (2013).

33. Di Paola, M. *et al.* Characterization of cervico-vaginal microbiota in women developing persistent high-risk Human Papillomavirus infection. *Sci. Rep.* **7**, 10200. https://doi.org/10.1038/s41598-017-09842-6 (2017).

34. Usyk, M. *et al.* Cervicovaginal microbiome and natural history of HPV in a longitudinal study. *PLoS Pathog.* **16**, e1008376. https://doi.org/10.1371/journal.ppat.1008376 (2020).

35. Lewis, W. G., Robinson, L. S., Gilbert, N. M., Perry, J. C. & Lewis, A. L. Degradation, foraging, and depletion of mucus sialoglycans by the vagina-adapted *Actinobacterium Gardnerella vaginalis. J. Biol. Chem.* **288**, 12067–12079. https://doi.org/10.1074/jbc.M113.453654 (2013).

36. Howe, L. *et al.* Mucinase and sialidase activity of the vaginal microflora: Implications for the pathogenesis of preterm labour. *Int. J. STD AIDS* **10**, 442–447. https://doi.org/10.1258/0956462991914438 (1999).

37. Lewis, W. G. *et al.* Hydrolysis of secreted sialoglycoprotein immunoglobulin A (IgA) in ex vivo and biochemical models of bacterial vaginosis. *J. Biol. Chem.* **287**, 2079–2089. https://doi.org/10.1074/jbc.M111.278135 (2012).

38. Brotman, R. M. *et al.* Interplay between the temporal dynamics of the vaginal microbiota and human papillomavirus detection. *J. Infect. Dis.* **210**, 1723–1733. https://doi.org/10.1093/infdis/jiu330 (2014).

39. Boskey, E. R., Telsch, K. M., Whaley, K. J., Moench, T. R. & Cone, R. A. Acid production by vaginal flora in vitro is consistent with the rate and extent of vaginal acidification. *Infect. Immun.* **67**, 5170–5175. https://doi.org/10.1128/IAI.67.10.5170-5175.1999 (1999).

40. O'Hanlon, D. E., Moench, T. R. & Cone, R. A. In vaginal fluid, bacteria associated with bacterial vaginosis can be suppressed with lactic acid but not hydrogen peroxide. *BMC Infect. Dis.* **11**, 200. https://doi.org/10.1186/1471-2334-11-200 (2011).

41. Alakomi, H. L. *et al.* Lactic acid permeabilizes Gram-negative bacteria by disrupting the outer membrane. *Appl. Environ. Microbiol.* **66**, 2001–2005. https://doi.org/10.1128/AEM.66.5.2001-2005.2000 (2000).

42. Ciapponi, A., Bardach, A., Glujovsky, D., Gibbons, L. & Picconi, M. A. Type-specific HPV prevalence in cervical cancer and high-grade lesions in Latin America and the Caribbean: Systematic review and meta-analysis. *PLoS ONE* **6**, e25493. https://doi.org/10.1371/journal.pone.0025493 (2011).

43. Zhou, C., Tuong, Z. K. & Frazer, I. H. Papillomavirus immune evasion strategies target the infected cell and the local immune system. *Front. Oncol.* **9**, 682. https://doi.org/10.3389/fonc.2019.00682 (2019).

44. Audirac-Chalifour, A. *et al.* Cervical microbiome and cytokine profile at various stages of cervical cancer: A pilot study. *PLoS ONE* **11**, e0153274. https://doi.org/10.1371/journal.pone.0153274 (2016).

45. Lebeau, A. *et al.* HPV infection alters vaginal microbiome through down-regulating host mucosal innate peptides used by *Lactobacilli* as amino acid sources. *Nat. Commun.* **13**, 1076. https://doi.org/10.1038/s41467-022-28724-8 (2022).

46. Moscicki, A. B., Shi, B., Huang, H., Barnard, E. & Li, H. Cervical-vaginal microbiome and associated cytokine profiles in a prospective study of HPV 16 acquisition, persistence, and clearance. *Front. Cell Infect. Microbiol.* **10**, 569022. https://doi.org/10.3389/fcimb.2020.569022 (2020).
47. Schwebke, J. R., Muzny, C. A. & Josey, W. E. Role of *Gardnerella vaginalis* in the pathogenesis of bacterial vaginosis: A conceptual model. *J. Infect. Dis.* **210**, 338–343. https://doi.org/10.1093/infdis/jiu089 (2014).
48. Ferreira, C. S. T., Marconi, C., Parada, C. M. G. L., Ravel, J. & da Silva, M. G. Sialidase activity in the cervicovaginal fluid is associated with changes in bacterial components of *Lactobacillus*-deprived microbiota. *Front. Cell Infect. Microbiol.* **11**, 813520. https://doi.org/10.3389/fcimb.2021.813520 (2021).
49. Marconi, C. *et al.* Sialidase activity in aerobic vaginitis is equal to levels during bacterial vaginosis. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **167**, 205–209. https://doi.org/10.1016/j.ejogrb.2012.12.003 (2013).
50. Santos-Greatti, M. M. V., da Silva, M. G., Ferreira, C. S. T. & Marconi, C. Cervicovaginal cytokines, sialidase activity and bacterial load in reproductive-aged women with intermediate vaginal flora. *J. Reprod. Immunol.* **118**, 36–41. https://doi.org/10.1016/j.jri.2016.08.005 (2016).
51. Briselden, A. M., Moncla, B. J., Stevens, C. E. & Hillier, S. L. Sialidases (neuraminidases) in bacterial vaginosis and bacterial vaginosis-associated microflora. *J. Clin. Microbiol.* **30**, 663–666. https://doi.org/10.1128/jcm.30.3.663-666.1992 (1992).
52. Olmsted, S. S., Meyn, L. A., Rohan, L. C. & Hillier, S. L. Glycosidase and proteinase activity of anaerobic gram-negative bacteria isolated from women with bacterial vaginosis. *Sex Transm. Dis.* **30**, 257–261. https://doi.org/10.1097/00007435-200303000-00016 (2003).
53. Muzny, C. A. & Schwebke, J. R. Pathogenesis of bacterial vaginosis: Discussion of current hypotheses. *J. Infect. Dis.* **214**(Suppl 1), S1-5. https://doi.org/10.1093/infdis/jiw121 (2016).
54. Swidsinski, A. *et al.* An adherent *Gardnerella vaginalis* biofilm persists on the vaginal epithelium after standard therapy with oral metronidazole. *Am. J. Obstet. Gynecol.* **198**(97), e91-96. https://doi.org/10.1016/j.ajog.2007.06.039 (2008).
55. Marconi, C., Duarte, M. T., Silva, D. C. & Silva, M. G. Prevalence of and risk factors for bacterial vaginosis among women of reproductive age attending cervical screening in southeastern Brazil. *Int. J. Gynaecol. Obstet.* **131**, 137–141. https://doi.org/10.1016/j.ijgo.2015.05.016 (2015).
56. Goire, N., Nissen, M. D., LeCornec, G. M., Sloots, T. P. & Whiley, D. M. A duplex Neisseria gonorrhoeae real-time polymerase chain reaction assay targeting the gonococcal porA pseudogene and multicopy opa genes. *Diagn. Microbiol. Infect. Dis.* **61**, 6–12. https://doi.org/10.1016/j.diagmicrobio.2007.12.007 (2008).

Author contributions
C.M. and M.G. conceived and supervised the study. A.N.B. and G.V.S.P. recruited volunteers. J.N., R.F., A.N.B. and G.V.S.P. performed the experiments. J.N. performed the statistical analysis and interpreted the data. J.N. wrote the manuscript under supervision of C.M. All authors read, reviewed and approved the final manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-41469-8.

Correspondence and requests for materials should be addressed to J.N.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.