Journal of Medical Microbiology (2014), **63**, 162–175 **DOI** 10.1099/jmm.0.066407-0

Identification, quantification and subtyping of *Gardnerella vaginalis* in noncultured clinical vaginal samples by quantitative PCR

Sergey V. Balashov, Eli Mordechai, Martin E. Adelson and Scott E. Gygax

Femeris Women's Health Research Center, Medical Diagnostic Laboratories, L. C., Genesis Biotechnology Group, Hamilton, NJ, USA

Correspondence
Sergey V. Balashov
sbalashov@femeris.com

Received 30 July 2013
Accepted 6 November 2013

Gardnerella vaginalis is an important component of the human vaginal microflora. It is proposed to play a key role in the pathogenesis of bacterial vaginosis (BV), the most common vaginal condition. Here we describe the development, validation and comparative analysis of a novel molecular approach capable of *G. vaginalis* identification, quantification and subtyping in noncultured vaginal specimens. Using two quantitative PCR (qPCR) assays, we analysed *G. vaginalis* bacterial loads and clade distribution in 60 clinical vaginal-swab samples. A very high pathogen prevalence was revealed by species-specific qPCR not only among BV patients (100 %), but also in healthy women (97 %), although the *G. vaginalis* concentration was significantly lower in non-BV samples. *G. vaginalis* clades identified in vaginal specimens by subtyping multiplex qPCR, which targets four clade-specific genetic markers, had frequencies of 53 % for clade 1, 25 % for clade 2, 32 % for clade 3 and 83 % for clade 4. Multiple clades were found in 70 % of samples. Single *G. vaginalis* clades were represented by clade 1 and clade 4 in 28 % of specimens. A positive association with BV was shown for clade 1 and clade 3, while clade 2 was positively associated with intermediate vaginal microflora, but not with BV. Clade 4 demonstrated no correlation with the disorder. The presence of multiple clades had a high positive association with BV, whereas *G. vaginalis* identified as a single clade was negatively linked with the condition. Polyclonal *G. vaginalis* infection may be a risk factor for BV.

INTRODUCTION

Gardnerella vaginalis is a facultatively anaerobic, catalase- and oxidase-negative bacterium. *G. vaginalis* cells are pleomorphic, Gram-negative to Gram-variable, non-encapsulated and non-motile rods with a mean size of 0.5 to 1.5 μm (Greenwood & Pickett, 1980). Initially named *Haemophilus vaginalis* by discoverers, the micro-organism was later referred to as *Corynebacterium vaginale* and taxonomically assigned as *G. vaginalis* (Gardner & Dukes, 1955; Piot *et al.*, 1980). Currently it is the only species within the genus *Gardnerella* (List of Prokaryotic Names with Standing in Nomenclature – www.bacterio.net). *G. vaginalis* is mainly considered a part of the lower female genital tract microflora (Catlin, 1992). It can also be routinely isolated from the male urogenital tract as was reported in the first publication on this micro-organism and a number of other studies (Briselden & Hillier, 1990; Eren *et al.*, 2011; Leopold, 1953; Swidsinski *et al.*, 2010). The presence of *G. vaginalis* in the oral cavity and anal samples has also been described (El Aila *et al.*, 2011; Holst, 1990; Marrazzo *et al.*, 2012). Along with disorders in the urinary and genital tracts, *G. vaginalis* has been identified as a causative agent of bacteraemia, septicaemia with infective endocarditis, vertebral osteomyelitis, acute hip arthritis and retinal vasculitis (Graham *et al.*, 2009; Lagacé-Wiens *et al.*, 2008; Neri *et al.*, 2009; Sivadon-Tardy *et al.*, 2009; Yoon *et al.*, 2010).

Since the moment of discovery *G. vaginalis* has been linked with bacterial vaginosis (BV), the most common vaginal condition. BV is characterized by the elevation of vaginal pH and clinical symptoms, such as malodorous vaginal discharge and the presence of clue cells visualized on wet mount, although a large percentage of women with BV can be asymptomatic (Amsel *et al.*, 1983; Schwebke & Desmond, 2007). The disorder is associated with a dramatic change in vaginal microflora, when protective *Lactobacillus* species are depleted and replaced by fastidious anaerobic bacteria including *Prevotella*, *Atopobium*, *Megasphaera*, *Sneathia* and *G. vaginalis* (Eschenbach, 1993; Forsum *et al.*, 2005; Nugent *et al.*, 1991). Described as 'a single etiological agent' of BV by the discoverers, *G. vaginalis*, however, failed to fulfil Koch's postulates for microbial infection in multiple subsequent studies (Catlin, 1992; Holst, 1990; Menard *et al.*,

Abbreviations: ARDRA, amplified ribosomal DNA restriction analysis; ATCC, American Type Culture Collection; BV, bacterial vaginosis; CI, confidence interval; C_t, cycle threshold; LOD, limit of detection; qPCR, quantitative PCR.

2008; Zozaya-Hinchliffe *et al.*, 2010). The complex polymicrobial nature of BV and the presence of *G. vaginalis* in the vaginal milieu of healthy individuals argue against the definition of *G. vaginalis* as a sole disease-causative species. The micro-organism, nevertheless, remains a prime suspect in the pathogenesis of BV since it is present in 95–100 % of BV patients, although the association of *G. vaginalis* with *Atopobium* or *Megasphaera* has been shown to display better predictive value for this condition (Fredricks *et al.*, 2009; Menard *et al.*, 2008; Muzny & Schwebke, 2013; Schellenberg *et al.*, 2011). It was proposed that G. *vaginalis* might serve as an initial colonizer critical for the development of BV (Patterson *et al.*, 2010). Some studies suggest that *G. vaginalis* dominance in the vaginal microflora can represent a distinct transitional or intermediate state between healthy microflora occupied by lactobacilli and BV-type microflora populated with fastidious anaerobes (Schellenberg *et al.*, 2011). After more than half a century the role of *G. vaginalis* in such an 'enigmatic' condition as BV remains obscure.

As a species *G. vaginalis* displays a wide range of metabolic phenotypes. Although the majority of clinical *G. vaginalis* isolates are catalase- and oxidase-negative, and capable of human blood β-haemolysis, a great variability exists in carbohydrate fermentation, sialidase production, hydrolysis of hippurate, lipase and β-galactosidase activities (Catlin, 1992; Piot *et al.*, 1982). The three latter enzymic characteristics were used for the development of a biotyping scheme facilitating studies of *G. vaginalis* epidemiology (Piot *et al.*, 1984). Strains belonging to eight possible biotypes have been isolated with different frequencies depending on geographical location and patient health status (Aroutcheva *et al.*, 2001; Briselden & Hillier, 1990; Pleckaityte *et al.*, 2012; Udayalaxmi *et al.*, 2011). Multiple biotype carriage was determined in a number of studies and considered fairly common. The distribution of biotypes was found to be different among women with and without BV; however, there exists a considerable disagreement about associating particular biotypes with BV between studies (Briselden & Hillier, 1990; Numanović *et al.*, 2008; Piot *et al.*, 1984). Other studies showed no statistically significant differences between the pathogen biotypes in BV patients and subjects with normal vaginal microflora, which led authors to conclude that no specific phenotype or genotype of *G. vaginalis* is pathogenic and causes BV (Aroutcheva *et al.*, 2001; Tosun *et al.*, 2007). Absence of a correlation between *G. vaginalis* biotypes/genotypes and virulence factors, such as adherence to vaginal epithelial cells, biofilm production, surface hydrophobicity, vaginolysin production, and phospholipase C and protease activities, was also demonstrated (Pleckaityte *et al.*, 2012; Udayalaxmi *et al.*, 2011). Conflicting results on *G. vaginalis* biotypes and BV association might be partially explained by deficiencies of the phenotypic approach, which is prone to erroneous biotype identification (Moncla & Pryke, 2009).

Limitations of the phenotypic *G. vaginalis* biotyping scheme were addressed in subsequent studies suggesting modified techniques, though still based on bacterial culture (Benito *et al.*, 1986). A number of molecular approaches, such as random amplified polymorphic DNA (RAPD) and amplified ribosomal DNA restriction analysis (ARDRA), were shown to be useful for the identification of three to four different *G. vaginalis* genotypes (Ingianni *et al.*, 1997). Although a link with sialidase production was found in two distinct genotypes, a correlation between BV and any of the genotypes could not be identified (Santiago *et al.*, 2011). More recently, advances in next generation sequencing technology have allowed the differentiation of *G. vaginalis* strains and subgroups according to sequence variations in 16S rRNA and the *cpn60* genes, as well as whole genome sequencing analysis (Ahmed *et al.*, 2012; Eren *et al.*, 2011; Paramel Jayaprakash *et al.*, 2012). Genome-level comparisons demonstrated significant divergence between *G. vaginalis* strains isolated from BV patients and healthy women (Harwich *et al.*, 2010; Yeoman *et al.*, 2010). The comprehensive analysis of the genomes of 17 *G. vaginalis* clinical isolates revealed profound differences in genome size, G+C DNA content and genic composition, allowing their separation into four clades with distinct gene pools and genomic properties (Ahmed *et al.*, 2012). Sets of unique genetic biomarkers were identified for each clade, suggesting differences in metabolic capabilities and virulence potentials between clades.

Captivated by the extraordinary high level of genomic variability between four distinct *G. vaginalis* clades revealed by genomic analysis, we explored the utility of conserved clade-specific genetic markers for a new molecular culture-independent typing scheme. Here we report the development of two quantitative PCR (qPCR) assays for species-specific identification of *G. vaginalis* and differentiation of four genetic clades. The assay validation included cross-reactivity, accuracy, PCR efficiency, interference, limit of detection (LOD), and analytical specificity and sensitivity determination. qPCR assays were evaluated using *G. vaginalis* clinical isolates, other microbial cultures, human chromosomal DNA, plasmid DNA and vaginal swabs. The new qPCR-based typing approach was shown to be specific for *G. vaginalis* clades and capable of micro-organism identification, quantification and subtyping in noncultured vaginal samples. Using a set of characterized vaginal specimens from patients with BV and healthy individuals, we demonstrated significant differences in the association of four *G. vaginalis* clades with BV.

METHODS

Patient specimens, BV diagnoses, DNA samples and microbial strains. One hundred vaginal-swab specimens were collected at the Vaginitis Clinic, Wayne State University School of Medicine, Detroit, MI, USA, as a part of a multidisciplinary study on BV pathogenesis. Samples were shipped to Medical Diagnostic Laboratories, L. C., frozen in UTM-RT universal transport medium (Copan Italia), from August to October of 2012.

Clinical and laboratory assessments of BV condition were performed in 60 vaginal-sample donors at the Vaginitis Clinic according to Amsel criteria and Nugent scores (Amsel *et al.*, 1983; Nugent *et al.*,

1991). Nugent scores were based on the enumeration of different vaginal bacterial morphotypes by Gram stain and microscopy of a vaginal sample, in which a scoring range is associated with normal flora (score 0–3), intermediate flora (score 4–6) or abnormal flora (score 7–10). Amsel criteria were based on a clinical evaluation that includes an elevation of the vaginal pH >4.5, the presence of clue cells using wet prep microscopy, homogeneous white vaginal discharge and fishy odour (positive KOH amine test). A patient was required to have three of the four criteria, including the mandatory presence of elevated pH and the presence of clue cells, to be classified as positive for BV. Forty vaginal-sample donors were not assessed for BV conditions. Vaginal swabs from these 40 patients were used as uncharacterized specimens.

Informed consent was obtained from all subjects. Inclusion criteria included premenopausal women older than 18 years of age. Exclusion criteria included women who had received any vaginal anti-inflammatory or anti-histamine treatment, received antibiotics within 30 days, or who were pregnant, immunocompromised or undergoing treatment for any urogenital infectious disease other than BV. The median age of the participants was 32.5 years (range 20 to 46 years). Human experimentation guidelines of the US Department of Health and Human Services, as well as those of the investigators' institutions, were strictly followed.

Micro-organisms from three different sources were used in the study. Twenty-four *G. vaginalis* strains (Table 1) were isolated from nine characterized and nine uncharacterized vaginal-swab specimens by plating aliquots of UTM-RT transport medium on brain heart infusion agar (Becton Dickinson) supplemented with 5 % sheep blood and *G. vaginalis* selective supplement (Oxoid) and incubating them anaerobically at 37 °C for 7 days. Presumptive identification of the isolates as *G. vaginalis* was based on β-haemolysis on human blood–Tween bilayer agar plates (Becton Dickinson) and Gram-variable staining under microscopy examination. Twelve *G. vaginalis* strains described previously were kindly provided by Dr Garth D. Ehrlich, Center for Genomic Sciences, Allegheny-Singer Research Institute, Pittsburgh, PA, USA (Ahmed *et al.*, 2012). Three *G. vaginalis* strains (Table 1), and thirty-eight strains of bacterial, fungal and protozoan species (Table 2), were purchased from the American Type Culture Collection (ATCC) (Manassas, VA, USA). Human chromosomal DNA was purchased from Promega.

DNA extraction and sequencing. DNA from clinical samples was extracted using the QIAamp Mini kit (Qiagen) in combination with pretreatment with proteinase K and mechanical homogenization. Vaginal specimens were thawed at room temperature and vortexed vigorously. A 600 µl aliquot of UTM-RT transport medium was pipetted into a lysing matrix E tube (MP Biomedicals) and combined with 60 µl proteinase K (Qiagen). Mechanical disruption was performed on a FastPrep-24 homogenizer (MP Biomedicals) for 1 min at 4 m s^{-1} speed. After homogenization, 400 µl UTM-RT transport medium was transferred to a new 1.5 ml microcentifuge tube and subjected to a QIAamp Mini kit DNA extraction procedure according to manufacturer's recommendations. DNA from bacterial and fungal cultures, including ATCC lyophilized cultures, was extracted by the same method following rehydration in distilled water if needed. Synthetic *tuf* control plasmid DNA was extracted from bacterial cultures using a Plasmid Mini kit (Qiagen). DNA sequencing of PCR products was performed on a CEQ8000 automated DNA analyser using a GenomeLab DTCS Quick Start kit (Beckman Coulter). The M13 −20 forward sequencing primer 5′-GTAAAACGACGG-CCAGT-3′ and the M13 −26 reverse sequencing primer 5′-AACAGC-TATGACCATG-3′ were used for PCR and sequencing when synthetic control plasmid served as a template. Primers Gv_tuf_S6 and Gv_tuf_AS7 (Table 3) were used for PCR and sequencing when PCR fragments were generated by amplifying the *G. vaginalis tuf* gene.

***G. vaginalis* conventional biotyping and ARDRA genotyping.** *G. vaginalis* phenotypic biotyping was performed using standard biochemical tests: hydrolysis of hippurate, and lipase and β-galactosidase activities as described elsewhere (Piot *et al.*, 1984). Original egg yolk agar prepared with egg yolk 50 % enrichment (Becton Dickinson), but not the 4-methylumbelliferyl-oleate spot test, was used for lipase activity determination (Moncla & Pryke, 2009). *G. vaginalis* strains ATCC 14018 and B512 were used as positive and negative controls, respectively, for all three tests in the biotyping experiments (Ahmed *et al.*, 2012; Paramel Jayaprakash *et al.*, 2012). For the ARDRA, *G. vaginalis* 16S rRNA gene fragments were amplified with primers GV10F 5′-GGTTCGATTCTGGCTCAG-3′ and ωMB 5′-TACCTTGTTACGACTTCGTCCCA-3′ as described elsewhere (Santiago *et al.*, 2011). PCR fragments were digested with *Hpa*II restriction endonuclease (New England Biolabs) according to the manufacturer's recommendations. Restriction digest fragments were analysed using a 2100 Bioanalyser (Agilent Technologies) and compared with the original genotyping patterns (Ingianni *et al.*, 1997).

qPCRs and qPCR controls. qPCR primers and TaqMan probes were designed using the deposited chromosomal sequences of *G. vaginalis* strains ATCC 14019, 00703C2mash, 00703Dmash and 409-05, with corresponding GenBank accession numbers NC_014644, ADEU01000000, ADEV01000000 and NC_013721 (Ahmed *et al.*, 2012). Oligonucleotides (Table 3) were purchased from Integrated DNA Technologies. Multiplex and uniplex TaqMan qPCRs developed in this study were performed in 25 µl reactions containing 1× PerfeCTa qPCR UNG SuperMix (Quanta BioSciences), 800 nM each DNA primer, 100 nM each TaqMan probe and 2.5 µl DNA. Uniplex SYBR Green qPCRs were performed in 25 µl reactions containing 1× PerfeCTa SYBR Green UNG SuperMix (Quanta BioSciences), 400 nM each DNA primer and 2.5 µl DNA. The cycling parameters were 45 °C for 5 min UNG treatment, 95 °C for 3 min initial denaturation, 40 cycles denaturation at 95 °C for 15 s, and annealing plus extension at 60 °C for 45 s with fluorescence acquisition at the end of each cycle. Previously published uniplex TaqMan qPCRs targeting 16S rRNA and *cpn60* genes were performed as described (Fredricks *et al.*, 2009; Menard *et al.*, 2008). MX300P instruments (Agilent Technologies) were used for all qPCR experiments.

Synthetic control plasmid was constructed by cloning PCR fragments of the *tuf* gene of *G. vaginalis* ATCC isolate 14018 (Table 1), amplified with primers Gv_tuf_S4 and Gv_tuf_AS3 (Table 3), into pCR2.1 vector using a TOPO TA cloning kit (Life Technologies). Plasmid preparations were tested for DNA concentration using a Nanodrop 2000 instrument (Thermo Fisher Scientific), diluted serially and used as quantitative standards in all qPCR experiments. A concentration range of 0 to 10^7 copies per reaction was used for LOD determination. For quantitative assessment of *G. vaginalis* bacterial loads in clinical samples, four synthetic *tuf* control plasmid concentrations 10, 10^3, 10^5 and 10^7 copies per µl were applied in duplicate to eight separate qPCRs containing primers Gv_tuf_S4, Gv_tuf_AS3 and TaqMan probe Gv_tuf_TM1 (Table 3).

Data analysis. qPCR results were analysed with the MX300P software version 4.10. The interassay analytical sensitivity and analytical specificity were calculated using InStat software (GraphPad Software) version 3.06 using a two-way contingency table analysis and Fisher's exact test with 95 % confidence interval (CI). The association analysis was performed with InStat software applying the Pearson *r* correlation test. MEGALIGN software version 6.1 (DNASTAR) and the Web-based National Center for Biotechnology Information Basic Local Alignment Search Tool (BLAST) were used for DNA alignment.

Table 1. Thirty-nine *G. vaginalis* strains and their characterization by phenotypic biotyping, ARDRA genotyping and qPCRs

Pairs of *G. vaginalis* strains of different clades isolated from vaginal specimens 015, 028, 052, 076, 077 and 090 are in bold. ND, Not determined.

G. vaginalis strain	Source/sample	Clade by qPCR	Biotype	Genotype	*tuf* qPCR	16S rRNA qPCR	*cpn*60 qPCR
14018	ATCC	1	1	1	+	+	+
14019	ATCC	1	4	1	+	+	+
49145	ATCC	1	4	1	+	+	+
B472	ASRI*	1	1	1	+	+	+
B473	ASRI*	1	1	1	+	+	+
B474	ASRI*	1	1	1	+	+	+
B475	ASRI*	4	5	3	+	+	−
B476	ASRI*	4	5	3	+	+	−
B477	ASRI*	1	1	1	+	+	+
B478	ASRI*	1	4	ND	+	+	+
B479	ASRI*	3	5	3	+	+	+
B482	ASRI*	ND	5	4	+	+	−
B483	ASRI*	3	7	3	+	+	+
B512	ASRI*	3	7	3	+	+	+
B513	ASRI*	2	5	4	+	+	−
S0012	001	1	1	1	+	+	+
S0070	007	4	5	3	+	+	−
S0151	015	1	4	ND	+	+	+
S0154	015	4	5	3	+	+	−
S0191	019	1	ND	1	+	+	+
S0244	024	1	4	1	+	+	+
S0282	028	1	1	1	+	+	+
S0284	028	2	5	ND	+	+	−
S0380	038	4	5	ND	+	+	−
S0432	043	1	1	1	+	+	+
S0523	052	4	5	3	+	+	−
S0524	052	1	1	1	+	+	+
S0741	074†	1	4	1	+	+	+
S0762	076†	2	7	3	+	+	−
S0763	076†	1	1	1	+	+	+
S0771	077†	2	5	4	+	+	−
S0775	077†	4	5	3	+	+	−
S0822	082†	1	ND	1	+	+	+
S0835	083†	1	1	1	+	+	+
S0841	084†	2	7	3	+	+	−
S0853	085†	1	4	1	+	+	+
S0901	090†	2	7	3	+	+	−
S0903	090†	1	4	1	+	+	+
S0925	092†	1	4	1	+	+	+

*Twelve *G. vaginalis* strains were kindly provided by Dr Garth D. Ehrlich, Center for Genomic Sciences, Allegheny-Singer Research Institute (ASRI) (Ahmed *et al.*, 2012).

†Vaginal specimens 074, 076, 077, 082, 083, 084, 085, 090 and 092 have not been characterized by the conventional techniques and qPCRs. They are not included into the list of characterized samples in Table 4.

RESULTS

G. vaginalis species-specific qPCR development and validation

Using GenBank-deposited *G. vaginalis* genomic sequences, we searched for chromosomal genes conserved in all four clades. Among multiple conserved housekeeping gene candidates, including *tuf*, *vly*, *gyrA*, *polB*, *groL*, *gyrB*, *recA* and others, the *tuf* gene, encoding translation elongation factor Tu, was chosen for the development of the *G. vaginalis* species-specific qPCR. qPCR primers and TaqMan probes were designed targeting *tuf* sequences conserved among the *G. vaginalis* clades and distinct from species of the closely related genus *Bifidobacterium* by BLAST analysis. Different primers and TaqMan probe combinations were tested in conventional PCRs and qPCRs using

Table 2. Thirty-eight microbial ATCC strains used in the study

Microbial species	ATCC strain
Atopobium vaginae	BAA-55
Bacteroides ureolyticus	33387
Bifidobacterium angulatum	27535
Bifidobacterium animalis	25527
Bifidobacterium bifidum	29521
Bifidobacterium dentium	27534
Bifidobacterium longum	15707
Candida albicans	90028
Chlamydia trachomatis	VR-901B
Corynebacterium genitalium	33030
Cryptococcus neoformans	32045
Enterobacter aerogenes	13048
Enterococcus faecalis	700221
Enterococcus faecium	19434
Escherichia coli	11303
Klebsiella oxytoca	13182
Lactobacillus crispatus	33197
Lactobacillus gasseri	19992
Lactobacillus iners	55195
Lactobacillus jensenii	25258
Lactococcus lactis	19435
Leptotrichia buccalis	14201
Listeria monocytogenes	7644
Moraxella catarrhalis	25238
Mycoplasma hominis	15488
Neisseria gonorrhoeae	19424
Peptococcus niger	27731
Peptostreptococcus anaerobius	27337
Prevotella bivia	29303
Proteus mirabilis	29906
Pseudomonas aeruginosa	BAA427
Salmonella typhimurium	49416
Staphylococcus aureus	25923
Staphylococcus epidermidis	12228
Streptococcus agalactiae	A909
Streptococcus pyogenes	BAA595
Trichomonas vaginalis	30246
Ureaplasma urealyticum	27618

chromosomal DNA of *G. vaginalis* type strain ATCC 14018 as a template. Rejection criteria included PCR specificity, amplification efficiency, fluorescence output and primer dimer formation. Primers Gv_tuf_S4 and Gv_tuf_AS3, and fluorescently labelled TaqMan probe Gv_tuf_TM1, demonstrated the best performance (Table 3).

G. vaginalis species-specific qPCR was tested for cross-reactivity, accuracy, LOD and PCR efficiency. Human chromosomal DNA plus chromosomal DNA extracted from two sets of micro-organisms in concentrations of 25 ng per reaction were used for cross-reactivity determination: 15 strains of *G. vaginalis* comprising 3 ATCC isolates and 12 previously described strains (Table 1), and 38 strains of bacterial, fungal and protozoan species (Table 2) indigenous to the vaginal environment (Ahmed *et al.*, 2012). No cross-reactivity was demonstrated for the *tuf*-targeting *G. vaginalis* qPCRs against all species tested. All 15 qPCRs containing *G. vaginalis* DNAs produced positive signals with similar cycle threshold (C_t) scores. For determining accuracy, amplification products of the *tuf* gene generated from the PCR using *G. vaginalis* ATCC 14018 DNA as template and primers Gv_tuf_S4 and Gv_tuf_AS3 were cloned into the pCR2.1 plasmid and sequenced. The control *tuf* plasmid was quantified, serially diluted and used to determine the qPCR amplification efficiency and LOD using a template concentration range of 0 to 10^7 copies per reaction. The lowest concentration in which the assay demonstrated 100 % positivity in triplicate qPCRs was 10 copies per reaction, which was considered the LOD. Amplification efficiency (E) was 110.4 %, the coefficient of determination (R^2) was 0.999 and linear regression was $y=-3.096\log x+37.86$.

Interassay sensitivity and specificity of the *G. vaginalis* species-specific qPCR were determined by comparative analysis of 60 DNA samples extracted from noncultured clinical vaginal-swab specimens. Two previously described qPCR assays targeting *G. vaginalis* 16S rRNA and *cpn60* genes were used in this experiment (Fredricks *et al.*, 2009; Menard *et al.*, 2008). Out of 60 specimens tested, 59 were positive for *G. vaginalis* by the *tuf* qPCR described here, 58 samples were positive by the 16S rRNA qPCR and 44 samples were positive by the *cpn60* qPCR. The interassay sensitivity and specificity against 16S rRNA qPCR were calculated as 100 % (95 % CI 0.938 to 1.00) and 50 % (95 % CI 0.013 to 0.987), respectively. Interassay sensitivity and specificity against *cpn60* qPCR were calculated as 100 % (95 % CI 0.920 to 1.00) and 6 % (95 % CI 0.002 to 0.302), respectively. Results for the three qPCRs used for analytical sensitivity and specificity determination are shown in Table 4. Quantitative assessment of *G. vaginalis* DNA by the *tuf* qPCR revealed a wide range of concentrations from single copies to 10^7 genomic copies μl^{-1} in the extracted DNA samples. Median concentrations in BV-negative and BV-positive samples defined by Amsel criteria were 1.9×10^4 and 1.6×10^7 copies μl^{-1}, respectively. When three sample groups were divided according to Nugent scores 0–3, 4–6 and 7–10 corresponding *G. vaginalis* median DNA concentrations were 2.0×10^3, 2.4×10^6 and 1.7×10^7 copies μl^{-1}. The quantitative results of bacterial loads in clinical samples are summarized in Fig. 1.

To confirm the species specificity of the developed *tuf* qPCR for *G. vaginalis*, we amplified a larger fragment of the *tuf* gene using primers Gv_tuf_S6 and Gv_tuf_AS7 (Table 3). The same DNA samples from noncultured clinical vaginal-swab specimens used in the interassay sensitivity and specificity experiments were used as templates. The resulting PCR fragments were sequenced using Gv_tuf_S6 forward primer. BLAST analysis revealed 95–100 % similarity with the deposited *G. vaginalis tuf* gene sequences. Lower than 100 % identity was due to sequence heterogeneity corresponding to clade-specific variation within the *tuf* gene (data not shown). Sequence analysis therefore

Table 3. *G. vaginalis* qPCR primers and probes

PCR primers and TaqMan probes used in *G. vaginalis* clade-specific multiplex qPCR are in bold. BHQ1, Black Hole Quencher®-1; BHQ2, Black Hole Quencher®-2; Cy5, Cy5™ fluorophore; FAM, 6-carboxyfluorescein fluorophore; MAXN, MAX™ fluorophore NHS ester; ND, not determined; ROXN, carboxy-x-rhodamine fluorophore NHS ester.

G. vaginalis clade	Primer/probe	Sequence	Amplicon (bp)	Gene	Protein	Database ID
1, 2, 3, 4	Gv_tuf_S4	5′-TCCCAACCCCAACTCACGATCTT-3′	149	*tuf*	Translation elongation factor Tu	GI:311114364
	Gv_tuf_AS3	5′-RCGCAAACCAACRATCTCAACTGG-3′				
	Gv_tuf_TM1	5′-FAM-CCATCTCCGGTCGTGGTACCGTTG-BHQ1-3′				
	Gv_tuf_S6	5′-GAGGGCTCGCTGACCTACCG-3′	344			
	Gv_tuf_AS7	5′-GGCACTCGCACACACCAAGG-3′				
1	**Gv1_fuc1_S**	5′-CCAGTCATAAGTTTGCGTTTTACC-3′	139	*fuc1*	Putative α-L-fucosidase	GI:311113989
	Gv1_fuc1_AS	5′-TGGCACTGGCAAAGTTTACAAC-3′				
	Gv1_fuc1_TM	5′-FAM-CTCGCCGCAAGCACCATCAAGCCA-BHQ1-3′				
	Gv1_galK_S	5′-TTCTAGATTATTCGCCGCCAAATC-3′	108	*galK*	Galactokinase	GI:311115066
	Gv1_galK_AS	5′-TTGCGATGTGTTGAAGGTAATGC-3′				
2	**Gv2_hyp_S**	5′-GCAAAGCAGACTGAGCGTATTAG-3′	124	ND	Hypothetical protein	GI:388060098
	Gv2_hyp_AS	5′-GTAATAATCAGGCTCCTCATCGC-3′				
	Gv2_hyp_TM	5′-5MAXN-CGCAGGCGCTCGCATAACAGTGCA-BHQ1-3′				
	Gv2_cel_S	5′-GCTTGGGGTTCATATGGTGATGG-3′	137	ND	Cellulosome anchoring protein	GI:388059846
	Gv2_cel_AS	5′-TCTTTATCAGACACGCCCTTAGC-3′				
3	**Gv3_thi_S**	5′-TTCTGCTTCTTCTGCTATTTGCTG-3′	142	ND	Thioredoxin	GI:388062216
	Gv3_thi_AS	5′-TTCGTTGACTTTTGGGCAACATG-3′				
	Gv3_thi_TM	5′-ROXN-CGGTCCGTGCCGTTCATTTGGTCC-3BHQ2-3′				
	Gv3_a-b_S	5′-TGATTACGCTCACGCTCTCG-3′	149	ND	α/β Hydrolase fold protein	GI:388063058
	Gv3_a-b_AS	5′-CGGCAACAGCTTTAGGAAGAAG-3′				
4	**Gv4_cic_S**	5′-CCTACGCAAGCTCCAGACGAC-3′	74	ND	Chloride transporter, CIC family	GI:283783343
	Gv4_cic_AS	5′-ACAAGTTGCACTCTTCGAGCTGG-3′				
	Gv4_cic_TM	5′-Cy5-ACTCGGCTGAAGCACACCACCACT-BHQ2-3′				
	Gv4_all_S	5′-CACGCTGGCACAACAATGATG-3′	139	ND	Allantoate amidohydrolase	GI:283783238
	Gv4_all_AS	5′-TTGGAACTACGCTGATTCTACCG-3′				

Table 4. Sixty vaginal-swab specimens and their characterization by Nugent scores, Amsel criteria, *G. vaginalis* species-specific qPCRs and multiplex clade-specific qPCR

Specimens used for *G. vaginalis* strain isolation are in bold. ND, not determined.

Sample	Nugent score	Amsel criteria	*tuf* qPCR C_t	16S rRNA qPCR C_t	*cpn60* qPCR C_t	Clade by qPCR
001	10	4	14.9	16.3	16.1	1, 3, 4
002	0	0	35.3	37.3	35.5	1
003	8	4	13.4	16.8	17.9	1, 2, 4
004	0	0	35.5	39.0	–	4
005	10	4	13.1	15.9	18.3	1, 4
006	0	0	25.4	25.4	28.5	1, 4
007	1	1	23.5	25.9	–	4
008	8	4	13.4	16.6	17.6	1, 4
009	4	0	29.6	32.4	35.9	4
010	0	1	27.1	29.1	32.0	1, 4
011	3	0	25.0	28.2	26.0	1
012	8	4	13.0	16.0	–	4
013	8	3	15.0	18.9	15.7	1, 4
014	6	0	14.4	17.4	21.5	2, 4
015	10	4	14.8	17.1	17.1	1, 3, 4
016	2	0	14.4	17.2	–	2, 4
017	0	0	35.0	35.6	–	4
018	8	4	14.1	16.9	19.8	2, 4
019	10	4	14.2	16.1	18.6	1, 3, 4
020	10	4	15.0	16.7	16.0	3, 4
021	7	4	13.1	16.5	14.0	1
022	8	4	15.1	17.3	19.2	1, 3, 4
023	0	0	23.8	23.8	23.6	1, 3, 4
024	10	4	15.6	17.3	18.5	3, 4
025	ND	0	19.4	22.3	24.8	1, 2, 4
026	8	4	14.0	15.8	16.1	1, 3, 4
027	4	0	20.0	22.6	25.3	3, 4
028	4	0	16.2	19.7	17.9	1, 2, 3, 4
029	0	0	24.6	27.0	25.4	1, 3
030	2	1	29.5	37.7	–	4
031	4	1	16.7	19.6	21.2	1, 4
032	0	0	26.5	31.8	–	4
033	0	0	–	–	–	–
034	0	0	26.1	28.6	–	4
035	0	0	36.1	39.3	–	4
036	10	4	16.1	18.3	17.8	1, 3
037	0	0	36.1	39.3	–	4
038	0	0	21.1	23.5	–	2, 4
039	0	0	20.9	23.5	26.6	3, 4
040	10	4	17.5	19.5	24.2	2, 4
041	7	4	15.7	17.0	17.7	3, 4
042	4	1	19.7	23.3	–	2, 4
043	7	4	14.9	18.8	15.8	1, 2, 4
044	3	0	30.9	32.9	–	4
045	10	4	15.4	18.1	18.9	1, 4
046	5	4	17.8	20.8	20.4	2, 3
047	0	0	29.3	29.7	30.6	3, 4
048	0	0	18.1	20.8	–	4
049	4	3	15.4	18.2	18.0	1, 3
050	8	4	14.3	16.8	19.4	1, 4
051	0	1	32.0	34.5	32.4	1
052	7	3	13.6	16.6	16.0	1, 3, 4
053	4	0	14.7	18.2	15.8	1, 4

Table 4. cont.

Sample	Nugent score	Amsel criteria	*tuf* qPCR C_t	16S rRNA qPCR C_t	*cpn60* qPCR C_t	Clade by qPCR
054	4	1	17.3	20.0	21.0	2, 3
055	4	1	15.2	19.2	16.5	1, 4
056	7	4	13.6	16.8	15.5	1, 2, 4
057	4	0	37.6	–	–	4
058	0	0	30.8	33.0	35.5	1, 4
059	6	0	17.1	19.8	24.5	1, 2, 4
060	0	0	18.7	20.4	20.2	1, 2, 4

confirmed the specificity of the developed *G. vaginalis* species-specific qPCR and also demonstrated the presence of multiple *G. vaginalis* clades in 15 out of 21 analysed vaginal-swab specimens.

G. vaginalis multiplex clade-specific qPCR development and validation

A set of eight unique *G. vaginalis* clade-specific genes identified in a previous genome sequencing study was used for the development of multiplex clade-specific qPCR (Ahmed *et al.*, 2012). Genetic markers, corresponding proteins and database identification numbers are listed in Table 3. The specificity of eight pairs of primers to four *G. vaginalis* clades was confirmed by BLAST analysis and in the SYBR Green qPCR experiments using chromosomal DNA samples from 12 previously characterized *G. vaginalis* isolates representing all four clades (Ahmed *et al.*, 2012). All eight genetic targets were shown to be specific for the corresponding clades with one exception: *G. vaginalis* isolate B482, identified previously as clade 2, produced no signals in any of the SYBR Green qPCRs (Table 1). In cross-reactivity experiments non-specific amplification was observed for primers Gv4_all_S and Gv4_all_AS targeting the putative allantoate amidohydrolase encoding gene when DNA from *Bifidobacterium* species was used as a template (Table 2). Based on the cross-reactivity and amplification specificity results, we chose four pairs of clade-specific primers out of the eight for further analysis.

Four fluorescently labelled TaqMan probes were designed to target amplicons generated with the clade-specific primers (Table 3). The resulting multiplex qPCR incorporated eight primers and four probes for four clade-specific genes encoding a putative α-L-fucosidase (clade 1), a hypothetical protein (clade 2), thioredoxin (clade 3) and a chloride transporter (clade 4). Cross-reactivity, interassay sensitivity and specificity experiments were performed as described for the *G. vaginalis* species-specific qPCR. No cross-reactivity was detected for human DNA or the bacterial, protozoan and fungal species listed in Table 2. There was 100 % concordance between the clade-specific multiplex qPCR and the species-specific *tuf* qPCR: *G. vaginalis* clades in different combinations were identified in 59 out of 60 analysed vaginal-swab DNA samples. Sample number 033 was negative by both qPCRs. The interassay sensitivity and specificity were therefore determined as 100 %. *G. vaginalis* clade distribution in 60 vaginal specimens is shown in Table 4. For determining interference, we used chromosomal DNA samples from four

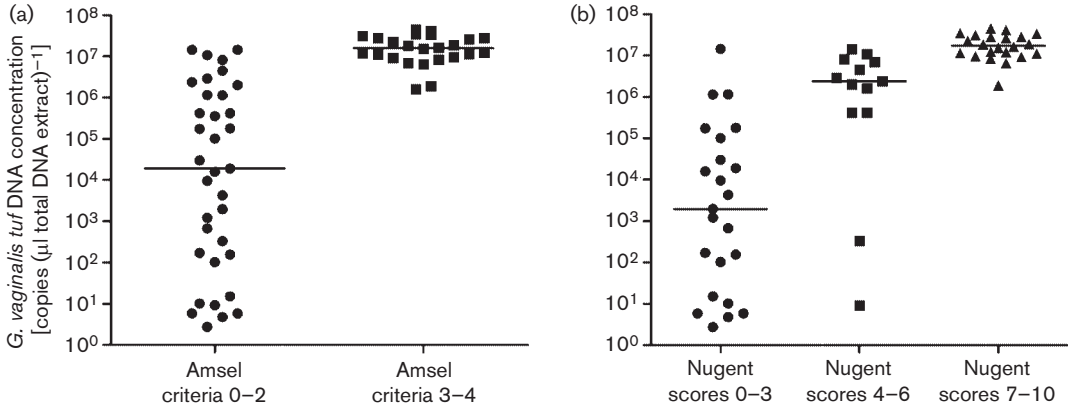

Fig. 1. Concentrations of *G. vaginalis* DNA in 60 clinical vaginal specimens characterized according to (a) Amsel criteria and (b) Nugent scores. qPCRs targeting *G. vaginalis* chromosomal *tuf* gene (encoding translation elongation factor Tu) were used for quantitative assessments. Concentrations are shown as the number of *G. vaginalis* genomic copies (µl total DNA extract) $^{-1}$. Bars represent median values.

Table 5. Association between *G. vaginalis* clades in 60 vaginal-swab specimens and BV diagnoses by Nugent scores and Amsel criteria

The association analysis was performed using the Pearson *r* correlation test; *r* values greater than 0.25 indicative of positive association and lower than – 0.25 indicative of negative association are in bold.

Clade	Clade 1 (n=32)	Clade 2 (n=15)	Clade 3 (n=19)	Clade 4 (n=50)	Nugent scores 0–3 (n=24)	Nugent scores 4–6 (n=13)	Nugent scores 7–10 (n=22)	Amsel criteria $\geqslant$3 (n=24)
Single clade (n=17)	**−0.376**	**−0.363**	**−0.428**	−0.116	**0.468**	−0.151	**−0.325**	**−0.362**
Two or more clades (n=42)	**0.408**	**0.378**	**0.446**	0.195	**−0.505**	0.168	**0.347**	**0.386**
Clade 1 (n=32)		−0.077	0.062	−0.149	**−0.259**	−0.076	**0.296**	**0.286**
Clade 2 (n=15)			−0.145	0.052	−0.236	**0.257**	−0.040	0.000
Clade 3 (n=19)				−0.176	**−0.263**	0.077	0.226	**0.322**
Clade 4 (n=50)					−0.091	−0.090	0.155	0.000

previously described *G. vaginalis* strains: B472 (clade 1), B513 (clade 2), B483 (clade 3) and B475 (clade 4) (Ahmed *et al.*, 2012). DNA extracts were combined in different combinations of two, three or all four clades together in concentrations of 10 ng per PCR and used as templates for amplification. No interference between the detection of different clades was observed. The accuracy of the multiplex clade-specific qPCR was determined by a comparison of qPCR results with *tuf* DNA sequencing data produced for 21 vaginal-swab specimens during species-specific qPCR validation. The *tuf* DNA sequencing analysis revealed that all sequenced samples corresponded to the most abundant *G. vaginalis* clades identified by the lowest C_t scores in the clade-specific qPCRs.

G. vaginalis clinical strain isolation and subtyping

In order to confirm clade-specific multiplex qPCR results and demonstrate multi-clade *G. vaginalis* infections, we isolated bacterial cultures from vaginal specimens. Nine samples characterized with Amsel criteria and Nugent scores, and nine uncharacterized specimens listed in Table 1 were used. More than 80 initial isolates presumtively identified as *G. vaginalis* by β-haemolysis of human blood and Gram-stain microscopy were tested with species-specific and clade-specific qPCRs. Based on the qPCR results, we chose 1 *G. vaginalis* strain per clade per swab sample, which resulted in the 24 independent isolates listed in Table 1. Most of the isolated clinical strains belonged to clade 1 (n=14). Clade 2 (n=5) and clade 4 (n=5) strains were less frequently isolated. We did not find any clade 3 strains among the *G. vaginalis* clinical isolates.

With a single exception, the clade specificity of *G. vaginalis* isolates corresponded to the results of the multiplex clade-specific qPCR assay performed on total DNA samples extracted from the noncultured vaginal specimens. *G. vaginalis* strain S0244 subtyped as clade 1 was isolated from sample number 024, which contained clade 3 and clade 4 according to the qPCR. All the other *G. vaginalis* isolates from characterized vaginal specimens belonged to the clades identified in the corresponding vaginal swabs by qPCR and listed in Table 4. A sample, number 007, harbouring a single *G. vaginalis* clade 4 according to the clade-specific qPCR yielded an S0070 strain of the same clade. *G. vaginalis* isolates of two different clades were isolated from characterized samples 015, 028 and 052, and uncharacterized specimens 076, 077 and 090, therefore confirming qPCR results and demonstrating a presence of multiple *G. vaginalis* clades in the vaginal microflora.

Twenty-four *G. vaginalis* strains isolated within this study were combined with three ATCC isolates and twelve *G. vaginalis* strains described previously, and subjected to phenotypic biotyping and genotyping by ARDRA (Ahmed *et al.*, 2012; Ingianni *et al.*, 1997; Piot *et al.*, 1984). Only four biotypes, biotype 1, biotype 4, biotype 5 and biotype 7, were found among the 39 *G. vaginalis* strains analysed. There was a tight association between *G. vaginalis* clades

and biotypes. Twenty strains of clade 1 were subtyped as biotypes 1 or 4. Two clade 1 strains S0191 and S0822 were not biotyped due to poor growth on egg yolk agar. Clade 2 isolates (n=6) were identified as biotypes 5 or 7. Clade 3 strains (n=3) were also assigned as biotypes 5 or 7. Seven isolates of clade 4 belonged to biotype 5. The ARDRA assay revealed only three genotypes: genotype 1, genotype 3 and genotype 4. We did not see genotype 2 among 39 analysed *G. vaginalis* isolates. Similar to the biotyping approach, we observed a correlation between *G. vaginalis* clades and genotypes. Twenty strains of clade 1 belonged to genotype 1. Five clade 2 isolates belonged to either genotype 3 or genotype 4. Three clade 3 strains were identified as genotype 3. Six isolates of clade 4 were also subtyped as genotype 3. *G. vaginalis* strains B478 (clade 1), S0151 (clade 1), S0284 (clade 2) and S0380 (clade 4) generated ambiguous restriction fragment patterns intermediate between genotype 1 and genotype 3 (Table 1).

Association of *G. vaginalis* clades with BV

The distribution of *G. vaginalis* clades in 60 vaginal-swab specimens characterized by the described multiplex clade-specific qPCR was compared with patient BV diagnoses based on either clinical BV symptoms or conventional microbiological evaluation (Table 4). There were notable differences between the four clades and their associations with BV (Table 5). Clades 1 and 3 were both negatively associated with normal vaginal microflora defined by Nugent scores 0–3 and positively associated with high Nugent scores 7–10 and high Amsel criteria. Presence of *G. vaginalis* clade 2 was negatively associated with normal microflora characterized by Nugent scores 0–3 and positively linked with intermediate microflora with Nugent scores 4–6. However, there was no correlation between clade 2 and either BV symptoms or high Nugent scores 7–10. There was no association between clade 4 and BV defined by either Amsel criteria or Nugent scores.

We divided 59 *G. vaginalis* positive vaginal samples according to the number of detected *G. vaginalis* clades into two groups: a group with a single clade present (n=17) and a group with two or more different clades (n=42). There was a substantial difference in association of these two groups with BV. The presence of a single *G. vaginalis* clade (either clade 1 or clade 4) was negatively associated with BV as defined by both Nugent score or Amsel criteria and positively associated with normal vaginal microflora with Nugent score of 0–3. Conversely, multi-clade *G. vaginalis* communities were positively associated with clinical BV symptoms and abnormal microflora with Nugent scores 7–10 and negatively associated with Nugent scores 0–3 (Table 5).

DISCUSSION

Despite being the first micro-organism associated with BV by Gardner and Dukes, *G. vaginalis* remains the most controversial species, with disparate phenotypic and genetic features, as well as an elusive role in BV pathogenesis (Gardner & Dukes, 1955). The ambiguous nature of this species begins from its very basic characteristics, such as Gram stain. Gram-positive in its ultrastructural characteristics and chemical composition, the *G. vaginalis* cell wall fails to retain crystal violet during the conventional Gram staining procedure, which causes bacterial cells to be identified as both Gram-positive and Gram-negative depending on growth conditions such as medium used and growth phase (Greenwood & Pickett, 1980; Piot *et al.*, 1980; Sadhu *et al.*, 1989). The great phenotypic variability observed between *G. vaginalis* isolates was used for epidemiological subtyping of this species, resulting in eight biotypes on the basis of β-galactosidase, lipase and hippurate hydrolysis reactions (Piot *et al.*, 1984). None of the biotypes, however, has been reliably associated with BV (Aroutcheva *et al.*, 2001; Tosun *et al.*, 2007). The most profound differences between *G. vaginalis* clinical strains were discovered only recently using a comparative genomic sequencing approach (Ahmed *et al.*, 2012; Harwich *et al.*, 2010; Yeoman *et al.*, 2010). The level of diversity between the genomes of 17 *G. vaginalis* isolates comprising four distinct clades described by Ahmed *et al.* (2012) was extraordinarily high and far beyond that observed within any other bacterial species, which provided a strong argument for separating these four clades into individual species.

In this work we have described the development and validation of a new qPCR genotyping assay capable of the identification and quantification of *G. vaginalis*, and the subtyping of four distinct subtypes or clades revealed elsewhere by genomic analysis (Ahmed *et al.*, 2012). The proposed approach addresses a number of deficiencies of the existing typing methods. The phenotypic biotyping scheme introduced by Piot *et al.* (1984) is based on enzymic reactions, the efficiencies of which might vary significantly depending on conditions. Hippurate hydrolysis, for instance, is pH-sensitive and inaccurate at a pH level higher than 6.4, whereas the accuracy of the lipase reaction greatly depends on the substrate used with 4-methylumbelliferyl oleate frequently producing erroneous biotypes (Moncla & Pryke, 2009; Piot *et al.*, 1982). The ARDRA genotyping approach developed by Ingianni and colleagues is less error-prone (Ingianni *et al.*, 1997; Santiago *et al.*, 2011). Both methods require isolation of pure *G. vaginalis* cultures. Preferential isolation of strains of a certain biotype or genotype is possible given differences in their metabolic potentials, which in turn might introduce bias to an assessment of *G. vaginalis* subtype distribution in clinical specimens. Limitations of current subtyping techniques might, at least in part, explain discordances in the frequencies of *G. vaginalis* biotypes among different studies and conflicting results of associations of certain biotypes with BV (Moncla & Pryke, 2009). The clade subtyping by qPCR proposed here does not require culture and therefore can be applied to noncultured clinical samples, allowing quantitative assessment of bacterial loads and qualitative identification of *G.*

vaginalis clades in a given specimen. Moreover, since no isolation of live bacterial cells is involved, the assay can be performed on a variety of samples, including archived DNAs or vaginal specimens collected and stored at suboptimal conditions.

The quantitative assessment of *G. vaginalis* in vaginal specimens by qPCR has proven to be important for the molecular diagnosis of BV (De Backer *et al.*, 2007; Fredricks *et al.*, 2009; Menard *et al.*, 2008; Zozaya-Hinchliffe *et al.*, 2010). The recent discovery of a high level of genetic variability among *G. vaginalis* clades, however, raises a question of the specificity of such assays. As was demonstrated during the validation process, the previously described *G. vaginalis* species-specific *cpn60* qPCR used for interassay sensitivity and specificity determination was capable of detecting only two out of four now recognized *G. vaginalis* clades, clade 1 and clade 3 (Table 1), which was accounted for by *cpn60* sequence heterogeneity among clades (Menard *et al.*, 2008). In the screening of 60 vaginal specimens *cpn60* qPCR produced false-negative results for a number of samples populated by *G. vaginalis* clade 2 and clade 4 (Table 4). The *tuf* gene-based assay described here, and another species-specific 16S rRNA qPCR described elsewhere, targeted conserved sequences of these genes and were specific to all four clades (Fredricks *et al.*, 2009). Among multiple housekeeping genes tested in this study and used as conserved genetic markers in other bacterial molecular phylogenetic studies, the *tuf* gene demonstrated the highest level of homology between *G. vaginalis* clades (Paradis *et al.*, 2005; Ventura & Zink, 2003; Wertz *et al.*, 2003). Similar to other studies on different bacterial taxonomic groups, including the closely related genus *Bifidobacterium*, the utility of *tuf* for species-specific *G. vaginalis* identification was demonstrated in this work (Chavagnat *et al.*, 2002; Sheu *et al.*, 2010; Ventura & Zink, 2003). Since there is only one copy of the translation elongation factor Tu encoding gene per *G. vaginalis* genome, *tuf* is a more appropriate target for quantitative assessment of bacterial loads by qPCR compared with the 16S rRNA gene, for which copy number per genome varies greatly among bacterial species (Farrelly *et al.*, 1995; Klappenbach *et al.*, 2001).

One of the unusual findings observed during the assay validation was the very high positivity rate of *G. vaginalis* in healthy women with no clinical symptoms of BV. Out of 36 analysed BV-negative vaginal specimens, 35 (97 %) were positive by *tuf* qPCR and 34 (94 %) were positive by 16S rRNA qPCR (Table 4). The high prevalence of *G. vaginalis* reaching 100 % in BV patients is a known phenomenon, confirmed by this and other studies. The reported frequency of *G. vaginalis* among healthy women is, however, significantly lower: 47–85 %, including 62.5 % positivity in post-treatment patients by 16S rRNA qPCR described by Fredricks and colleagues (Fredricks *et al.*, 2009; Menard *et al.*, 2008; Zozaya-Hinchliffe *et al.*, 2010). The higher *G. vaginalis* prevalence rate might be explained by the patient cohort enrolled for this study. BV diagnoses were made at the time of admission on the basis of either Amsel criteria or Nugent scores. Undocumented previous episodes of BV might have contributed to the *G. vaginalis* infections. Another possible explanation might lie in an enhanced DNA extraction efficiency. The pretreatment of the clinical specimens with proteinase K and mechanical homogenization might have resulted in more efficient cell lysis and DNA extraction, therefore improving detection of low levels of the micro-organism. Quantitative assessment of *G. vaginalis* loads in vaginal specimens (Fig. 1) demonstrated lower and more variable concentrations of pathogen DNA in non-BV samples, similar to the dose-dependent association reported by other researchers (De Backer *et al.*, 2007; Fredricks *et al.*, 2009; Menard *et al.*, 2008; Zozaya-Hinchliffe *et al.*, 2010).

The novel qPCR-based *G. vaginalis* typing assay described in this work is based on the division of this species between four genetically distinct clades proposed by Ahmed *et al.* (2012). In their recent study these authors elucidated a core *G. vaginalis* genome consisting of only 746 genes common between 17 strains and sets of distributed clade-specific genes. In our work, we have confirmed the utility of analysed conserved genetic targets for subtyping clinical *G. vaginalis* isolates. The developed multiplex single-tube qPCR targeting genes encoding putative α-L-fucosidase, a hypothetical protein, thioredoxin and CIC family chloride transporter (Table 3) was proven to be clade-specific and capable of strain typing and identification of the four *G. vaginalis* clades in noncultured clinical vaginal specimens. With a single exception, 39 *G. vaginalis* strains from different sources (Table 1) were successfully typed by qPCR and identified as harbouring just one out of four targeted conserved clade-specific genes in line with reported genetic isolation and the absence of horizontal transfer of distributed genes among the four *G. vaginalis* clades (Ahmed *et al.*, 2012). If taxonomic assignment of the *G. vaginalis* clades is revised towards their separation into four species of the genus *Gardnerella*, as suggested by Ahmed *et al.* (2012), the described multiplex qPCR can be applied for species differentiation rather than subtyping purposes.

The clade-specific qPCR genotyping assay demonstrated a strong association between certain *G. vaginalis* clades, biotypes and genotypes (Table 1). As can be seen from the comparative analysis of different typing methods, the *G. vaginalis* strains of clade 1 composed a very distinct group, which was identified as different from other clades by all three typing techniques. This might indicate greater phylogenetic distance between clade 1 and the other clades of *G. vaginalis*, which in turn might imply a need for its taxonomic reassignment.

The analysis of *G. vaginalis* clade distribution in 60 vaginal specimens by a multiplex clade-specific qPCR has revealed a dominance of clade 4 (n=50) followed by less frequent clade 1 (n=32), clade 3 (n=19) and clade 2 (n=15). There was no association between different clades, suggesting the absence of mutualistic or antagonistic relationships in

multi-clade communities (Table 5). Polyclonal *G. vaginalis* communities composed of more than one clade (n=42), however, prevailed in analysed samples. The rate of co-colonization by multiple *G. vaginalis* clades (70 %) was significantly higher than the previously described presence of multiple biotypes and genotypes in vaginal specimens (Aroutcheva *et al.*, 2001; Briselden & Hillier, 1990; Piot *et al.*, 1984; Santiago *et al.*, 2011). In 17 vaginal swabs *G. vaginalis* was represented by a single clade: either clade 1 (n=4) or clade 4 (n=13). Clade 2 and clade 3 were found only as parts of multi-clade communities, which might indicate their dependence on activities important for survival in the vaginal environment provided by other clades. *G. vaginalis* strains of these two clades, however, demonstrated no growth deficiencies when cultured in laboratory setting.

A major advantage of the proposed clade-specific qPCR typing method is its independence from culture. This feature, along with the multiplex single-tube format, makes it a molecular typing tool meeting the requirements for a high-throughput diagnostic assay. The clinical significance of the described assay with respect to BV pathogenesis was demonstrated by evident differences in association of *G. vaginalis* clades and the disorder. A high positive correlation with BV defined by both microbiological criteria and clinical symptoms was observed for clade 1 and clade 3, whereas clade 4 had no association with the condition. Clade 2 exhibited a negative association with healthy vaginal microflora and a positive relationship with intermediate microflora rather than BV (Table 5). The new molecular approach proposed here might, at least in part, explain the controversy of association of different conventional biotypes with BV reported in a number of studies. As shown by Briselden & Hillier (1990), the lipase-positive biotypes 1, 2, 3 and 4 were recovered more frequently from women with BV than from women with normal vaginal flora. In our study, clade 1, composed of biotype 1 and biotype 4, demonstrated a positive association with the disorder. Numanović *et al.* (2008) reported that biotype 7, along with biotypes 2 and 3, was more often isolated from women who suffered from BV. Clade 3, highly associated with BV, included biotype 7, as shown by the typing methods comparison. Aroutcheva *et al.* (2001) found that biotype 5 was dominant among patients with normal vaginal microflora. In our work, biotype 5 isolates were subtyped to three different clades: clade 2, clade 3 and clade 4. Clade 4 was most frequently identified in vaginal specimens and demonstrated no link with the disorder.

The most significant difference in association with BV was, however, discovered not for individual clades, but for clade groups. Co-colonization by multiple clades had substantially higher association with BV compared with *G. vaginalis* present as a single clade (Table 5). One of the possible reasons for such disparity might be that multi-clade communities were often composed of BV-associated clades, whereas specimens containing a single clade were most frequently colonized by *G. vaginalis* clade 4, which had the lowest association to BV. The presence of numerous clades in the vaginal milieu may be a result of having multiple sexual partners. As shown by both the biotyping approach and analysis of nucleotide variations within the 16S rRNA gene, monogamous women and their male partners tend to be colonized by the same *G. vaginalis* strain or biotype (Eren *et al.*, 2011; Piot *et al.*, 1984). Unprotected sex with a new sexual partner is a recognized risk factor of BV (Fethers *et al.*, 2008; Muzny & Schwebke, 2013). Finally, polyclonal *G. vaginalis* infections might possess greater population-level virulence potential in line with a distributed genome hypothesis stating that subspecies genetic diversity improves population survival and maximizes its fitness to diverse environmental conditions as proposed by Ehrlich and colleagues (Ahmed *et al.*, 2012; Ehrlich *et al.*, 2008). Other findings, such as the prevalence of multi-clade communities and their structures, can also be explained from this point of view.

The major limitation of this study was the small number of clinical specimens. Even though 60 analysed vaginal samples were well characterized by both conventional and molecular methods, other patient cohorts might produce different results of *G. vaginalis* clade distribution or association with BV depending on their health status, race, age or geographical location. The thorough validation of the method described here, however, ensures its application for *G. vaginalis* identification, quantification and subtyping in a laboratory setting aiming to improve pathogen diagnosis and epidemiological surveillance. The longitudinal studies of clade distribution in unaffected women and patients with recurrent BV or undergoing antibiotic therapy are required to further elucidate the complexity of *G. vaginalis* as a species and its contribution to vaginal health.

In conclusion, our study describes the development and validation of a novel culture-independent *G. vaginalis* subtyping molecular approach based on qPCR. Two proposed assays are capable of *G. vaginalis* identification, quantification and subtyping in noncultured vaginal specimens. The utility of the approach was demonstrated by identifying *G. vaginalis* bacterial loads and clade distribution in 60 clinical swab samples omitting the pathogen culture. The clade-specific association with BV and prevalence of polyclonal infections revealed here might shed some light on the role of *G. vaginalis* in BV pathogenesis, advancing our knowledge of this enigmatic condition and promoting better diagnostic and treatment options.

ACKNOWLEDGEMENTS

We thank Dr Jack D. Sobel from the Vaginitis Clinic, Wayne State University School of Medicine, Detroit, MI, USA, for providing clinical specimens and their characterization, and Dr Garth D. Ehrlich from the Center for Genomic Sciences, Allegheny-Singer Research Institute, Pittsburgh, PA, USA, for providing *G. vaginalis* strains.

REFERENCES

Ahmed, A., Earl, J., Retchless, A., Hillier, S. L., Rabe, L. K., Cherpes, T. L., Powell, E., Janto, B., Eutsey, R. & other authors (2012). Comparative genomic analyses of 17 clinical isolates of *Gardnerella vaginalis* provide evidence of multiple genetically isolated clades consistent with subspeciation into genovars. *J Bacteriol* **194**, 3922–3937.

Amsel, R., Totten, P. A., Spiegel, C. A., Chen, K. C., Eschenbach, D. & Holmes, K. K. (1983). Nonspecific vaginitis. Diagnostic criteria and microbial and epidemiologic associations. *Am J Med* **74**, 14–22.

Aroutcheva, A. A., Simoes, J. A., Behbakht, K. & Faro, S. (2001). *Gardnerella vaginalis* isolated from patients with bacterial vaginosis and from patients with healthy vaginal ecosystems. *Clin Infect Dis* **33**, 1022–1027.

Benito, R., Vazquez, J. A., Berron, S., Fenoll, A. & Saez-Neito, J. A. (1986). A modified scheme for biotyping *Gardnerella vaginalis*. *J Med Microbiol* **21**, 357–359.

Briselden, A. M. & Hillier, S. L. (1990). Longitudinal study of the biotypes of *Gardnerella vaginalis*. *J Clin Microbiol* **28**, 2761–2764.

Catlin, B. W. (1992). *Gardnerella vaginalis*: characteristics, clinical considerations, and controversies. *Clin Microbiol Rev* **5**, 213–237.

Chavagnat, F., Haueter, M., Jimeno, J. & Casey, M. G. (2002). Comparison of partial *tuf* gene sequences for the identification of lactobacilli. *FEMS Microbiol Lett* **217**, 177–183.

De Backer, E., Verhelst, R., Verstraelen, H., Alqumber, M. A., Burton, J. P., Tagg, J. R., Temmerman, M. & Vaneechoutte, M. (2007). Quantitative determination by real-time PCR of four vaginal *Lactobacillus* species, *Gardnerella vaginalis* and *Atopobium vaginae* indicates an inverse relationship between *L. gasseri* and *L. iners*. *BMC Microbiol* **7**, 115.

Ehrlich, G. D., Hiller, N. L. & Hu, F. Z. (2008). What makes pathogens pathogenic. *Genome Biol* **9**, 225.

El Aila, N. A., Tency, I., Saerens, B., De Backer, E., Cools, P., dos Santos Santiago, G. L., Verstraelen, H., Verhelst, R., Temmerman, M. & Vaneechoutte, M. (2011). Strong correspondence in bacterial loads between the vagina and rectum of pregnant women. *Res Microbiol* **162**, 506–513.

Eren, A. M., Zozaya, M., Taylor, C. M., Dowd, S. E., Martin, D. H. & Ferris, M. J. (2011). Exploring the diversity of *Gardnerella vaginalis* in the genitourinary tract microbiota of monogamous couples through subtle nucleotide variation. *PLoS ONE* **6**, e26732.

Eschenbach, D. A. (1993). Bacterial vaginosis and anaerobes in obstetric-gynecologic infection. *Clin Infect Dis* **16** (Suppl. 4), S282–S287.

Farrelly, V., Rainey, F. A. & Stackebrandt, E. (1995). Effect of genome size and *rrn* gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. *Appl Environ Microbiol* **61**, 2798–2801.

Fethers, K. A., Fairley, C. K., Hocking, J. S., Gurrin, L. C. & Bradshaw, C. S. (2008). Sexual risk factors and bacterial vaginosis: a systematic review and meta-analysis. *Clin Infect Dis* **47**, 1426–1435.

Forsum, U., Holst, E., Larsson, P. G., Vasquez, A., Jakobsson, T. & Mattsby-Baltzer, I. (2005). Bacterial vaginosis – a microbiological and immunological enigma. *APMIS* **113**, 81–90.

Fredricks, D. N., Fiedler, T. L., Thomas, K. K., Mitchell, C. M. & Marrazzo, J. M. (2009). Changes in vaginal bacterial concentrations with intravaginal metronidazole therapy for bacterial vaginosis as assessed by quantitative PCR. *J Clin Microbiol* **47**, 721–726.

Gardner, H. L. & Dukes, C. D. (1955). *Haemophilus vaginalis* vaginitis: a newly defined specific infection previously classified non-specific vaginitis. *Am J Obstet Gynecol* **69**, 962–976.

Graham, S., Howes, C., Dunsmuir, R. & Sandoe, J. (2009). Vertebral osteomyelitis and discitis due to *Gardnerella vaginalis*. *J Med Microbiol* **58**, 1382–1384.

Greenwood, J. R. & Pickett, M. J. (1980). Transfer of *Haemophilus vaginalis* Gardner and Dukes to a new genus, *Gardnerella*: *G. vaginalis* (Gardner and Dukes) comb. nov. *Int J Syst Bacteriol* **30**, 170–178.

Harwich, M. D., Jr, Alves, J. M., Buck, G. A., Strauss, J. F., III, Patterson, J. L., Oki, A. T., Girerd, P. H. & Jefferson, K. K. (2010). Drawing the line between commensal and pathogenic *Gardnerella vaginalis* through genome analysis and virulence studies. *BMC Genomics* **11**, 375.

Holst, E. (1990). Reservoir of four organisms associated with bacterial vaginosis suggests lack of sexual transmission. *J Clin Microbiol* **28**, 2035–2039.

Ingianni, A., Petruzzelli, S., Morandotti, G. & Pompei, R. (1997). Genotypic differentiation of *Gardnerella vaginalis* by amplified ribosomal DNA restriction analysis (ARDRA). *FEMS Immunol Med Microbiol* **18**, 61–66.

Klappenbach, J. A., Saxman, P. R., Cole, J. R. & Schmidt, T. M. (2001). rrndb: the ribosomal RNA operon copy number database. *Nucleic Acids Res* **29**, 181–184.

Lagacé-Wiens, P. R., Ng, B., Reimer, A., Burdz, T., Wiebe, D. & Bernard, K. (2008). *Gardnerella vaginalis* bacteremia in a previously healthy man: case report and characterization of the isolate. *J Clin Microbiol* **46**, 804–806.

Leopold, S. (1953). Heretofore undescribed organism isolated from the genitourinary system. *U S Armed Forces Med J* **4**, 263–266.

Marrazzo, J. M., Fiedler, T. L., Srinivasan, S., Thomas, K. K., Liu, C., Ko, D., Xie, H., Saracino, M. & Fredricks, D. N. (2012). Extravaginal reservoirs of vaginal bacteria as risk factors for incident bacterial vaginosis. *J Infect Dis* **205**, 1580–1588.

Menard, J. P., Fenollar, F., Henry, M., Bretelle, F. & Raoult, D. (2008). Molecular quantification of *Gardnerella vaginalis* and *Atopobium vaginae* loads to predict bacterial vaginosis. *Clin Infect Dis* **47**, 33–43.

Moncla, B. J. & Pryke, K. M. (2009). Oleate lipase activity in *Gardnerella vaginalis* and reconsideration of existing biotype schemes. *BMC Microbiol* **9**, 78.

Muzny, C. A. & Schwebke, J. R. (2013). *Gardnerella vaginalis*: still a prime suspect in the pathogenesis of bacterial vaginosis. *Curr Infect Dis Rep* **15**, 130–135.

Neri, P., Salvolini, S., Giovannini, A. & Mariotti, C. (2009). Retinal vasculitis associated with asymptomatic *Gardnerella vaginalis* infection: a new clinical entity. *Ocul Immunol Inflamm* **17**, 36–40.

Nugent, R. P., Krohn, M. A. & Hillier, S. L. (1991). Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. *J Clin Microbiol* **29**, 297–301.

Numanović, F., Hukić, M., Nurkić, M., Gegić, M., Delibegović, Z., Imamović, A. & Pasić, S. (2008). Importance of isolation and biotypization of *Gardnerella vaginalis* in diagnosis of bacterial vaginosis. *Bosn J Basic Med Sci* **8**, 270–276.

Paradis, S., Boissinot, M., Paquette, N., Bélanger, S. D., Martel, E. A., Boudreau, D. K., Picard, F. J., Ouellette, M., Roy, P. H. & Bergeron, M. G. (2005). Phylogeny of the *Enterobacteriaceae* based on genes encoding elongation factor Tu and F-ATPase β-subunit. *Int J Syst Evol Microbiol* **55**, 2013–2025.

Paramel Jayaprakash, T., Schellenberg, J. J. & Hill, J. E. (2012). Resolution and characterization of distinct *cpn*60-based subgroups of *Gardnerella vaginalis* in the vaginal microbiota. *PLoS ONE* **7**, e43009.

Patterson, J. L., Stull-Lane, A., Girerd, P. H. & Jefferson, K. K. (2010). Analysis of adherence, biofilm formation and cytotoxicity suggests a greater virulence potential of *Gardnerella vaginalis* relative to other bacterial-vaginosis-associated anaerobes. *Microbiology* **156**, 392–399.

Piot, P., van Dyck, E., Goodfellow, M. & Falkow, S. (1980). A taxonomic study of *Gardnerella vaginalis* (*Haemophilus vaginalis*) Gardner and Dukes 1955. *J Gen Microbiol* **119**, 373–396.

Piot, P., Van Dyck, E., Totten, P. A. & Holmes, K. K. (1982). Identification of *Gardnerella* (*Haemophilus*) *vaginalis*. *J Clin Microbiol* **15**, 19–24.

Piot, P., Van Dyck, E., Peeters, M., Hale, J., Totten, P. A. & Holmes, K. K. (1984). Biotypes of *Gardnerella vaginalis*. *J Clin Microbiol* **20**, 677–679.

Pleckaityte, M., Janulaitiene, M., Lasickiene, R. & Zvirbliene, A. (2012). Genetic and biochemical diversity of *Gardnerella vaginalis* strains isolated from women with bacterial vaginosis. *FEMS Immunol Med Microbiol* **65**, 69–77.

Sadhu, K., Domingue, P. A., Chow, A. W., Nelligan, J., Cheng, N. & Costerton, J. W. (1989). *Gardnerella vaginalis* has a Gram-positive cell-wall ultrastructure and lacks classical cell-wall lipopolysaccharide. *J Med Microbiol* **29**, 229–235.

Santiago, G. L., Deschaght, P., El Aila, N., Kiama, T. N., Verstraelen, H., Jefferson, K. K., Temmerman, M. & Vaneechoutte, M. (2011). *Gardnerella vaginalis* comprises three distinct genotypes of which only two produce sialidase. *Am J Obstet Gynecol* **204**, e1–e7.

Schellenberg, J. J., Links, M. G., Hill, J. E., Dumonceaux, T. J., Kimani, J., Jaoko, W., Wachihi, C., Mungai, J. N., Peters, G. A. & other authors (2011). Molecular definition of vaginal microbiota in East African commercial sex workers. *Appl Environ Microbiol* **77**, 4066–4074.

Schwebke, J. R. & Desmond, R. (2007). Natural history of asymptomatic bacterial vaginosis in a high-risk group of women. *Sex Transm Dis* **34**, 876–877.

Sheu, S. J., Hwang, W. Z., Chiang, Y. C., Lin, W. H., Chen, H. C. & Tsen, H. Y. (2010). Use of *tuf* gene-based primers for the PCR detection of probiotic *Bifidobacterium* species and enumeration of bifidobacteria in fermented milk by cultural and quantitative real-time PCR methods. *J Food Sci* **75**, M521–M527.

Sivadon-Tardy, V., Roux, A. L., Piriou, P., Herrmann, J. L., Gaillard, J. L. & Rottman, M. (2009). *Gardnerella vaginalis* acute hip arthritis in a renal transplant recipient. *J Clin Microbiol* **47**, 264–265.

Swidsinski, A., Doerffel, Y., Loening-Baucke, V., Swidsinski, S., Verstraelen, H., Vaneechoutte, M., Lemm, V., Schilling, J. & Mendling, W. (2010). *Gardnerella* biofilm involves females and males and is transmitted sexually. *Gynecol Obstet Invest* **70**, 256–263.

Tosun, I., Alpay Karaoğlu, S., Ciftçi, H., Buruk, C. K., Aydin, F., Kiliç, A. O. & Ertürk, M. (2007). [Biotypes and antibiotic resistance patterns of *Gardnerella vaginalis* strains isolated from healthy women and women with bacterial vaginosis]. *Mikrobiyol Bul* **41**, 21–27 (in Turkish).

Udayalaxmi, J., Bhat, G. K. & Kotigadde, S. (2011). Biotypes and virulence factors of *Gardnerella vaginalis* isolated from cases of bacterial vaginosis. *Indian J Med Microbiol* **29**, 165–168.

Ventura, M. & Zink, R. (2003). Comparative sequence analysis of the *tuf* and *recA* genes and restriction fragment length polymorphism of the internal transcribed spacer region sequences supply additional tools for discriminating *Bifidobacterium lactis* from *Bifidobacterium animalis*. *Appl Environ Microbiol* **69**, 7517–7522.

Wertz, J. E., Goldstone, C., Gordon, D. M. & Riley, M. A. (2003). A molecular phylogeny of enteric bacteria and implications for a bacterial species concept. *J Evol Biol* **16**, 1236–1248.

Yeoman, C. J., Yildirim, S., Thomas, S. M., Durkin, A. S., Torralba, M., Sutton, G., Buhay, C. J., Ding, Y., Dugan-Rocha, S. P. & other authors (2010). Comparative genomics of *Gardnerella vaginalis* strains reveals substantial differences in metabolic and virulence potential. *PLoS ONE* **5**, e12411.

Yoon, H. J., Chun, J., Kim, J. H., Kang, S. S. & Na, D. J. (2010). *Gardnerella vaginalis* septicaemia with pyelonephritis, infective endocarditis and septic emboli in the kidney and brain of an adult male. *Int J STD AIDS* **21**, 653–657.

Zozaya-Hinchliffe, M., Lillis, R., Martin, D. H. & Ferris, M. J. (2010). Quantitative PCR assessments of bacterial species in women with and without bacterial vaginosis. *J Clin Microbiol* **48**, 1812–1819.