Characterisation of vaginal *Prevotella* strains from a cohort of South African adolescent girls and young women with and without bacterial vaginosis

Kirsten Welp

Supervisor: Dr Brian Kullin

Co-supervisors: Professor Jo-Ann Passmore & Dr Lynthia Paul

Dissertation presented for the degree of MSc (Med) in Medical Virology

Division of Medical Virology

Department of Pathology

Faculty of Health Science

University of Cape Town

February 2024

Declaration

I, Kirsten Welp, hereby declare that the work on which this dissertation is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever

Signature: Signed by candidate

Date: 8 February 2024

Table of contents

Acknowledgements

I would like to offer my sincerest thanks to my supervisor Dr Brian Kullin for all the assistance in lab work, bioinformatics, and writing, and for being a great support overall throughout all stages of my Master's journey. I am very lucky to have had such a present and supportive supervisor, available whenever he was needed. Additionally, a huge thanks to Dr Lynthia Paul for her co-supervision and for the valuable insight and input in compiling this dissertation.

Additionally, I extend thanks to Drs Remy Froissart and Marie Vasse for the contribution of *Prevotella* isolates for this project, and to all my fellow MIG post-graduate students for their support both in the lab and out of it.

I would also like to thank Professor Jo-Ann Passmore for the opportunity to remain as part of the MIG group for my MSc, and for all the growth and development opportunities that came with it. The entire MIG group had a role in how enjoyable my Master's years was, but a special thanks to Jo-Ann, Tanya, and Allison for their assistance in organising funding and admin. I am very grateful to have received funding from the Calestous Juma Scientific Leadership Fellowship, awarded to Professor Jo-Ann Passmore via the Bill and Melinda Gates Foundation, making it possible for me to pursue this degree.

Abbreviations

AAM	Antibiotic assay medium
Abi	Abortive infection
AMR	Antimicrobial resistance
ANI	Average nucleotide identity
BHI	Brain heart infusion
BLAST	Basic Local Alignment Search Tool
BV	Bacterial vaginosis
CARD	Comprehensive antibiotic resistance database
CDC	Centers for Disease Control and Prevention
CFU	Colony forming units
CIN	Cervical intraepithelial neoplasia
CLSI	Clinical and Laboratory Standard Institute
Cpn60	Chaperonin-60
CST	Community state types
CVM	Cervicovaginal mucus
DDH	DNA-DNA hybridisation
EPS	Exopolysaccharide
FGT	Female genital tract
HIV	Human immunodeficiency virus
HPV	Human papillomavirus
hrHPV	High-risk human papillomavirus
HSV-2	Herpes-simplex virus
IL-6/ IL-8	Interleukin-6/ interleukin-8
LAL	Limulus amoebocyte lysate
LC	Langerhans cells
LKV	Laked Brucella blood with kanamycin and vancomycin
LPS	Lipopolysaccharide
MGE	Mobile genetic elements
MIC	Minimum inhibitory concentration

MLS	Macrolides, lincosamides and streptogramins
NCBI	National Center for Biotechnology Information
OTU	Operational taxonomic unit
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PID	Pelvic inflammatory disease
PPROM	Preterm premature rupture of membranes
PTB	Pre-term birth
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SNP	Single nucleotide polymorphism
SOC	Standard of Care
STI	Sexually transmitted infection
TLR4	Toll-like receptor 4
UT	Universal target
UTI	Urinary tract infection
VFDB	Virulence factor database
WGS	Whole genome sequence
WHO	World Health Organisation

List of Figures and Tables

Figures

Supplementary Figures

Tables

Supplementary Tables

Abstract

A common condition in reproductive age women, bacterial vaginosis (BV), is characterised by a dysbiosis in the vaginal microbiome, and frequently associated with abundance of the genus *Prevotella*. Prevalence of BV is particularly high in sub-Saharan Africa and comprehensive characterisation of South African BV-associated bacteria is vital for BV management in the region. However, to date, most research has focussed on BV-associated *Prevotella* species from the global North.

A total of 69 *Prevotella* isolates were purified from vaginal samples from the UChoose cohort study, which was designed to assess the effect of hormonal contraceptives on HIV risk in adolescent girls and young women (16-19 years old). Isolate genus was confirmed by *Prevotella* genus-specific PCR for all isolates. A subset was speciated by 16S rRNA gene sequencing to be *P. bivia* (n=36) and a putative *P. melaninogenica* (n=1) and were carried forward for further phenotypic characterisation, including antimicrobial resistance testing by Etest methodology, and quantification of biofilm formation by crystal violet staining. South African *Prevotella* isolates displayed 2.7% and 8.3% resistance to metronidazole and clindamycin, respectively, with higher rates of resistance against non-BV antibiotics (19.44%-38.89%). Several strains exhibited co-resistance to two (12/36), three (6/36) or four (2/36) antibiotics. Biofilm formation was significantly greater than the *P. bivia* ATCC 29303^T type-strain in over a quarter of strains ($p>0.05$), with some intra-sample diversity in biofilm-forming ability.

Additionally, 36 strains (35 *P. bivia* and 1 *P. melaninogenica*) were whole genome sequenced using Illumina MiSeq technology. Genomes of these isolates, and of published sequences on the NCBI RefSeq database, were annotated and screened for antimicrobial resistance and virulence genes, as well as for the presence of prophages or anti-phage defence mechanisms. Phylogenetic analysis by *cpn60* gene sequencing and core genome alignment revealed that intra-sample diversity was generally low, but at least one sample had two distinct strains at the same time point. The putative *P. melaninogenica* isolate was most closely related to another vaginal *Prevotella* spp., but both were potential novel species based on average nucleotide identity score (<95%). Isolation location (South African versus international) was not reflected in the phylogenetic trees and there were no differential genes associated with isolate origin. Antimicrobial resistance genes relevant to metronidazole (*nimK*) and clindamycin resistance (*ermF*) were identified in 2.27% (1/44) and 29.09% (4/44) strains, respectively. The variants of the β-lactamase gene, *cfxA*, were differentially associated with resistance to a β-lactam antibiotic ($p=0.000212$).

Lipopolysaccharide (LPS) production was assessed in a subset of *Prevotella* strains using a commercially available isolation and quantification kit. LPS production was 24-fold higher in the *P. melaninogenica*-related strain, compared to five *P. bivia* isolates tested ($p<0.05$), and was structurally different to *P. bivia* LPS as observed by SDS-PAGE.

Overall, isolates, added to a *Prevotella* biorepository, were phenotypically and genotypically characterised, which will enable regionally relevant BV research into future intervention and treatment plans.

Chapter 1: Literature review

1.1 Vaginal microbiome

1.1.1 Introduction

Within the human body, much research over the past decade has elucidated the role of the microbiota in host systemic health and highlighted the complexity of these microbial ecosystems (Peterson et al., 2009). Historically, a large proportion of microbiome-based research has focused on the gut, oral cavity and skin, with other body sites only more recently garnering increased attention (Abdill et al., 2022). One of these 'other', only more recently investigated microbiomes, is the vaginal tract, despite its public health and reproductive health importance.

1.1.2 Vaginal microbiota

In a paper published in 1892, the German gynaecologist Albert Döderlein identified what he called the 'Döderlein bacillus' being the dominant species in the healthy vagina – a Gram-positive bacillus, now believed to be a species of *Lactobacillus*, which he noted was found in greater abundance in low pH vaginas. Since these early days, *Lactobacillus* spp. have been thought to be associated with better vaginal health and reduced vaginal symptoms, an idea that is still central today (Chee et al., 2020; Thomas, 1928). The vaginal microbiota is optimally characterized by a low diversity, low pH (<4.5) ecosystem that is almost singly colonized by *Lactobacillus* species (primarily *Lactobacillus crispatus*, *Lactobacillus gasseri* and *Lactobacillus jensenii*) (Boskey et al., 1999; Ravel et al., 2011). These *Lactobacillus*-dominant communities engage in a mutualistic relationship with the host, using nutrients available in the vagina and in turn, providing protection to the host. For instance, vaginal free glycogen has been shown to be an energy source for the lactobacilli, promoting their growth (Mirmonsef et al., 2014). In return, lactobacilli produce pH-lowering compounds such as lactic acid, as well as target-specific bacteriocins, which protect against colonisation by opportunistic pathogens (Boskey et al., 2001; Karaoglu et al., 2003). Thus, the host provides an

anaerobic, nutritious vaginal niche for the lactobacilli in exchange for the antimicrobial, anti-inflammatory and protective properties of the bacteria (Chetwin et al., 2019).

Defined by Ravel et al. in 2011, vaginal microbial communities can broadly be categorized into five community state types (CSTs), which represent five core vaginal microbiomes, each of which is dominated by hallmark bacterial species (Ravel et al., 2011). Four of the five core vaginal communities are characterised by *Lactobacillus* spp. dominance – CST-I by *L. crispatus,* CST-II by *L. gasseri,* CST-III by *Lactobacillus iners* and CST-V by *L. jensenii*. CST-IV is, in contrast, not characterised by lactobacilli, but rather by a heterogenous community of obligate anaerobic bacteria, including *Gardnerella* spp.*, Prevotella* spp.*, Fannyhessea vaginae* (previously *Atopobium vaginae*) and *Sneathia* spp., among others.

1.1.3 Factors impacting the composition of the vaginal microbiota

While some post-pubescent individuals maintain a stable vaginal microbiome over time, longitudinal studies have shown that there can be rapid alterations in the microbiome, shifting from one CST to another (Gajer et al., 2012). Typically, vaginal communities described as CST-I are reported to be most stable, while CST-III is reported to transition to other CSTs more readily (Munoz et al., 2021). Multiple factors, both internal and external, have been investigated as having the ability to cause shifts in the vaginal microbiota.

One of the endogenous factors that greatly impacts the microbial composition is hormone fluctuations. Throughout a woman's life, major hormonal events greatly shape the vaginal microbiome. Increased oestrogen levels at puberty stimulate the production of glycogen by the vaginal epithelial cells, facilitating the growth of lactobacilli and the production of lactic acid (Mirmonsef et al., 2014). Prior to puberty, and post-menopause, when oestrogen levels are lower, the lactobacilli are markedly lower in number (Hillier & Lau, 1997). This role of oestrogen in modulating the growth of lactobacilli is reinforced by studies assessing the effect of oestrogen hormone therapy in postmenopausal women, where supplementation with exogenous oestrogen can shift the microbiome from *Lactobacillus*-deficient to *Lactobacillus*-dominant (Galhardo et al., 2006; Pabich et al., 2003). Thus, it stands to reason that the menstrual cycle, which is associated with short-

term fluctuations in sex hormones, may also have an impact on the microbiome of the vagina. Several studies have shown that during the menstrual cycle, *Lactobacillus* abundance increases, until menses, which results in a decrease in lactobacilli and an increase of other, non-*Lactobacillus* species (Eschenbach et al., 2000; Gajer et al., 2012; Lopes dos Santos Santiago et al., 2011). This period of menses is highly disruptive to the vaginal microbiota, reducing stability, partly due to the specific changes in oestrogen and progesterone levels experienced at this time, or due to menstrual blood directly. In addition to the endogenous factors that impact the microbial composition of the vagina, exogenous factors, such as douching, antibiotics and sexual practices, can also result in community alterations (Mayer et al., 2015; Noyes et al., 2018; Onderdonk et al., 1992; Schwebke et al., 1999).

Studies such as these have tried to identify factors that impact intra-individual temporal changes in the vaginal microbiome, but importantly, there are also inter-individual differences in microbial composition. Factors such as race and ethnicity have been proposed to affect the composition of the vaginal microbiota. A study in a cohort of four self-described ethnic groups (white, Asian, black, and Hispanic) revealed a significant difference in the microbial composition amongst the groups. Asian and white women were primarily dominated by lactobacilli (CST-I, -II, -III, -V) in 80.2% and 89.7% of cases, respectively, while black and Hispanic women in only 61.9% and 59.6% of cases, respectively (Ravel et al., 2011). These differences have also been reported in other studies in different regions (Fettweis et al., 2014; Zhou et al., 2007, 2010). In general, white and Asian women are reported to have non-*iners Lactobacillus*-dominant vaginal communities more commonly, while Hispanic or black individuals (particularly in Africa) more frequently have non-*Lactobacillus* communities (CST-IV) or *L. iners* dominated CST-III communities (Balle, Konstantinus, et al., 2020; Jespers et al., 2017; Wang et al., 2023). Whether this is due to genetic or behavioural and socioeconomic factors is not yet known, but it does suggest that researchers need to critically look at what is defined as a 'healthy' vaginal microbiota, as it may not be applicable for all population groups (McKinnon et al., 2019).

1.2 Bacterial vaginosis (BV)

1.2.1 Introduction

BV is a commonly described condition in reproductive age women, with an incredibly high reported prevalence of 34-58% in some South African populations (Kenyon et al., 2013). BV is defined as the depletion of beneficial, acid-producing vaginal lactobacilli, accompanied by a bloom of anaerobic bacteria (Eschenbach et al., 1989; Muzny et al., 2018). These alterations are characterised by an increased vaginal pH and a subsequent loss of protection associated with a more acidified vagina. Some of these BV-associated anaerobes include *Gardnerella* spp.*, Prevotella* spp.*, Megasphaera* spp.*, F. vaginae* and *Sneathia* spp. (Zozaya-Hinchliffe et al., 2010). BV is a highly heterogenous condition. While the optimal vaginal state is one with low microbial diversity, BV is often associated with a complex community of different bacteria. However, certain bacteria are reported to play key roles in the initiation and progression of the condition, as well as the negative sequelae associated with it. Schwebke et al. (2014) proposed a model showing that *Gardnerella* spp. may be responsible for initiating the BV state. However, *Gardnerella* spp. have been identified in women without BV, so their presence alone is not always sufficient for BV development (Machado et al., 2016). The ability of *Gardnerella* spp. to sometimes result in BV may be due to virulence differences in *Gardnerella* spp. strains or subclades within the genus, but may also be dependent on the synergistic action of other bacterial species present along with *Gardnerella* spp. in complex BV polymicrobial environment (Fredricks et al., 2005; Nisha et al., 2019).

1.2.2 Diagnosis, risk factors and treatment of BV

<u>Diagnosis</u>

While not always presenting with symptoms, BV is a leading cause of abnormal vaginal discharge and vaginal malodour (Bradshaw et al., 2006; Majigo et al., 2021). BV is typically diagnosed based on one of two standardized clinical tests – Amsel criteria or Nugent scoring (Amsel et al., 1983; Eschenbach et al., 1989; Nugent et al., 1991). In 1983, Amsel et al. developed a set of clinical criteria for diagnosis of BV based on the presence of at least 3 of 4 well-defined vaginal symptoms associated with BV – a thin, homogenous

white or yellow discharge, a vaginal pH of over 4.5 (observed by testing vaginal discharge on litmus paper), a fishy odour after addition of 10% (w/v) potassium hydroxide to the wet mount (known as a 'whiff test'), and microscopically observing stippled epithelial cells covered in bacteria (known as clue cells). The presence of three of the four specified conditions is indicative of a positive BV diagnosis according to this method. Alternatively, Nugent scoring can be used for diagnosis (Nugent et al., 1991). This method is microbiological in nature, defined by scoring based on the presence or absence of specific bacterial morphotypes in vaginal smears examined under the microscope. Scores are assigned based on the absence of *Lactobacillus* morphotypes (large Gram-positive bacilli), the presence of curved, Gram-negative rods (indicative of *Mobiluncus* spp.) and the presence of *Gardnerella/ Bacteroides* spp. (Table 1.1). Adding the scores will result in a value between 0 and 10 which determines BV diagnosis (0-3 as BV-negative, 4-6 as BV-intermediate and 7-10 as BV-positive).

Table 1.1**: Criteria for bacterial vaginosis diagnosis by Nugent scoring**

Morphotype	**Number of cells / oil immersion field**	**Score**
Large, Gram-positive bacilli (*Lactobacillus*-like)	30+	0
	5-30	1
	1-4	2
	<1	3
	0	4
Tiny, Gram-variable coccobacilli (*Gardnerella*-like)	>30	4
and/or Gram-negative bacilli (*Bacteroides*-like)	5-30	3
	1-4	2
	<1	1
	0	0
Curved Gram-negative rods (*Mobiluncus*-like)	>5	2
	1-4	1
	0	0

Risk factors for BV

While the exact 'causes' are not clear, several risk factors have been identified which may leave an individual more susceptible to BV onset or BV recurrence. Sexual behavioural

practices are thought to affect BV acquisition risk. For example, number of lifetime sex partners, being a woman who has sex with women, frequency of intercourse, lack of use of condoms, and history of sexually transmitted infections (STIs) have been shown to be associated with BV risk (Fethers et al., 2009; Guédou et al., 2013; Marrazzo et al., 2010). However, while BV transmission dynamics sometimes resemble aspects of STI transmission, there is controversy surrounding the concept of BV as an STI (Morris et al., 2001). Some studies report no incidence of BV in sexually inexperienced individuals, such as that of Fethers and colleagues (Fethers et al., 2008, 2009), while Koumans and colleagues reported a significant prevalence of 18.8% in individuals who had never had sex (Koumans et al., 2007). These conflicts may suggest that BV is more likely, or enhanced, with sexual experience, however more, large cohort studies are needed to determine how prevalent BV is in sexually inexperienced individuals. Other identified non-sexual risk factors for BV include douching, use of a copper intrauterine device (IUD) and smoking (Achilles et al., 2018; Guédou et al., 2013; Nelson et al., 2018; Ranjit et al., 2018). These factors are thought to impact the vaginal microbiota in such a way that can result in a shift toward a BV-like ecosystem, but the exact mechanisms of these shifts need further investigation.

Standard of care BV treatment

Following a positive BV diagnosis, recommended treatment involves the use of broad-spectrum antibiotics, including metronidazole or clindamycin. The Centers for Disease Control and Prevention (CDC) and the World Health Organization (WHO) have the same guidelines for management of BV, recommending: 1) 400-500mg oral metronidazole, twice daily for 7 days; 2) 0.75% metronidazole intravaginal gel, once daily for 5 days; or 3) 2% clindamycin intravaginal gel, once daily for 7 days (*Guidelines for the Management of Sexually Transmitted Infections*, 2021; *Sexually Transmitted Infections Treatment Guidelines*, 2021). The desired outcome following treatment is a return to a more optimal vaginal microbial composition and a disappearance of symptoms. A cure rate of 70-80% has been reported by multiple studies (Lugo-Miro et al., 1992; Oduyebo et al., 2009), however, the standard of care (SOC) antibiotics are not always able to resolve BV long term. Studies conducting long-term follow up after BV treatment have shown a recurrence rate of 50% within 12 months of initial BV clearance (Bradshaw et al., 2006).

Theories behind the frequency of recurrence include reintroduction of the BV-associated microbes from a partner, antimicrobial resistance (AMR) to metronidazole and clindamycin or the protection of certain bacteria by the biofilm scaffold associated with BV (Beigi et al., 2004; Faught & Reyes, 2019; Swidsinski et al., 2008). Currently, management of recurrent BV is with repeated prescription of the SOC antibiotics, but this poses a significant risk for emergence of antimicrobial resistant bacteria, vaginally and systemically, and highlights a need for research into alternative treatment options (Eschenbach, 2007).

1.2.3 Biofilms in BV

Biofilms have been identified as an important feature in BV pathogenesis and persistence, forming an important part of BV ecosystems (Swidsinski et al., 2005). Biofilms are defined as complex communities of different microbes, attached to a biotic or abiotic surface, and are coated in a matrix of extracellular polysaccharides, proteins and nucleic acids (Høiby et al., 2011). In line with a study by Swidsinski and colleagues in 2005, a conceptual hypothesis was proposed by Schwebke and colleagues proposing that *Gardnerella* spp. are required for initiation of the biofilm on the vaginal epithelium, laying down a 'framework' before the biofilm is colonized by other BV-associated bacteria, such as *P. bivia* and *F. vaginae* (Muzny et al., 2019; Schwebke et al., 2014; Swidsinski et al., 2005). *Gardnerella* spp. have been shown to form biofilms more readily than other BV-associated microbes and compared to free-living bacteria, the biofilm-resident cells display a 5-fold increased tolerance to hydrogen peroxide and a 4- to 8-fold increased tolerance to lactic acid, both of which are protective compounds produced by lactobacilli (Patterson et al., 2007; Swidsinski et al., 2005). Some other BV-associated bacteria, such as *P. bivia* and *F. vaginae,* can adhere to the vaginal epithelium and have potential to form biofilms, but these strains are more likely to be outcompeted by *L. crispatus* than *Gardnerella* spp., and may rely on the latter to properly establish in a biofilm (Machado et al., 2013).

Most biofilm studies have been undertaken *in vitro*, with a pre-determined group of bacteria, which does not reflect the extremely complex, multi-species communities that may exist in the vagina, with host and external factors playing a role. However, the

formation of this biofilm, with the incorporation of BV-associated anaerobes into the *Gardnerella*-established biofilm, is an important component of BV pathogenesis, providing a scaffold for adhesion of BV-microbes and is speculated to contribute to frequently reported BV recurrence following the recommended treatment plan (Bradshaw et al., 2006; Swidsinski et al., 2008). The study conducted by Swidsinski and colleagues reported that biofilms consisting mainly of *Gardnerella* spp. and *F. vaginae* were only suppressed temporarily following treatment with metronidazole, but biofilm activity was resumed within 3 weeks of treatment with SOC BV antibiotics. As such, disruption of these BV biofilms is a major research objective for trying to effectively eliminate BV-associated bacteria in the long-term.

1.2.4 Adverse sequelae associated with BV

For those individuals with asymptomatic BV, diagnosis and treatment may still be beneficial (Muzny & Schwebke, 2020). BV is not only undesirable due to the uncomfortable vaginal symptoms but is associated with several adverse health outcomes. Some of the negative outcomes commonly associated with BV are obstetric and gynaecologic sequelae, including pre-term birth (PTB), low infant birth weight, infertility, endometritis, pelvic inflammatory disease (PID) and early miscarriage (Al-Memar et al., 2020; Haggerty et al., 2004, 2016; Hillier et al., 1995; Jacobsson et al., 2002).

PTB, defined as birth at less than 37 gestational weeks, is a major cause of global neonatal death worldwide and as such, has garnered a lot of attention in BV research (Liu et al., 2012). A meta-analysis of 18 studies, consisting of more than 20 000 participants, determined that risk of PTB was twice as likely in BV-positive versus BV-negative individuals (Leitich et al., 2003). The presence and abundance of specific bacterial taxa have been associated with the onset of PTB and other obstetric outcomes, but the mechanisms of initiating these adverse events is not known (Fettweis et al., 2019; Nelson et al., 2014). One hypothesis is that uterine ascension of BV bacteria may put women at a higher risk of PTB, as the detection of microbes in amniotic fluid is associated with PTB (Onderdonk et al., 2008; Watts et al., 1992). However, this is a disputed hypothesis, as the existence of a placental microbiome is not universally accepted. A second

hypothesis underscores inflammation associated with BV as a driver of PTB. BV can be a highly inflammatory condition, especially in the upper genital tract, and may contribute to lower gestational age (Simhan et al., 2003).

Another gynaecological complication that is associated with BV is PID, characterised by an infection of the reproductive organs, which can lead to inflammation of the cervix (cervicitis), fallopian tubes, ovaries, and uterine lining (endometriosis) (Haggerty et al., 2004). This is primarily thought to be attributed to microorganism ascension, particularly ascension of *Neisseria gonorrhoeae, Chlamydia trachomatis* and *Mycoplasma genitalium*. However, BV-associated bacteria have also been identified in PID cases without gonococcal or chlamydial infection (Haggerty et al., 2004). Thus, due to the association of BV with PID, there may be an increased risk of infertility or ectopic pregnancy in women with BV, especially since many women are asymptomatic and undiagnosed for both BV and PID (Taylor et al., 2013).

Additionally, BV-positivity is associated with increased risk of acquiring several STIs. Multiple studies assessing risk factors for herpes simplex virus 2 (HSV-2) have shown that a BV-like microbiota is associated with a positive HSV-2 diagnosis in South African women, and otherwise (Abbai et al., 2016; Cherpes et al., 2003; Nagot et al., 2007). *Chlamydia trachomatis* and *N. gonorrhoeae* are 4 and 3.4 times more prevalent in individuals with BV, respectively (Wiesenfeld et al., 2003). This same study showed that the higher the 'abnormality' of the vaginal microbiota (i.e., the higher the Nugent score), the increased likelihood there was of a positive test for chlamydial or gonorrhoeal infection. BV has also been shown to associate with nearly a 2-fold increased risk of *Trichomonas vaginalis* infection (Abbai et al., 2016; Seña et al., 2021) and associated with human papillomavirus (HPV) infection and non-clearance (Dareng et al., 2016; Gillet et al., 2011; Guo et al., 2012).

Of relevance to Sub-Saharan Africa, BV-positivity is reportedly associated with an increased risk of human immunodeficiency virus (HIV) acquisition (Gosmann et al., 2017). A meta-analysis of 23 published studies showed a consistent association between BV and HIV infection risk (Atashili et al., 2008). Sub-Saharan Africa has the highest incidence of HIV infection worldwide, with 63% of all HIV infections in 2022 occurring in Sub-Saharan African women (UNAIDS Fact Sheet, 2023). The reported

disproportionately high BV incidence in the region makes these associations with increased HIV acquisition more concerning for women (Kenyon et al., 2013). The mechanism behind this increased risk is not fully understood, but it is speculated that specific BV-associated microbes result in inflammation, and thus an increased recruitment of HIV target cells to the genital mucosa, increasing the probability of HIV uptake by these cells (Anahtar et al., 2015; Gosmann et al., 2017; Masson et al., 2015).

It has been suggested that these adverse health outcomes are a risk for women with even asymptomatic BV, who may not be aware of their positive BV status. Treating asymptomatic BV is not currently recommended according to South African, nor CDC or WHO guidelines, however taking these reproductive health impacts into account, this may be a viable option to reduce the burden of STIs and pregnancy complications in the future (*Guidelines for the Management of Sexually Transmitted Infections*, 2021; *Sexually Transmitted Infections Management Guidelines 2018*; *Sexually Transmitted Infections Treatment Guidelines*, 2021).

1.3 *Prevotella* spp.

1.3.1 Introduction

Named after French bacteriologist, A. R. Prévot, *Prevotella* spp. are common inhabitants of various human and ruminal microbiomes (Tett et al., 2021). *Prevotella* are rod-shaped Gram-negative, non-sporulating, obligate anaerobes, first described with the identification of *Bacteroides melaninogenicus* (now *Prevotella melaninogenica*) in 1921 (Oliver & Wherry, 1921). *Prevotella* species were initially classed as members of the genus *Bacteroides,* with the reclassification of the bacteria to the genus *Prevotella* based on the susceptibility of strains to bile, their moderate saccharolytic activity and differences in enzyme production (Shah & Collins, 1990). In healthy humans, *Prevotella* spp. can be found in relatively low abundance at the mucosal surfaces of the oral cavity (Könönen, Jousimies-Somer, et al., 1994; Könönen et al., 2022) and urogenital tract (Modena et al., 2017; Ravel et al., 2011), as a commensal in a balanced microbiome. *Prevotella* spp. can also be found as a commensal in high abundance in the gut, associated with consumption of a ‘non-Westernised’, plant-based diet (as opposed to

higher *Bacteroides* abundance in protein-rich and high animal fat diets) (Wu et al., 2011). However, these commensals can transition to pathobionts when there is a bloom in abundance of the *Prevotella* spp.*,* particularly in the mouth and urogenital tract. Microbial dysbiosis in these body sites can result in an alteration of host immune response, leading to an imbalance in immune homeostasis and in turn, can allow a further outgrowth of dysbiotic strains, which can grow and establish more easily (Larsen, 2017). This transition can result in several localized conditions, such as that of the mouth (periodontitis), the respiratory tract (pneumonia) and the vagina (BV) (Dahlén, 1993; Ravel et al., 2011; Yamasaki et al., 2013). In the context of BV, *P. bivia* is one of the most prominent BV-associated organisms, found in one study to be 10-fold more abundant in BV-afflicted individuals compared to BV-unafflicted individuals (Zozaya-Hinchliffe et al., 2010).

Prevotella bivia is a hallmark of BV diagnosis and pathogenesis, but is present with other BV-associated species, such as *Gardnerella* spp. (Ravel et al., 2011). Symbiosis has been reported between these species, where growth of each is enhanced by the mutual exchange of nutrients (Pybus & Onderdonk, 1997). Amino acids produced by *Gardnerella* spp. metabolism are used by *P. bivia*, resulting in ammonia that is in turn used by *Gardnerella* spp. *Prevotella bivia* has been shown to incorporate into *G. vaginalis* biofilms, and compared to mono-culture *G. vaginalis* biofilms, glycosyltransferase genes required for biofilms were up-regulated when *P. bivia* was incorporated (Castro et al., 2021; Yeoman et al., 2010). This suggests a symbiosis within the BV biofilm between *P. bivia* and *Gardnerella* spp. that may promote biofilm formation and maintenance and promote the growth of each species within the biofilm. There have also been studies suggesting that *Gardnerella* spp. presence in the vagina may promote the ascension of *P. bivia* into the uterus (Gilbert et al., 2019).

This symbiosis highlights a complex inter-species relationship between *P. bivia* and *Gardnerella* spp. that may explain their frequent dual presence in BV. Symbiotic relationships such as these may exist amongst other BV-associated species, as has been reported for *F. vaginae* and *G. vaginalis* in biofilms (Castro et al., 2021) and *P. bivia* with *Peptostreptococcus anaerobius* (Pybus & Onderdonk, 1998), but most research has focused on the primary BV-species, *P. bivia* and *G. vaginalis*.

1.3.2 The role of *Prevotella* spp. in adverse BV sequelae

Some of the negative health outcomes associated with BV, including obstetric, gynaecological, and STI outcomes, are associated with *Prevotella* spp. Below are some of the BV-associated sequelae that *Prevotella* spp. are thought to promote.

Sexually transmitted infections

Human Immunodeficiency Virus

Several factors have been shown to impact susceptibility to HIV, including cervicovaginal mucus, epithelial integrity and immune cells present at the site of infection. Target HIV cells in the female genital tract (FGT) are likely $CD4^+$/ $CCR5^+$ cells, with the virus either infecting these cells at the vaginal mucosa directly, or indirectly by presentation to $CD4^+$ T cells by dendritic cells or vaginal Langerhans cells (LC) which have internalized HIV (Belyakov & Berzofsky, 2004; Hladik et al., 2007). *Prevotella* spp. abundance in the vaginal microbiota has been associated with increased risk of HIV acquisition. One study showed that *P. bivia* and *P. melaninogenica* were significantly more abundant in women who acquired HIV compared to those who didn't (Eastment & McClelland, 2018). In some cases, this may be due to increased mucosal inflammation that leads to a disruption of epithelial barrier function and recruitment of HIV target cells (Arnold et al., 2016; Gosmann et al., 2017). However, *Prevotella* spp. have also been shown to greatly increase HIV susceptibility risk by promoting uptake of HIV-1 by vaginal LCs, increasing the risk of presentation to HIV target cells and subsequent HIV infection (van Teijlingen et al., 2022). *Prevotella* spp. may also contribute to thinning of vaginal mucus by production of sialidases (Briselden et al., 1992; Nunn et al., 2015), which facilitates viral migration to the underlying epithelial layer. Finally, *Prevotella* spp. have been implicated in the degradation of HIV microbicides, such as tenofovir, decreasing their efficacy *in vivo* (Klatt et al., 2017).

Human Papillomavirus

It is reported that women with BV are more likely to have high risk HPV (hrHPV) types at the cervix than women without BV (Dareng et al., 2016).Persistent infection with hrHPV can lead to the development of cervical intraepithelial neoplasia (CIN) and cervical

cancer (Walboomers et al., 1999). A two-year, prospective cohort study showed that the abundance of *Prevotella* spp. was significantly higher in women with persistent hrHPV than in women without persistent HPV (Dong et al., 2022). Additionally, this study showed that HPV-16, and HPV-18, the HPV subtypes that cause ~70% of cervical cancers, were associated with *Prevotella* spp. abundance. Persistence and reduced regression of CIN (i.e. the growth of abnormal cells on the cervical epithelial tissue which can progress into cervical cancer) is associated with *P. timonensis* presence in the vaginal microbiota (Mitra et al., 2020). The mechanisms by which *Prevotella* spp. are associated with hrHPV may be the same as the proposed mechanisms of increased HIV acquisition, with altered mucus viscosity and reduced epithelial barrier integrity allowing better access to the vaginal epithelium. Additionally, studies have reported that there is a relationship between hrHPV (particularly HPV-16 and HPV-18) and toll-like receptor 4 (TLR4) and NFkB signalling in vaginal epithelial cells (Dong et al., 2022) . Since *Prevotella* spp. can activate NFkB signaling, this might be one mechanism by which its abundance can promote HPV progression and persistent infection (Nasu & Narahara, 2010).

Chlamydia

Apart from the chronic STIs, HPV and HIV, *Prevotella* spp. are also associated with chlamydial infections. Increased susceptibility to chlamydia is proposed to be due to sialidase-induced changes to the vaginal mucosa and epithelial damage, but there is also a role of the species in the metabolism of the pathogen. *Prevotella* spp. are a source of vaginal indole and the *C. trachomatis* enzyme, TrpBA, enables the conversion of this indole to tryptophan, which promotes persistence of chlamydial infection (Wang et al., 2022; Ziklo et al., 2016). Interestingly, while *Prevotella* spp. are capable of producing indole (via the *tnaA* tryptophanase gene), there are within-genus variations which can affect indole availability in the FGT and may impact how *Prevotella* spp. interact with and promote persistence of *C. trachomatis* (Aiyar et al., 2014).

<u>Obstetric and gynaecologic effects</u>

Prevotella spp. have been specifically associated with PTB, one of the most commonly reported BV-associated obstetric outcomes (Mikamo et al., 1998). One study identified that signature *Prevotella* species associated with PTB are *P. timonensis, P. bivia, P.*

corporis and *Prevotella buccalis* (Freitas et al., 2018). Another study showed an association between *P. buccalis* abundance and PTB, which also correlated with increased levels of what are believed to be labour-inducing cytokines, interleukin-6 (IL-6) and interleukin-8 (IL-8) (Kumar et al., 2021). One third of premature births are attributed to preterm premature rupture of membranes (PPROM), which is a contributor to neonatal death, foetal skeletal deformities, and neurodevelopmental impairment (Morris et al., 2016; Paramel Jayaprakash et al., 2016). *Prevotella* spp. are often, and in some studies, always, present at the onset of PPROM and can persist throughout treatment (Baldwin et al., 2015; Yan et al., 2022), with the study by Yan and colleagues suggesting that *P. timonensis* in particular could be used as a biomarker for PPROM. Additionally, *P. bivia* abundance is associated with severe pre-eclampsia in pregnant women, a condition characterised by increased maternal blood pressure and may result in PTB (Lin et al., 2020). The researchers also noted significantly higher plasma levels of tumour necrosis factor (TNF) in individuals with pre-eclampsia, but it was not determined whether this was due to *Prevotella* presence or not.

More research is needed into determining if treatment of *Prevotella* spp., even in asymptomatic BV cases, would reduce the burden of these conditions, or if identification of *Prevotella* spp. could be used as a biomarker indicative of a high-risk individual. While many of these BV associated adverse sequelae cannot be explained by the abundance of *Prevotella* spp. alone, the association of *Prevotella* spp. with many of the health risks highlights why it is regarded as a species of BV research interest. While the exact mechanisms of the promotion of these conditions by *Prevotella* spp. is not known, one common hypothesis for both the STI and gynaecologic outcomes is the propensity of *Prevotella* spp. to induce vaginal inflammation. Different *Prevotella* spp. strains may have variations in their potential to contribute to disease, with different phenotypic displays of virulence.

1.3.3 Factors playing a role in *Prevotella* spp. colonization

Biofilm formation

Formation of biofilms can be of importance to bacteria attempting to colonise host tissue, playing a vital role in many infections, including BV. Biofilms are highly complex

and highly organized functional communities of microbes that self-aggregate and produce an extracellular polysaccharide (EPS) as a matrix, allowing for attachment to either biotic or abiotic surfaces (Yin et al., 2019). This community arrangement also enables resistance to harsh environmental conditions, including antibiotic presence, extreme pH, and extreme temperature, and a potentially ideal environment for the exchange of AMR genes between bacterial cells (Michaelis & Grohmann, 2023).

For *Prevotella* species, the most well described biofilms are those that form in the oral cavity by oral-resident *Prevotella* species, such as *Prevotella intermedia, Prevotella nigrescens, Prevotella pallens and P. melaninogenica*. *Prevotella* species can be present as commensals in the oral microbiome, colonizing the mouths of infants from early in life, however, some of these commensals are able to transition to a pathogenic state (Könönen, Asikainen, et al., 1994). The current hypothesis is that this may be due to the formation of dysbiotic biofilms, driving diseases such as periodontitis (Jiao et al., 2014). Formation of oral biofilms is best described for the oral pathobiont *P. intermedia*, which is positively associated with increased abundance in subgingival tissue of individuals with periodontitis (Socransky et al., 1998).

As a virulence factor, biofilms are often thought to be advantageous to the bacteria by protecting the cells within the EPS and decreasing antibiotic penetration. There are, however, several studies that suggest that biofilms not only physically protect the bacteria, but that there are several changes in gene expression and phenotype depending on the lifestyle of the microbe (whether it is planktonic or biofilm-resident). A study undertaken by Karched et al. in 2022 analysed the proteomic profiles of the molecules secreted by planktonic and biofilm *P. intermedia* cultures (Karched et al., 2022). Surprisingly, 40-50% of all proteins secreted were unique to either planktonic or biofilm cells and biofilm cells stimulated marked increases in pro-inflammatory expression of macrophage inflammatory protein-1α (MIP-1 α), MIP-1β, IL-1β and IL-8 by human whole blood cells. In parallel to this, other studies have suggested that *Escherichia coli* and *Pseudomonas aeruginosa* andother Gram-negative anaerobes, also display different inflammatory potential on human cells depending on their growth mode, which was modulated by differential modifications to the lipid A subunit of lipopolysaccharide (LPS) (Chalabaev et al., 2014; Ciornei et al., 2010). While it is not yet known if the *P. intermedia*

biofilm cells which induce higher inflammation are also due to LPS modifications, it is becoming clear that mode of growth does impact bacterial interaction with the host and that the biofilm may play a more intricate role in BV pathogenesis than simply preventing antibiotic penetration.

There are also biofilm interactions with the host in the other direction, with host hormones impacting biofilm formation. First identified due to the increase in gingival infections in pregnant women, it has been shown that an increase in sex hormones (oestrogen and progesterone) promotes biofilm formation and inflammation of the gingival tissue (Cornejo Ulloa et al., 2021). One of the hypotheses that has been put forward is that a bloom in species such as *P. intermedia* and *P. melaninogenica*, which are able to use the sex hormones as an essential growth factor instead of vitamin K (Kornman & Loesche, 1982) is responsible, in part, for the observed inflammation. Again, while this has been reported for the oral anaerobes, the question remains if this same phenomenon is seen in other niches with other species. Most relevantly to BV, whether this is seen for *P. bivia* in the vaginal microbiome, where oestrogen and progesterone levels may be in higher abundance than the oral cavity.

Little work has been performed on *Prevotella* biofilm formation in other human niches, thus these studies on *P. intermedia* form much of the knowledge surrounding the formation of biofilms by the species. Nevertheless, these studies provoke questions or ideas about *Prevotella* biofilms in other body sites and highlight that there may be an importance to these biofilms in other contexts, such as BV. *Prevotella bivia* is reported to form smaller biofilms than its BV-counterpart, *G. vaginalis*, but more research is needed to elucidate these in *vivo* BV biofilms and the role *P. bivia* may play in their formation (Machado, Jefferson, et al., 2013).

Antimicrobial resistance

Another factor contributing to the inability of BV to be cleared long-term is the potential of acquired AMR in BV associated bacteria. Recent trends of AMR in *Prevotella* spp. have shown that a large proportion are resistant to penicillin (60-80% prevalence in Europe), ampicillin (51-78% prevalence) and clindamycin (30-40% prevalence in Europe and the US) (Reissier et al., 2023). Emergence of AMR is often attributed to the acquisition of AMR

genes, which can be transferred horizontally from one bacterium to another, facilitating spread of the genes. The occurrence of many well documented AMR genes, such as β-lactam (*cfxA*), clindamycin (*erm*), metronidazole (*nim*) and tetracycline (*tet)* resistance genes, on mobile genetic elements (MGEs) pose a concern for increased AMR resistance in other species (Sood et al., 2023).

Clindamycin and metronidazole are the SOC antibiotics for treatment of BV. Metronidazole resistance amongst *Prevotella* spp. is generally considered rare, although there have been some reports of resistance (Alauzet et al., 2010; Veloo et al., 2019). When present, resistance is often associated with carriage of *nim* genes, which encode nitroimidazole reductase enzymes that convert the metronidazole prodrug into an inactive form (Leiros et al., 2004). *Prevotella* spp. resistance rates against clindamycin are higher, with reports of 36.7%-89% resistance in different studies in different regions (Sood et al., 2023; Veloo et al., 2020), suggesting that certain regions may have a greater burden of AMR than others. Clindamycin resistance in *Prevotella* spp. is often associated with the presence of *erm* genes that encode 23S rRNA erythromycin resistance methylases (Veloo et al., 2019). These enzymes methylate the specific antibiotic target site on the 23S ribosomal subunit, thus preventing efficient binding by the antibiotic (Weisblum, 1995).

Resistance to non-SOC BV antibiotics is also relatively common in *Prevotella* species. This is of relevance because their use for non-BV applications might favour the selection of resistant strains in the vaginal environment. The tetracycline class of antibiotics is commonly used for both Gram-positive and Gram-negative bacteria. It is a bacteriostatic, broad-spectrum antibiotic that has been widely used since the mid-20th century (Grossman, 2016). *Prevotella* spp. are often reported to have the ribosomal protection mechanism of resistance, encoded by *tetQ*. Expression of ribosomal protection proteins prevents binding of the antibiotic to the 30S ribosomal subunit, abrogating the activity of the antibiotic. The *tetQ* gene has been shown to be transferred on MGEs in *Prevotella* species, with MGEs being identified in *P. nigrescens* and *P. intermedia* that are genetically distinct from the well-described transposons in *Bacteroides fragilis* (Tribble et al., 2010). Studies performed on isolates from diseased lungs show there is an association between the presence of *tetQ* and phenotypic tetracycline (particularly doxycycline) resistance in

Prevotella species, compared to strains lacking *tetQ* (Sherrard et al., 2014; Webb et al., 2022). The prevalence of *tetQ* (54%) reported by Webb and colleagues in 2022 was higher than has been previously reported for lung-derived strains in a 2019 Dutch study (30.3%) and in a UK study in 2014 (19%). It isn't clear whether these differences are due to regional diversity, or indicative of an increasing occurrence of this resistance gene over time, but it is clear that *tet* gene prevalence is widespread in the phylum *Bacteroidetes* and in *Prevotella* spp. more specifically (Sherrard et al., 2014; Veloo et al., 2019).

Another class of antibiotic often used for infections with Gram-negative bacteria is the β-lactams, for which resistance is often attributed to β-lactamase production. β-lactamase genes, such as *cfxA*, are reported to be harboured on MGEs and may be horizontally transferred between organisms. In *Bacteroides* species, Ferreira and colleagues detected *cfxA* genes associated with the presence of Tn*4555* mobilizable transposons (12.2kB), which contain a mobilization protein (*mobA*) and *oriT,* and could play a role in disseminating this gene to other bacteria (Ferreira et al., 2007). In *P. nigrescens* and *P. intermedia*, another study also reported an association between the *cfxA* genes and Tn4555 *mobA*, suggesting that these transposons can also be found in *Prevotella* species (Fernández-Canigia et al., 2015).

While β-lactam antibiotics and tetracyclines are not recommended for BV treatment, the global effects of the antibiotics may impact the vaginal microbiome and the presence of mobilizable AMR genes could lead to increased antimicrobial resistance to these broad-spectrum antibiotics. The increasing occurrence of AMR and the high BV recurrence rates highlights the need for the development of novel therapeutics and a movement away from only having antibiotic-based approaches. Funding and research is being directed toward the possibility of using probiotics or vaginal microbiota transplantation (VMT) as a supplement, or replacement, to metronidazole and/or clindamycin treatment (Cohen et al., 2020; Lev-Sagie et al., 2019). However, until these other promising avenues for treatment have been rigorously investigated, developed, and approved, understanding the AMR landscape in BV-associated bacteria is important to investigate, guiding decisions regarding selection of antimicrobials and preventing unnecessary treatment with potentially ineffective drugs

Bacteriophages

Bacteriophages (phages) are ubiquitous members of different microbiomes, playing a role in the structure and stability of the microbiota. They are capable of re-shaping microbial ecosystems by strain-specific targeting of bacterial populations, by facilitating horizontal gene transfer and by lysing individual bacterial cells, allowing for redistribution of nutrients (Zuppi et al., 2022). For example, a reduced gut abundance of *Prevotella* was associated with a subsequent, rapid increase in putative *Prevotella*-targeting *Crassvirales* observed by qPCR, suggesting a drastic microbiota change due to the infectivity and activity of phages (Shkoporov et al., 2024). Most phages exhibit one of two main lifecycles. Virulent phages adopt the lytic lifecycle, whereby they infect the bacterial cell and hijack the host machinery to produce additional phage particles. The cell is then lysed, killing the host and the phage particles go on to infect more cells. Temperate phages, on the other hand, can enter the lysogenic lifecycle. Here, the phage genome becomes integrated into the genome after infection of a bacterial cell. The integrated phage genome (prophage) then replicates along with the host genome until an environmental signal (usually host stress-related) causes the prophage to excise from the genome and enter the lytic cycle. Some phages, in the lysogenic life cycle harbour virulence or AMR determinants (Brown-Jaque et al., 2015; Penadés et al., 2015) and can potentially facilitate spread of clinically relevant genes to their hosts.

To date, bacteriophages have not been successfully isolated from *Prevotella* spp. in culture, however evidence of functional *Prevotella* phages has been identified by sequencing of individual strains and by metagenomic sequencing of specific niches (Devoto et al., 2019). *Prevotella* phage genomes have been identified in the oral microbiome, associated with *P. histicola*, *P. scopos* and other unnamed *Prevotella* spp. (Shkoporov et al., 2024). In this study, while 18/173 predicted *Prevotella* phage genomes were jumbo phages (>200kbp), five had predicted genomes of over 300kbp, significantly larger than the average phage genome size. It appears that *Prevotella* spp. phages are generally large, as another metagenomic study identified megaphages (>540kbp) predicted to target gut *Prevotella* (Devoto et al., 2019). These phages were identified in one Bangladeshi cohort initially, but then also identified in other cohorts in Bangladesh,

Tanzania, and India, as well as in baboons and pigs from Kenya and Denmark, respectively (Devoto et al., 2019). These phages, defined as Lak phages by Devoto et al. (2019), were approximately 5 times larger than *Prevotella* crAss-like phages and only 4.6 times smaller than the *Prevotella* spp. host genome itself. The widespread identification of these Lak phages was a significant finding, but no studies have reported if there is a role of these phages in *Prevotella* functionality in the gut, and none have yet been isolated for further characterisation, thus all current *Prevotella* spp. phage knowledge is based on bioinformatic analyses.

Virulence factors

For a bacterial cell to effectively colonize and establish in the environment it encounters, certain bacterial structures, systems, or molecules (so-called virulence factors) are employed to adhere, survive, and out-compete, often enabling the bacterium to cause disease. In general, most research into *Prevotella* spp. virulence has been undertaken in the context of the oral microbiota, where *P. intermedia* and *P. nigrescens* are especially well studied. These species are implicated in inflammatory periodontal diseases, such as acute necrotizing ulcerative gingivitis and are present in the subgingival complexes of adults with periodontitis (Dufty et al., 2017; Socransky et al., 1998). These studies, on non-vaginal *Prevotella* species, may not reflect virulence factors in *P. bivia* or other vaginal *Prevotella* spp., but may guide research and hypotheses. There have been some reports of potential virulence factors for *P. bivia* which will be discussed here, with a specific focus on the role of these virulence factors in BV pathogenesis.

Lipopolysaccharides (LPS)

One of the most important proposed contributions of *Prevotella* spp. to BV symptoms and health outcomes is promotion of vaginal inflammation. This is likely due to LPS production by *Prevotella* spp., as *P. bivia* has been shown to be a significant source of LPS in the vagina of women with BV (Aroutcheva et al., 2008). LPS is an essential component of the Gram-negative cell wall and functions as a surface antigen, making it highly antigenic to the host immune system. Vaginal LPS can trigger immune stimulation by activating NF-κB pathways through binding to epithelial-, monocyte- and macrophage TLR4 and CD14 receptors (Nasu & Narahara, 2010).

The abundance of LPS in the BV FGT has been shown to be associated with inflammation, but the extent to which *Prevotella* spp. LPS causes high-level inflammation directly has not been well studied. In other niches, *Prevotella* spp. LPS production is also clinically important, playing a role in diseases such as periodontitis (Kim et al., 2007). However, this is not always the case. Studies in the lung microbiome have shown that *Prevotella* spp. LPS stimulation of toll-like receptors (TLRs) is somewhat limited, which may allow *Prevotella* spp. to be well tolerated in the lung (Larsen et al., 2015). Variation in LPS structure may also play a role in determining the degree of immune stimulation by LPS from different organisms (Hashimoto et al., 2003). Thus, while *Prevotella* spp. may be one of the primary sources of LPS in the FGT of women with BV, LPS produced by other species in the FGT (such as species in the taxon *Proteobacteria*) may also contribute significantly to inflammation related to BV, and further research is needed to understand the individual contributions.

Proteases

Proteolytic activity, or the ability to degrade specific proteins, has been reported in numerous *Prevotella* species. Different forms of these enzymes may play a role in tissue destruction, evasion of the host immune response, inflammatory signalling and metabolism by the bacterium (Sharma et al., 2022).

The first barrier to colonization in the FGT is the mucosal barrier. The mucous membrane is primarily made up of water, glycoproteins (mucins) and ions, providing antibacterial functions by containing lactoferrin, lysozymes and immunoglobulins (Wiggins et al., 2001). The thickness of the cervicovaginal mucus (CVM) is linked to the ability of the mucin molecules to interact with one another, which is dependent on the specific carbohydrate residues on the glycoprotein (Moncla et al., 2016). Mucinases are often reported to be secreted by hallmark BV bacteria, and these enzymes specifically target the carbohydrates in the mucin glycoproteins, affecting the composition of the CVM.

The most well-studied of the mucinases in the context of BV is sialidase. Sialidases are produced by *Gardnerella* spp. and *P. bivia*, with studies reporting detection of the enzyme in 75% of women with BV (Marconi et al., 2013). Indeed, sialidase presence in vaginal fluid is often used as a biomarker for BV diagnosis and several commercially available

tests that diagnose BV based on the detection of the enzyme have been developed, which show sensitivity and specificity comparable to Nugent scoring for BV diagnosis (Myziuk et al., 2003). The hydrolytic enzymes can cleave the α-ketosidic linkage between glycosyl residues of sialic acids and glycoproteins or glycolipids (Briselden et al., 1992). A role of the enzyme in degradation of host defense molecules, such as *G. vaginalis*-vaginolysin-specific immunoglobulin-A (IgA), has been postulated, suggesting a role for sialidase in downregulation of the innate immune response (Lewis et al., 2012). While a specific gene for sialidase production has not been identified in *P. bivia* (Robinson et al., 2019), the organism is believed to be a significant source of sialidase in the vagina (Ferreira et al., 2021). Briselden et al.analyzed vaginal bacteria from numerous BV-positive women and showed that *P. bivia*, followed by *Prevotella disiens*, had the highest sialidase producing ability compared to other vaginal bacteria (Briselden et al., 1992). Apart from better enabling colonization of the FGT, production of this enzyme is also associated with early spontaneous PTB, with a directly proportional relationship between sialidase activity and the likelihood of an adverse pregnancy outcome (Cauci et al., 2002; Cauci & Culhane, 2011).

Adhesins

While destruction of the mucous membrane by proteases allows the bacteria better access to the underlying host cells, the ability to adhere to these cells, by means of adhesins, is another bacterial virulence factor. To date, most of the research on adhesins in *Prevotella* spp. has focused on *P. intermedia* isolated from the oral microbiome. Surface structures, such as fimbriae, have been identified in *P. intermedia* and demonstrated to result in hemagglutinating activity with human erythrocytes (Leung et al., 1999). Haemagglutination and haemolysis have also been detected in the presence *P. nigrescens* (Okamoto et al., 1999). The haemin released from the degradation of erythrocytes by these activities could act as a source of exogenous iron, supporting the growth of *Prevotella* spp. (Leung et al., 1998). Haemagglutination is correlated with the presence of the *phg* gene in *P. intermedia* and *P. nigrescens* (Lee et al., 2021), however, related systems have not yet been described for vaginal *Prevotella* spp.

Metabolic products: biogenic amines and organic acids

One of the hallmark features of BV diagnosis based on Amsel criteria is a positive 'whiff' test, or a fishy odour (Amsel et al., 1983). This malodour is mostly due to the production of biogenic polyamines by BV-associated bacteria and genes for the biosynthesis of biogenic amines have been detected in several *Prevotella* species, including *P. bivia*.

Biogenic amines, such as putrescine, cadaverine and trimethylamine, may inadvertently allow for the expansion of BV-associated species. A model summarized by Nelson et al. demonstrates that the production of biogenic amines, by metabolism of amino acids, is coupled with the consumption of hydrogen ions and a raising of vaginal pH (Nelson et al., 2015). Such a shift toward a non-optimal state reduces the competitive advantage of the protective lactobacilli over low abundance pathogens and may allow them to bloom. Additionally, short-chain fatty acids (SCFAs) produced by *Prevotella* spp. during normal metabolism may impact the host immune response. One study showed that large amounts of acetic acid, succinic acid and volatile fatty acids are produced by *Prevotella* spp. compared to other BV-associated bacteria, which may allow the bacteria to avoid contact with polymorphonuclear leukocytes (PMNLs) by inhibiting chemotaxis, inhibiting immune cell migration and overall by preventing targeting by immune cells in the FGT (Eley et al., 2000).

1.3.4 Diversity in the genus *Prevotella*

Prevotella species are found in many niches in the human body but are not associated with disease in all of them. A 2021 review article reported on characterisation of more than 50 *Prevotella* species, a number which is likely higher today (Tett et al., 2021). *Prevotella* spp. have been isolated from many non-human hosts (mainly from animal rumens), where they play a role in carbohydrate metabolism (Betancur-Murillo et al., 2023). There have also been reported identifications from various environmental sites by metagenome assembled genome sequencing, however the majority of *Prevotella* spp. have been isolated from human hosts (Tett et al., 2021) A pangenome analysis by Gupta et al. showed that the genomes of *Prevotella* species from different human body sites have a different collection of genes, with different functional potential, which may be indicative of adaptations to those body sites (Gupta et al., 2015). Habitat-specific gene

presence and gene absence was observed in *Prevotella* spp. genomes, both between species and within-species. Thus, a *Prevotella* species or strain isolated from the gut, may have distinct site-enriched pathways to one from the vagina, reflecting evolutionary adaptations to that site.

Additionally, within each niche, intra-species diversity may also impact research outcomes. Single strains isolated from an individual tend to result in an underestimation of the metabolic potential of the species within that individual, since the *in vivo* reality is usually a consortium of strains belonging to that species, together providing a wider repertoire of functional genes (Lloyd-Price et al., 2017). *Prevotella* species are among those that have been investigated and shown to display diversity between strains in the vaginal microbiota, with multiple strains present within single individuals (Ma et al., 2020). The intra-species diversity of vaginal *Prevotella* spp. may provide a larger breadth of metabolic ability and may impact resilience and temporal stability in the vaginal microbiome.

Diversity of *Prevotella* spp. has also been observed between individuals, with diversity detected between ethnic groups, geographies, and ages (Yatsunenko et al., 2012). These variations may be associated with differences in lifestyle in different groups, as is the case for gut *Prevotella* spp., where abundance and species diversity are greater in those with a primarily 'non-Western' diet. Geography is also a determinant of the vaginal microbiota composition, with taxon abundance differing in different regions, but there are few studies assessing fine-scale, within-species variation across geographic regions (Lennard et al., 2019). Whether or not variations in BV-associated bacteria by location are associated with the geographic differences in BV prevalence is not known, but it has been shown that different strains may impact BV pathogenesis in different ways. Phenotypic differences in antibiotic response, biofilm forming ability and inflammation may exist in *Prevotella* spp., leading to some strains being more pathogenic than others (Mikamo et al., 1998). *Prevotella* spp. are frequently detected in the vaginal microbiota of women without BV, but factors that may drive the colonisation and bloom of some vaginal *Prevotella* spp. strains and not others still need to be elucidated.

1.4 Project rationale

There is a significant body of literature pointing to a central role for *Prevotella* spp. in both the development and persistence of BV and the negative health outcomes associated with the condition. However, reports of detailed characterisation of strains involved in BV are relatively rare and almost completely lacking for populations in developing parts of the world, including sub-Saharan Africa. Due to potential phenotypic differences in BV-associated bacteria across geographies, research into treatment and prevention plans that is relevant to the South African healthcare setting would benefit from representative strains from South African women. Establishment of a collection of South African *Prevotella* spp. strains would allow for phenotypic and genotypic characterisation, assessing how *Prevotella* spp. may contribute to BV, and allow for comparison to strains in international culture collections.

1.5 Aims and Objectives

The overall aim of this study was to generate a phenotypically and genotypically characterised biorepository of South African vaginal *Prevotella* species. This was achieved via three main objectives, each represented by a chapter in the thesis.

Chapter 2: To isolate vaginal *Prevotella* spp. from South African adolescent girls and young women and to characterise them phenotypically and in terms of high-level genetic diversity.

Chapter 3: To generate and use whole genome sequence data for *Prevotella* isolates to investigate their evolutionary relatedness, to identify carriage of clinically relevant genes, and to identify putative prophages and/or anti-phage resistance mechanisms.

Chapter 4: To quantify and structurally compare LPS produced by different South African *Prevotella* spp. isolates and LPS production by *Prevotella* spp. in biofilm versus planktonic growth states.

Chapter 2: Isolation and phenotypic characterisation of *Prevotella* spp.

2.1 Abstract

To date, little characterisation of *Prevotella* species isolated from South African women has been performed, limiting regionally relevant bacterial vaginosis (BV) research. To address this, 69 *Prevotella* spp. isolates were isolated from vaginal samples collected during a previously completed trial (UChoose), and a subset (n=37) were carried forward for phylogenetic interrogation and phenotypic characterisation. The majority of isolates (n=36) were identified as *Prevotella bivia*, while one was preliminarily identified as *Prevotella melaninogenica*, based on similarities to reference 16S rRNA and *cpn60* gene sequences. For isolates from the same sample, the *cpn60* gene diversity was generally low (0-4 SNPs), while longitudinal diversity was higher. Two consecutive isolates from one individual were identical, despite an 8-week interval between sampling times. Resistance rates to BV standard of care (SOC) antibiotics, metronidazole, and clindamycin, were low, at 2.7% (1/36) and 8.3% (3/36), respectively, while higher rates of resistance were observed to the non-BV genitourinary antibiotics doxycycline, amoxicillin, and azithromycin (16.66% [6/36] – 38.88% [14/36]). Over a quarter of novel isolates were able to form significantly larger biofilms than *P. bivia* ATCC 29303^T ($p<0.001$). Contrary to findings of previous studies, *Gardnerella vaginalis* ATCC 14018^T biofilm formation was not improved by the addition of *P. bivia* ATCC 29303^T .

2.2 Introduction

Prevotella spp. are frequently present in the microbiome of women with BV, where they play an important role in vaginal inflammation, a key driver of adverse BV outcomes (Anahtar et al., 2015; Aroutcheva et al., 2008; Zozaya-Hinchliffe et al., 2010). However, to date, they have been understudied and basic phenotypic data such as resistance to commonly used antimicrobials and ability to form biofilms, both of which play important roles in the persistence of BV, are lacking. Isolates from Africa, including South Africa, are further underrepresented in global strain banks and there have been no reports of detailed, culture-based characterisation of isolates from women with and without BV in the region. Therefore, in order to perform regionally relevant research, an important first step is the establishment of a strain collection including representative isolates.

To isolate vaginal *Prevotella* spp., having access to well characterised, stored samples from previous cohort studies is advantageous. Not only does using previously generated samples save time, preventing the need to recruit new participants, but extensive participant metadata is usually collected as part of these studies. The UChoose study was designed to investigate the effect of hormonal contraceptives on HIV risk in adolescent girls and young women, between the ages 16 and 19 (Balle, Konstantinus, et al., 2020). These adolescent girls and young women were enrolled between 2015 and 2018 at the Desmond Tutu Health Foundation Youth Centre in Masiphumulele, a peri-urban settlement near Cape Town, South Africa. As part of the study, participants were tested for HIV and BV and provided an extensive suite of swabs at three main study visits approximately 8 weeks apart. A subset of these samples underwent 16S rRNA gene sequencing, providing microbial community data (Balle, Konstantinus, et al., 2020). Using these samples to isolate and characterise *Prevotella* spp. strains would provide insight into genotypic and phenotypic diversity that may be relevant to BV research in South Africa.

Sequencing of the conserved bacterial 16S rRNA gene is a common tool used to identify bacteria as it is cost-effective and rapid (Deurenberg et al., 2017). However, the limited phylogenetic resolution of 16S rRNA gene sequences impacts the ability to use the gene as a means of investigating phylogenetic relatedness within species (Zeigler, 2003). As

such, it has been suggested that sequencing of another essential gene, the chaperonin-60 (*cpn60*) gene, is a better method for interrogating intra-species diversity (Links et al., 2012). The *cpn60* gene has been used for numerous studies investigating the microbial communities of the vaginal microbiota and has proved especially useful for classification of *G. vaginalis* into its sub-clades (Hill et al., 2005; Schellenberg et al., 2009, 2016). For *Prevotella* and *Bacteroides* spp., *cpn60* is regarded as an improved target compared to the 16S rRNA gene, providing better resolution and as such, can be used to infer strain diversity and evolution over time or geography (Sakamoto & Ohkuma, 2010).

Recent work has revealed unexpected diversity at a strain level in vaginal communities (France et al., 2022; Holm et al., 2023). Rather than existing as clonal populations within the overall microbial community, each species is made up of multiple strains with differing metabolic capabilities (Greenblum et al., 2015; Ma et al., 2020; Oh et al., 2014). Therefore, isolating, and characterising single representatives from a sample may miss some of the important phenotypic diversity present.

The aim of Chapter 2 was (1) to isolate vaginal *Prevotella* species from lateral vaginal wall swabs to create a biorepository of South African isolates, (2) to confirm the identity of these isolates based on the 16S rRNA gene, (3) to assess inter- and intra-sample diversity between the isolates based on *cpn60* universal target (UT) sequencing, (4) to determine the susceptibility of the South African isolates to BV-therapeutic and other genitourinary antibiotics, and (5) to assess the ability of the *Prevotella* spp. to form biofilms *in vitro*.

2.3 Materials and methods

2.3.1 Selection of target vaginal samples for isolation of *Prevotella* spp.

Lateral vaginal wall swabs from the UChoose study (Gill et al., 2020) were stored at -80°C in phosphate buffered saline (PBS) containing 20% (v/v) glycerol within 2 hours of sampling. Samples were collected from participants at three visit times (V1, V2 and V3), spaced approximately 8 weeks apart. Samples for culture were selected based on having the highest number of *Prevotella* assigned operational taxonomic units (OTUs) according to the UChoose 16S rRNA microbiome data (Balle, Konstantinus, et al., 2020). Additionally, *P. bivia* isolates (n=18), isolated from the same cohort by Dr Yann Dumont as part of another project, were gifted by Drs Rémy Froissart and Marie Vasse and included in the characterisation.

2.3.2 Isolation of *Prevotella* spp. from vaginal samples

Samples were inoculated onto pre-reduced Laked Brucella blood agar supplemented with kanamycin and vancomycin (LKV). The medium is composed of Brucella blood agar base supplemented with 5% (v/v) defibrinated sheep blood, haemin (0.1g/L), vitamin K_1 (0.01g/L), vancomycin (7.5mg/L) and kanamycin (0.1g/L)(Appendix). All experiments were performed under anaerobic conditions (85% N_2, 10% CO_2, 5% H_2) at 60% relative humidity and 37°C, in an anaerobic chamber (Baker Concept 1000 Anaerobic Workstation). Following inoculation and 48 hours incubation, up to five individual colonies were sub-cultured and inoculated into Brain Heart Infusion (BHI) broth (BD Bacto™, Franklin Lakes, NJ, USA), supplemented with haemin (5mg/ml), vitamin K_3 (0.5mg/ml), yeast extract (5mg/ml) and cysteine (0.5g/L)(Appendix). Putative *Prevotella* isolates were stored as 20% (v/v) glycerol stocks at -80°C.

2.3.3 *Prevotella* genus-specific PCR

Isolates presumed to be *Prevotella* spp. were confirmed by colony polymerase chain reaction (PCR) with previously described *Prevotella* genus-specific primers (Matsuki et al., 2004). DNA template for use in PCR was prepared by adding putative *Prevotella* colonies to Tris-EDTA (TE) buffer (10mM Tris, 0.1mM EDTA, pH 7.4) with proteinase K (200ng/ml) and incubating at 37°C for 20 minutes. Samples were then incubated at 80°C

for a further 20 minutes to inactivate the proteinase K before centrifugation (5000g for 3 minutes) to remove unlysed cells and debris. The resulting supernatant was used as template for PCR. Reactions (25µl) consisted of 2µl template, 0.5µM g-Prevo-F and g-Prevo-R primers (Table 2.1), 12.5µl 2× KAPATaq ReadyMix (Kapa Biosystems, Wilmington, MA, USA) and 8µl nuclease-free water. PCR cycle conditions were set at 94°C for 5 minutes, followed by 40 cycles of 20 seconds at 94°C, 20 seconds at 55°C and 30 seconds at 72°C, and finally, an extension step at 72°C for 7 minutes. Successful amplification of a 527-529bp product was confirmed by electrophoresis on 2% (w/v) agarose gels, stained with CondaSafe stain (Condalab, Torrejón de Ardoz, Spain) and visualised using a ChemiDoc gel imager (Bio-Rad, USA).

Table 2.1*:* **Primers used in PCR amplification**

Primer	Sequence (5'-3')	Primer type	Reference
27F	AGAGTTTGATCMTGGCTCAG	Universal 16S rRNA gene primer	(Weisburg et al., 1991)
1492R	TACCTTGTTACGACTT	Universal 16S rRNA gene primer	(Weisburg et al., 1991)
907R	CCGTCAATTCMTTTRAGTTT	Sequencing primer	(Lane et al., 1985)
g-Prevo-F	CACRGTAAACGATGGATGCC	*Prevotella* spp. specific 16S rRNA gene primer	(Matsuki et al., 2004)
g-Prevo-R	CCAGACGTTGGGCTGG	*Prevotella* spp. specific 16S rRNA gene primer	(Matsuki et al., 2004)
H279T7*	GCTAGTTATTGCTCAGCGGGAIIIIGCIGGIGAYGGIACIACIAC*	Universal *cpn60* UT primers	(Goh et al., 1996)
H1134T7*	TAATACGACTCACTATAGGGYKIYKITCICCRAAICCIGGIGCY*	Universal *cpn60* UT primers	(Goh et al., 1996)

*Underlined sequence represents landing sites for T7 sequencing primers

2.3.4 16S rRNA gene PCR and sequencing

Once confirmed to be *Prevotella* spp. by genus-specific PCR, 16S rRNA gene amplification and sequencing were performed for species level identification. Reactions

(50µl) contained 4µl template DNA, 0.5µM F27 and 1429R primers (Table 2.1), 25µl 2× KAPATaq ReadyMix (Kapa Biosystems, Wilmington, MA, USA), made up to 50µl with nuclease-free water. PCR conditions were set at 95°C for 5 minutes, followed by 35 cycles of 30 seconds at 95°C, 30 seconds at 55°C and 1 minute 45 seconds at 72°C, with a final extension step at 72°C for 7 minutes. Successful amplification of a ~1400bp product was confirmed by electrophoresis on a 1% (w/v) agarose gel, stained with CondaSafe (Condalab, Torrejón de Ardoz, Spain), and visualised with a ChemiDoc (Bio-Rad, USA) imager. PCR products were sent to Inqaba Biotec (Pretoria, South Africa) for Sanger sequencing using the 907R sequencing primer (Table 2.1), using the BigDye Terminator v3.1 standard kit and run on an ABI 3500XL machine. The resulting sequences were manually trimmed using Chromas v2.6.6 (Chromas, Technelysium Pty Ltd))and compared to sequences in the National Center for Biotechnology (NCBI) RefSeq database (O'Leary et al., 2016) with the Basic Local Alignment Search Tool (BLAST)(Altschul et al., 1990).

2.3.5 *cpn60* UT gene sequencing

The *cpn60* gene of selected isolates was amplified using the previously described H279T7 and H1134T7 primer pair (Table 2.1). Amplification reactions (50µl) consisted of 25µl 2× KAPATaq ReadyMix (Kapa Biosystems, Wilmington, MA, USA), 0.5µM of each primer, 4µl of template DNA and made up to a reaction volume of 50µl with nuclease-free water. The amplification reaction conditions were as follows: 95°C for 5 minutes, followed by 40 cycles of 94°C for 30 seconds, 50°C for 30 seconds and 72°C for 45 seconds, with a final 10-minute extension at 72°C. Amplification of ~550bp products was confirmed by electrophoresis on 2% (w/v) agarose gels, stained with CondaSafe (Condalab, Torrejón de Ardoz, Spain) and visualized with a ChemiDoc (Bio-Rad, USA) imager. PCR products were Sanger sequenced from both ends at Inqaba Biotec (Pretoria, South Africa), using T7 landing sites that were included in the H279T7 and H1134T7 primers. Pre-processing and assembly were performed using gap4 in Staden v2.0.0b11 (Staden, 1996).

Previously published *P. bivia* and *P. melaninogenica* reference *cpn60* sequences were obtained from cpnDB (Vancuren & Hill, 2019) and from genomes in the NCBI RefSeq

database (O'Leary et al., 2016). Additionally, sequences for the *cpn60* universal target (UT) were extracted from the *Prevotella* isolate genome assemblies, described in Chapter 3. Alignment of all *cpn60* sequences was undertaken using CLUSTAL_W v2.0.12 (Larkin et al., 2007) and trimmed in MEGAX v10.2.6 (Kumar et al., 2018). Phylogenetic trees were constructed using RAxML (Stamatakis, 2014) following best-fit model selection using ModelTest-NG_v0.1.7 (Darriba et al., 2020) and the final trees were generated with 500 bootstraps using a Generalised Time Reversible (GTR) model of substitution with the gamma model of rate heterogeneity. Trees were visualised using iTOL v6.8.1 (Letunic & Bork, 2007).

2.3.6 Antimicrobial susceptibility testing by Etest methodology

Susceptibility to SOC antibiotics, metronidazole and clindamycin, and to other commonly prescribed genitourinary antibiotics (amoxicillin, azithromycin, and doxycycline) was tested using MIC Test (Epsilometer test) strips (Liofilchem, Roseto degli Abruzzi, Italy). *Bacteroides fragilis* ATCC 25285^{T}, with known antibiotic susceptibility, was used as a quality control organism. Broth cultures were diluted in PBS to a 0.5 McFarland standard and streaked evenly across Antibiotic Assay Medium (AAM)(Appendix). Etest strips were placed centrally on the plate and incubated under anaerobic conditions for 48 hours. The minimum inhibitory concentration (MIC) was recorded as the minimum antibiotic concentration that inhibited bacterial growth and where breakpoints were available, interpreted according to the Clinical and Laboratory Standard Institute (CLSI) (CLSI, 2020) guidelines.

2.3.7 *In vitro* quantification of biofilm biomass by *Prevotella* spp.

Biofilms were prepared and the overall biofilm mass was quantified using a crystal violet method. To generate biofilms, bacterial cultures grown under anaerobic conditions for 16 hours at 37°C were diluted in fresh BHI (BD Bacto™, Franklin Lakes, NJ, USA) to a McFarland density of 0.5 using a spectrophotometer. Diluted cultures were inoculated into clear, flat-bottomed, 96-well microtiter plates (polystyrene plates, Greiner Bio-One, Monroe, NC, USA) to a final volume per well of 200µl and incubated under anaerobic

conditions at 37°C for 120 hours (5 days) to allow for biofilm formation. Quantification of the biofilm biomass was achieved using a well-described crystal violet assay (Peeters et al., 2008). Briefly, following removal of bacterial supernatant from each of the wells, 200µl PBS was used to wash each well twice. Thereafter, 100µl of 1% (w/v) crystal violet (Sigma-Aldrich, USA) was added to each well and left to stain for 30 minutes. Crystal violet was removed, followed by two washes with PBS and the addition of 100µl of 30% (v/v) acetic acid to solubilize the crystal violet. The solubilized crystal violet was further diluted 1:20 in 30% (v/v) acetic acid in a new, clear, flat-bottomed 96-well microtiter plate and quantified by measuring absorbance at 595nm with an ELx800 microplate reader (BioTek Inc, Winooski, VT, USA). The biofilm formation assays were performed in triplicate biological repeats, with twelve technical replicates per plate. Statistical analysis of biofilm biomass was performed by Kruskal-Wallis one-way ANOVA and Dunn's multiple comparison test in GraphPad Prism v9.2.0 (GraphPad Software, San Diego, CA, USA).

2.3.8 Co-culture *P. bivia* and *G. vaginalis* biofilm formation

Dual-species biofilms were prepared by inoculating *P. bivia* UC109_V3_2 (1×10^7 CFU/ml), determined to have a high propensity for biofilm formation in the preliminary analyses, or *G. vaginalis* ATCC 14018^T type strain (1×10^7 CFU/ml), in a final volume of 1ml, into 24-well flat microtiter plates (clear polystyrene plates, Greiner Bio-One, Monroe, NC, USA). Following anaerobic incubation for 24 hours at 37°C, planktonic cells were removed and replaced with either 1ml NYC III medium (Appendix), or 1×10^7 colony forming units (CFU)/ml of either *P. bivia* or *G. vaginalis* in fresh NYC III medium. After an additional 24 hours, bacterial supernatant was removed, followed by washing with 500µl pre-reduced PBS. A sterile toothpick was used to disrupt the biofilms, which were then resuspended in 200µl PBS. Serial dilutions of the biofilm cells were prepared and plated onto either NYCIII, which represents total bacterial cell count, or LKV plates, which facilitate growth of *P. bivia* only. For each replicate, biofilm crystal violet staining was also performed, as described previously, with biofilm biomass measured spectrophotometrically by absorbance at 595nm (ELx800, BioTek Inc.). This co-culture assay was performed in three biological repeats.

2.4 Results

2.4.1 Isolation and molecular confirmation of *Prevotella* isolates from samples obtained in the UChoose longitudinal study

Analysis of the UChoose 16S rRNA gene data, which was available for most samples from the original trial (Gill et al., 2020), revealed several OTUs classified as *Prevotella* spp. (Figure 2.1). The most commonly observed species both in prevalence across samples and with the highest relative proportion of reads per sample were *Prevotella amnii*, *P. timonensis* and *P. bivia*. When present, and after normalisation across samples, reads mapping to these organisms made up 23-37% of the reads from the sample. Minor species, making up less than 10% of sample reads when present, included *P. disiens, P. buccalis, Prevotella corporis, P. pallens* and *P. melaninogenica*. A number of taxa could not be resolved to the species level and were classified as 'Genus *Prevotella*'.

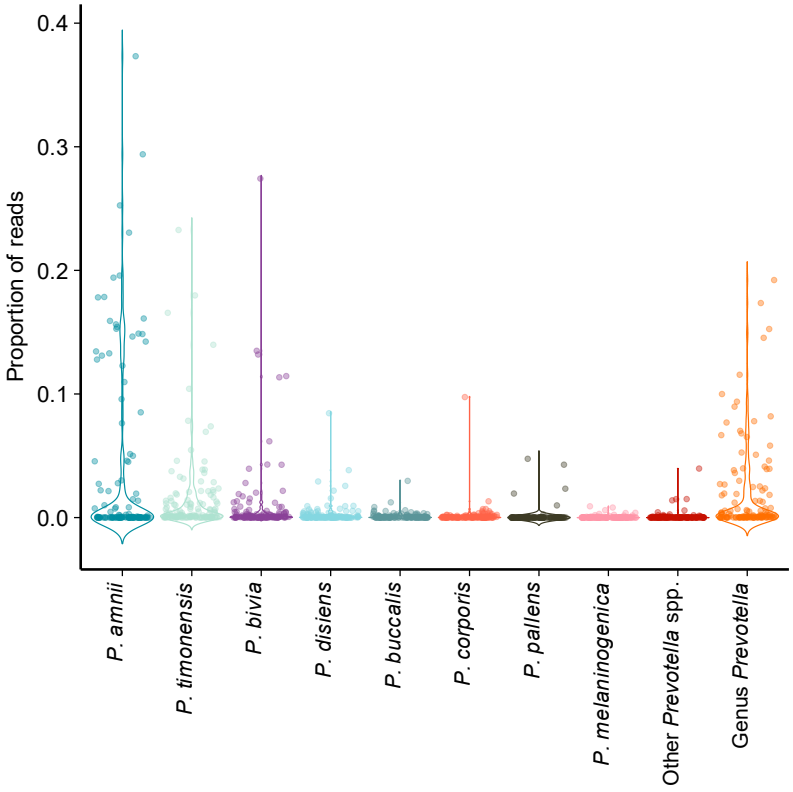


Figure 2.1: **Proportion of *Prevotella* spp. reads in the UChoose 16S rRNA gene data.** The proportion of reads assigned to *Prevotella* spp. operational taxonomic units (OTUs) out of the total number of bacterial reads in each sample. Genus *Prevotella* represents those *Prevotella* reads with no species level designation.

To isolate vaginal *Prevotella* spp. for further characterisation, samples from participants (n=15) with Nugent scores >3 were cultured anaerobically on selective media. An additional sample obtained from a separate study visit was included for three participants, for a total of eighteen samples. Up to five colonies were purified from each target sample, resulting in 48 putative *Prevotella* spp. isolates (Table 2.2). An additional 21 isolates were obtained from a previous project (Dumont et al. personal communication) to yield a final collection of 69 isolates from BV-positive (n=22), BV-negative (n=5) and BV-intermediate (n=4) individuals (Table 2.2). Approximately one fifth (7/31) of participants were treated for an STI prior to the sampling visit, with the most commonly prescribed antibiotics being azithromycin, ceftriaxone, and metronidazole.

All presumptive isolates were screened using *Prevotella* genus-specific primers. This primer set was used to confirm that all isolates belonged to the genus *Prevotella* (Figure 2.2). Isolates were then typed to the species level by partial sequencing of the 16S rRNA gene. Almost all (67/69) isolates were identified as *P. bivia* (≥99% sequence identity to the *P. bivia* JCM 6331^{T} type strain [accession = NR_113096.1] in all cases). The remaining two isolates matched most closely to the *P. melaninogenica* ATCC 25845^{T} type strain, albeit with a comparatively lower percentage sequence identity of 97.57% (query cover 98%). The putative *P. melaninogenica* strains were isolated from a single sample, UC143_V2, which contained 65 reads assigned to *P. melaninogenica* (OTU_17) according to the UChoose rRNA microbiome data.

Table 2.2: **UChoose samples selected for isolation of *Prevotella* spp. and their respective Nugent scores and antibiotic data.**

Sample ID	Nugent Score[1]	Antibiotics prior to sampling	n†
UC016_V1	5		1
UC024_V2	8		1
UC033_V1	8	Azithromycin, Ceftriaxone	4
UC037_V1	9		1
UC040_V1	0		1
UC040_V2	8		1
UC055_V2	8		2
UC055_V3	8		1
UC059_V1	5		5
UC064_V1*	7		1
UC065_V2	2		1
UC071_V1	7		6
UC074_V1	8		1
UC079_V2	8		1
UC095_V2	9		5
UC096_V1	8		4
UC104_V2	10		1
UC107_V1	8		5
UC107_V2	8	Azithromycin	4
UC109_V3	8	Azithromycin, Ceftriaxone, Metronidazole	5
UC113_V1	0		1
UC121_V1*	4		1
UC126_V1	8		4
UC132_V2	6	Azithromycin	1
UC136_V1	10		1
UC143_V2	8	Azithromycin, Ceftriaxone	5
UC146_V1*	8		1
UC149_V3	10	Azithromycin	1
UC158_V3	0	Azithromycin, Ceftriaxone, Metronidazole	1
UC162_V1	0		1
UC167_V3	7		1

*Samples for which no 16S rRNA data were available

†Number of isolates from each sample

[1]Nugent scores are categorised as negative (0-3), intermediate (4-6) or positive (7-10)

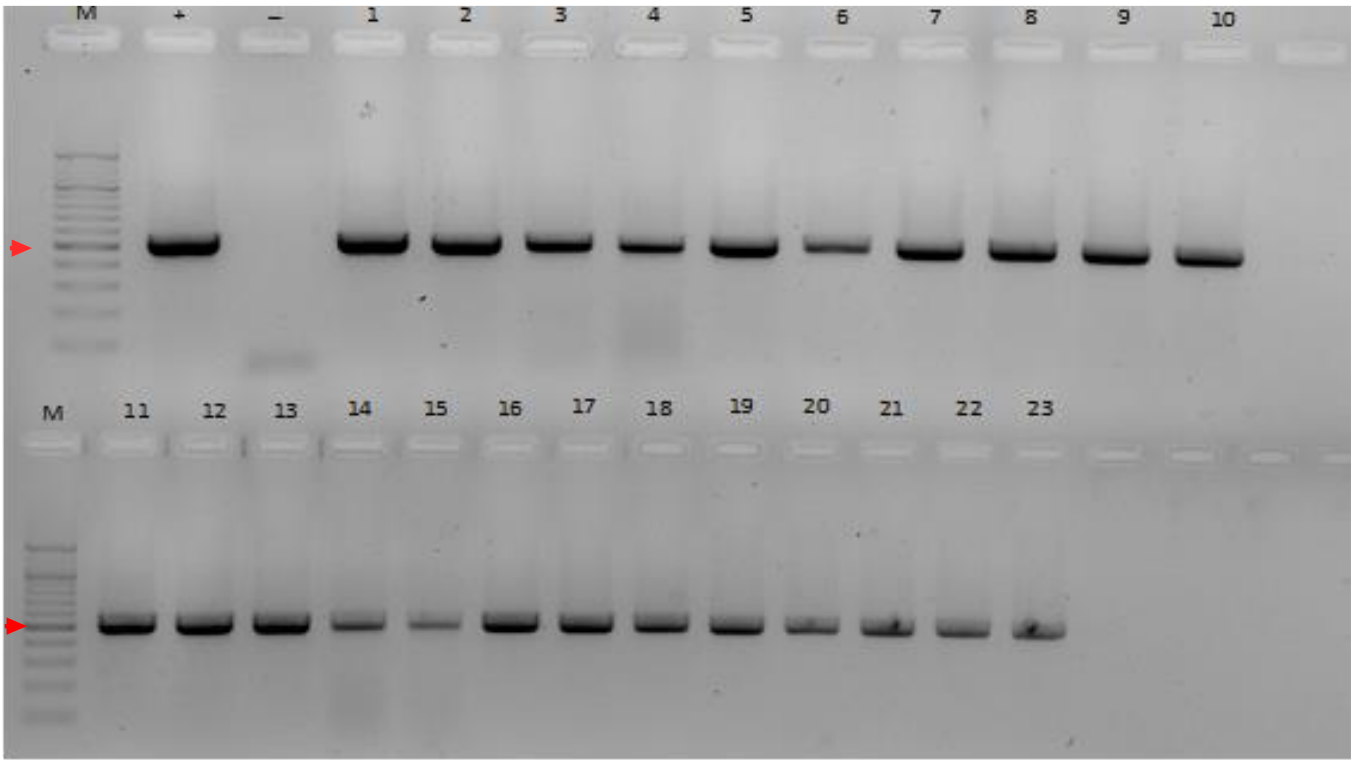


Figure 2.2: ***Prevotella* genus-specific PCR products.** Representative PCR products of 23 isolated strains, confirming genus identity of all strains. *P. bivia* ATCC 29303T was used as a positive control (+) and a no template control was included (-). DNA molecular weight marker (M), 1 Kb Plus DNA Molecular Weight Marker (Thermo

The community state types (CSTs) for the bacterial communities within each of the UChoose vaginal samples were determined by VALENCIA (VAginaL community state type Nearest CentroId clAssifier) (France et al., 2020) for another study (Anika Chicken, personal communication). Of the 31 samples from which isolates were obtained, excluding those without 16S data, 24 were classified as CST IV (having a low abundance of *Lactobacillus* spp.), four were CST-III (*L. iners* dominant) and only one was CST-I (*L. crispatus* dominant) (Figure 2.3). Interestingly, the CST-I sample, UC132, had an intermediate BV status (Nugent score of 6), despite its *L. crispatus* dominance. However, the community within this individual was not stably *L. crispatus* dominant, transitioning from CST-I at visit 2 to CST-III by visit 3. Eight individuals were stably CST-IV across the three visit times, however most of the other samples' CST transitions were between non-

Lactobacillus spp dominance and *L. iners* dominance. None of the individuals had a stably *L. crispatus*-dominant (CST-I) microbiota.

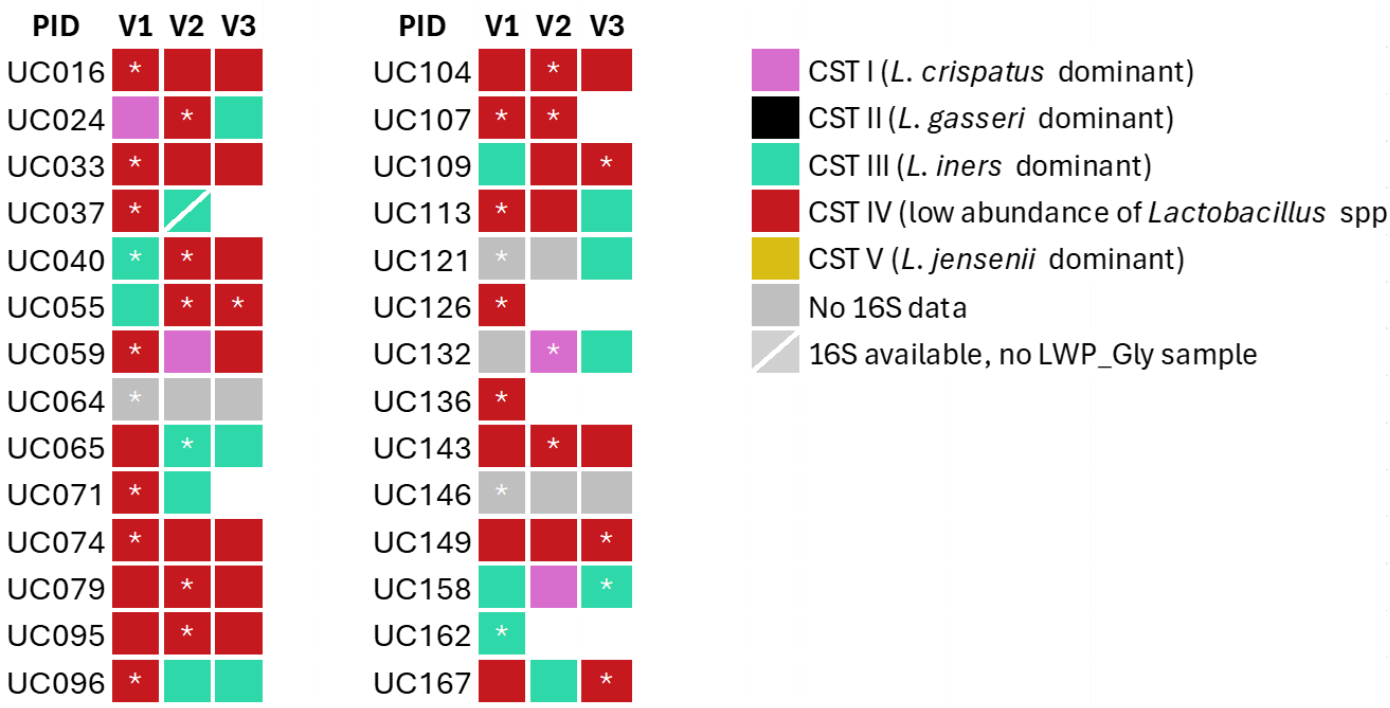


Figure 2.3: **Community state types (CSTs) for individuals from whom strains were isolated.** Where the data were available, vaginal communities were assigned CSTs by VALENCIA (France et al., 2020). The sample visit times from which stains were isolated are indicated by a '*'. LWP_Gly refers to the lateral wall PBS and glycerol swab sample.

2.4.2 Inter-strain and within sample diversity of select *Prevotella* clinical isolates

To assess initial high-level diversity within the isolate collection, *cpn60* gene PCR and sequencing were carried out on a subset of isolates from which DNA was available at the time (Figure 2.4 and 2.5). Additionally, *cpn60* gene sequences from whole genome assemblies generated later in the project (Chapter 3) were also included in this analysis. Isolates were selected for this analysis to represent both between-sample and within-sample variation. *Prevotella bivia* isolates from the same sample (i.e., same individual and same visit), were largely identical based on the *cpn60* UT region, with no single nucleotide polymorphisms (SNPs). However, while similar, there were 4 SNPs between the two isolates from UC071_V1. Isolates from a single individual at different visit times were identical in 1/3 cases (UC055_V3 and UC055_V2). However, the isolates from the other two samples (UC107 and UC040) differed between visits, with 6 and 8 SNPs, respectively. Some isolates from different individuals, UC096_V1_3 and UC055_V2_1,

UC109_V3 and UC096_V1, and UC071_V1_1 and UC095_V2, were highly similar, with 1 and 0 SNPs between them, respectively, and there were also cases where sequences from the NCBI RefSeq database were identical to *cpn60* sequences of South African isolates (GCF_001546565.2 and GCF_001065995.1). Overall, *P. bivia cpn60* sequences obtained from the cpnDB and NCBI RefSeq database, which were also of vaginal origin, did not cluster distinctly from the local isolates.

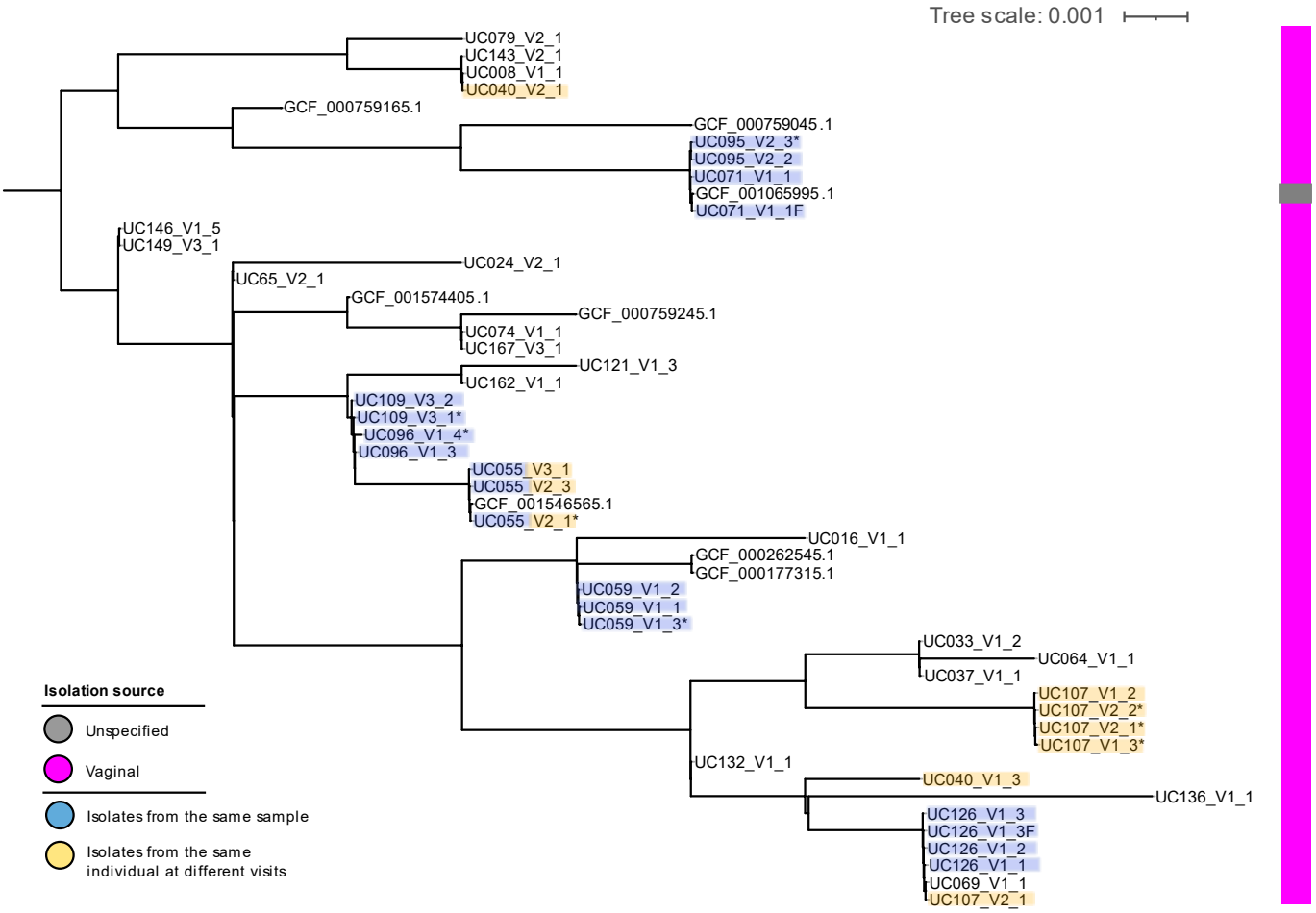


Figure 2.4*:* **Maximum likelihood tree of *P. bivia* clinical isolates.** The evolutionary relationship of isolates was inferred based on the *cpn60* gene using RAxML and the General Time Reversible model using 500 bootstrap replicates. The isolation source of the strains is indicated. The tree was generated using PCR-amplified *cpn60* genes (*), and the *cpn60* UT region from *P. bivia* genome assemblies and NCBI *P. bivia* reference strains (GCF_). The scale bar shows a phylogenetic distance of 0.001 nucleotide substitutions per site.

The *cpn60* gene from isolate UC143_V2_2, which was preliminarily identified as *P. melaninogenica*, was extracted and aligned to *P. melaninogenica* sequences obtained from the NCBI RefSeq database and cpnDB (Figure 2.5). These reference strains were isolated from oral, vaginal or lung (sputum) sources. The *cpn60* gene of UC143_V2_2 clustered with *P. melaninogenica* strain GAI 07411 (GCF_003609775.1), the only other *P melaninogenica* isolate in the reference databases from a vaginal source.

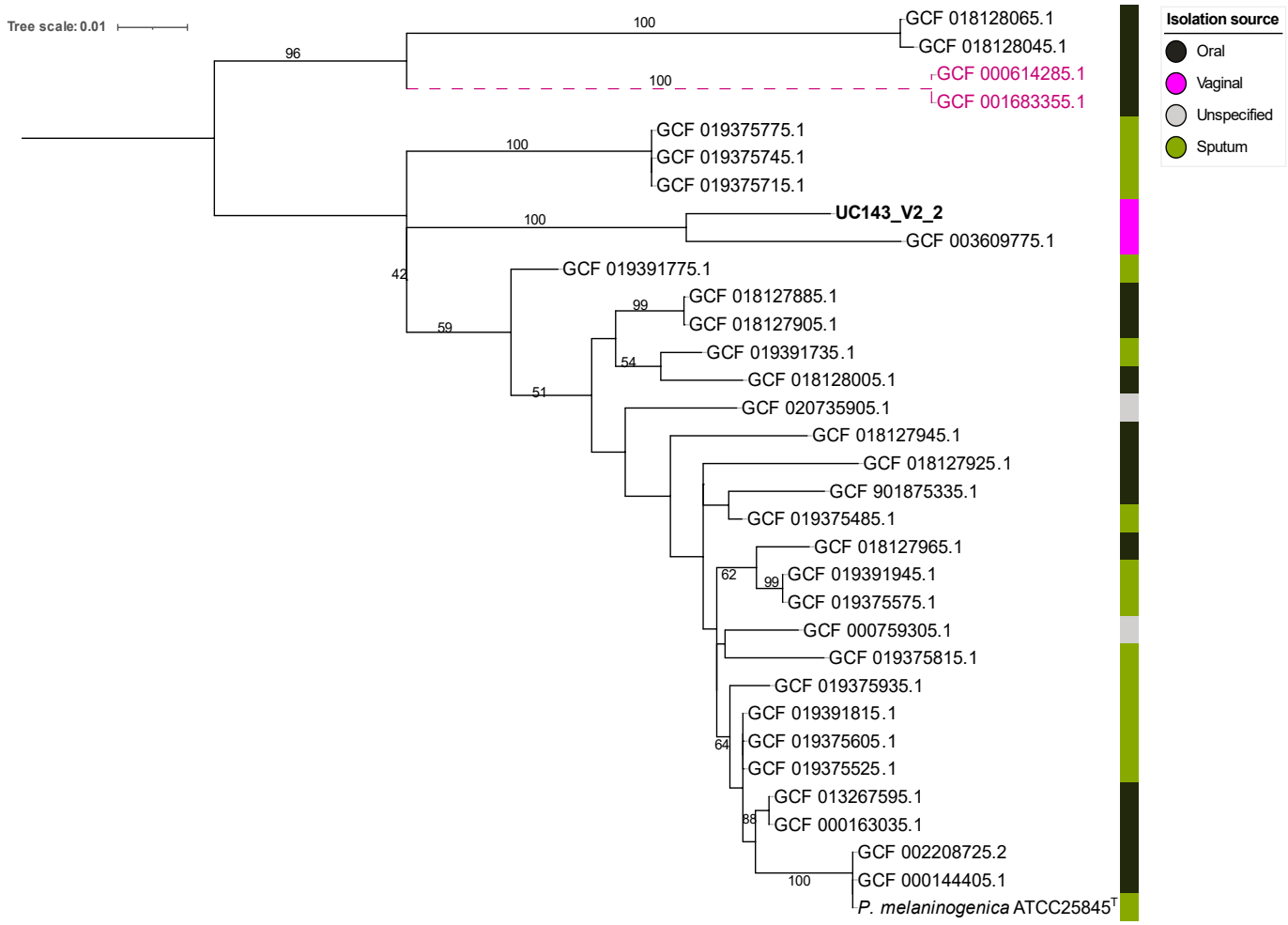


Figure 2.5: **Maximum likelihood tree of *P. melaninogenica* clinical isolates**. The *cpn60* gene of the putative *P. melaninogenica* isolate (n=1) was aligned with reference *P. melaninogenica* and *P. scopos* isolates. The evolutionary relationship of isolates was inferred based on the cpn60 gene using RAxML and the General Time Reversible model using 500 bootstrap replicates. The isolation source of the strains is indicated. The tree was generated using PCR-amplified *cpn60* genes (*), and the *cpn60* UT region from *P. bivia* genome assemblies and NCBI *P. bivia* reference strains (GCF_). The tree scale bar represents a phylogenetic distance of 0.01 nucleotide substitutions per site.

The 16S rRNA gene of the UC143_V2_2 isolate had a 97.67% sequence identity with the *P. melaninogenica* type strain, ATCC25845^{T}. UC143_V2_2 and GAI 07411 were clustered separately from most of the *P. melaninogenica* reference strains, which were derived from oral or sputum samples. Strains GCF_0181280651.1 and GCF_018128045.1, which were designated as *P. melaninogenica* on the NCBI RefSeq database, were clustered separately from the other *P. melaninogenica* strains and seemed to be more closely related to *P. scopos* based on the *cpn60* gene.

2.4.3 Selection of isolates for further characterisation

Following the initial isolation and screening, a final subset of 36 isolates was chosen for more detailed downstream analyses. Isolates were chosen to include at least one isolate per sample (n=26). Additionally, for three samples, a second isolate was also included (n=6) and for one sample, four isolates were selected (n=4). These latter isolates were chosen to assess broader genotypic and phenotypic variation within a sample that may be masked by the comparatively low resolution of the *cpn60* gene sequencing approach.

2.4.4 Susceptibility to SOC and general genitourinary antibiotics

To determine the antibiotic susceptibility of South African *Prevotella* isolates to SOC BV antibiotics as well as to other standard genitourinary antibiotics, the Etest methodology was used (Figure 2.6 and Table S1). Isolates were generally susceptible to metronidazole and clindamycin, with resistance rates of 2.7% (1/36) and 8.3% (3/36), respectively (Figure 2.6 A-C). A further 8.3% (3/36) of isolates showed intermediate susceptibility to metronidazole (MIC between 8-32 µg/ml). The susceptibility profiles of the clinical *Prevotella* strains to different genitourinary antibiotics showed higher rates of resistance for non-BV prescribed antibiotics, azithromycin, amoxicillin, and doxycycline (Figure 2.6 A, D-F), with 33.33% (12/36), 38.89% (14/36) and 19.44% (7/36) of isolates susceptible, respectively. Interestingly, isolates from the same sample sometimes showed phenotypic differences in susceptibility profiles. For example, one isolate from UC071_V1 was resistant to azithromycin, while a second isolate from the same sample was susceptible to the antibiotic.

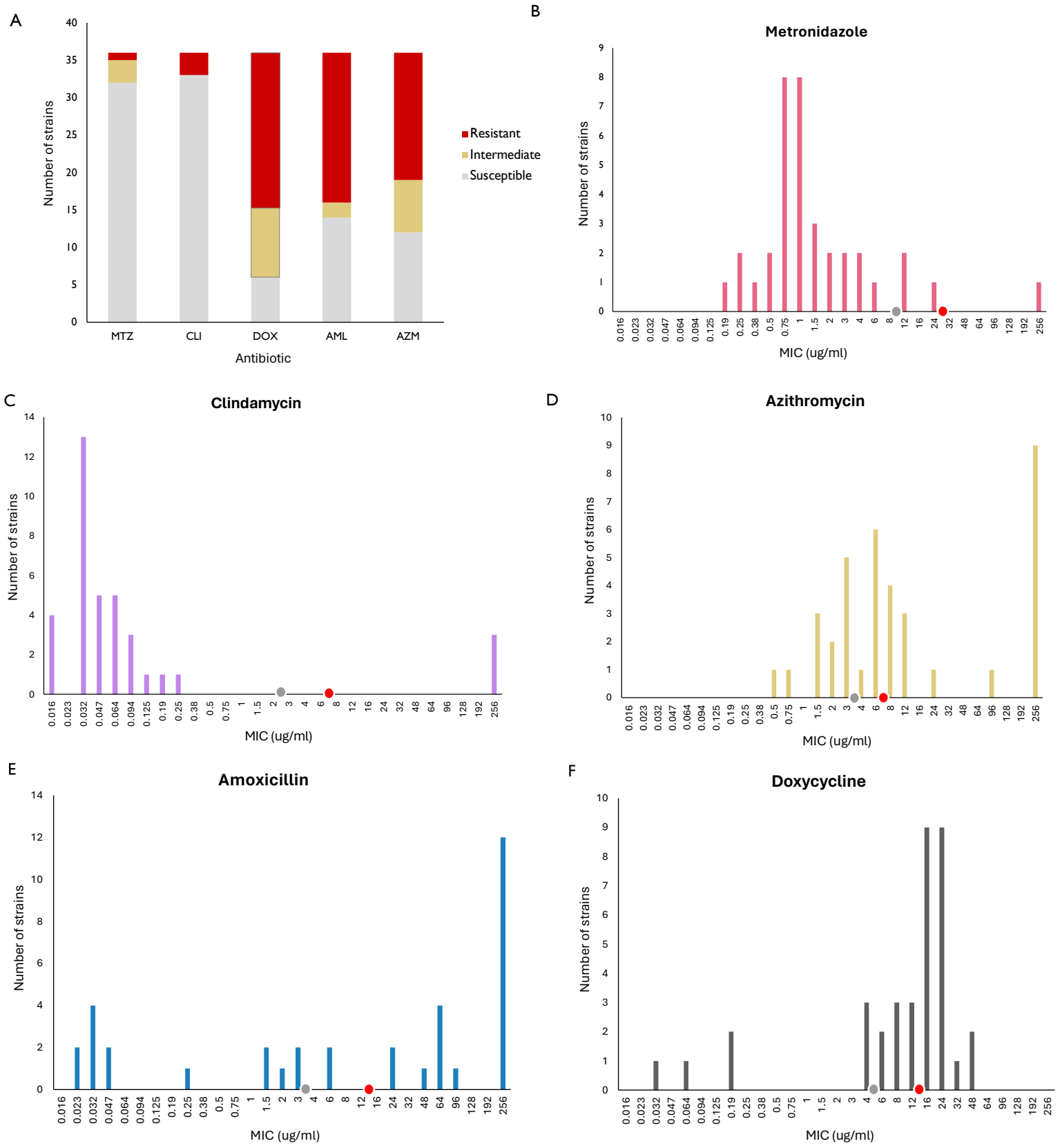


Figure 2.6: **Antimicrobial susceptibility testing to SOC BV and other genitourinary antibiotics**. (A) Antimicrobial susceptibility of clinical isolates (n=36) was tested by Etest methodology against five antibiotics: metronidazole (MTZ), clindamycin (CLN), doxycycline (DXT), amoxicillin (AML) and azithromycin (AZM). The AMR phenotype was classified based on the CLSI (CLSI, 2020) breakpoint guidelines (B-F). The MIC of BV SOC antibiotics to (B) metronidazole, (C) clindamycin, (D) azithromycin, (E) amoxicillin, and (F) doxycycline were plotted, with the intermediate and resistant cutoffs indicated on the x-axis by a grey and red dot, respectively.

Two isolates from UC143_V2_2 had very different phenotypes, with one isolate being resistant to azithromycin and doxycycline and having reduced susceptibility to amoxicillin, while the other isolate, the putative *P. melaninogenica* isolate, was susceptible to all tested antibiotics.

Resistance to multiple antibiotic classes was relatively common, with 36.11% (13/36) of isolates showing co-resistance to two antibiotic classes and a further 19.44% (7/36) resistant to three antibiotic classes. Two isolates, including the metronidazole resistant isolate, were resistant to four antibiotic classes.

2.4.5 Biofilm forming ability of clinical *Prevotella* isolates

To assess the ability of the isolates to produce biofilms *in vitro*, crystal violet staining and quantification of 5-day biofilms was performed. Formation of biofilms was variable by isolate (Figure 2.7). Just over a quarter of isolates (10/36, 27.8%) formed significantly larger biofilms than the *P. bivia* ATCC 29303^T type strain. There was no significant difference in biofilm formation between isolates from participants with BV (Nugent score ≥ 7) and those with BV-negative or BV-intermediate scores ($p>0.05$). In some cases, there were significant differences for isolates from the same sample. For example, UC059_V1_2 and UC126_V1_1 formed significantly larger biofilms than UC059_V1_1 and UC126_V1_3, respectively ($p<0.05$).

To determine the impact of the addition of *P. bivia* UC109_V3_2 and *G. vaginalis* ATCC 14018^T on biofilm formation *in vitro*, co-culture biofilm assays were done. For this preliminary analysis, early-stage (24-hour) biofilms of each single species were prepared, after which planktonic cells of the second species were added and the biofilm allowed to incubate for a further 24 hours before staining of the biofilm mass and colony counts. By visual inspection, the *G. vaginalis* monoculture biofilms appeared to be larger than those formed by *P. bivia*, however this difference was not statistically significant upon crystal violet quantification (Figure 2.8A).

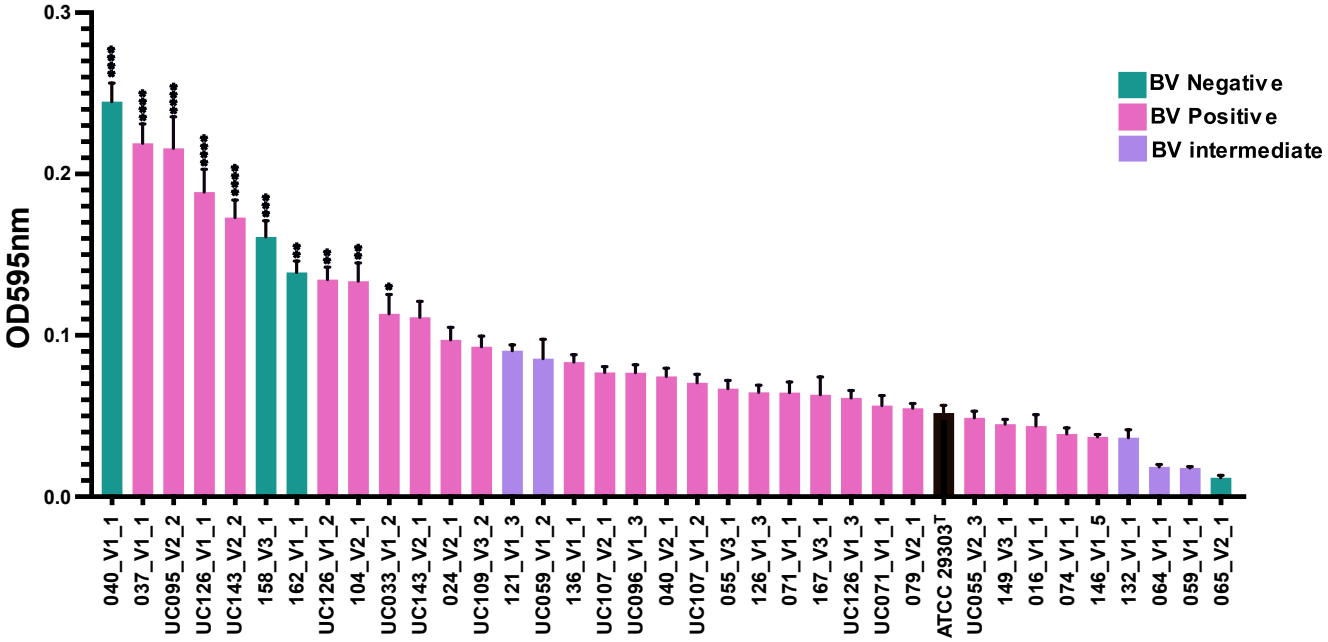


Figure 2.7: **Biofilm biomass of clinical isolates.** After incubation for five days under anaerobic conditions, crystal violet stained biofilms were quantified by measuring absorbance at 595nm. The mean absorbance was calculated from three biological repeats, with error bars representing standard error. Those isolates that differed significantly from the type-strain (*P. bivia* ATCC 29303^T) are indicated (Dunn's multiple comparison-corrected p-values: *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001).

Additionally, bacterial plate counts from each strain showed no significant difference in the number of viable bacteria in the monoculture biofilms. Addition of *P. bivia* to pre-established *G. vaginalis* biofilms did not enhance overall biofilm formation with similar biofilm masses observed for monoculture vs. co-culture biofilms. *Gardnerella vaginalis* numbers in monoculture vs. co-culture were also not statistically different. Based on bacterial counts, *P. bivia* did not appear to incorporate into the pre-formed *G. vaginalis* biofilm at all. When *G. vaginalis* was added to pre-formed *P. bivia* biofilms, *G. vaginalis* appeared to be abundant in this biofilm, while *P. bivia* numbers remained the same and there was no difference in the overall biofilm mass.

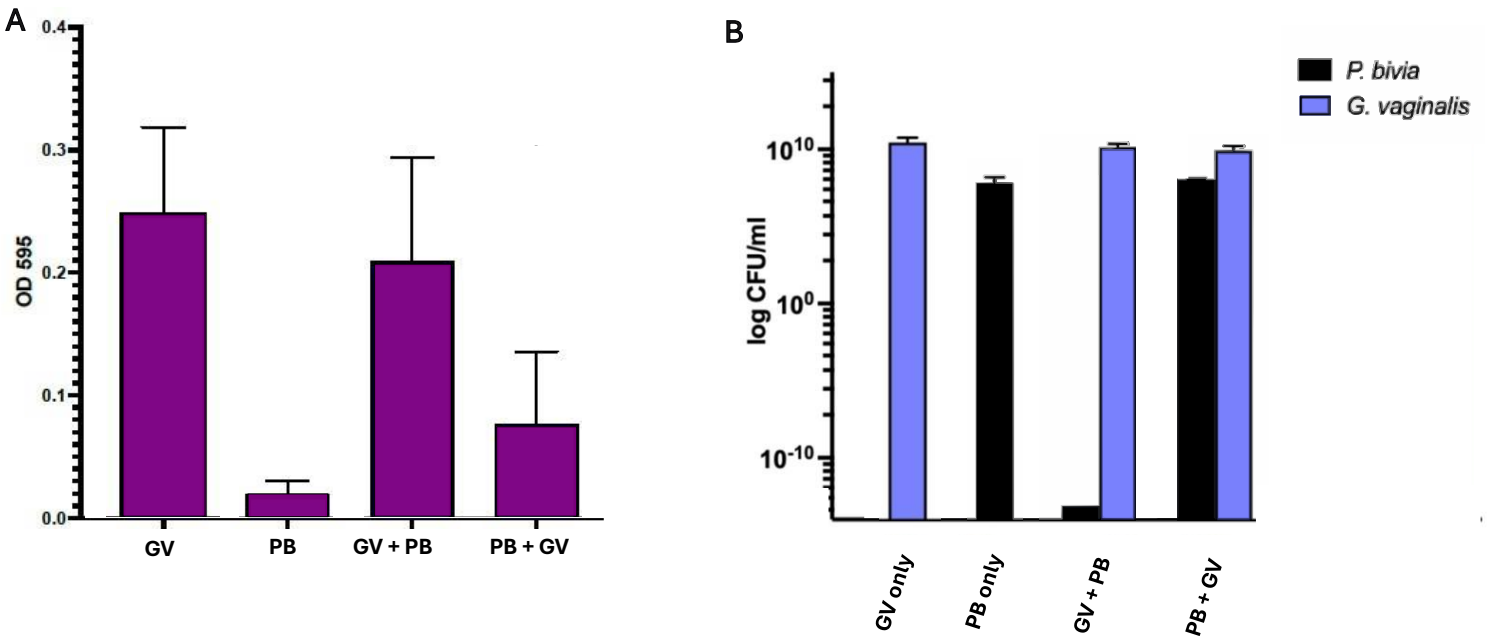


Figure 2.8: **Co-culture *P. bivia* and *G. vaginalis* biofilms**. (A) Monoculture and co-culture biofilms consisting of *P. bivia* (PB) and/or *G. vaginalis* (GV) were stained by crystal violet, followed by spectrophotometric quantification to infer biofilm biomass. No statistical significance was determined (Kruskal-Wallis one way ANOVA) (B) Colony counts (CFU/ml) of viable *P. bivia* and *G. vaginalis* are shown, represented on a logarithmic scale. No statistically significant difference was observed between monoculture and co-culture cell counts (2-way ANOVA with multiple comparisons, $p>0.05$).

2.5 Discussion

An increased abundance of diverse anaerobic bacteria, including *Prevotella* spp., in reproductive age women is associated with clinical diagnosis of BV (Zozaya-Hinchliffe et al., 2010). This condition is widespread worldwide, but is particularly common in Sub-Saharan Africa, with a prevalence of 42.1% reported in a meta-analysis by Torrone et al., and a prevalence of over 50% reported in another study (Kenyon et al., 2013; Torrone et al., 2018). The exact reasons for these disproportionately high rates in Sub-Saharan African are not well understood. Not only are vaginal *Prevotella* species understudied in comparison to their gut, oral and respiratory counterparts, but the relative lack of characterised *Prevotella* species from Africa compared to the global North also compounds the issues facing BV research in South Africa. *Prevotella* spp. have been shown to exhibit intraspecies diversity (Ma et al., 2020) and the lack of local *Prevotella* representatives limits regionally relevant research into the aetiology and pathogenesis of BV in South African women. To conduct locally relevant research and make

geographically appropriate progress in BV research, having a comprehensively characterised collection of South African vaginal *Prevotella* strains is important.

Availability of 16S rRNA gene data from the UChoose cohort (Gill et al., 2020) allowed for identification of which samples from the study contained *Prevotella* spp. The most common *Prevotella* species identified in these vaginal samples were *P. amnii, P. timonensis* and *P. bivia*. *Prevotella bivia* and its role in bacterial vaginosis has been described in Chapter 1, but *P. amnii* and *P. timonensis* have also been studied in the context of vaginal communities. *Prevotella amnii*, the most closely related *Prevotella* species to *P. bivia*, is a BV-associated organism and its abundance is associated with an elevated vaginal pH (Carter et al., 2022; Gottschick et al., 2017). *Prevotella amnii* has also been reported to increase susceptibility to PID and has been associated with PTB (Fettweis et al., 2019; Haggerty et al., 2020). *Prevotella timonensis* has likewise been significantly associated with BV (Srinivasan et al., 2012). Multiple studies have suggested an interaction between *P. timonensis* and host cells, with one study finding that exposure of vaginal Langerhans cells (LCs) to the bacterium causes increased uptake of HIV-1, turning normally protective LCs into reservoirs that increase transmission of HIV-1 (van Teijlingen et al., 2022). Another study by the same authors identified that dendritic cell maturation and inflammatory cytokine production strongly associated with the presence of *P. timonensis*, contributing to genital inflammation and thus some of the adverse sequelae associated with BV (van Teijlingen et al., 2020). In 3D endometrial cell models, *P. timonensis* also led to structural reorganisation of the cell surfaces, causing elongation of endometrial microvilli, which can enable improved bacterial adhesion to the cell surface (Ilhan et al., 2020).

Isolation and preliminary identification of *Prevotella* spp. isolates in the current project was achieved using a combination of a selective growth medium (LKV, selective for anaerobic Gram-negative bacteria) and colony PCR with a screening primer set [*Prevotella* genus-specific primers (Matsuki et al., 2004)]. This approach is relatively inexpensive and rapidly differentiates *Prevotella* spp. from off-target bacteria. Confirmed *Prevotella* spp. isolates can then be further classified using more expensive methods such as Sanger sequencing of the 16S rRNA gene. Interestingly, almost all bacteria isolated were identified as *P. bivia*, despite there being molecular evidence of significant

numbers of *P. amnii* and *P. timonensis* in the same samples, despite all three members of the genus being capable of growing on Brucella blood isolation medium (Srinivasan et al., 2016). It is possible that *P. bivia* cells are more resistant to stresses encountered during sampling and storage, and, therefore, more likely to be isolated in subsequent experiments. Indeed, it is important to note that samples used in the current study were convenience samples and were used for other analyses prior to the culture experiments. While all possible precautions were taken, it is possible that samples were compromised by freeze-thaw and oxygen exposure during processing for other analyses. Swabs were also not collected in pre-reduced media or stored in a reduced environment (there was no sodium thioglycolate or other reducing agents added). A combination of these factors may have contributed to the decreased viability of the more sensitive strains. Future studies in our group will collect dedicated culture samples directly into appropriate storage media such as sodium thioglycolate Eswab medium to optimise viability of oxygen sensitive microbes. Nevertheless, we were able to isolate 69 strains from 31 samples, to start a biorepository of South African *Prevotella* spp.

In addition to *P. bivia*, two isolates preliminarily identified as *P. melaninogenica* were obtained from one sample. *Prevotella melaninogenica* is the type species of the genus *Prevotella* and an early coloniser of the oral microbiome within the first months of life. It can persist in the mouths of those with no oral disease, but can also be an opportunistic pathogen (Könönen et al., 1999; Mager et al., 2003). It has also been identified as a key species in the cystic fibrosis lung, where it is high in abundance (Lamoureux et al., 2021). It is also often detected in culture-independent analyses of the vaginal microbiota. A handful of studies have investigated if there is a link between the oral microbiome and vaginal microbiome, with a small but significant relationship reported, that particularly involves *P. disiens* and *P. bivia* (Balle, Esra, et al., 2020; Persson et al., 2009; Zabor et al., 2010). This link may be due to host factors favouring *Prevotella* growth in both niches, but whether oral bacteria can be introduced, and survive, in the vagina remains a question. It is relatively rare, but not unheard of, to isolate viable *Prevotella melaninogenica* from vaginal samples (Duerden, 1993; Srinivasan et al., 2016). Interestingly, both isolates from the current study failed to produce the black pigment characteristic of members of this species, even after extended incubation. This, coupled with the ~97.6% 16S rRNA gene

sequence identity with the *P. melaninogenica* reference strain suggests they may be separate species or sub-species, however definitive assignment as a novel species or sub-species requires full genome sequence analysis.

Investigation of strain diversity for *Prevotella* spp. isolates in the collection using the *cpn60* gene did not suggest distinct differences between the South African isolates and those available in the reference databases (Vancuren & Hill, 2019). Whole genome sequencing (WGS) of the *Prevotella* spp. isolates allowed for extraction of the *cpn60* UT for additional isolates, which were not initially sequenced via *cpn60* PCR. The sequences obtained by PCR and by extraction from isolate genome assemblies were identical. Most isolates from the same sample were more similar to each other than to other isolates, suggesting within-sample diversity is low at the *cpn60* level. One exception was the two isolates from UC071_V1, isolated from the same visit, which had some nucleotide differences, reflecting that there can be some intrasample diversity in *Prevotella* spp. Most isolates from different sampling times showed diversity on a *cpn60* level. However, one pair were identical, despite the 8-week duration between visits (UC055). This individual was classified as BV positive, with a Nugent score of 8, at both time points, but whether persistence of the same *P. bivia cpn60* sequence type was associated with the persistent BV diagnosis would be interesting future work, where longitudinal metagenomic data would be able to show microbial dynamics over time for several samples and if there is an association of certain strains with BV positivity.

The putative *P. melaninogenica* isolate, UC143_V2_2, clustered with the only other vaginally derived *P. melaninogenica* strain in the *cpn60* phylogeny, with high bootstrap support and a relatively long branch length, suggesting that the vaginal strains are distinct from the oral- and sputum-derived strains, and not oral contaminants. To determine whether these strains are divergent enough from the rest of the *P. melaninogenica* strains to be classified as a different species, other methods could be employed. DNA-DNA hybridisation (DDH) has been regarded as the gold-standard method for determining species relatedness, with a 70% similarity cutoff to define a new species (Wayne et al., 1987). However, the time consuming and labour-intensive nature of this method, which requires availability of type strains, makes the use of sequence-derived parameters, such

as average nucleotide identity (ANI), which relies on *in silico* sequence comparisons, more attractive.

Phenotypic characterisation of newly isolated strains is important to provide insight into the pathogenesis of the strains and how they may respond to antibiotic treatment. Recurrence of BV post-antibiotic treatment is a major problem, with BV recurring in more than half of women within 12 months of initial clearance (Bradshaw et al., 2006). One study in South Africa showed recurrence rates of 42% and 30% in women living with and without HIV, respectively (Myer et al., 2006). The exact reasons why recurrence is so common are not clear, but one hypothesis is ineffective clearance of BV-associated bacteria by SOC BV antibiotics, potentially due to emergence of AMR in species such as *P. bivia* (Muzny & Sobel, 2022). Of the South African *Prevotella* isolates tested, almost all were susceptible to BV SOC antibiotics. For clindamycin, numerous studies have reported that approximately one third of isolated *Prevotella* spp. were resistant to clindamycin, including *Prevotella* spp. isolated from endometrial swabs (Petrina et al., 2017; Ulger Toprak et al., 2018). In some settings higher prevalences of clindamycin resistant *Prevotella* spp. have been observed. For example, in Ireland and Kuwait, studies have reported clindamycin resistance in 66.9% and 89.2% of isolates, respectively (Ali et al., 2022; Veloo et al., 2020). In all these studies, the *Prevotella* resistance rates to clindamycin were markedly higher than reported here (<10%). However, there have been reports of an increase in clindamycin resistance rates following therapy with the antibiotic. A study in 2005, looking at microbiological response to BV treatment, reported that 51% and 68% of isolates of *P. bivia* and black-pigmented *Prevotella* spp., respectively, were resistant to clindamycin following treatment with clindamycin, suggesting a selection of clindamycin-resistant strains following therapy (Austin et al., 2005). An additional concern relates to the underlying genetic mechanism of resistance. Clindamycin resistance in *Prevotella* spp. is often conferred by the carriage of *erm* genes and these are typically harboured on mobile genetic elements, raising the risk of spread by horizontal gene transfer. Since *erm* genes can encode cross resistance to other related antibiotics, such as azithromycin, it is possible that use of these antibiotics for non BV-applications may indirectly impact clindamycin efficacy in BV treatment plans by selecting for resistant isolates (Leclercq, 2002). Although exposure to antibiotics was

comparatively rare for participants enrolled in the parent study, five participants did receive azithromycin during the 8 weeks prior to sampling. Isolates from these participants were generally resistant to the antibiotic, although two azithromycin-susceptible isolates were observed. Two participants received metronidazole, but no metronidazole-resistant isolates were obtained from their samples. However, the resistance mechanisms in these isolates will need to be investigated further.

Clindamycin resistance is problematic amongst anaerobic bacteria, and thus not the preferred first choice for treatment of anaerobic infections. While studies find it as effective as metronidazole to treat BV, metronidazole is often regarded as the preferred drug for treatment, with the 7-day metronidazole course resulting in 80-90% efficacy in BV clearance (Koumans et al., 2002; Löfmark et al., 2010). Certain clades of *Gardnerella* spp. and *F. vaginae* may display intrinsic resistance to metronidazole, but *Prevotella* species have relatively low reports of resistance to date (Ferris et al., 2004; Schuyler et al., 2016). For example, Germany and Spain, which have among the highest reports of metronidazole resistance worldwide, with only 12% and 7% resistance rates, respectively (Cobo et al., 2019; Wolf & Stingu, 2022). However, despite these being 'high' resistance regions, the resistance to metronidazole is still comparatively lower than clindamycin resistance rates. Isolates in the current project were mostly susceptible to metronidazole, with only one isolate exhibiting resistance to the antibiotic, and three with an intermediate, reduced-susceptibility phenotype.

Other antibiotics commonly prescribed for urogenital infections include doxycycline, for chlamydia, gonorrhoea, *M. genitalium* and syphilis (Luetkemeyer et al., 2023), amoxicillin, for treatment of urinary tract infections (UTIs) (Tan & Chlebicki, 2016), and azithromycin, for chlamydia and gonorrhoeal infections (Peters et al., 2022). While not prescribed for BV directly, these antimicrobials may inadvertently impact the vaginal microbiota due to their use as broad-spectrum antimicrobials, and in the case of doxycycline, their increasing use as prophylaxis for sexual infections (Grant et al., 2020; Tamarelle et al., 2023). In the case of doxycycline, studies looking at endometrial and oral *Prevotella* isolates showed that 23% and 4.8%, respectively, were resistant to doxycycline (Petrina et al., 2019; Xie et al., 2014). The fact that more than two thirds of South African isolates were resistant to doxycycline is concerning. Doxycycline pre-

exposure prophylaxis (Doxy-PrEP) to prevent STIs has been suggested for some at-risk populations in South Africa and this may cause an increase in the acquisition and spread of resistance. The primary mechanism of resistance for tetracycline antibiotics is the presence of *tet* genes, which are often located on plasmids or transposons and therefore run the risk of horizontal gene transfer (Chopra & Roberts, 2001). So, while this antibiotic may not be prescribed for clearance of *P. bivia* or other BV-associated bacteria, the risk of transmission of AMR genes to other vaginal species, such as *N. gonorrhoeae* may be of concern.

Another hypothesised reason for the frequent recurrence of BV, and the inability of treatment to clear BV long-term, is the ability of the BV-associated strains to exist in protective biofilms on the vaginal epithelium, potentially protecting bacterial cells from killing by antimicrobials (Machado et al., 2016; Swidsinski et al., 2005). Biofilms are regarded as a hallmark feature of BV and *G. vaginalis*, proposed to be the primary biofilm-forming species involved in BV onset, is less susceptible to metronidazole and clindamycin when existing in a biofilm compared to when existing planktonically (Li et al., 2020; Swidsinski et al., 2008). While not classically regarded as primary biofilm-formers in BV, *Prevotella* species have been implicated in biofilms in the oral cavity, where their presence is associated with periodontal disease (Albaghdadi et al., 2021). Some *Prevotella* isolates in this study were able to form biofilms, with certain isolates forming significantly larger biofilms than the *P. bivia* ATCC 29393^{T} type-strain. Interestingly, *P. bivia* has been shown to affect biofilm formation by other species, or may exist within the polymicrobial biofilm community (Machado et al., 2013; Muzny et al., 2019). The current conceptual model of BV biofilms is that *Gardnerella* spp. are the primary colonizer of the vagina, forming the biofilm scaffolding, and that *P. bivia* may incorporate into this biofilm and promote its production (Muzny et al., 2019). The symbiotic relationship between *P. bivia* and *Gardnerella* spp., characterised by an exchange of ammonia and amino acids between the species, allows for enhancement of each other's growth (Pybus & Onderdonk, 1997). Thus, a synergistic relationship between *G. vaginalis* and *P. bivia* may also enhance the formation of biofilm and the onset of BV. Compared to a clinical isolate of *P. bivia*, *G. vaginalis* type-strain 48-hour biofilms were visibly larger when cultured individually. When a South African *P. bivia* isolate was co-inoculated with the *G. vaginalis*

ATCC 14018^T type strain, *G. vaginalis* abundance was comparable to its abundance in mono-culture biofilms. This suggests that while the addition of *P. bivia* did not impede the biofilm formation by *G. vaginalis*, it also did not enhance its formation. Contrary to the proposed model, *P. bivia* was unable to incorporate into the pre-established *G. vaginalis* biofilm using the methodology and growth media employed in this study. The model was proposed based on experiments done in continuous culture for three, thus different experimental conditions may explain the differences. It would be interesting to assess in future if co-culture of *G. vaginalis* with the top clinical *P. bivia* biofilm former (UC040_V1_1) would have had a more profound biofilm-promoting impact, as well as if the co-culture of strains isolated from the same source, which may have co-adapted together, may better interact with one another.

Assessing the biofilm forming ability of *P. bivia in vitro* has several limitations. Firstly, the *in vitro* nature of the assay does not necessarily mimic the *in vivo* ability of the strains to adhere to the epithelial surface and establish a mature biofilm, particularly in the context of continual replacement of vaginal epithelial cells during normal menstrual cycling. Additionally, it is difficult to distinguish a fully formed biofilm, which includes important structural components such as extracellular DNA and a polysaccharide matrix, from simple cell adherence to the surface of the microtiter plate using crystal violet alone. Further characterisation of biofilm structures using scanning electron or fluorescence microscopy are useful to confirm biofilm formation in *in vitro* models (Wilson et al., 2017). Relative species abundance in the co-culture biofilm assay was determined using serial dilution of the disrupted biofilm on selective media. More quantitatively precise methods, such as use of peptide nucleic acid fluorescent in situ hybridization (PNA FISH) with species-specific probes, can improve determination of microbes involved in multi-species biofilms (Sousa et al., 2023). These methods were beyond the scope of the current study. Finally, the impact of co-inoculation of *P. bivia* and *G. vaginalis,* rather than addition to pre-formed biofilms, should be investigated

This chapter described the isolation and basic characterisation of a collection of *Prevotella* spp. from samples provided by South African adolescent girls and young women with and without BV. This biorepository of strains represents a resource to begin regionally relevant research into the strains involved in BV in the country. However,

several questions remain. While *cpn60* sequencing suggested a relatively low within-sample diversity, a higher resolution approach, such as core genome alignments, would allow a more detailed analysis of strain relationships. Additional genomic characterisation is also necessary to determine the taxonomic status of the putative *P. melaninogenica* isolate. Finally, the genetic basis for the various antibiotic resistance phenotypes is unknown. Therefore, 36 isolates were subjected to whole genome sequencing, and the results are presented in the next chapter.

Chapter 3: Whole genome sequencing and genotypic characterisation of *Prevotella* spp. isolates

3.1 Abstract

Whole genome sequencing (WGS) technologies have become increasingly available to interrogate bacterial genomes and can reveal subtle evolutionary differences between strains and identify clinically relevant genes. South African *Prevotella* spp. isolate genomes (n=36) were sequenced, and the evolutionary diversity of isolate genomes compared to reference *Prevotella* sequences was investigated. Assembled genomes were also screened for antimicrobial resistance genes, virulence genes, the presence of prophages and anti-phage defence mechanisms. Core genome alignment and pangenome analyses revealed that South African isolates were not distinct from non-SA strains. However, the UC143_V2_2 isolate (preliminarily described as *P. melaninogenica* in Chapter 2), had a smaller genome size than *P. melaninogenica* reference sequences and clustered separately in core genome phylogenetic trees. Further analysis using average nucleotide identity scores suggested that *P. melaninogenica* genome sequences in the NCBI RefSeq database represent up to four genomospecies, with UC143_V2_2 belonging to one separate genomospecies (ANI score <95%). Genes proposed to confer resistance to metronidazole and clindamycin, *nimK* and *ermF*, respectively, were identified in 2.27% (1/36) and 8.33% (3/36) of genome assemblies, with the *nimK* gene located on a previously described Tn6456 transposon. Additional AMR determinants were identified, including *tetQ* (tetracycline resistance), which in some isolates was present on a CTnDOT-like family transposon, and different variants of the β-lactamase gene, *cfxA*, which were differentially associated with amoxicillin susceptibility (p=0.00021). High quality viral complete candidate prophage sequences were identified in several *P. bivia* genomes (n=13). Finally, anti-phage defence mechanisms, including predicted CRISPR/Cas, abortive infection and restriction modification systems were prevalent across all *Prevotella* genomes.

3.2 Introduction

While low-resolution information regarding strain relatedness can be obtained from sequencing of conserved bacterial amplicons, such as the *cpn60* gene, these methods are limited in their ability to detect subtle evolutionary differences and to identify genes that may be clinically relevant. As such, another alternative to amplicon sequencing, is WGS. Use of WGS, as the name implies, uses the entire genome of the bacterium, and diversity is assessed not only based on a single gene, but rather the genome in its entirety. Thus, WGS provides a more accurate measure of inter-strain diversity, reflecting differences that 16S rRNA or *cpn60* gene sequencing cannot (Brumfield et al., 2020).

Another advantage of WGS is the ability to detect and track certain strains or mutations of clinical relevance in a population, allowing for monitoring of evolution and transmission of a bacterium of interest. This was seen to be vital information during the COVID-19 pandemic for SARS-CoV-2 (Lorenzo-Redondo et al., 2021), but has also been used to investigate bacterial strains. Bacterial pangenomes (i.e. the entire gene repertoire) can be interrogated, providing insight into specific genes that may play a role in disease severity and complicate treatment (Deurenberg et al., 2017). One example of the utility of WGS-based investigations for *Prevotella* spp. is the analysis of oral *Prevotella intermedia,* and the finding that disease-causing strains encode markedly different metabolic processes and more virulence factors compared to strains from 'non-diseased' individuals, identifying potential biomarkers or targets for oral disease (Zhang et al., 2017). This information also informs public health agencies on the prevalence of certain genes of interest in different populations, allowing for tracking of the spread of potential AMR determinants (Köser et al., 2014). However, to date, few studies have been published on fully sequenced *Prevotella bivia* strains, particularly from South African women.

Modern large scale sequencing projects, such as the National Institute of Health (NIH) Human Microbiome Project (NIH HMP Working Group et al., 2009) have provided reference genomes for many host-associated microbial species, including several *Prevotella* spp. The genome of *P. bivia* is approximately 2.5Mbp in length and has recently been proposed to consist of two circular chromosomes in at least some strains (Peng et

al., 2023). *Prevotella melaninogenica* has a larger genome, which also comprises two chromosomes, together totalling 3.1Mbp in length. While there are certain proposed virulence factors utilized by *P. bivia* for vaginal colonisation, such as sialidase production, a specific gene for this virulence factor has not yet been identified (Ferreira et al., 2021). *P. melaninogenica*, however, has defined virulence genes such as haemolysin-encoding *phyA* and the type IX secretion system gene *porK*, which aid in its colonisation of the oral cavity (Kondo et al., 2018; Okamoto et al., 1999). With regards to AMR, both *P. melaninogenica* and *P. bivia* have been reported to contain genes encoding different variants of the β-lactamase gene *cfxA*, as well as *ermF*, *tetQ* and *nim* genes, which may confer resistance to β-lactam-, macrolide-, tetracycline- and nitroimidazole-antibiotics, respectively (Castillo et al., 2022; Yokoyama et al., 2023).

Prevotella spp. are predicted to be the hosts of a class of megaphages, known as Lak phages, that have genomes of up to >540Kbp in size (Devoto et al., 2019). However, these have primarily been identified in gut ecosystems and very little is known about bacteriophages in vaginal strains. WGS provides the opportunity to interrogate bacterial genomes for the presence of integrated prophages to begin to understand the interaction between bacteriophages and their hosts in the vaginal ecosystem.

The aim of the work presented in Chapter 3 was to (1) interrogate the evolutionary relatedness and identity of South African strains using core- and pangenome-based approaches, to (2) identify antimicrobial resistance gene determinants and correlate their presence with the phenotypic antibiotic resistance data from Chapter 2, to (3) quantitate the carriage of putative virulence genes and to (4) identify integrated prophage sequences and the presence of proposed anti-phage resistance mechanisms. Strains were selected to represent the diversity of samples (at least one strain per participant) and also included multiple strains from single samples (to investigate within sample diversity), as well as strains isolated from the same individual at different times (to investigate longitudinal persistence).

3.3 Methods

3.3.1 Genomic DNA extraction and sequencing

Whole genome sequencing (WGS) was performed on 38 *Prevotella* isolates – some strains were selected to assess intra-sample diversity (n=10) and some to assess longitudinal diversity in the same individual at different visit times (n=6). The remaining 22 strains were isolated from individual samples. Genomic DNA was extracted from the *Prevotella* spp. isolates (n=38) using the PureLink™ Microbiome DNA Purification Kit (Invitrogen, Thermo Scientific Inc, MA, USA), following the advised protocol with final elution in 60µl of the elution buffer provided in the kit. The concentration and quality of the extracted DNA were determined using a Nanodrop 1000 Instrument (Thermo Scientific Inc, MA, USA) and the integrity of the DNA was confirmed in 1% (w/v) agarose gels by electrophoresis. Sequencing of the DNA was performed by the Centre for Proteomic and Genomic Research (CPGR, Observatory, Cape Town, South Africa) on the Illumina MiSeq Instrument, using the Illumina MiSeq Reagent Micro Kit v2, for 300 cycles.

3.3.2 Genome assembly and annotation

Following sequencing, FastQC v0.12.1 (Andrews, 2010) was used for a primary quality control check on the raw sequences and the results summarised with multiQC v1.17 (Ewels et al., 2016). Raw reads were trimmed using trim-galore v0.6.10 with a minimum Phred quality score cutoff of 15 in order to remove adapters and low quality sequences (Martin, 2011). SPAdes v13.15.5 (Prjibelski et al., 2020) with the '-isolate' option was used for genome assembly and QUAST v5.2 (Mikheenko et al., 2018) was used to assess the assembly metrics. Genome assemblies failing initial QC were removed from further analyses. CheckM (Parks et al., 2015) was used to estimate overall completeness and contamination of the assembled sequences. To annotate the assembled contigs, Prokka v1.14.6 (Seemann, 2014) was used, with *P. bivia* DSM 20514^T (accession NZ_JH660658) and *P. melaninogenica* ATCC 25845^T (accessions NC_014370 and NC_014371) reference sequences as trusted protein sets for the *P. bivia* and *P. melaninogenica* annotations, respectively.

3.3.3 Core- and pan-genome analysis of clinical *Prevotella* isolates

Core genome identification was performed using Panaroo v1.3.4 (Tonkin-Hill et al., 2020) in 'strict' mode ('-clean-mode strict') with the core genome threshold set at 1 (-core_threshold 1). Alignments based on concatenated core gene sequences were undertaken using PRANK v170427 (Löytynoja, 2014). Maximum likelihood trees based on the core genome alignments were generated using iq-Tree v2.2.2.6 (Minh et al., 2020) using the best fit partitioned models suggested by the program (pre-run using the '-m MFP' option). Trees were visualised and annotated using iTOL v6.8.1 (Letunic & Bork, 2007). The output of Panaroo ('gene-presence-absence') and iq-Tree phylogenetic tree were also used to visualise the pangenome of *P. bivia* isolates using Phandango v1.3.0 (Hadfield et al., 2018), followed by an analysis with Scoary v1.6.16 (Brynildsrud et al., 2016) to determine if any genes were specifically associated with South African versus non-South African isolates, or with BV positivity (Nugent score ≥ 7) versus BV negativity (Nugent score < 7).

3.3.4 Genetic distance between *Prevotella* isolates

To determine the genetic relatedness of the *Prevotella* isolates and those obtained from the NCBI RefSeq resource (O'Leary et al., 2016), the ANI pairwise distances were calculated with Pyani v0.2.x (Pritchard et al., 2015), using MUMmer3 (ANIm) to align the sequences (Richter & Rosselló-Móra, 2009). The output was used to generate a matrix of strain similarity scores.

3.3.5 Identification of antimicrobial resistance determinants and virulence genes in *Prevotella* isolates

Assembled isolate genomes were screened for the presence of genes associated with antimicrobial resistance using ABRicate v1.0.1 (Seemann, 2015), with multiple databases, including ARG-ANNOT (Gupta et al., 2014), The Comprehensive Antibiotic Resistance Database (CARD) (Alcock et al., 2023), NCBI Antimicrobial Resistance Gene Finder Plus (AMRFinderPlus) (Feldgarden et al., 2021) and Resfinder. To screen for potential known virulence factors, the Virulence Factor Database (VFDB) (Chen et al., 2005) was used. Any genes or gene clusters of interest not included in the databases of

the tools above were obtained from the NCBI Reference Gene catalogue, and manually searched for using BLAST against all the assembled genomes (Camacho et al., 2009; Feldgarden et al., 2021).

3.3.6 Identification of prophages and anti-phage resistance mechanisms

Potential integrated prophages were identified using Virsorter2 v2.2.4 (Guo, Bolduc, et al., 2021) using the suggested standard operating procedure (SOP) published by the creators of the program [Version 3 (Guo, Vik, et al., 2021)]. In short, potential prophage sequences were first identified by Virsorter2, with lenient criteria. This was followed with a quality control check using checkV v1.0.1 (Nayfach et al., 2021) to assign quality scores and trim host sequence from either end of each potential prophage. Thereafter, trimmed sequences were reassessed by a second round of Virsorter2 to prepare them for the DRAMv v1.4.6 pipeline, which was used to annotate the predicted viral sequences (Shaffer et al., 2020). Following the DRAMv annotation, manual quality control checks were used to remove sequences that may not be viral. In order to determine if any of the complete or high quality identified prophage sequences were similar to previously characterised phages, Dashing v1.0.2.4-g0635 (Baker & Langmead, 2019) was utilised, to determine Jaccard similarity scores to every phage in the NCBI RefSeq database for *Caudoviricetes* (downloaded 2023-07-17). Phage genomes were re-annotated with Pharokka v1.5.1 (Bouras et al., 2023), and input into clinker v0.0.28 (Gilchrist & Chooi, 2021) for genome comparisons. Anti-phage resistance mechanisms were identified from the annotated *Prevotella* spp. isolate genomes using DefenseFinder v1.2.0 (Abby et al., 2014; Tesson et al., 2022).

3.4 Results

3.4.1 Assembly metrics of *Prevotella* sequences obtained by whole genome sequencing

A total of 38 *Prevotella* spp., described in Chapter 2, were available for WGS. Of these, two assemblies failed initial quality assessment (UC104_V2_1 and UC158_V3_1) due to high levels of genome fragmentation (N50 values <2 Kbp) and potential contamination and/or mis-assembly (Table S2). These assemblies were excluded from further analyses. The median genome size for the remaining *P. bivia* isolates was approximately 2.5 Mbp (Table 3.1), which is within the range of genome sizes commonly reported for *P. bivia* isolates (Table S2). Conversely, the putative *P. melaninogenica* assembly size was slightly smaller than the genome sizes of *P. melaninogenica* isolates in the NCBI RefSeq database (2.9 Mbp versus 3.1 Mbp – 3.4 Mbp), but most close in genome size to the other vaginal *P. melaninogenica* isolate (Table S2). All isolate genomes had <41% GC content, which is typical of members of the phylum *Bacteroidota* (Teng et al., 2023). All genome N50 values ranged between 23 Kbp and 161 Kbp, with most strains having at least 50% if their genome content on contigs of >70 Kbp.

Table 3.1 **Summary assembly metrics from whole genome sequencing for *P. bivia* isolates (n=35) and *P. melaninogenica* (UC143_V2_2).**

	P. bivia*	***P. melaninogenica***
Genome size (bp)	2 491 510(2 361 601-2 698 246)	2904179
N50 (bp)	70 372 (22 779-117 289)	161629
L50	11 (7-31)	6
%GC	39.70 (39.20-39.98)	40.67
No. contigs	109 (63-464)	51
No. CDS	2 140 (1 987-2 333)	2307
No. tRNAs	50 (48-53)	48
Completeness (%)	89.19 (87.11-93.59)	88.61
Contamination (%)	3.96 (2.07-14.97)	2.45
Coding density (%)	86.73 (86.09-87.22)	82.25

* Median and range for n=35 *P. bivia* isolates

3.4.2 Core genome alignment of *Prevotella* spp. to reveal high-resolution strain relatedness

While sequencing of the *cpn60* gene is often used as a proxy for interrogating strain diversity, WGS enables the alignment of the entire core genome of the strains for a holistic view of inter-strain diversity that may not be reflected by differences in the *cpn60* gene. A maximum likelihood tree was generated from core genome alignments for all 35 *P. bivia* isolates, along with reference sequences obtained from the NCBI RefSeq database (Figure 3.1A). Strains of *P. bivia* isolated from the same sample were generally very closely related (1-180 core genome SNPs), although 1 pair of isolates (UC071_V1_1 and UC071_V1_1F) were markedly different, with 9193 core genome SNPs. Strains from the same individual but isolated at different study visits clustered together in one case (UC055_V3_1 and UC055_V2_3) and were similar at a core genome level (308 SNPs). However, in the remaining two cases, they clustered separately in the core genome tree and were more distantly related (8931 and 8962 core genome SNPs for the UC107 and UC040 isolates, respectively). Sequences obtained from the NCBI RefSeq database did not cluster separately from the South African isolates.

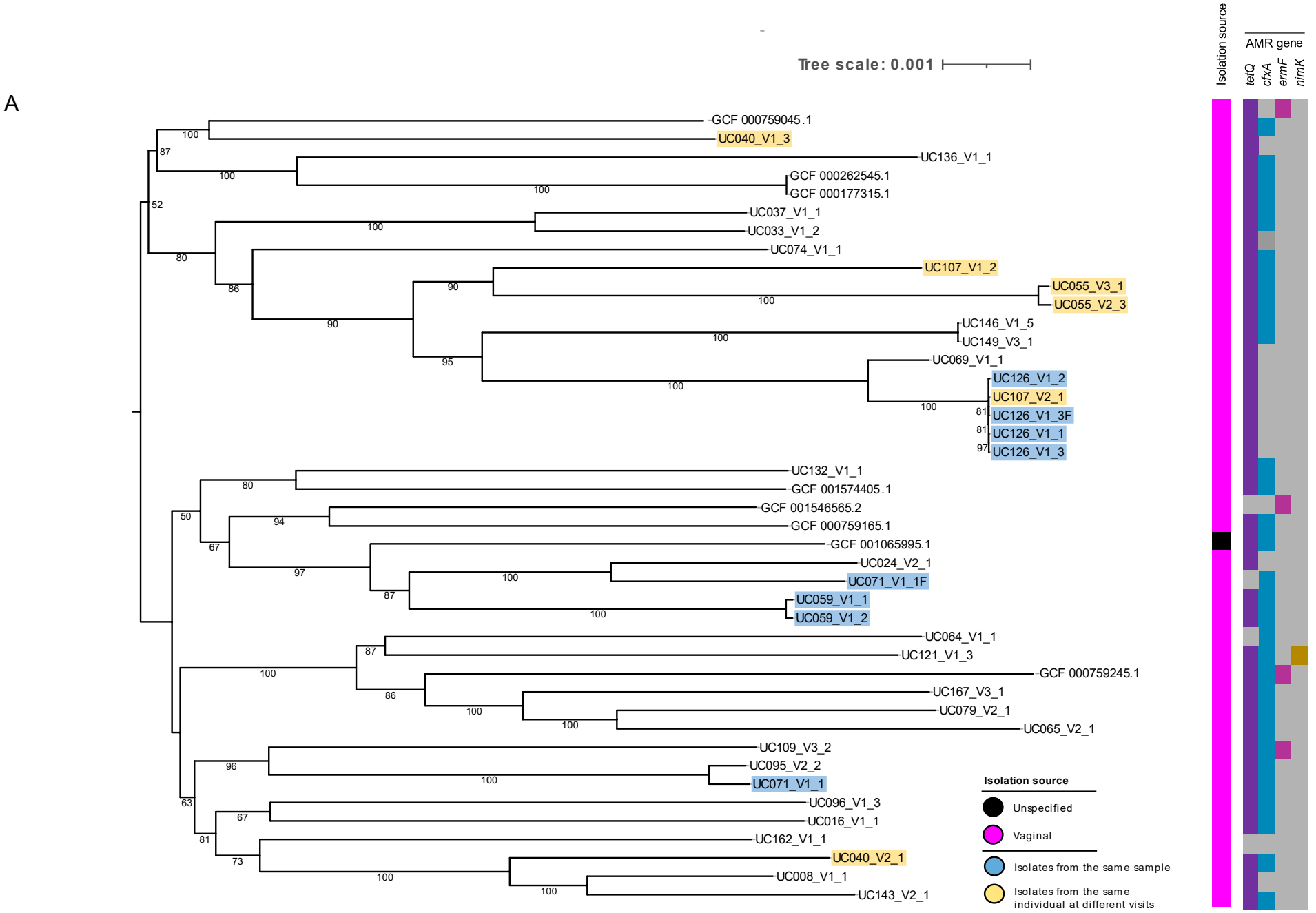
The putative *P. melaninogenica* isolate (UC143_V2_2) clustered separately from most of the *P. melaninogenica* reference sequences (Figure 3.1B). UC143_V2_2 did not cluster with GCF_03609775.1, the only other vaginal isolate. These two isolates were in fact distantly related, having 100 286 core genome SNPs. Additionally, two *P. melaninogenica* sequences obtained from the NCBI RefSeq database, GCF_018128045.1 and GCF_018128065.1, also clustered entirely separately from the other *P. melaninogenica* strains and, when *Prevotella scopos* isolates were included in the alignment, appeared more closely related to the latter (Figure 3.1B, for the original core genome tree without *P. scopos* sequences, see Figure S1).

3.4.3 Genetic distance and similarity between *Prevotella* isolates

To determine the whole-genome-based relatedness of the *Prevotella* isolates within their respective species-groups, ANI pairwise comparisons for all sequences were calculated (Figure 3.2). All *P. bivia* strains were highly similar to each other, with ANI scores ranging from 98.58%-99.99%, which is above the proposed 95% species identity cutoff and

suggests that all isolates have correctly assigned taxonomies (Figure 3.2A). The three strains from UC126_V1 had the highest ANI scores ranging between 99.994% and 99.999% and were also very similar to UC069_V1_1 and UC107_V2_1 (99.99% ANI score), supporting the results obtained by core genome alignments and suggesting that these isolates are very closely related. Similarly, the two isolates from participant UC107 (UC107_V2_1 and UC107_V1_2), which appeared to be more distantly related based on core genome alignments, had a pairwise ANI score of only 98.965%.

The *P. melaninogenica* as a group had much lower ANI pairwise values compared to *P. bivia* (Figure 3.2B). Based on the proposed <95% species cutoff, the *P. melaninogenica* strain group appeared to comprise up to four different genomospecies. The 'main' *P. melaninogenica* group contained the type-strain, ATCC 25845^{T} (GCF_000144405.1). The two strains with the lowest similarity to the core *P. melaninogenica* strains were GCF_018128045.1 and GCF_018128065 and were quite different to the main *P. melaninogenica* group (ANI range 88.52%-88.91%). The UC143_V2_2 isolate had ANI scores ranging from 91.48% to 91.76% compared to the main *P. melaninogenica* group and was most closely related to the other vaginal isolate, GCF_03609775.1 (ANI score 91.90%). Finally, the ANI scores of GCF_003609775.1 versus the main *P. melaninogenica* group ranged between 90.82% and 91.34%. The remaining sequences within the main group were all closely related to one another and to the *P. melaninogenica* type strain, *P. melaninogenica* ATCC 29845^{T} (ANI scores 96.13%-99.99%)


A
Tree scale: 0.001
Isolation source
AMR gene
tetQ
cfxA
ermF
nimK
GCF 000759045.1
UC040_V1_3
UC136_V1_1
GCF 000262545.1
GCF 000177315.1
UC037_V1_1
UC033_V1_2
UC074_V1_1
UC107_V1_2
UC055_V3_1
UC055_V2_3
UC146_V1_5
UC149_V3_1
UC069_V1_1
UC126_V1_2
UC107_V2_1
UC126_V1_3F
UC126_V1_1
UC126_V1_3
UC132_V1_1
GCF 001574405.1
GCF 001546565.2
GCF 000759165.1
GCF 001065995.1
UC024_V2_1
UC071_V1_1F
UC059_V1_1
UC059_V1_2
UC064_V1_1
UC121_V1_3
GCF 000759245.1
UC167_V3_1
UC079_V2_1
UC065_V2_1
UC109_V3_2
UC095_V2_2
UC071_V1_1
UC096_V1_3
UC016_V1_1
UC162_V1_1
UC040_V2_1
UC008_V1_1
UC143_V2_1
Isolation source
Unspecified
Vaginal
Isolates from the same sample
Isolates from the same individual at different visits

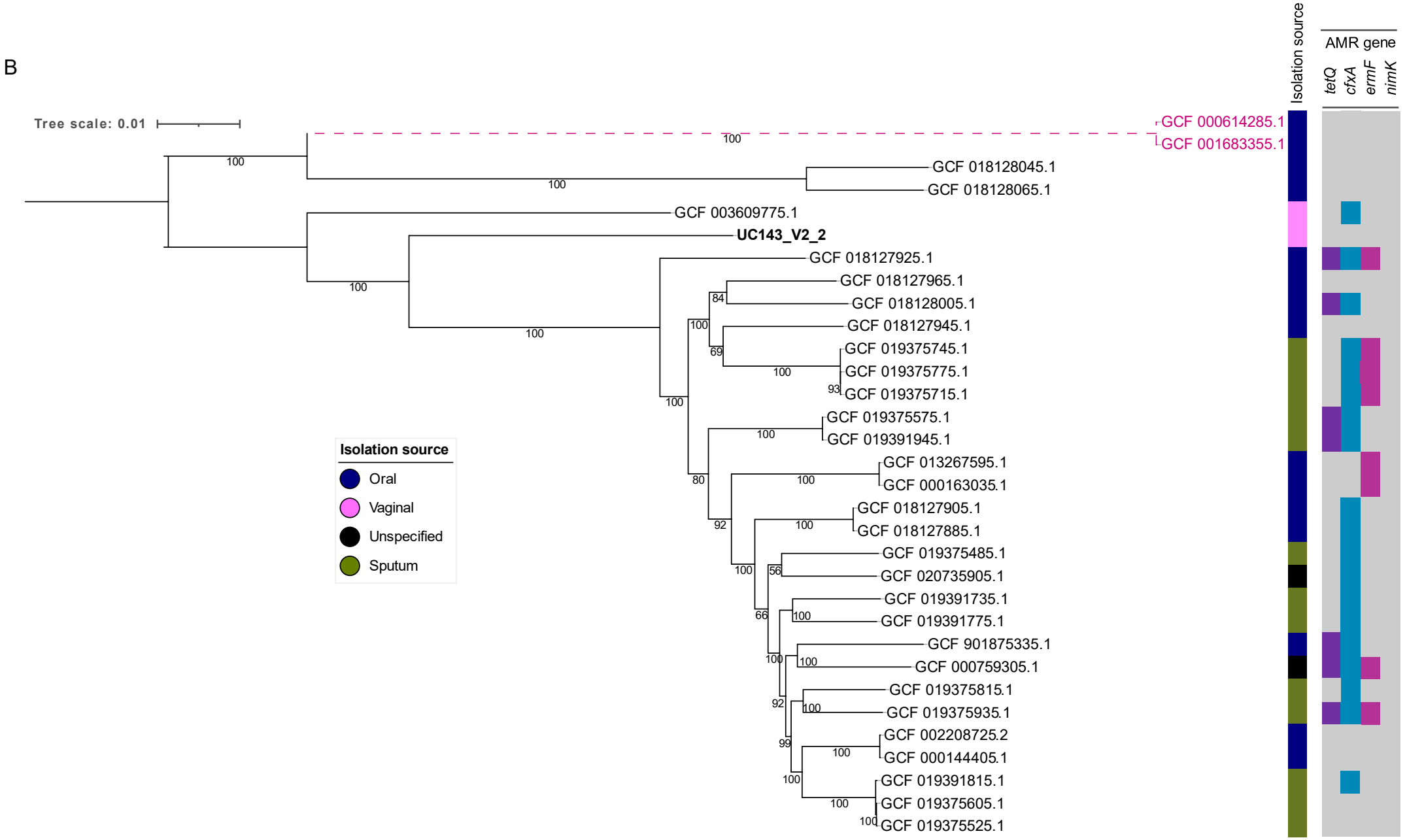


Figure 3.1: **Core genome alignment of *Prevotella* isolates.** *Prevotella* spp. whole genome sequences downloaded from the NCBI RefSeq database were aligned with South African isolates to generate maximum likelihood trees for (A) *P. bivia* and (B) *P. melaninogenica* and *P. scopos* (in pink). Bootstrap values of over 40 are displayed on the branches. Strains from the same sample or individual are coloured accordingly and isolation source is indicated. Presence or absence of four antimicrobial resistance determinant genes (*tetQ, cfxA, ermF* and *nimK*) is also indicated. The tree scale represents 0.01 nucleotide substitutions per site.

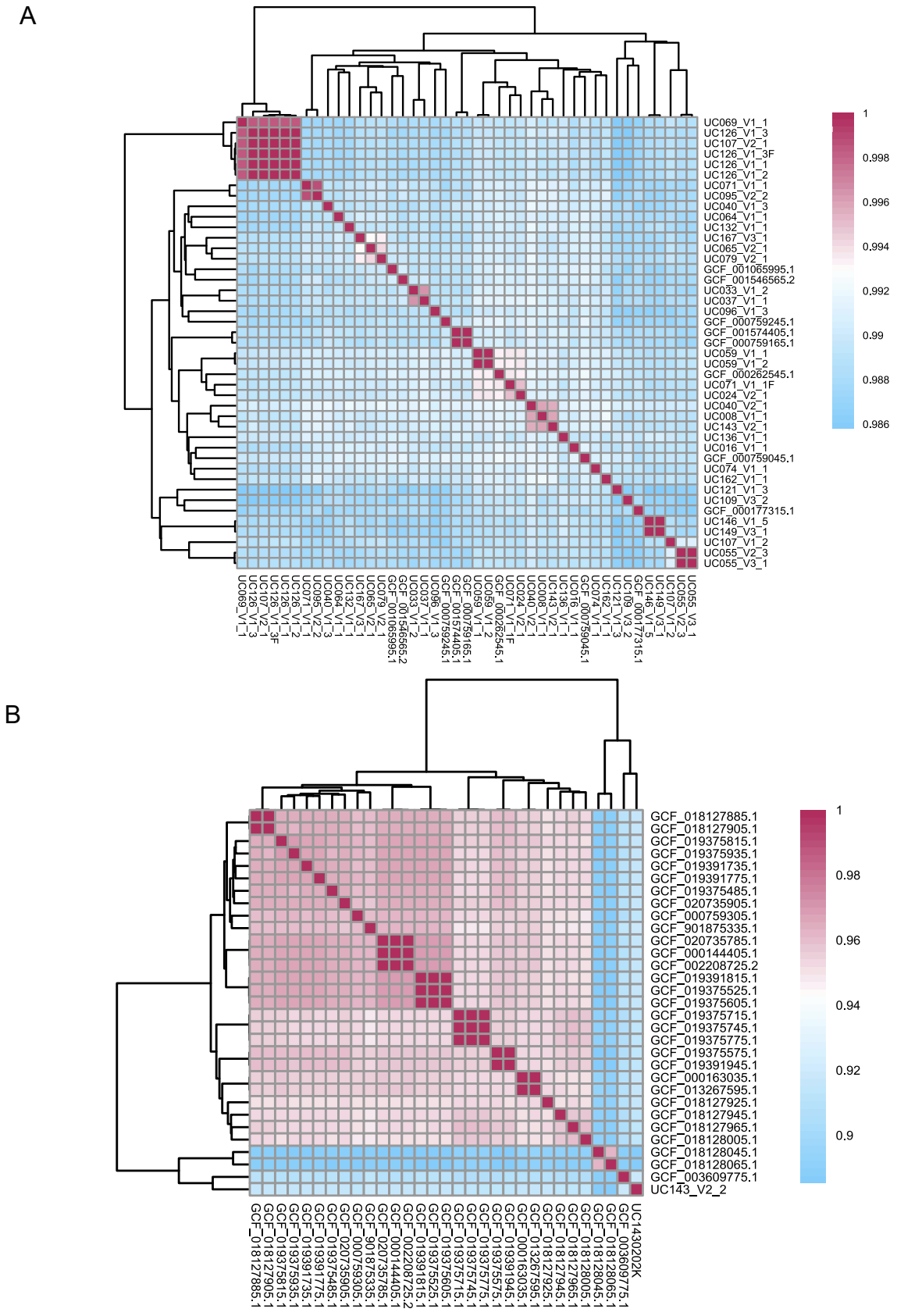


Figure 3.2*:* **Heatmaps of average nucleotide identity (ANI) of *Prevotella* isolates.** Pairwise ANI scores were calculated for (A) *P. bivia* and (B) *P. melaninogenica* isolate genomes and their respective NCBI RefSeq sequences and similarity was represented with heatmaps.

3.4.4 Pangenome analysis of *Prevotella* spp.

The *P. bivia* pangenome comprised a total of 4244 genes, of which a third (1434, 33.8%) were core genes (in 99%-100% of strains), 135 soft core genes (in 95%-99% of strains), and 2695 accessory genes (Figure 3.3A). Patterns of gene presence/absence were similar for strains that were closely related at a core genome level, suggesting that the accessory genomes were not significantly different in these strains. However, in at least one case, there were differences in the accessory, with an additional 40 genes present in UC055_V2_3 vs UC055_V3_1. This was mainly due to the absence of an 18.6 Kbp contig containing 22 genes in the latter isolate. While the majority of annotations for this contig were hypothetical, genes encoding proteins with homology to conjugal transfer proteins (*traG* and *traO*) commonly found on plasmids were present. A similar contig with high sequence identity was present in a distantly related isolate, UC079_V2_1, from a different study participant. BLAST comparisons to the NCBI nucleotide database revealed no significant matches for either potential plasmid, suggesting that they are novel. The *P. melaninogenica* group had a much larger pangenome (8615 genes), of which only 14.7% were core genes (1263 genes), along with soft core genes and 7137 accessory genes (Figure 3.3B). Removal of the four 'inconclusive' *P. melaninogenica* strains resulted in a somewhat reduced pangenome (7214 genes), of which 19.4% were core genes (1416).

No *P. bivia* genes were significantly associated with strain origin (SA vs non-SA) based on Bonferroni-corrected significance values in a Scoary pangenome-wide association analysis. Similarly, when stratifying strains by host BV status (BV positive [Nugent score >7] vs BV negative [Nugent score <7]), no genes were significantly associated with BV. However, all isolates from samples from BV negative women contained *asnA*, a predicted aspartate ammonia ligase gene, whereas this gene was only present in 35% of isolates from BV positive women (naïve $p < 0.0001$, Bonferroni-corrected $p = 0.474$).

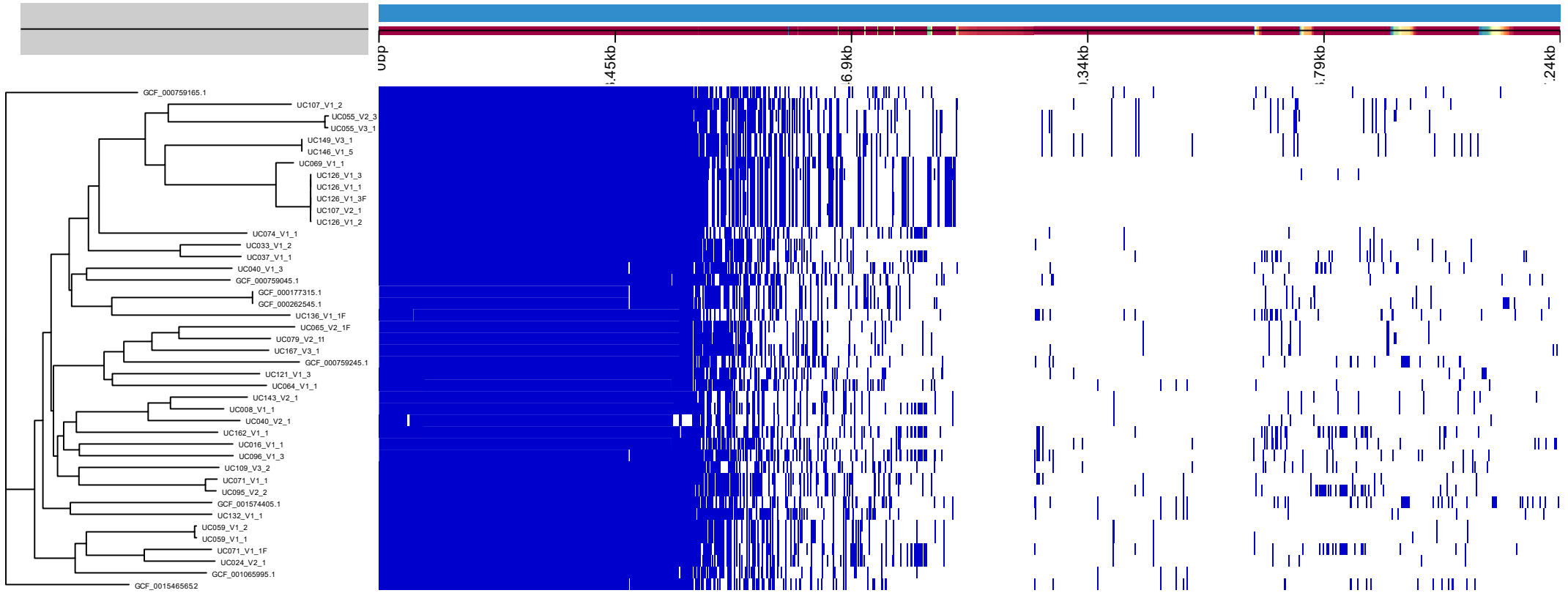
3.4.5 Antimicrobial resistance determinants present in the *Prevotella* spp. sequences

To determine the presence of genes potentially involved in antimicrobial resistance (AMR), annotated isolate and reference genomes were screened against various AMR databases (Seemann, 2015). Genes encoding potential resistance for tetracyclines

(*tetQ*), β-lactam antibiotics (*cfxA*) and lincosamides (*ermF)* were found in 90.91% (40/44), 70.45% (31/44) and 9.09% (4/44) of *P. bivia* strains, respectively (Figure 3.1A). The *P. melaninogenica* isolates had a lower frequency of *tetQ* and *cfxA* carriage, with 23.3% (7/30) and 63.33% (19/30) prevalence, respectively. The *ermF* gene was more commonly identified, present in 30% (9/30) of *P. melaninogenica* strains (Figure 3.1B). Dual carriage of *tetQ* and *cfxA* was seen in 63.63% (28/44) and 23.33% (7/30) of *P. bivia* and *P. melaninogenica* strains, respectively. A further six strains had three AMR genes concurrently (3 *P. melaninogenica* and 3 *P. bivia* strains). Five triple-AMR-gene strains contained *tetQ, cfxA* and *ermF* (Figure 3.1A and 3.1B).

The *tetQ* gene is known to be associated with transposable elements such as CTnDOT (Waters & Salyers, 2013). Therefore, the genomic location of the *tetQ* gene was investigated further for the presence of previously described transposable elements. Regions with homology to the CTn341 transposon (accession AY515263.1), previously identified in *B. fragilis* to contain *tetQ*, were identified in 65.90% (29/44) of *P. bivia* genomes (Figure 3.4). The *rteA-rteB-rteC* operon, part of the conjugative apparatus of the *Bacteroides* CTnDOT transposon, were identified with *tetQ* in all but 4 isolates. However, only 11/44 isolate genomes had co-carriage of this operon and the *mobA-C* genes, required for initiating mobilisation.

A


GCF_000759165.1
UC107_V1_2
UC055_V2_3
UC055_V3_1
UC149_V3_1
UC146_V1_5
UC069_V1_1
UC126_V1_3
UC126_V1_1
UC126_V1_3F
UC107_V2_1
UC126_V1_2
UC074_V1_1
UC033_V1_2
UC037_V1_1
UC040_V1_3
GCF_000759045.1
GCF_000177315.1
GCF_000262545.1
UC136_V1_1F
UC065_V2_1F
UC079_V2_11
UC167_V3_1
GCF_000759245.1
UC121_V1_3
UC064_V1_1
UC143_V2_1
UC008_V1_1
UC040_V2_1
UC162_V1_1
UC016_V1_1
UC096_V1_3
UC109_V3_2
UC071_V1_1
UC095_V2_2
GCF_001574405.1
UC132_V1_1
UC059_V1_2
UC059_V1_1
UC071_V1_1F
UC024_V2_1
GCF_001065995.1
GCF_001546565.2
6.9kb
24kb

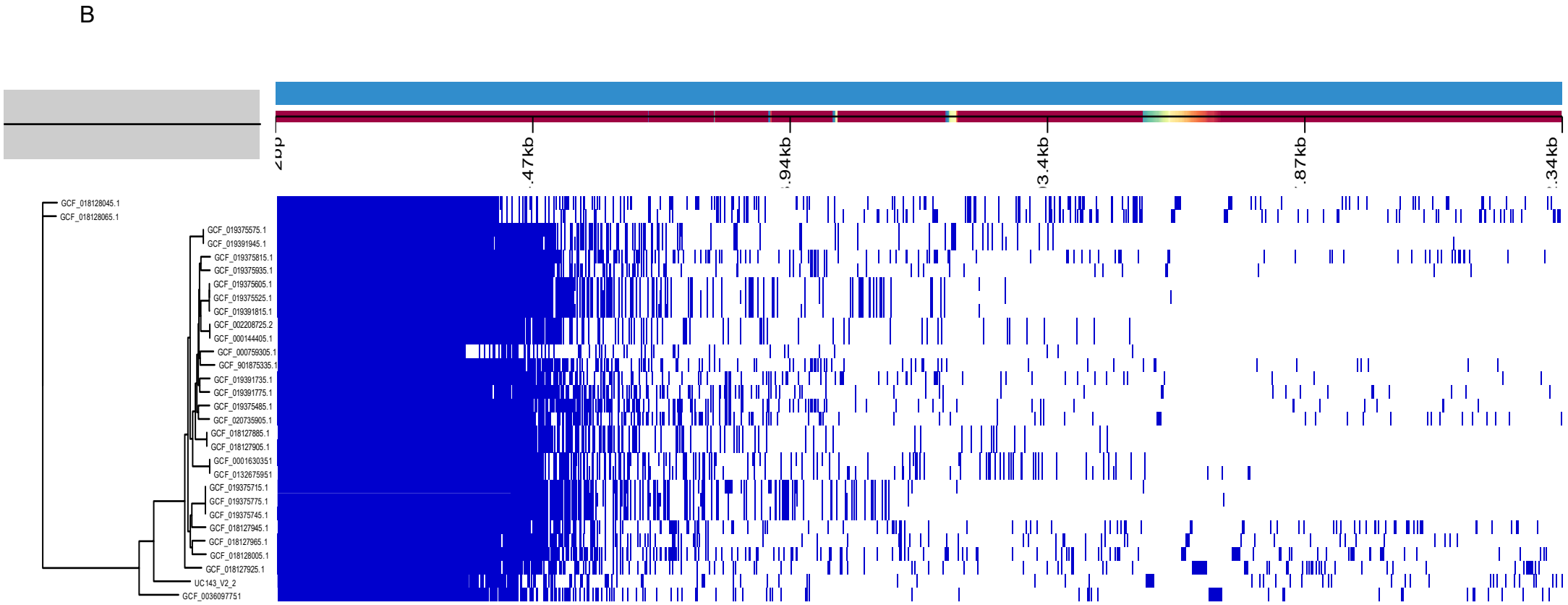


Figure 3.3: **The pangenome of *Prevotella* isolates with the corresponding phylogenetic tree.** The tree of genomes, determined by iq-Tree, with corresponding accessory genes in the pangenome of (A) *P. bivia* and (B) *P. melaninogenica* represented using Phandango (Hadfield et al., 2018).

Since the AMR databases do not contain every known resistance gene, the genomes were screened for additional 5-nitroimidazole reductase genes, *nimA-K*, which have previously been associated with metronidazole resistance in *Bacteroides* spp. and *Prevotella* spp. One isolate, UC121_V1_3, contained a putative *nimK* gene, located in a gene region with high similarity to the previously described Tn6456 transposon (Figure 3.5). The regions of the transposon flanking the *nimK* gene included genes encoding an integrase, a mobilisation protein, an insertion sequence (IS1380), *qacE* (encoding a putative efflux pump), a Crp/Fnr regulator and identical repeat regions. The *nimK* gene was not detected in any of the *P. melaninogenica* strains. No other potential *nim* genes were identified in any of the isolates.

Correlations between AMR gene presence and phenotypic resistance were determined for 34 of the South African isolates (Figure 3.6). Most strains encoding *tetQ* were resistant to doxycycline (25/34), although five strains harboured the gene without phenotypic resistance and one strain was resistant without the gene being present. Carriage of *ermF* was observed in only one of the South African isolates which was clindamycin resistant, while three others were phenotypically resistant to clindamycin without harbouring the gene. Similarly, while the strain with the identified *nimK* gene showed reduced susceptibility to metronidazole, two additional isolates with reduced susceptibility and one fully resistant isolate did not contain any *nim* gene. Mere presence of the gene for β-lactamase production, *cfxA*, did not always correlate with resistance with ~20% of the *cfxA* not displaying phenotypic resistance to amoxicillin. However, different *cfxA* gene variants were detected in the different strains and the presence of *cfxA3/A4* was correlated with significantly higher amoxicillin MIC values than the presence of *cfxA/A2* ($p = 0.000212$), the latter being associated with low-intermediate susceptibility to the antibiotic (Figure 3.7).

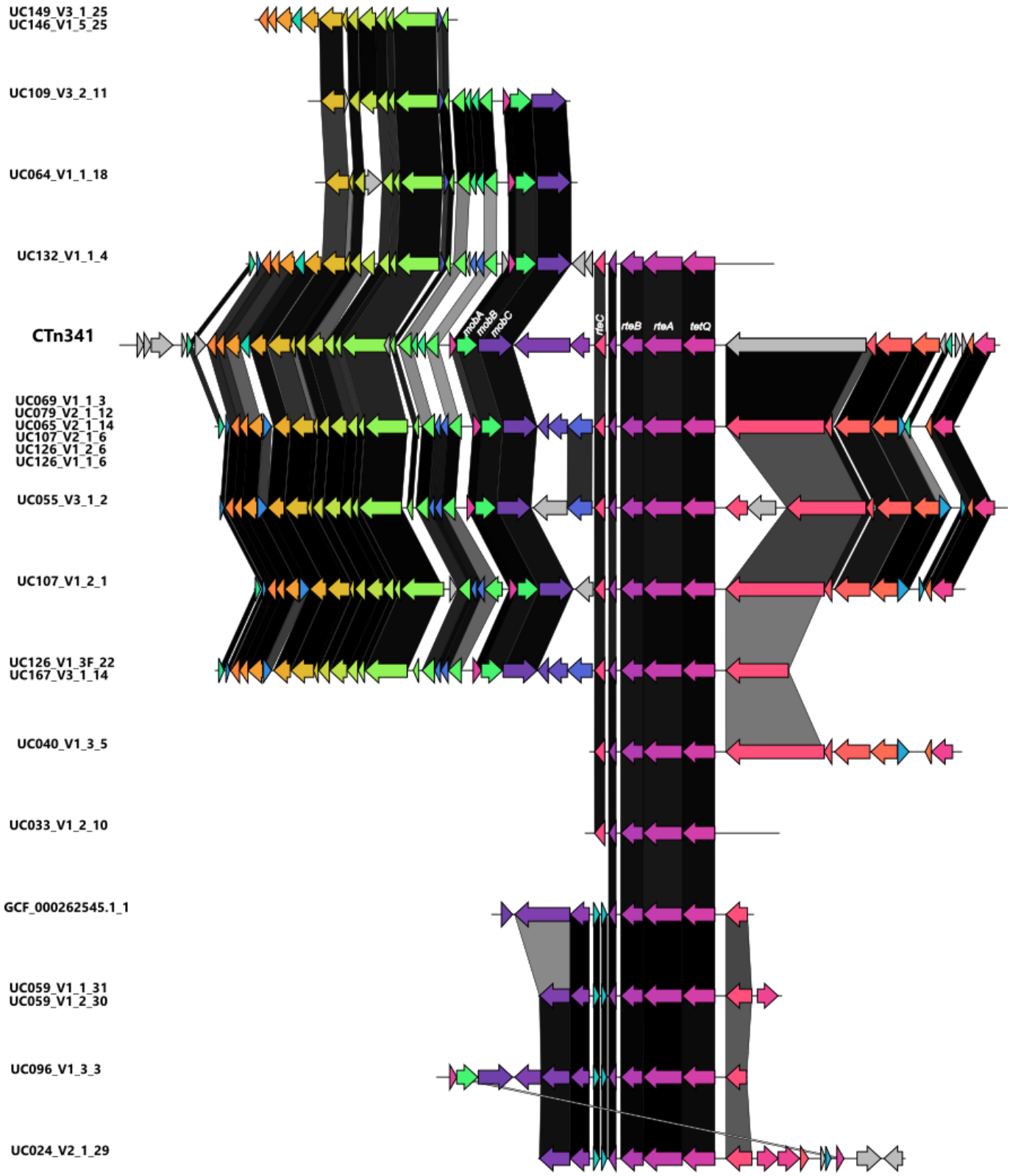


Figure 3.4: **Identification of CTn*DOT*-related transposon carriage in *Prevotella* spp. isolate genomes**. South African and reference *P. bivia* genomes were screened for homology to the previously published CTnDOT-family transposon, CTn341 (AY512263.1) and visualised using clinker. Genes involved in the transfer of the transposon are labelled on the reference CTn341 alignment.

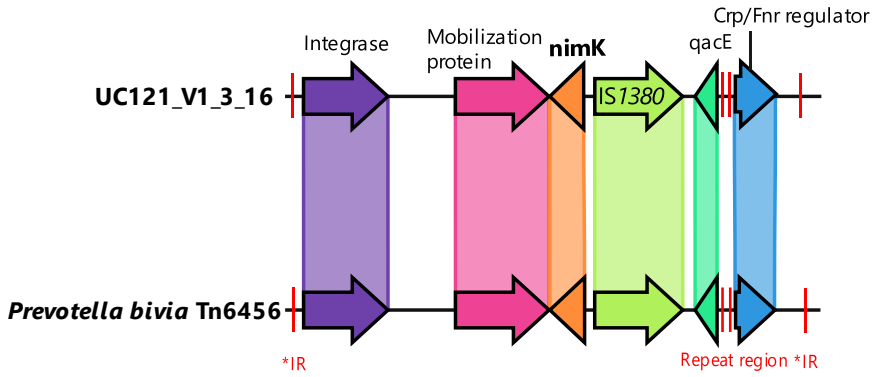


Figure 3.4: **The *nimK* transposon in UC121_V1_3 in comparison to the Tn6456 transposon.** Gene clusters in the regions of the *nimK* gene were compared to a previously published *P. bivia nimK* transposon with Clinker. The different gene regions are labelled and inverted repeats (IRs) are indicated.

Figure 3.5: **Correlation between presence of AMR genes and phenotypic antibiotic resistance.** Four antimicrobial resistance genes were detected in the *Prevotella* genomes which may confer resistance to tetracyclines (i.e. doxycycline), lincosamides (i.e. clindamycin), β-lactam antibiotics (i.e. amoxicillin) and nitroimidazoles (i.e. metronidazole). 'Phenotype observed' refers to the strain having phenotypic resistance to the antibiotic in the specific class.

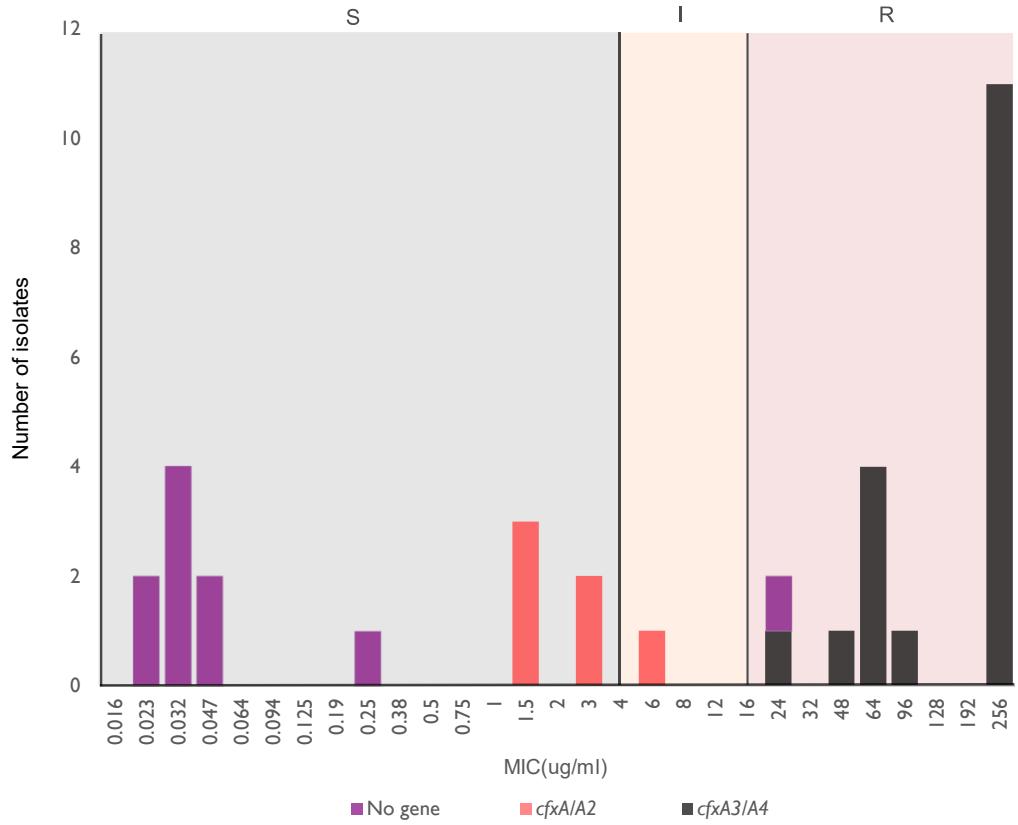


Figure 3.6: ***cfxA* gene variants and its effect on antimicrobial susceptibility.** The MIC (µg/ml) values of the isolates are indicated with their respective absence (No gene) or presence (*cfxA/A2* or *cfxA3/A4*) of the *cfxA* gene. The CLSI breakpoint for susceptibility (S), intermediate phenotype (I) and resistance (R) to amoxicillin is indicated on the graph.

3.4.6 Previously described virulence factors identified in the *Prevotella* spp. sequences

In order to assess for any genes potentially involved in virulence, annotated isolate genomes, as well as reference genomes, were also screened against the virulence factor database (VFDB) (Chen et al., 2005), as well as a list of putative virulence factors described previously(Sharma et al., 2022). While none of the genomes had genes with homology to those in the VFDB, genes with homology (>60% identity) to previously described *Prevotella* spp. virulence genes were identified in the *Prevotella* spp. isolate genomes (Table 3.2). The majority of *P. bivia* isolate genomes (>95%) carried genes encoding putative quorum sensing molecules (*luxS)*, haemolysins (including *hlyA* and *hly*-III) and proteases (*dpp)*. Additionally, a region of 75% homology to the undecaprenyl diphosphatase, *uppP*, gene, involved in bacitracin resistance, was identified in 97.73%

(43/44) of *P. bivia* isolates. Quorum sensing AI-2 molecules, haemolysins and a dipeptidyl peptidase protease were identified in all *P. melaninogenica* isolate genomes, while 93.33% (28/30) carried a gene proposed to encode a capsular exopolysaccharide biosynthesis protein, which was not as abundant in the *P. bivia* isolate genomes.

Table 3.2: **Putative virulence factors identified in Prevotella spp. isolate genomes**

	Protein name	**Gene name**	**Locus**	***Prevotella* spp. source**	**Percentage of strains carrying gene (n)**	
					P. bivia	***P. melaninogenica***
Quorum sensing AI-2 molecules	S-ribosyl homocysteine lyase	*luxS*	B129_RS0101260	*P. intermedia*	97.73 (43)	100 (30)
	AI-2E family transporter	-	B129_RS102810	*P. intermedia*	97.73 (43)	100 (30)
AMR	Undecaprenyl-diphosphatase	*uppP*	B129_RS0101320	*P. intermedia*	97.73 (43)	100 (30)
	Tetracycline resistance protein TetQ	*tetP*	HMPREF0650_RS03660	*P. buccalis*	88.64 (39)	23.33 (7)
	β-lactamase	-	HMPREF9144_RS08425	*P. pallens*	70.45 (31)	73.33 (22)
	TetR/ AcrR family transcriptional regulator	-	HMPREF9144_RS10840	*P. pallens*	75.00 (33)	0 (0)
	Pyridoxamine 5'-phosphate oxidase family	*nimB*	HMPREF9136_RS01340	*P. dentalis*	18.18 (8)	10 (3)
Haemolysins	Haemolysin family protein	-	B129_RS0103455	*P. intermedia*	95.45 (42)	100 (30)
	HlyD family secretion protein	-	B129_RS0104030	*P. intermedia*	97.73 (43)	100 (30)
	Haemolysin	*hlyA*	HMPREF6485_RS01780	*P. buccae*	95.45 (42)	100 (30)
	Haemolysin-III	*hly-III*	HMPREF9136_2207	*P. dentalis*	97.73 (43)	100 (30)
Proteases	Dipeptidyl peptidase	*dpp*	HMPREF0659_RS00470	*P. melaninogenica*	97.73 (43)	100 (30)
EPS	Capsular exopolysaccharide biosynthesis protein	-	HMPREF9136_RS10305	*P. dentalis*	68.18 (30)	93.33 (28)

3.4.7 Identification of prophage sequences in South African *Prevotella* spp. genomes

A primary prophage analysis found 73 potential prophage sequences from 33/36 of the *Prevotella* spp. genomes. Of these, 48 were categorised as low quality/genome fragments, 12 as medium quality/genome fragments and 13 as high quality/complete prophages, according to checkV and MIUViG classifications, respectively (Figure 3.8A). While multiple low and medium quality genome fragments were present in most bacterial genomes, no isolate harboured more than one predicted complete prophage sequence. The complete prophage sequences had predicted genome sizes ranging from 6 391 – 64 987 bp, with a median genome size of 39 952 bp (Figure 3.8B).

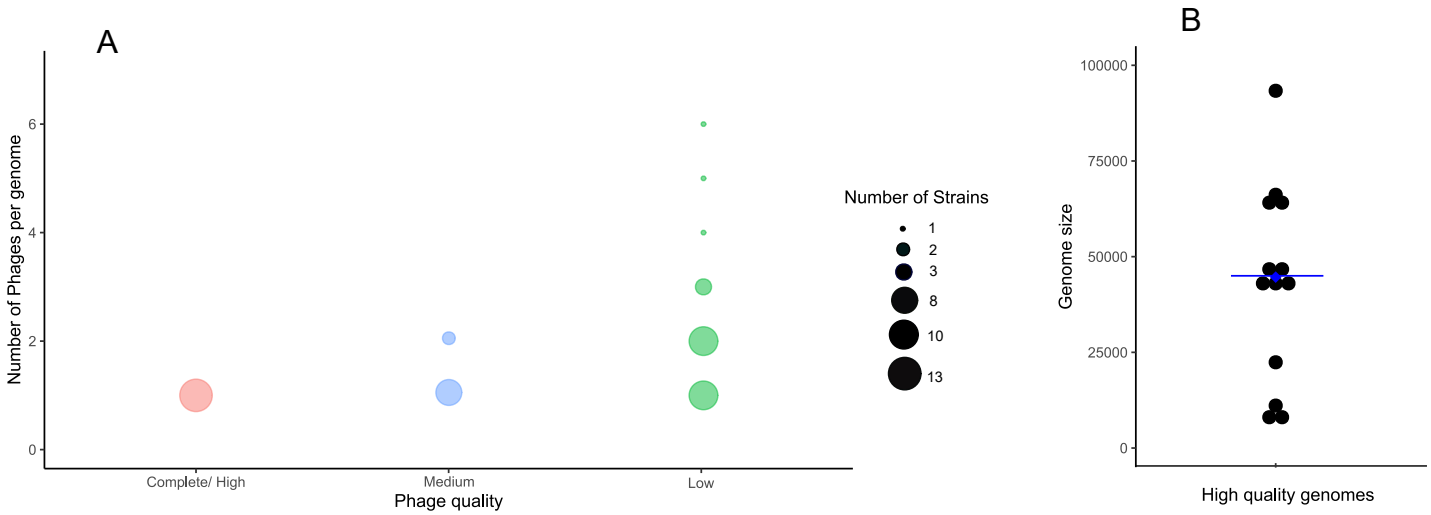


Figure 3.7: **Identification of putative prophage genomes** (A) Total numbers of prophages identified per strain, stratified by predicted prophage genome quality (MIUViG). (B) Predicted genome size of high-quality prophage genomes.

To determine if the predicted complete prophages had any close relatives amongst previously characterised bacteriophages, pairwise Jaccard similarities were determined to all *Caudoviricetes* (i.e. tailed bacteriophages) sequences in the NCBI RefSeq database (Table S3). Jaccard similarities to the closest relatives in the database were extremely low for all sequences, ranging between 0.0118-0.0211 (~12%-20% similarity). Best-hit matches were generally predicted to belong to the previously defined *Siphoviridae*-morphotype (long, non-contractile tails) (11/13), with the remaining two sequences

showing low levels of similarity to bacteriophages of the *Myoviridae* (short, contractile tails) and *Podoviridae* (very short, non-contractile tails) morphotypes. Visualisation of the complete and high-quality prophage sequences identified that there were two main clusters of phages, with 4 phages in one cluster and 6 phages in the second cluster (Figure 3.9). Of the individual, 'unique' phages of this dataset, UC143_V2_1 and UC040_V1_3 shared a small degree of genomic synteny but were mostly distinct from one another. Potential phage sequences from isolates from the same sample, UC059_V1_2 and UC059_V1_1, were identical to each other, but also similar to phages from strains from 4 separate individuals.

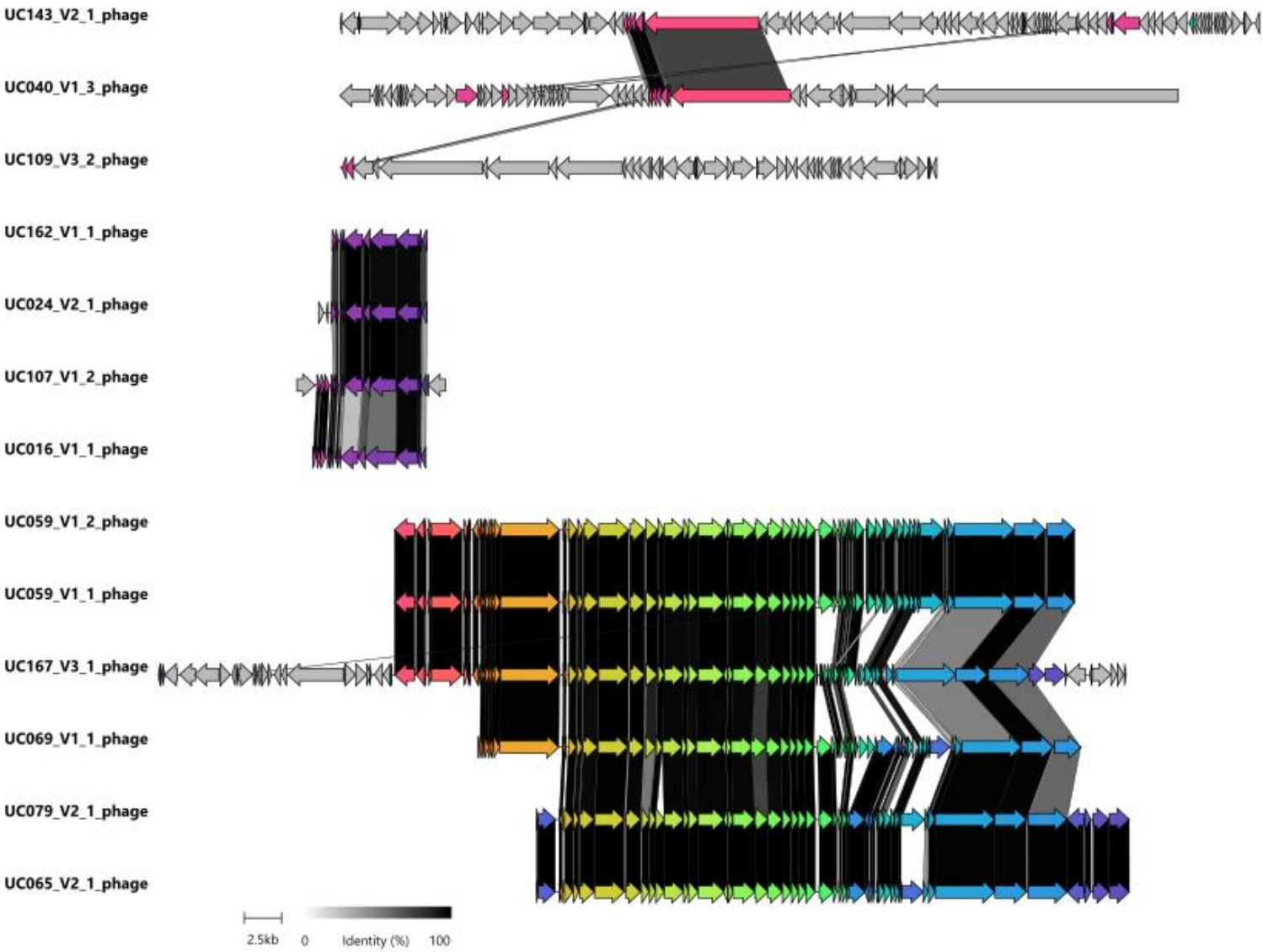


Figure 3.9: **Visualisation of prophage genomes.** High quality viral genomes were visualised using Clinker, with identity between phage genomes indicated.

3.4.8 Identification of potential anti-phage defence mechanisms in South African *Prevotella* spp.

In order to investigate the anti-phage arsenal of local South African isolates, all genomes were screened for known and predicted anti-phage defence mechanisms. CRISPR-Cas elements were present in all genomes, however only 30/36 of the strains contained complete CRISPR-Cas systems, with all these stains predicted to have the Class 2 Subtype IIC CRISPR-Cas system. The remaining six isolate assemblies did not have all the necessary components for CRISPR functionality. The most identified non-CRISPR-Cas-based anti-phage defence mechanism was the phage anti-restriction-induced system (PARIS), a particular class of abortive infection (Abi) mechanism, found in all 36 strains (Figure 3.10). Additional general 'Abi' mechanisms, belonging to the four subtypes (AbiD, AbiE, AbiH and SanaTA) were predicted for 31 strains, with multiple subtypes sometimes present in a single strain (for example, UC040_V1_3, which contained all three Abi subtypes) (Table S4). Putative restriction modification (RM) systems were predicted for all but one strain. As for the Abi systems, some strains contained multiple of the same RM type as well as different RM types, simultaneously. The only other commonly predicted defence mechanism belonged to the PrrC/RloC class, that are responsible for pausing translation during bacterial infection and were present in 11 strains. The remainder of the predicted defence mechanisms were only rarely identified, present in five or less isolates.

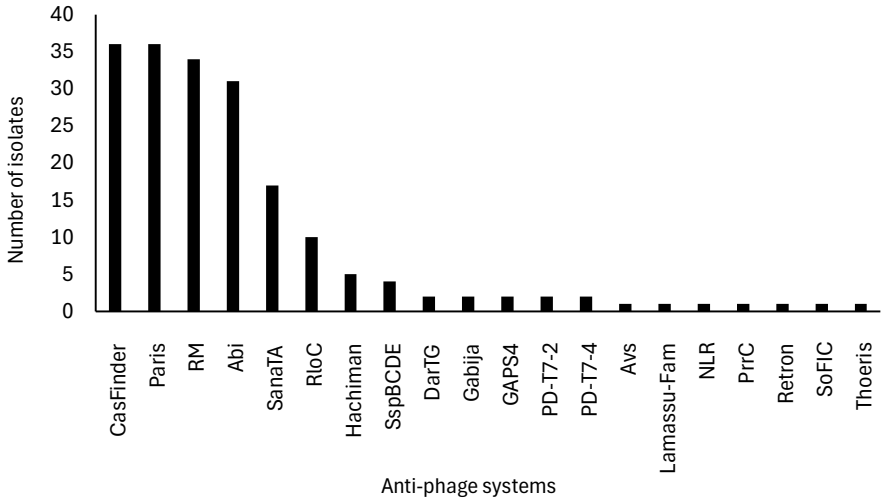


Figure 3.10: **Anti-phage defence mechanisms in *Prevotella* spp.** Detection of anti-phage defence systems present in all *Prevotella* isolate sequences (n=36) was performed using DefenseFinder.

3.5 Discussion

With the increased affordability and availability of NGS technologies, characterisation of bacterial isolates has begun to rely not only on phenotypic or lab-based techniques, but also on genotypic analyses, which provide valuable insights into evolutionary relationships and gene carriage. Following genome assembly and quality control checks of the strains in this study, 36 strains were carried forward for further genomic analyses. The overall genome metrics (including genome size and %GC content) of South African *P. bivia* strains were comparable to previously sequenced strains in the NCBI RefSeq database and there was little evidence of contamination within the sequenced genomes. However, the UC143_V2_2 strain, which was preliminarily identified as *P. melaninogenica* based on its 16S rRNA gene sequence, had a slightly smaller genome size than is typical for *P. melaninogenica* strains in the NCBI RefSeq database. This, along with evidence discussed later in this chapter, lends support to the idea that this isolate may represent a novel bacterial species. *Prevotella melaninogenica* has been shown to possess two chromosomes, which explains, in part, the larger genome size (Gupta et al., 2015). Interestingly, a recent complete genome assembly for *P. bivia* PLW0727 suggests that *P. bivia* may also possess two chromosomes (Peng et al., 2023). The use of short read sequencing technology in this study did not allow for an analysis of chromosome structure in the South African *Prevotella* spp. isolates. However, future work including a long-read approach, such as PacBio or Oxford Nanopore technologies (ONT) sequencing, could be carried out to generate complete, closed assemblies for selected isolates to answer this question, especially as some isolate sequences in this study had contamination levels greater than 10%. Additionally, the two *P. bivia* strains that did not pass the quality control analyses and were excluded due to their high contiguity will be re-sequenced as part of future work.

Comparative genomics approaches have been used to analyse the fine-grained molecular diversity of bacterial populations. For this study, a subset of isolates from both the same sample and the same individual at different time points were included to begin to examine the intra-sample and longitudinal diversity of *Prevotella* strains. While a much larger set of isolates is required to explore these questions fully, the data presented here provide some interesting preliminary insights. For example, two strains from the same

sample (UC071_V1_1 and UC071_V1_1F, which were identical based on 16S rRNA gene and *cpn60* sequencing), were significantly different at the genomic level with over 9000 core genome SNPs and ~650 accessory genes unique to one or the other. Previous metagenomic-based studies have revealed that the vaginal community is not necessarily clonal, with relatively high levels of intra-sample diversity for certain species. Indeed, Ma et al. showed that *Prevotella* gene counts in vaginal communities were higher than in any individual species genomes, suggesting that vaginal communities consist of species populations made up of individual strains having differing gene repertoires (Ma et al., 2020). Other strains isolated from the same sample were more closely related. It is possible that the diversity within these samples was lower, however isolation and sequencing of more strains and/or shotgun metagenomics would be necessary to confirm this. Interestingly, even isolates that were very closely related at a core genome level displayed some phenotypic variation. For example, *P. bivia* strains isolated from the first visit of participant UC126 as a group had 1-20 core genome SNPs, yet differed in their ability to form biofilms and only one of them was phenotypically resistant to azithromycin (Chapter 2). It is likely that differences in the accessory genome were responsible for these phenotypic differences.

Strains of *P. bivia* isolated at different time points in two of three cases, were phylogenetically distinct from one another. Due to the possibility of strain variation within a sample, it cannot be determined whether in these cases the earlier strain was lost or replaced by another *P. bivia* strain. The third pair of strains were genetically similar at a core genome level despite being isolated approximately 8 weeks apart, suggesting persistence between samplings. Interestingly, the accessory genomes of the two isolates were different, with the later isolate lacking a contig that may have been a plasmid. The presence of a similar potential plasmid in another distantly related isolate from another study participant suggests possible horizontal gene transfer within the population.

Core genome alignments for *P. melaninogenica* reference strains and the UC143_V2_2 isolate revealed extensive genomic variation. The only two vaginal *P. melaninogenica* sequences [*P. melaninogenica* GAI07411 (accession GCF_003609775.1) and the South African isolate UC143_V2_2] were genetically distinct from oral-, sputum-, and gut-derived strains and likely did not appear simply due to introduction of an oral bacterium

to the vagina pre-sampling. These strains had similar genome sizes, which may be due to adaptations to the vaginal environment. These isolates were also distinct from one another and clustered separately. The reference sequence for *P. melaninogenica* GAI 07411 currently has a taxonomic status of 'inconclusive'. Nevertheless, the strain has been fairly well characterised and has been used in *in vitro* knockout experiments (Kondo, 2021; Kondo et al., 2018). Interestingly, two other *P. melaninogenica* strains clustered more closely to the outgroup, *P. scopos*. *Prevotella scopos,* first described in 2011, was identified as being closely related to, but distinct from, *P. melaninogenica* (Downes & Wade, 2011). The two strains were both oral isolates and, similarly to *P. melaninogenica* GAI 07411, despite being present in the NCBI RefSeq database, their taxonomic status is recorded as 'inconclusive'.

Based on the core genome and ANI results, it appears that the currently recognised *P. melaninogenica* species consists of up to four genomospecies. The term 'genomospecies' has been used to refer to 'organisms for which data from nucleic acid comparisons indicate that they constitute a new species, but for which no distinguishing phenotypic properties have yet been found' (Schloter et al., 2000). ANI represents the average identity of nucleotides between any two genomes, reflecting the DDH concept *in silico*, with a 95% ANI identity having been determined to be equivalent to the 70% identity in DDH (Goris et al., 2007; Konstantinidis & Tiedje, 2005). The ANI analysis on strain UC143_V2_2 revealed that it does appear to be a distinct genomospecies from the rest of the previously published *P. melaninogenica* strains, with an ANI score below the 95% cutoff. The other vaginal strain, *P. melaninogenica* GAI 07411 also appears to constitute a separate genomospecies. A valuable aspect of future work would be understanding how prevalent both of these potential novel species are in vaginal populations. Searching for them in a well-established vaginal metagenome database, such as the human vaginal non-redundant gene catalogue (VIRGO), would determine whether this species has been detected in vaginal microbiomes previously (Ma et al., 2020). The remaining two strains, *P. melaninogenica* F0054 and F0091 (accession numbers GCF_0180128045.1 and GCF_0180128065.1, respectively), likely form another genomospecies, possibly of *P. scopos*. One further piece of evidence supporting the proposal to reassign some *P. melaninogenica* strains to separate genomospecies is the

very large species pangenome, with almost 85% made up of accessory genes. This would be expected as the current species designation included a set of strains that have diverged significantly from one another.

There are some caveats to the above interpretation. Palmer et al. carried out a large-scale comparison of the use of different ANI calculation methods for the delineation of bacterial species (Palmer et al., 2020). While the method employed in this study (ANIm) generally correlated well with ANIb scores, the authors suggest determining taxon-specific cutoffs, which may vary for different species. Alternative methodologies like dDDH may overcome these limitations. Furthermore, extensive additional *in vitro* characterisation based on cellular morphology, optimal growth conditions, and various biochemical and metabolic tests is needed before reclassification as a novel species is possible (Baron, 1996).

Intra-species investigation of evolutionary relationships and genetic distances amongst strains is only part of the utility of WGS for genome interrogation. Analysis of the specific genes present or absent in strains can identify features of different strains and provides the opportunity to investigate intra-species adaptation to particular ecological niches or geographic regions (Anani et al., 2020; Rouli et al., 2015). For example, pangenome profiling of *Prevotella copri* isolates revealed differentiation between strains in Western and non-Western individuals by principal coordinate analysis (PCA). However, these differences were proposed to be due to dietary differences and not solely due to geography (De Filippis et al., 2019). While a basic analysis did not identify genes that were significantly more frequently present or absent in South African isolates and they did not cluster separately from previously published sequences, results from another study in the group suggested some phenotypic differences with respect to susceptibility to probiotics (Anika Chicken, personal communication). Isolates of *P. bivia* from South African women tended to be inhibited better by *L. crispatus* strains of South African origin than non-SA *P. bivia* isolates. Future work, including the sequencing of greater numbers of *P. bivia* isolates from different geographic locations is required to identify whether geographic differentiation is a feature of the species.

Additional analysis revealed no genes significantly associated with host BV-positivity in the study participant. However, one gene, *asnA,* was identified in all strains from BV-

negative women, while only being present in less than half of isolates from women with BV. The enzyme encoded by this gene, aspartate-ammonia ligase, converts L-aspartate to L-asparagine, in an ammonia-dependent reaction (Schleicher et al., 2021). Ammonia is a common metabolic bioproduct of *P. bivia* growth, where it is produced by the degradation of L-asparagine and aspartate and can become inhibitory to cell growth if it accumulates in the environment. *Gardnerella* spp. are proposed to utilize the ammonia produced by *P. bivia*, and synergistically provide amino acids which are used in *P. bivia* amino acid metabolism, enhancing each other's growth (Pybus & Onderdonk, 1997). It is tempting to speculate that in a non-BV environment, where *Gardnerella* spp. cell numbers are significantly reduced or absent, the presence of an ammonia ligase whose action results in the removal of excess ammonia would be a competitive advantage. Future work into the role of aspartate-ammonia ligase-mediated ammonium consumption in *P. bivia* growth would be required to test this hypothesis.

An important component of pathogen characterisation is determining whether strains harbour genes associated with AMR and virulence. AMR genes that may confer resistance to BV antibiotics, metronidazole, and clindamycin, were identified in some of the South African isolates. Overall, resistance to metronidazole was rare amongst the South African isolates. However, a *nimK* homologue was identified in one isolate, UC121_V1_1, which had intermediate susceptibility to metronidazole. The mechanism of resistance conferred by *nim* genes is the production of a nitroimidazole reductase, converting the drug to aminoimidazole, negating its antibacterial activity (Carlier et al., 1997). Metronidazole resistance is not very common in *Prevotella* spp. in general, and while *nim* genes have been reported in the species, it is at quite a low frequency (Veloo et al., 2020). Nevertheless, of the 12 previously described *nim* genes (*nimA-L*), five have been reported in *Prevotella* species. Lubbe et al. detected *nimA* in a *P. bivia* strain in South Africa in 1999, Katsandri and colleagues identified *nimC* (*Prevotella oralis*) and *nimE* (*P. oralis and Prevotella buccalis*) in strains from Greece, and *nimI* was identified in *Prevotella baroniae* strains from France (Alauzet et al., 2010; Katsandri et al., 2006; Lubbe et al., 1999). In 2018 a novel *nimK* gene was described in three metronidazole-resistant *P. bivia* isolates from the Netherlands, with localisation on a mobile genetic element (Veloo et al., 2018). The *nimK* gene identified in strain UC121_V1_3 shared 100% nucleotide identity with the

genes in the previously described *nimK* transposon, which also contained an *IS1380* family transposase. Insertion sequences (IS) elements such as *IS1380* are often found upstream of *nim* genes and are thought to be responsible for enhancing their expression in metronidazole-resistant isolates, although this is not always the case (Alauzet et al., 2019). Interestingly, the *nimK* transposon in both the SA and the Netherlands isolates additionally contained a *qacE* gene encoding a protein with similarity to efflux small MDR (SMR) transporters. Although SMR efflux pumps have been implicated in resistance to biocides (He et al., 2011), their role in *Prevotella* spp. remains unknown. Nevertheless, it is noteworthy that the Netherlands strains and the UC121_V1_3 isolate were all phenotypically resistant to clindamycin, despite (in the case of UC121_V1_3 at least) lacking an *ermF* gene. From these analyses, it cannot be determined whether or not the lack of *ermF* is due to incomplete genome sequences, due to the gene falling on the beginning or end of a contig or because there is indeed no *ermF* gene. Unfortunately, full genome data for the Netherlands strains were not available for the current study, but it would be interesting to see whether these strains also lacked previously published clindamycin resistance determinants. Furthermore, given the fact that *nimK* and *qacE* genes appeared to be on a mobile genetic element, should they be shown to be functionally responsible for the phenotypic resistance observed, it would be useful to determine whether this element poses a risk for horizontal gene transfer between members of the vaginal microbiota. In addition to the isolate harbouring the *nimK* gene, two other strains had reduced susceptibility to metronidazole and a third was fully resistant to the antibiotic. The mechanism of resistance in these strains is unknown. Alternative mechanisms have been proposed such as reduced uptake of the drug (Liu et al., 2000), efflux of the drug from the cell (Pumbwe et al., 2007) or overexpression of the DNA repair protein RecA (Steffens et al., 2010). Future work such as the use of efflux pump inhibitors and proteomic comparisons of resistant and susceptible strains would be necessary to elucidate the metronidazole-resistance mechanisms in these isolates.

Resistance to clindamycin, a lincosamide antibiotic, is much more common in *Prevotella* species than resistance to metronidazole. The most well described mechanism for resistance to clindamycin, and other macrolides, lincosamides and streptogramins (MLS), is encoded by erythromycin resistance methylase (*erm)* genes, which encode a

rRNA methylase. The primary *erm* gene variant in *Prevotella* and *Bacteroides* spp. is *ermF,* which was detected in one South African strain, and whose presence appears to predict resistance to clindamycin (Veloo et al., 2019). This methylase prevents the action of MLS antibiotics, which all share overlapping binding regions on the 50S ribosomal subunit. Therefore, besides clindamycin, azithromycin effectiveness is also impacted by the presence of *erm* genes, affecting treatment outcomes for chlamydial and gonorrhoeal infections (Peters et al., 2022). Presence of *ermF* in lung derived *Prevotella* spp. has been associated with both azithromycin and clindamycin resistance (Sherrard et al., 2014; Webb et al., 2022). In this study, of three clindamycin-resistant strains, two were isolated from individuals that received azithromycin prior to sampling, and one harboured the *ermF* gene. The mechanism of resistance in the isolates that lacked an *erm* gene may have been related to the presence of efflux pumps that were not detected in the AMR determinant screen and could be tested by using efflux pump inhibitors (e.g. reserpine, verapamil, CCCP). Another proposed lincosamide inactivating enzyme, LinB, has been identified in *Enterococcus faecium*, but to date, it has not been identified in *Prevotella* spp. (Morar et al., 2009).

Interestingly, in *Bacteroides*, *ermF* is often found on a conjugative transposon (CTnDOT) accompanying the tetracycline-resistance gene, *tetQ* (Waters & Salyers, 2013). The hypothesis is that low levels of tetracycline facilitate several steps that result in the transfer of the transposon, transmitting not only tetracycline resistance, but also potentially resistance to erythromycin, azithromycin, and clindamycin. Almost all the South African isolates contained *tetQ*, and co-occurrence of *tetQ* with *ermF* was detected in the clindamycin resistant isolate, UC109_V3_2, and two of three other *ermF* containing isolates. CTnDOT has not been described in *Prevotella* spp. yet, but the co-occurrence of *ermF* and *tetQ* in *Prevotella* isolates, and the evolutionary closeness to *Bacteroides*, makes the transposon worth looking for in the South African isolates.

Regions homologous to a previously described CTnDOT-family transposon, CTn341 (Bacic et al., 2005), were identified in some of the South African *Prevotella* spp. isolates, with many of them almost full length, and the large majority carrying the *tetQ* gene. Contiguity of the genomes, as is characteristic with short-read genome sequencing, can be a limitation in trying to identify large gene clusters or transposons, as genes may be on

separate contigs. However, the homologous regions that appeared to be 'truncated' versions of CTn341 were not found near the beginning or end of contigs. Thus, unless these genomes were misassembled, it is likely these isolates harbour some genes without the full complement of genes in the transposon. Future work would involve determining whether these putative transposons are mobilizable and can confer *tetQ* to other *Prevotella* spp.

While UC143_V2_2, the *P. melaninogenica* relative, had no resistance genes, the prevalence of *ermF* was much higher in *P. melaninogenica* reference sequences than in *P. bivia*. However, since phenotypic AMR data for most of the sequenced strainswas not available, whether this translated to phenotypic MLS resistance is unknown. Similar to what was observed here, a previous study on oral isolates of *P. melaninogenica* detected a 20% prevalence of the *ermF* gene, compared to the reported prevalence *ermF* in *P. bivia* at 5.9% in a different study, thus *erm* gene presence appears to be more common in *P. melaninogenica* (Castillo et al., 2022; Veloo et al., 2019).

Another class of antimicrobials broadly used for various infections, including those of the female genital tract, is β-lactam antibiotics. While penicillin is mostly used for the treatment of Gram-positive bacteria, β-lactams such as amoxicillin can also be used for some Gram-negative bacteria. In the urogenital tract, they may be prescribed for treatment of UTIs (Tan & Chlebicki, 2016), and while not specifically prescribed for BV, may impact the vaginal microbiota indirectly. In Gram-negative anaerobes, such as *Prevotella* spp., the most prevalent resistance mechanism to β-lactams is production of β-lactamases, most frequently attributed to *cfxA,* which has been associated with phenotypic resistance (Toprak et al., 2020). There are numerous variations of the *cfxA* gene that have been found in *Bacteroides* and *Prevotella* spp., *cfxA, -A2, -A3, -A4, -A5* and *-A6*, defined by several point mutations (García et al., 2008; Yokoyama et al., 2023). The *cfxA* gene was identified in approximately 70% of the South African isolates, with two groups of *cfxA* variants, *cfxA/A2* and *cfxA3/A4* identified. This prevalence of the *cfxA* gene was similar to the result of another study of 45 *Prevotella* spp. isolates, where 77.8% had the *cfxA* gene and were actively producing β -lactamases (Yokoyama et al., 2023). Analysis of the effect of carriage of the gene variants on the MIC for South African isolates against amoxicillin revealed that presence of the *cfxA3/A4* gene was significantly

associated with resistant MIC values compared to *cfxA/A2*. Another study has shown the same result, with CfxA3-producing isolates showing higher MICs compared to CfxA2-producing isolates (Yokoyama et al., 2023). Amoxicillin sensitivity in the absence of its clavulanic acid inhibitor was tested in this study, enabling detection of strains still sensitive to this β-lactam antibiotic.

Screening of isolate genomes against the VFDB (Chen et al., 2005), a reference database of virulence factors, identified no regions in *Prevotella* spp. isolate genomes with high homology (>60%) to virulence genes in the database. Thus, the isolate genomes were manually screened for previously described *Prevotella* spp. virulence genes, summarised in a review by Sharma et al. (2022). The most commonly identified putative virulence-associated genes in the *Prevotella* spp. isolates included those encoding quorum sensing molecules, haemolysins, a protease, and exopolysaccharides. Interestingly, a large number of these proteins are predicted to play a role in biofilm formation. The protein encoded by *luxS* is an autoinducer-2 (AI-2), which is a component of the AI-2 quorum sensing mechanism present in both Gram-positive and Gram-negative bacterial species and is thought to facilitate communication on an intra- and interspecies level within various environments, including biofilms. The synthesis of AI-2, an autoinducer (Surette et al., 1999) has been identified in oral *P. intermedia* spp. (Frias et al., 2001) and ruminal *Prevotella* spp. (Gorenc et al., 2008). However, further experiments are needed to confirm a role in multispecies biofilm formation. Additionally, a gene predicted to encode a capsular exopolysaccharide (EPS) biosynthesis protein was identified in more than 70% and 90% of *P. bivia* and *P. melaninogenica* genomes, respectively. This protein is also thought to be involved in biofilm, forming the extracellular matrix (Koo et al., 2013). The production of EPS has been shown to reduce phagocytosis of *P. intermedia* and suggests that EPS can allow *Prevotella* spp. to evade host immune response (Yamanaka et al., 2011). It has also been shown to resist the host immune complement system and facilitate decreased antibiotic susceptibility (Geisinger & Isberg, 2015).

Other putative virulence factor genes were *hlyA* and *hly*-III, which encode haemolysins that are responsible for lysing red blood cells by forming pores in their cell membrane. Iron is an essential element required for *Prevotella* spp. growth, mainly shown for *P.*

intermedia in the oral cavity (Gibbons & Macdonald, 1960; Suzuki et al., 2012). In niches where free iron is low, but red blood cells are present, production of haemolysins may allow *Prevotella* spp. to acquire iron from the red blood cells. Thus, in the vaginal canal, menstrual bleeding may act as an iron source for those producing haemolysins, and may explain why *Prevotella* spp. abundance may increase following menses (Oerlemans et al., 2022). The other abundant putative virulence gene identified was *dpp*, encoding a dipeptidyl peptidase (DPP), which plays a role in the pathogenicity of oral *Prevotella* spp. by impacting host immune regulation and signal transduction (Kahne et al., 1999). *Prevotella* spp. have been shown to have high activity of DPP (Gazi et al., 1995), but whether the presence of the gene is directly indicative of DPP activity is not known, and production of the protease *in vitro* would need to be studied in order to confirm its activity.

To date, no virulent bacteriophages (phages) with activity against vaginal *Prevotella* spp. have been identified. These could be useful in phage therapy applications to help treat bacterial vaginosis and other *Prevotella*-related vaginal conditions by specifically targeting the bacterium responsible for disease without impacting other members of the vaginal microbiota, as is the case for antibiotics. However, WGS allows for identification of phages integrated in the genomes of bacteria (prophages). While these temperate phages cannot directly be used in phage therapy, due to their potential to lysogenise into the bacterial host rather than lysing it, they can be valuable sources of compounds with therapeutic potential. For example, prophage elements encoding endolysins in the *G. vaginalis* genome are being researched as part of combination therapies to specifically kill target *G. vaginalis*, and disrupt its biofilms, aiding in the treatment of BV (Arroyo-Moreno et al., 2022; Landlinger et al., 2021). Another way these prophages can be manipulated to kill target bacteria is by inducing the integrated phages, and then removing the genetic modules used in lysogeny to convert them to virulent phages (Gibb et al., 2021).

In the South African *Prevotella* isolates, only thirteen of the identified potential prophage sequences were classified as high-quality based on MIUViG guidelines (Roux et al., 2019). Two of the prophage sequences had predicted genomes <7.5kb in size, which is at the lower end of published phage genome sizes. Whether these prophages harbour the entire set of genes required for functionality will still need to be determined. An important

limitation to prophage identification in short-read assembled genomes is the resulting contiguity, which makes it difficult to identify phages if they are split across multiple contigs. This is another reason why generation of full-length, complete genomes for a subset of isolates in this study would be useful. All the identified phages appeared to be novel and clustered into two groups of similar phages, with some unique individual sequences. Future work would involve attempting to induce the predicted prophages using classic prophage inducers (i.e. mitomycin C or antibiotics) and visualisation of the resulting phage particles using electron microscopy. Determination of the host range for any induced phages would also be useful in downstream therapeutic applications (for example, if they could be re-engineered to become virulent). Finally, phage genomes could be mined for the presence of genes encoding useful functions (e.g. endolysins).

Bacterial anti-phage mechanisms have evolved to combat phage infection in a constant back and forth between phage and bacterial defence evolution (Georjon & Bernheim, 2023). CRISPR/Cas systems have evolved as a primary line of defence and function as a bacterial analogue of the immune system (Barrangou et al., 2007). The basic premise is that upon infection of a bacterium with a phage or mobile genetic element, the CRISPR/Cas system degrades the foreign DNA and incorporates segments of the foreign DNA into its genome (where it is called a 'spacer'). These spacers are separated by repeat regions and are indicative of prior infections of the bacterium. The CRISPR/Cas system is divided into two classes, based on the complexity of their effector complexes. The CRISPR/Cas systems in the South African *Prevotella* spp. isolates were all of the Class 2 subtype 2C, suggesting that they had function via a single, multidomain effector protein (Class 2), which in this case, is Cas9 (subtype 2C) (Mir et al., 2018). CRISPR/Cas Class 2C systems have not been extensively described in *Prevotella* spp., but have been reported in *P. intermedia* strains in one study (Johnston et al., 2017), and in 40 strains of *B. fragilis* in another study (Tajkarimi & Wexler, 2017), suggesting that these systems are plausibly identified in *Prevotella* spp. All the South African *Prevotella* spp. isolates contained CRISPR/Cas elements, but six did not appear to have all the necessary components for a complete, functional system. However, complete closed genome sequences will be required to confirm the absence of specific components. Future work

will include further investigations into the spacer sequences in the CRISPR arrays of the South African *Prevotella* spp. to determine previous infection history.

Non-CRISPR/Cas anti-phage systems were also identified. PARIS, a phage anti-restriction-induced system, was the most frequently detected anti-phage mechanism in the South African *Prevotella* isolates. This mechanism is a form of Abi, which is specifically triggered by a phage anti-restriction protein, killing the host (Rousset et al., 2022). In contrast to other Abi mechanisms that are more generally triggered by any phage component, PARIS is only triggered after the first round of defence has been breached by the invading phage. Some other Abi-like mechanisms detected include PrrC and its distant homologue, RloC (which inhibit bacterial translation), SanaTA (a toxin-antitoxin system), Thoeris (which degrades cellular NAD^+), DarTG (a toxin-antitoxin system) and bacterial retrons (Davidov & Kaufmann, 2008; Dy et al., 2014; Millman et al., 2020). Restriction modification (RM) systems were also common in the South African *Prevotella* spp. isolates. These systems function by attacking foreign (differentially methylated) DNA in the cytoplasm. Some of the other antiviral defence systems identified have not been fully characterised yet, including Hachiman and Gabija (Egido et al., 2021). Whether these anti-phage defence mechanisms provide functional protection against phages would need to be determined in future work.

The results presented in this chapter illustrate the usefulness of isolating more than one strain from a particular vaginal sample and highlight the potential for whole genome sequencing for holistically interrogating inter-strain relatedness and picking up evolutionary differences that may not be detected using the 16S rRNA or *cpn60* genes. In addition, it allowed the identification of clinically relevant AMR- and virulence-related genes. The result is a well characterised, both phenotypically and genotypically, collection of South African primary isolates that can be used as a basis for further analyses. One of the key factors in *Prevotella* spp. virulence, and one of their reported roles in BV pathogenesis, is the production of LPS, which can promote inflammation in the vagina (Aroutcheva et al., 2008). However, there have been very few studies comparing LPS production by vaginal *Prevotella* spp. isolates. Therefore, LPS production was compared between isolates in the South African strain collection and the results of these analyses are presented in Chapter 4.

Chapter 4: Investigation of lipopolysaccharide production by *Prevotella* spp.

4.1 Abstract

Lipopolysaccharide (LPS) is proposed to be a major driver of BV-associated inflammation, contributing to some adverse BV health outcomes. *Prevotella* spp. produce LPS, with *P. bivia* particularly identified as a major source of LPS in the female genital tract. To compare quantity and macromolecular structure of LPS produced by different South African *Prevotella* spp., LPS was isolated and quantified, using commercial kits, followed by visualisation of LPS by SDS-PAGE. The *P. bivia* isolates produced comparable concentrations of LPS, however UC143_V2_2, the *P. melaninogenica*-relative, produced 24-fold greater concentrations of LPS compared to the *P. bivia* ATCC 29303^T type strain ($p<0.05$). LPS macromolecular structure appeared similar for all five *P. bivia* isolates regardless of growth mode (planktonic vs biofilm), with all strains having evidence of the O-antigen. The UC143_V2_2 isolate, however, had a different banding pattern to *P. bivia* isolates, suggesting a different LPS structure between the two species. Screening of 36 *Prevotella* spp. isolate genomes for putative LPS biosynthetic genes identified the presence of genes encoding homologues of lipid A biosynthetic genes for penta-acylated LPS in all genomes. Finally, a potential gene cluster involved in core oligosaccharide biosynthesis was identified in the *P. bivia* ATCC 29303^T genome.

4.2 Introduction

All Gram-negative bacteria have important cell wall components that distinguish them from Gram-positive bacteria and play a role in their virulence. The Gram-negative cell outer membrane contains LPS, which contributes to the functioning of the bacterial outer membrane and prevents penetration of small hydrophobic molecules into the cell. Crucially, however, it also acts as a surface antigen (Kellenberger & Ryter, 1958; Mühlradt & Golecki, 1975) and can be highly antigenic to the host, resulting in inflammation when it is recognised by host pattern recognition receptors (PRRs). In fact, as little as 1 to 2µg of purified LPS can be lethal to humans when injected, illustrating the potential of the compound to act as a potent endotoxin (Sauter & Wolfensberger, 1980).

Prevotella spp., particularly *Prevotella bivia*, are proposed to play a role in BV-associated obstetric and gynaecological outcomes, including pre-eclampsia and PID (Lin et al., 2020; Mikamo et al., 1998). The mechanisms by which *P. bivia* impact reproductive health are not certain, however, *P. bivia* is a major source of LPS in the vagina and its presence it thought to induce inflammation that may drive these complications. In support of this, *Prevotella* spp. abundance in the vagina has been shown to be associated with vaginal inflammation and the production of specific pro-inflammatory cytokines (Anahtar et al., 2015; Eade et al., 2012). Inflammation, which may be elicited by *P. bivia* LPS, could trigger early onset of labour (Kumar et al., 2021) and may play a role in increased susceptibility to STIs, for example increased HIV susceptibility due to recruitment of HIV target $CD4^+$ T cells to the vagina (Gosmann et al., 2017; Masson et al., 2015). LPS produced by vaginal *Prevotella* spp. is proposed to activate NF-κB via TLR4 and CD14 receptors on vaginal epithelial cells, monocytes, and macrophages (Nasu & Narahara, 2010), highlighting the role of LPS as the driver of *P. bivia*-mediated inflammation.

Mature LPS usually consists of three main components (Emiola et al., 2015). The innermost component is called lipid A and forms a lipid anchor that tethers the mature LPS to the cell membrane. This is also the most immunostimulatory component of the LPS (Raetz et al., 2009). Studies in various Gram-negative bacteria have shown that lipid A biosynthesis is an eight- or nine-step enzyme-catalysed process that begins with a UDP-GlcNAc (UDP-N-acetyl-D-glucosamine) precursor, which is first acylated (LpxA) and

then deacetylated (LpxC) (Emiola et al., 2015). Another acylation step follows (LpxD), after which the resulting UDP-2,3-diacyl-GlcN is cleaved to form 2,3-diacyl-GlcN-1 (LpxH) and condensed with another molecule of UDP-2,3-diacyl-GlcN to form lipid X (LpxB). This is followed by phosphorylation (LpxK) to form lipid IV_A and incorporation of two molecules of 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) by KdtA (sometimes called WaaA) to form Kdo_2-lipid-IV_A. In bacteria with penta-acylated lipid A (all *Bacteroides* and *Prevotella* spp. studied to date), addition of a secondary laurate chain to the lipid A by LpxL is the final step, whereas in species with hexa-acylated lipid A, an additional myristate chain is also added by LpxM.

Although the core structure is generally conserved, the acylation pattern of lipid A affects the inflammatory effects of the LPS. Studies on the difference acylation patterns have shown that penta-acylated LPS may be approximately 100x less immune stimulating than hexa-acylated LPS (Park et al., 2009). Another structural variation that has been identified to impact LPS-induced inflammation is lipid-A palmitoylation, which is associated with existing in a biofilm in *Escherichia coli* and *Pseudomonas aeruginosa* (Chalabaev et al., 2014; Ciornei et al., 2010). In these studies, biofilm and planktonic cells of the same strain induced different inflammatory responses due to these structural differences of lipid A. Whether or not *P. bivia* LPS differs structurally depending on growth mode has not been investigated.

The remaining two components of mature LPS are a core oligosaccharide region and a polysaccharide repeat region called the O-antigen. Lipid A is covalently linked to the core oligosaccharide region, which is usually made up of Kdo and heptoses (inner core) and hexoses (outer core) and may play a role in regulation of membrane permeability and biofilm formation (Clifton et al., 2016). The full-length O-antigen is only present in bacteria that have ‘smooth’ LPS and consists of repeating oligosaccharide units, whereas in bacteria with ‘rough’ LPS, this region is lacking or reduced. The structure of the O-antigen region is incredibly diverse and can play a role in a range of functions (Whitfield et al., 2020). Unlike the genes involved in lipid A biosynthesis, which are typically scattered around the genome, the genes encoding the enzymes involved in the biosynthesis of the core oligosaccharide and O-antigen regions are usually located in defined gene clusters in bacteria that have been studied thus far (Jacobson et al., 2018).

Purification of LPS historically relied on hot aqueous-phenol extraction methodology developed by Westphal and Jann, but this technique is time consuming and makes use of dangerous chemicals (Westphal & Jann, 1965). More rapid, easy-to-use protocols have been developed, that do not require the use of phenol, and can be purchased as a commercial kit, such as the Pierce Chromogenic Endotoxin Quantification Kit (ThermoFisher Scientific, USA). One common method to quantify LPS in samples is the limulus amoebocyte lysate (LAL) assay. This assay uses amoebocytes from the horseshoe crab, which undergoes a clotting cascade upon exposure to bacterial endotoxins, such as LPS (DuBose et al., 1980). Pairing this reaction with a chromogenic reaction, and the generation of a standard curve, allows for spectrophotometric quantification of LPS from bacterial samples. The purified LPS can undergo sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) to assess purity of the final sample and detect any macromolecular structural differences.

The aim of the work presented in Chapter 4 was to, (1), isolate LPS from South African *Prevotella* spp., existing planktonically and in a biofilm, (2), quantify LPS with a commercial LAL assay, (3), investigate structural differences by SDS-PAGE of LPS purified from *Prevotella* spp., and (4), identify candidate genes involved in the biosynthesis of LPS.

4.3 Methods

4.3.1 Purification of lipopolysaccharide (LPS) from *Prevotella* spp.

To purify LPS from different strains of *Prevotella* spp. a commercial LPS Isolation Kit (MAK339, Sigma-Aldrich, MO, USA) was used. LPS from *P. bivia* (n=5) and UC143_2_2, the *Prevotella melaninogenica*-related strain (n=1), were isolated. For each strain, 48-hour broth cultures, normalised to an OD_{600} of 1, were pelleted in a pre-weighed tube and the BHI was removed before the bacterial pellet was weighed on an analytical balance. Thereafter, the manufacturer's instructions were followed (Sigma-Aldrich), including addition of the lysis buffer, sonication at 2-10 watts with a continuous pulse of 3 x 30 seconds using a Q125 Sonicator (Qsonica, CT, USA) and finally addition of proteinase K (0.1mg/ml) and digestion for 60 minutes at 60°C. Lysate was centrifuged for 10 minutes at 2500 × g and stored at -80°C until further analysis.

Additionally, to isolate LPS from strains in different growth modes (biofilm or planktonic), strains (n=3) that had previously been shown to form robust biofilms (Chapter 2, section 2.4.5) were allowed to form biofilms in 24-well, flat-bottomed tissue culture plates for 120 hours (polystyrene plates, Greiner Bio-One, Monroe, NC, USA), with a starting OD_{600} of 0.2. Following incubation, wells were washed with PBS and the biofilm mass was resuspended in 1ml PBS, by scraping the wells with sterile pipette tips. Following dilution to OD_{600} of 1, the prescribed protocol was followed, as for the planktonic cells, with weighing of biofilm cell pellet following centrifugation and removal of PBS. LPS was purified from three biological repeats of each sample, in each condition (biofilm or planktonic).

4.3.2 Quantification of LPS production by *Prevotella* biofilm and planktonic cells

Quantification of LPS was undertaken using a commercial kit (Pierce Chromogenic Endotoxin Quantification Kit, ThermoFisher Scientific, USA), using the LAL methodology, following the prescribed protocol. In short, 50µl of LAL (ThermoFisher, USA) was added to the LPS samples, which had been diluted 1/10, 1/100 and1/1000 in endotoxin free water (EFW), as well as a set of *E. coli* LPS standards supplied with the kit. Following incubation at 37°C for the time indicated on the specific vial, 100µl of Chromogenic Substrate (ThermoFisher) was added to each sample, followed by an additional incubation at 37°C for 6 minutes. The reaction was then stopped with 25% (v/v) acetic acid and the optical density at 405nm measured after assay completion in a plate reader (ELx800, BioTek Inc.). The quantification was performed with three technical repeats for each of the three biological replicates of the isolated LPS. The absorbance of blank replicates was subtracted from the mean absorbance per sample. The endotoxin standards, derived from *E. coli* O111:B4, were prepared in high (0-1EU/ml) and low (0-0.1EU/ml) concentration sets and used to calculate final LPS concentration from the samples, adjusted according to their dilution factor.

4.3.3 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) for visualisation of purified LPS

The macromolecular structures of the different LPS samples were visualised using SDS-PAGE, using the PROTEAN II system (Bio-Rad, USA). Electrophoresis was conducted

using precast 4-20% Mini-PROTEAN™ TGX gels, with 1x Tris-glycine SDS running buffer, in a Mini-PROTEAN Tetra cell. To prepare samples for electrophoresis, 2x Laemmli sample buffer was added 1:1 to each sample and heated for 3 minutes at 95°C. A total of 10µl of each prepared LPS was loaded onto the gel, along with 3µl of an SDS-PAGE molecular weight marker (Precision Plus standard [BBRD1610363], BioRad, USA). Electrophoresis was conducted at 200V for 40 minutes. Gels were stained using a commercial silver staining kit (PROT-SIL2, Sigma-Aldrich, USA), using the manufacturer's recommended protocol and visualised using a GelDoc Go System (Bio-Rad, USA).

4.3.4 Identification of candidate genes involved in the biosynthesis of lipid A and the core oligosaccharide region

Translated gene products that have previously been identified to play a role in the biosynthesis of lipid A and the core oligosaccharide region in *Bacteroides thetaiotamicron* (Jacobson et al., 2018) were aligned to the annotated South African isolate genome assemblies described in Chapter 3 using blastx (Camacho et al., 2009) with a minimum query coverage cutoff of 80%. For each candidate gene identified, predicted protein products from all SA isolate genomes were aligned and the mean pairwise amino acid differences across the alignment ('p-distance') calculated using MEGA (Kumar et al., 2018).

4.4 Results

4.4.1 Quantification of LPS isolated from *Prevotella* spp.

To determine whether there were strain-level differences in total LPS production, LPS was extracted and quantified from five *P. bivia* isolates and compared to the *P. bivia* ATCC 29303^T type strain. Concentrations determined from the standard curve were normalised to the amount of bacterial pellet obtained after centrifugation in the purification protocol and the results expressed as EU/mg of bacterial mass. LPS isolated from *P. bivia* strains existing planktonically averaged 0.375 EU/mg (0.20 – 0.58 EU/mg) (Figure 4.1A). However, UC143_V2_2, the *P. melaninogenica*-relative, produced approximately 24x greater amounts of LPS than the amount produced by *P. bivia* strains (8.99 EU/mg [SE ± 0.26 EU/mg]) and significantly more than that of the *P. bivia* ATCC 29303^T type strain (Dunn's

multiple comparison test p<0.05). To examine possible differences in LPS production for planktonic vs biofilm resident cells, LPS was quantified for three isolates under both growth modes (Figure 4.1B). LPS from biofilm-resident cells, also normalised to pellet weight (mg), was significantly higher than that produced by their planktonic counterparts in all three strains (Šídák multiple comparisons test, p<0.0001). The mean concentration was 8.623 EU/mg (SE ± 0.512 EU/mg), approximately 20-fold larger than the planktonic mean concentration.

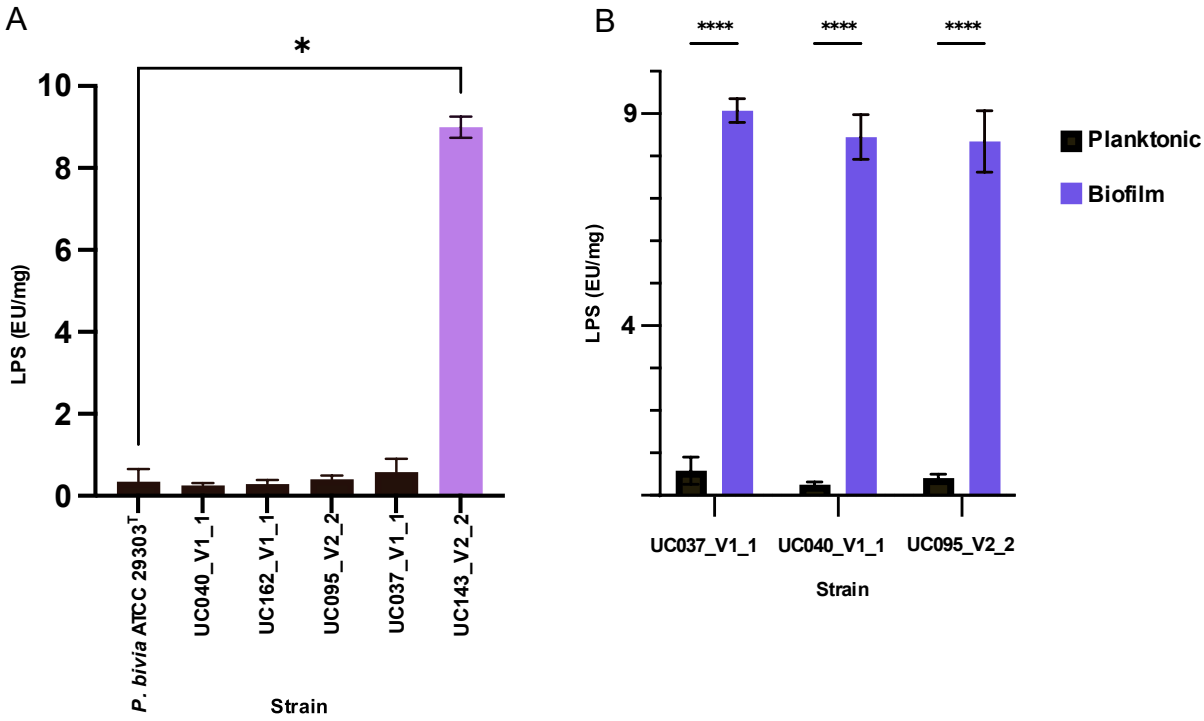


Figure 4.2*:* **Quantification of LPS by the LAL assay.** LPS production by a subset of *Prevotella* spp. (n=5 *P. bivia* and UC143_V2_2). LPS was isolated from three biological repeats and normalised according to the mg of bacterial pellet (A) LPS was quantified from planktonic cultures. Statistical analysis was performed by a Kruskal Wallis and a Dunn's Multiple comparison test to compare strain LPS production. (B) LPS was isolated from planktonic- and biofilm-existing cells from strong biofilm forming *P. bivia* strains (n=3). Statistical analysis was undertaken by 2-way ANOVA and Šídák multiple comparisons test to compare planktonic to biofilm LPS production. *p<0.05, ****p<0.0001.

4.4.2 LPS structural variations

Large-scale structural differences of LPS, as well as presence of O-antigen, can be visualised using SDS-PAGE with silver staining (Figure 4.2 and 4.3). LPS isolated from planktonic cultures had clear banding patterns near the bottom of each well, with consistent banding at ~10kD, ~20kD and ~25kD for all *P. bivia* strains tested (lanes 1-5,

Figure 4.2). LPS extracted from *P. bivia* isolates displayed similar banding patterns regardless of strain, with the exception of a ~17kD band for one isolate (UC095_V2_2) and a ~35kD band in the *P. bivia* ATCC 29303^{T} strain, which may not have been present in the other *P. bivia* isolates. However, the *P. melaninogenica*-relative, UC143_V2_2, showed a markedly different banding pattern, suggestive of a different macromolecular structure. All isolates lacked the higher MW repeat banding pattern characteristic of the O-antigen.

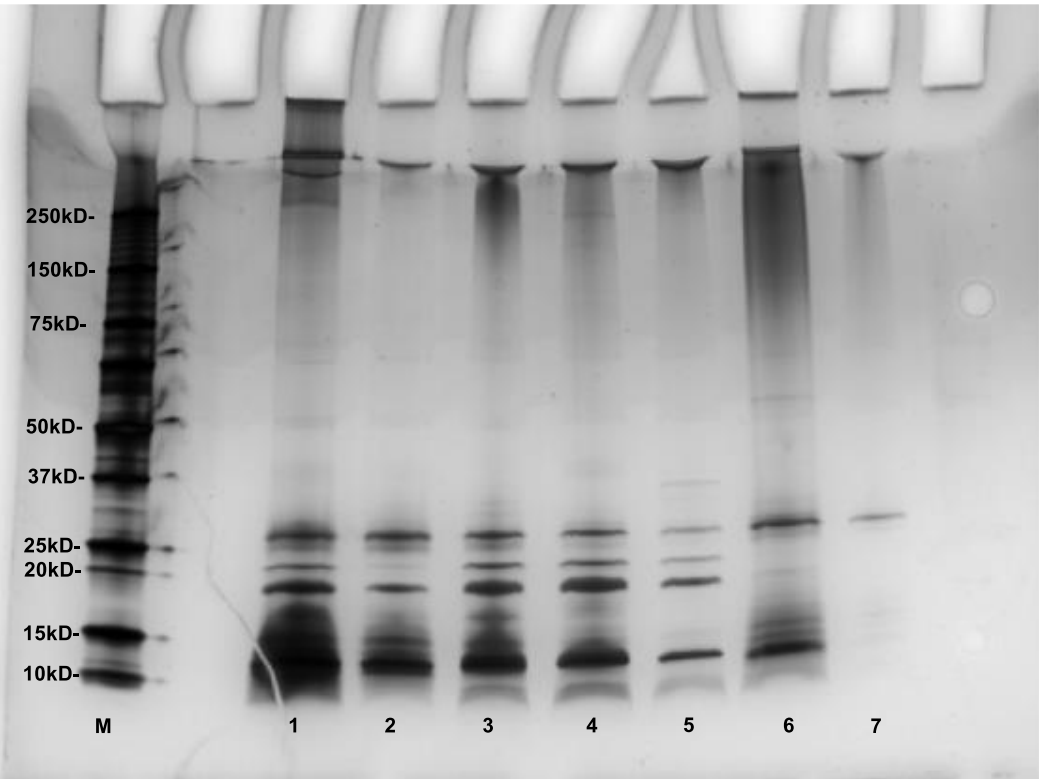


Figure 4.3: **LPS isolated from planktonic *Prevotella* spp.** SDS-PAGE was performed on *Prevotella* spp. from planktonic cultures, with a Precision Plus Protein™ Unstained standard (Bio-Rad, USA). Samples were loaded at a 1/5-1/20 dilution of the purified LPS, optimised for equal loading prior to the final gel. Samples were loaded as follows: M = Protein standard, 1 = UC037_V1_1, 2 = UC040_V1_1, 3 = UC095_V2_2, 4 = UC162_V1_1, 5 = *P. bivia* ATCC 29303^{T} type strain, 6 = UC143_V2_2 (1/10 dilution). 7 = UC143_V2_2 (1/100 dilution).

Banding patterns for LPS extracted from planktonic and biofilm-associated cells displayed similarities regardless of growth mode (Figure 4.3). Interestingly, however, biofilm LPS from strain UC095_V2_2 appeared to lack a ~20kD band, despite it being present in the planktonic LPS sample from the same strain.

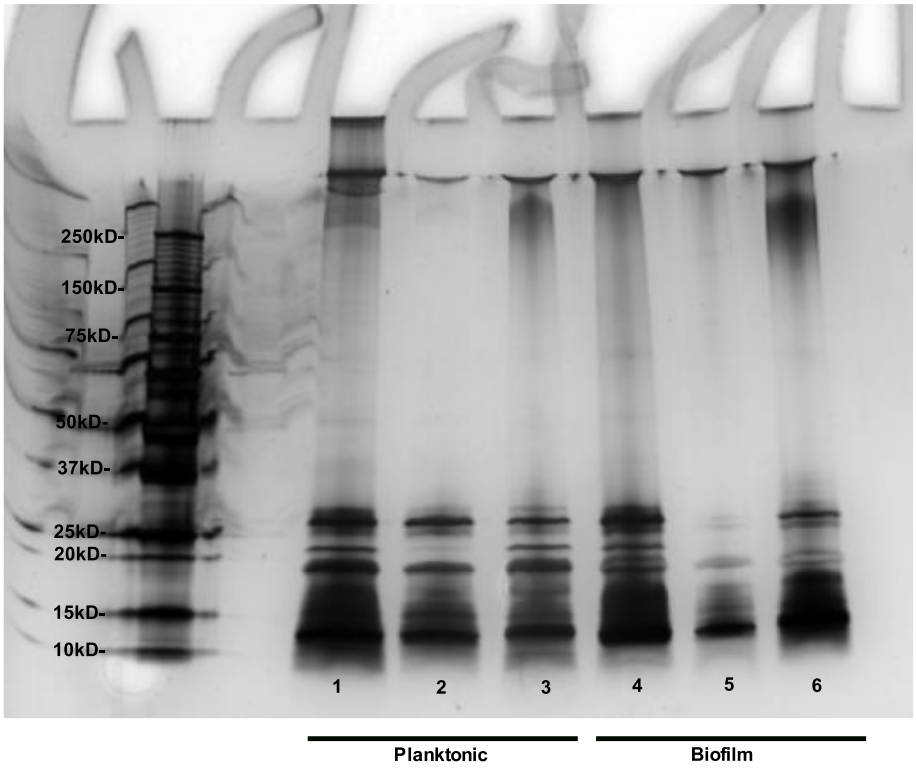


Figure 4.4: **LPS isolated from planktonic versus biofilm resident *Prevotella* spp.** SDS-PAGE was performed to compare planktonic (lanes 1-3) and biofilm (lanes 4-6) *P. bivia* LPS production. Lanes 1 & 4 = UC037_V1_1, lanes 2 & 5 = UC040_V1_1, and lanes 3 & 6 = UC095_V2_2. Electrophoresis was performed with the Precision Plus Protein™ Unstained standard (Bio-Rad, USA).

4.4.3 Identification of potential LPS biosynthetic gene clusters

The genetic basis for LPS biosynthesis has been examined previously in *Bacteroides thetaiotaomicron* (Jacobson et al., 2018). Using the genes identified in that study, genes predicted to encode proteins involved in the biosynthesis of lipid A were identified in all South African *Prevotella* spp. genomes, as well as the *P. bivia* ATCC 29303^T and *P. melaninogenica* ATCC 25845^T type strain genomes (Table 4.1). Amino acid identities of the encoded proteins to the *B. thetaiotaomicron* homologues ranged from 50-75%. In all cases, the predicted proteins were highly conserved, with little genotypic variation observed between isolates of the same species, with pairwise distance for the identified sequence products ranging between 0.00 and 0.01.

Table 4.1: **Genes encoding candidate proteins involved in the biosynthesis of lipid A**

Gene	Locus tag in reference genome*		Amino acid identity to *B. thetaiotaomicron* protein (%)		Number of SA *Prevotella* spp. isolates harbouring gene	Mean p-distance†
	P. bivia	*P. melaninogenica*	*P. bivia*	*P. melaninogenica*		
lpxA	PREBIDRAFT_RS10495	HMPREF0659_RS04090	63.14	65.49	35/35	0.00
lpxC	PREBIDRAFT_RS10500	HMPREF0659_RS04085	70.07	73.48	35/35	0.01
lpxD	PREBIDRAFT_RS10505	HMPREF0659_RS04080	63.95	64.24	35/35	0.00
lpxH	PREBIDRAFT_RS05190	HMPREF0659_RS06125	68.5	71.15	35/35	0.00
lpxB	PREBIDRAFT_RS07360	HMPREF0659_RS05385	64.29	66.22	35/35	0.00
lpxK	PREBIDRAFT_RS04285	HMPREF0659_RS04200	50.68	52.22	35/35	0.01
kdtA	PREBIDRAFT_RS06480	HMPREF0659_RS01420	58.27	57.21	35/35	0.01
lpxL	PREBIDRAFT_RS08810	HMPREF0659_RS07735	63.19	66.59	35/35	0.00
lpxM	N.D.	N.D.	N.D.	N.D.	0/35	N.D.

* Assembly accessions: *P. bivia* ATCC 29303^T (GCF_000262545.1), *P. melaninogenica* ATCC 25845^T (GCF_000144405.1)

† The mean proportion of pairwise amino acid differences for all identified sequence products

Alignment of the gene cluster (BT3362-BT3380) identified by Jacobson et al. (2018) to be involved in the biosynthesis of the LPS oligonucleotide core with *P. bivia* and *P. melaninogenica* genomes did not yield candidate gene clusters with high homology. However, one *B. thetaiotaomicron* gene product (BT3376), predicted to encode a tailoring enzyme, did show similarity to a predicted DegT/DnrJ/EryC1/StrS family aminotransferase (PREBIDRAFT_RS07545) in the *P. bivia* ATCC 29303^T genome. Investigation of the genomic context around this gene identified a gene cluster with several genes predicted to encode glycosyltransferases, acyltransferases, and lipopolysaccharide biosynthesis proteins (Figure 4.4).

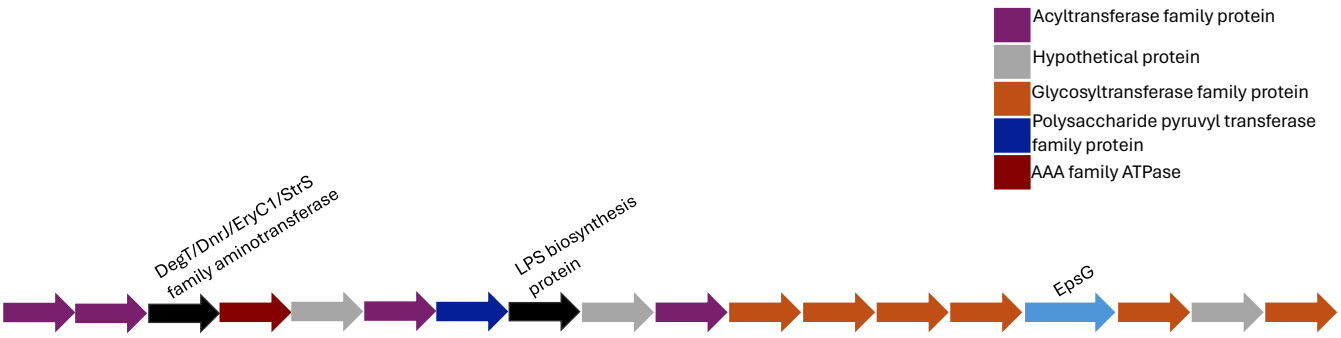


Figure 4.5: **Schematic of a putative core oligosaccharide biosynthetic gene cluster.** Putative LPS core oligosaccharide biosynthesis genes were identified in the annotated genome for the *P. bivia* ATCC 29303^{T} type strain. Genes are coloured according to putative protein function.

4.5 Discussion

Lipopolysaccharide is a highly antigenic component of the *Prevotella* spp. cell wall that is proposed to play a role in BV-mediated vaginal inflammation. The BV microbiota may contain several Gram-negative species, including *Sneathia* spp. and *Megasphaera* spp., but *Prevotella* spp. are regarded as a prominent source of vaginal LPS (Aroutcheva et al., 2008; Ravel et al., 2011). Therefore, in this final chapter, we aimed to analyse LPS production by vaginally derived *Prevotella* spp. isolates.

Prevotella bivia species isolated in this project produced comparable amounts of LPS *in vitro*, with no significant differences in LPS production amongst the strains. Additionally, the quantity of the biofilm LPS was consistent among the different strains, suggesting that intra-species LPS production is similar, in both the planktonic and biofilm growth states. There was, however, an inter-species difference when comparing LPS from *P. bivia* and UC143_V2_2, the potentially novel *P. melaninogenica*-related strain with the latter isolate producing almost 24 times more LPS than the *P. bivia* strains. This may imply that this strain and/or species has the potential to contribute to vaginal inflammation based on the quantity of LPS. However, it is important to note that differences in LPS structure significantly affect inflammation potential. Interestingly, *P. melaninogenica* LPS has been shown to be less inflammatory than LPS produced by other species, such as *Porphyromonas gingivalis, Fusobacterium nucleatum, Pseudomonas aeruginosa* or

Salmonella typhimurium (Council, 2013; Rossano et al., 1993). Therefore, cell culture experimentation is required to assess inflammation induced by UC143_V2_2.

Prolific biofilm-forming strains were selected to compare the LPS produced between biofilm and planktonic *P. bivia* strains. In this project, the quantity of LPS produced per mg weight from biofilm strains was significantly higher than the quantity produced by planktonic cells. However, there are important factors that may have affected these calculations. Biofilm growth is fundamentally different to planktonic growth, and it is estimated that microbial cells only constitute 10-25% of the mature biofilm, with the primary structural component being extracellular matrix (Rather et al., 2021). Therefore, it is difficult to normalise LPS production using pellet weight alone. Attempts to normalise using plate counts of viable cells are also challenging, due to the difficulties associated with separating individual cells in the biofilm and the fact that non-viable cells still possess LPS. Finally, qPCR-based normalisation approaches are complicated by the presence of larger amounts of extracellular DNA within the biofilm. With these limitations in mind, it is nevertheless interesting that LPS production appeared to be up to 20-fold higher in biofilm-associated growth. Previous studies in *P. aeruginosa* have suggested that LPS can play a vital role in adhesion of Gram-negative bacteria to both abiotic and biotic surfaces (Abu-Lail et al., 2007). Thus, the presence of larger amounts of endotoxin may assist biofilm establishment.

Studies investigating LPS and its relationship to biofilms further have shown that certain LPS structural modifications can lead to more stable biofilm formation. For example, palmitate acyl chains added to lipid A in *E. coli* and *P. aeruginosa* may increase hydrophobicity of LPS and encourage adherence to surfaces (Chalabaev et al., 2014; Ciornei et al., 2010; Nakao et al., 2012). Structural modifications akin to these may also result in altered immune stimulation, affecting the impact of LPS on inflammation (Ciornei et al., 2010). These structural variations can be detected on SDS-PAGE of the extracted LPS (Chalabaev et al., 2014). Therefore, since it is not known whether there are LPS structural differences between biofilm and planktonic LPS in *Prevotella* spp., a preliminary analysis of LPS structure between LPS from these growth modes was conducted. Considering the apparently similar banding pattern of the planktonic and biofilm-sourced LPS by SDS-PAGE, there was no clear evidence that growth mode is

associated with distinct structural variations in the LPS of strains analysed in this study. LPS extracted from one strain (UC095_V2_2) appeared to lack a ~20kD band that was present in its planktonic form, which may suggest some structural dissimilarity. However, incorporation of palmitoylate into lipid A usually results in the addition of SDS-PAGE bands, rather than their loss. Therefore, further analysis using more specific tools, such as mass spectrometry, are required to determine the exact modification. Interestingly, the *P. melaninogenica*-related strain, UC143_V2_2, had a different banding pattern to the *P. bivia* isolates, indicating that structural variations may exist between the different *Prevotella* species.

The difference between "smooth"(S)- and "rough"(R)-type LPS structures lies in the presence or absence of the O-antigen component (Kilár et al., 2013). Of the *Bacteroidetes*, only *Bacteroides vulgatus* has been identified as potentially harbouring S-type LPS with previously characterised *P. bivia* and *P. melaninogenica* strains being shown to have R-type LPS (Eidhin & Mouton, 1993). Similarly, the SDS-PAGE results here suggested R-type LPS for all *Prevotella* spp. isolates included in the analysis as shown by the absence of the characteristic ladder-patterning observed for S-type LPS. The R- or S-state of the LPS can impact signalling with the host cells. A study in *Salmonella* has shown that R-LPS is less inflammatory than its smooth counterpart, although other studies have suggested that they have the same inflammatory effect *in vivo*, but act by different signalling pathways (Sali et al., 2019; Zanoni et al., 2012).

Another structural variation in LPS that is proposed to alter immunogenicity is the acylation status of lipid A. Studies have shown that LPS with hexa-acylated lipid A elicits a stronger innate immune response than penta-, tetra- or tri-acylated lipid A LPS (Li et al., 2013). Genes encoding products with similarity to each of the enzymes involved in the biosynthesis of the core lipid A moiety were identified in all isolate genomes. However, no genomes harboured putative *lpxM* homologues, suggesting that, as with other *Bacteroides* and *Prevotella* spp., the isolates in this study had penta-acylated LPS (Fathy Mohamed et al., 2017). A candidate gene cluster encoding enzymes involved in the biosynthesis of the core oligosaccharide was also identified in *P. bivia* isolate genomes. This cluster harboured a LPS biosynthetic gene, glycosyltransferases, acyltransferases, and a polysaccharide pyruvyl transferase family protein, with similar functions to the

oligosaccharide gene cluster previously identified in *B. thetaiotaomicron* (Jacobson et al., 2018). In order to characterise this gene cluster further and confirm a role in the biosynthesis of LPS, individual genes could be inactivated via mutagenesis and the effect on LPS and core oligosaccharide explored (Jacobson et al., 2018). Regions of homology to the putative gene cluster detected in the *P. bivia* type strain were detected in some *P. bivia* clinical isolate genomes, however due to assembly contiguation and the presence of some of the genes at the end of a contig, full-length gene clusters were difficult to identify in many of the strains.

One important aspect for further research, which was beyond the scope of the present study, is examining LPS-induced inflammation in vaginal cell culture models. Measurement of cytokine production in cells stimulated using equimolar amounts of LPS purified from different *Prevotella* spp. isolates, as well as LPS extracted from biofilm and planktonic grown cells, will help to elucidate their role in inflammatory processes. Additionally, since multiple LPS-producing bacterial strains co-exist in the vagina, investigating the effect of multiple LPS types on inflammation could be valuable. Furthermore, as mentioned above, there are important follow up experiments that can be performed to confirm the structural variations predicted by SDS-PAGE analysis. Nevertheless, the work presented in this chapter provides useful preliminary insights into the structure and variation in LPS produced by different *Prevotella* spp. isolates. Overall, the results suggest that biofilm formation does not lead to obvious structural differences in LPS and that at a species level, LPS structure is quite similar. Additionally, it appears that *P. bivia* and the *P. melaninogenica*-related strain have R-type LPS that is penta-acylated and that the latter isolate may produce significantly more LPS than *P. bivia* strains in the planktonic state.

Chapter 5: Conclusion

BV is highly prevalent in South African women, and its association with STI and HIV acquisition, as well as with obstetric and gynaecologic complications, makes it an important research focus. *Prevotella* is one genus that is commonly isolated from women with BV. One barrier to BV treatment and prevention plans is that little research has focused on isolation and characterisation of BV bacteria, including *Prevotella* spp., from sub-Saharan Africa and there is a marked lack of South African *Prevotella* spp. sequence data in publicly available genome databases.

The aim of this project was to isolate and characterise *Prevotella* spp. from South African adolescent girls and young women. In total, 69 isolates were obtained (67 *Prevotella bivia* and 2 *Prevotella melaninogenica* isolates) and identified by 16S rRNA and *cpn60* gene sequencing. Amplicon sequencing provided some evidence that the *cpn60* gene in South African strains did not differ greatly from the gene in international reference databases. As expected, strains from the same sample were largely more similar to each other than to strains from other samples or other individuals, but at least one sample had distinctly different strains at the same time, suggesting heterogeneity in this individual and supporting previous studies suggesting intra-vaginal strain diversity (Ma et al., 2020). Further investigation of strains across visit times could be valuable to determine the level of persistence of bacterial strains longitudinally and if specific strains play a role in BV persistence. Interestingly, the preliminarily assigned *P. melaninogenica* isolate (UC143_V2_2) had a *cpn60* gene most similar to the only other publicly available vaginal *P. melaninogenica*, both of which clustered separately from the other *P. melaninogenica* strains, providing further evidence that UC143_V2_2 may be distinct from the 'core' *P. melaninogenica* group.

Resistance to antimicrobials has been proposed as a factor contributing to BV treatment failure and post-treatment relapse. This study showed that AMR to SOC BV antibiotics, metronidazole, and clindamycin, was low in South African *Prevotella* spp. isolates. While metronidazole resistance is generally low for the genus *Prevotella,* clindamycin resistance was markedly lower in this study than has been previously reported for vaginal *Prevotella* spp. in other studies (Petrina et al., 2019). These results are promising for

continuing use of SOC BV antibiotics in South Africa, but since this is a small sample size, responsible use of the antibiotics should still be advocated.

Biofilm analysis showed variable patterns in biofilm formation, suggesting strain-level differences, even amongst strains isolated from the same individual. Addition of *P. bivia* to a *G. vaginalis* biofilm did not result in an enhancement of biofilm biomass, contrary to previous reports. The basic *in vitro* biofilm assay approach used in this study has been widely employed in other studies, however, it is important to note that the method is not without limitations. Alternatives such as the Calgary device (Ceri et al., 1999) and even microfluidic systems (Jung et al., 2015) have been proposed to overcome some of the limitations. Future work to examine the ultrastructure of biofilms described in this study using scanning electron microscopy are currently underway.

WGS has the potential to provide valuable insights into the genotypic diversity of bacterial pathobionts. Core- and pangenome analyses revealed fine-grained diversity between *Prevotella* spp. isolates, detecting strain-differences that were not reflected in *cpn60* amplicon sequencing. Systematic pan-genomic differences were not observed between South African and non-South African isolates, and there were no significant gene differences between isolates from BV-positive and BV-negative women. To date, a detailed, strain-level analysis of *P. melaninogenica* genome diversity has not been reported. The work in this study indicates that *P. melaninogenica* genomes in the NCBI RefSeq database, along with the UC143_V2_2 strain isolated in this study, may represent up to four separate genomospecies and might explain the very large pangenome of *P. melaninogenica,* since removal of these isolates from the analysis resulted in a smaller pangenome. Future WGS investigation may include a long-read sequencing approach (like PacBio or Oxford Nanopore sequencing) to generate more complete, closed assemblies.

AMR gene analysis revealed previously identified antimicrobial markers for metronidazole, clindamycin, tetracyclines and β-lactams. Importantly, one strain contained a potential metronidazole resistance gene (*nimK*) on a mobilizable transposon, which has been identified previously in *P. bivia* of Dutch origin. The fact that this gene was found in a strain with reduced metronidazole susceptibility tends to support a functional role for it in 5-nitroimidazole resistance, however additional

experimentation could confirm this by ruling out other mechanisms of nitroimidazole resistance. A clindamycin-resistance determinant gene, *ermF,* was also detected in a clindamycin-resistant strain isolated from a woman that received azithromycin prior to sampling. It would be interesting to determine whether this strain was present before the treatment and persistent throughout, however there were no earlier timepoints available for this woman. A potential determinant for doxycycline resistance, *tetQ*, was detected at high rates, and while this may not impact *Prevotella* spp. directly, the potential carriage of this gene on mobile genetic elements poses a risk of transfer, with *Prevotella* spp. acting as a reservoir. Regions of homology to a CTnDOT-family transposon, CTn341, were identified in several strains that harboured the *tetQ* gene. This is especially concerning considering the increasing use of doxycycline as a pre-exposure prophylaxis for many STIs and determining whether *tetQ* is transposable would be interesting future work. Interestingly, different variants of another AMR gene, *cfxA*, were detected in >70% of strains, with the *cfxA3/A4* variant associated with significantly higher MIC values, similar to what has been reported recently (Yokoyama et al., 2023).

The South Africa *Prevotella* spp. genomes contained 13 high-quality potential prophages. To-date, vaginal *Prevotella* spp. phages have not been induced *in vitro* and future work will involve attempts to do this using prophage inducers to characterise them further. Diverse anti-phage defence mechanisms were predicted for all genomes. Almost all strains contained CRISPR/Cas Class 2 subtype IIC systems, with the next most common anti-phage mechanism being PARIS (a form of abortive infection) and restriction modification systems. These systems have not been well described in *Prevotella* spp. previously and it would be interesting to determine whether they are functional in these strains.

Finally, investigation of the most notorious *Prevotella* spp. virulence factor, LPS, was performed, as it is proposed to play a primary role in the vaginal inflammation associated with BV. LPS was isolated from a subset of *Prevotella* spp. and, contrary to previous studies in other Gram-negative bacteria, the structure of the LPS did not appear different between growth modes (pelagic vs biofilm). Whether this is due to species-level differences between the Gram-negative bacteria in this study versus others, or due to the different environmental sources of the species is not known. The novel *P.*

melaninogenica-relative produced significantly more LPS than *P. bivia* strains and appeared structurally different. More work will need to be done to explore LPS inflammatory and structural differences in different *Prevotella* spp. Additionally, all *Prevotella* spp. isolate genomes appeared to produce R-type LPS, based on the marked lack of the characteristic O-antigen ladder patterning on SDS-PAGE. Lipid A biosynthetic genes, and one core oligosaccharide gene cluster, was identified in the *Prevotella* spp. isolate genomes, and in line with prior studies, the LPS was predicted to be penta-acylated due to the absence of a *lpxM* gene homologue.

In conclusion, the phenotypic and genotypic characterisation of South African *P. bivia* isolates, and a potentially novel species, described in this thesis, contributes to the creation of a local *Prevotella* spp. biorepository that will be available for future BV research both within the region and internationally. Isolate genomic information will be made publicly available on PubMed and isolates will be stored in-lab at UCT Faculty of Health Sciences. The findings of this study provide several interesting avenues for future research, some of which are actively being pursued, to better understand the role of *Prevotella* in BV and the adverse health outcomes associated with the condition.

References

Abbai, N. S., Reddy, T., & Ramjee, G. (2016). Prevalent bacterial vaginosis infection—A risk factor for incident sexually transmitted infections in women in Durban, South Africa. *International Journal of STD & AIDS*, *27*(14), 1283–1288. https://doi.org/10.1177/0956462415616038

Abby, S. S., Néron, B., Ménager, H., Touchon, M., & Rocha, E. P. C. (2014). MacSyFinder: A Program to Mine Genomes for Molecular Systems with an Application to CRISPR-Cas Systems. *PLoS ONE*, *9*(10), e110726. https://doi.org/10.1371/journal.pone.0110726

Abdill, R. J., Adamowicz, E. M., & Blekhman, R. (2022). Public human microbiome data are dominated by highly developed countries. *PLoS Biology*, *20*(2), e3001536. https://doi.org/10.1371/journal.pbio.3001536

Abu-Lail, L. I., Liu, Y., Atabek, A., & Camesano, T. A. (2007). Quantifying the adhesion and interaction forces between *Pseudomonas aeruginosa* and natural organic matter. *Environmental Science & Technology*, *41*(23), 8031–8037. https://doi.org/10.1021/es071047o

Achilles, S. L., Austin, M. N., Meyn, L. A., Mhlanga, F., Chirenje, Z. M., & Hillier, S. L. (2018). Impact of contraceptive initiation on vaginal microbiota. *American Journal of Obstetrics and Gynecology*, *218*(6), 622.e1-622.e10. https://doi.org/10.1016/j.ajog.2018.02.017

Aiyar, A., Quayle, A. J., Buckner, L. R., Sherchand, S. P., Chang, T. L., Zea, A. H., Martin, D. H., & Belland, R. J. (2014). Influence of the tryptophan-indole-IFNγ axis on human genital *Chlamydia trachomatis* infection: Role of vaginal co-infections. *Frontiers in Cellular and Infection Microbiology*, *4*, 72. https://doi.org/10.3389/fcimb.2014.00072

Alauzet, C., Lozniewski, A., & Marchandin, H. (2019). Metronidazole resistance and nim genes in anaerobes: A review. *Anaerobe*, *55*, 40–53. https://doi.org/10.1016/j.anaerobe.2018.10.004

Alauzet, C., Mory, F., Teyssier, C., Hallage, H., Carlier, J. P., Grollier, G., & Lozniewski, A. (2010). Metronidazole Resistance in *Prevotella* spp. And Description of a New nim Gene in *Prevotella baroniae*. *Antimicrobial Agents and Chemotherapy*, *54*(1), 60–64. https://doi.org/10.1128/AAC.01003-09

Albaghdadi, S. Z., Altaher, J. B., Drobiova, H., Bhardwaj, R. G., & Karched, M. (2021). In vitro Characterization of Biofilm Formation in *Prevotella* Species. *Frontiers in Oral Health*, *2*, 724194. https://doi.org/10.3389/froh.2021.724194

Alcock, B. P., Huynh, W., Chalil, R., Smith, K. W., Raphenya, A. R., Wlodarski, M. A., Edalatmand, A., Petkau, A., Syed, S. A., Tsang, K. K., Baker, S. J. C., Dave, M., McCarthy, M. C., Mukiri, K. M., Nasir, J. A., Golbon, B., Imtiaz, H., Jiang, X., Kaur, K., … McArthur, A. G. (2023). CARD 2023: Expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Research*, *51*(D1), D690–D699. https://doi.org/10.1093/nar/gkac920

Ali, S., Dennehy, F., Donoghue, O., & McNicholas, S. (2022). Antimicrobial susceptibility patterns of anaerobic bacteria at an Irish University Hospital over a ten-year period (2010–2020). *Anaerobe*, *73*, 102497. https://doi.org/10.1016/j.anaerobe.2021.102497

Al-Memar, M., Bobdiwala, S., Fourie, H., Mannino, R., Lee, Y. S., Smith, A., Marchesi, J. R., Timmerman, D., Bourne, T., Bennett, P. R., & MacIntyre, D. A.

(2020). The association between vaginal bacterial composition and miscarriage: A nested case–control study. *Bjog*, *127*(2), 264. https://doi.org/10.1111/1471-0528.15972

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

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. *The American Journal of Medicine*, *74*(1), 14–22. https://doi.org/10.1016/0002-9343(83)91112-9

Anahtar, M. N., Byrne, E. H., Doherty, K. E., Bowman, B. A., Yamamoto, H. S., Soumillon, M., Padavattan, N., Ismail, N., Moodley, A., Sabatini, M. E., Ghebremichael, M. S., Nusbaum, C., Huttenhower, C., Virgin, H. W., Ndung'u, T., Dong, K. L., Walker, B. D., Fichorova, R. N., & Kwon, D. S. (2015). Cervicovaginal Bacteria Are a Major Modulator of Host Inflammatory Responses in the Female Genital Tract. *Immunity*, *42*(5), 965–976. https://doi.org/10.1016/j.immuni.2015.04.019

Anani, H., Zgheib, R., Hasni, I., Raoult, D., & Fournier, P.-E. (2020). Interest of bacterial pangenome analyses in clinical microbiology. *Microbial Pathogenesis*, *149*, 104275. https://doi.org/10.1016/j.micpath.2020.104275

Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data* [Computer software]. http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Arnold, K. B., Burgener, A., Birse, K., Romas, L., Dunphy, L. J., Shahabi, K., Abou, M., Westmacott, G. R., McCorrister, S., Kwatampora, J., Nyanga, B., Kimani, J., Masson, L., Liebenberg, L. J., Abdool Karim, S. S., Passmore, J.-A. S., Lauffenburger, D. A., Kaul, R., & McKinnon, L. R. (2016). Increased levels of inflammatory cytokines in the female reproductive tract are associated with altered expression of proteases, mucosal barrier proteins, and an influx of HIV-susceptible target cells. *Mucosal Immunology*, *9*(1), 194–205. https://doi.org/10.1038/mi.2015.51

Aroutcheva, A., Ling, Z., & Faro, S. (2008). *Prevotella bivia* as a Source of Lipopolysaccharide in the Vagina. *Anaerobe*, *14*(5), 256–260. https://doi.org/10.1016/j.anaerobe.2008.08.002

Arroyo-Moreno, S., Cummings, M., Corcoran, D. B., Coffey, A., & McCarthy, R. R. (2022). Identification and characterization of novel endolysins targeting *Gardnerella vaginalis* biofilms to treat bacterial vaginosis. *Npj Biofilms and Microbiomes*, *8*(1), Article 1. https://doi.org/10.1038/s41522-022-00285-0

Atashili, J., Poole, C., Ndumbe, P. M., Adimora, A. A., & Smith, J. S. (2008). Bacterial vaginosis and HIV acquisition: A meta-analysis of published studies. *AIDS (London, England)*, *22*(12), 1493–1501. https://doi.org/10.1097/QAD.0b013e3283021a37

Austin, M. N., Beigi, R. H., Meyn, L. A., & Hillier, S. L. (2005). Microbiologic Response to Treatment of Bacterial Vaginosis with Topical Clindamycin or Metronidazole. *Journal of Clinical Microbiology*, *43*(9), 4492–4497. https://doi.org/10.1128/JCM.43.9.4492-4497.2005

Bacic, M., Parker, A. C., Stagg, J., Whitley, H. P., Wells, W. G., Jacob, L. A., & Smith, C. J. (2005). Genetic and Structural Analysis of the Bacteroides Conjugative Transposon CTn341. *Journal of Bacteriology*, *187*(8), 2858–2869. https://doi.org/10.1128/JB.187.8.2858-2869.2005

Baker, D. N., & Langmead, B. (2019). Dashing: Fast and accurate genomic distances with HyperLogLog. *Genome Biology*, *20*(1), 265. https://doi.org/10.1186/s13059-019-1875-0

Baldwin, E. A., Walther-Antonio, M., MacLean, A. M., Gohl, D. M., Beckman, K. B., Chen, J., White, B., Creedon, D. J., & Chia, N. (2015). Persistent microbial dysbiosis in preterm premature rupture of membranes from onset until delivery. *PeerJ*, *3*, e1398. https://doi.org/10.7717/peerj.1398

Balle, C., Esra, R., Havyarimana, E., Jaumdally, S. Z., Lennard, K., Konstantinus, I. N., Barnabas, S. L., Happel, A.-U., Gill, K., Pidwell, T., Lingappa, J. R., Gamieldien, H., Bekker, L.-G., Passmore, J.-A. S., & Jaspan, H. B. (2020). Relationship between the Oral and Vaginal Microbiota of South African Adolescents with High Prevalence of Bacterial Vaginosis. *Microorganisms*, *8*(7), Article 7. https://doi.org/10.3390/microorganisms8071004

Balle, C., Konstantinus, I. N., Jaumdally, S. Z., Havyarimana, E., Lennard, K., Esra, R., Barnabas, S. L., Happel, A.-U., Moodie, Z., Gill, K., Pidwell, T., Karaoz, U., Brodie, E., Maseko, V., Gamieldien, H., Bosinger, S. E., Myer, L., Bekker, L.-G., Passmore, J.-A. S., & Jaspan, H. B. (2020). Hormonal contraception alters vaginal microbiota and cytokines in South African adolescents in a randomized trial. *Nature Communications*, *11*, 5578. https://doi.org/10.1038/s41467-020-19382-9

Baron, E. J. (1996). Classification. In S. Baron (Ed.), *Medical Microbiology* (4th ed.). University of Texas Medical Branch at Galveston. http://www.ncbi.nlm.nih.gov/books/NBK8406/

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science (New York, N.Y.)*, *315*(5819), 1709–1712. https://doi.org/10.1126/science.1138140

Beigi, R. H., Austin, M. N., Meyn, L. A., Krohn, M. A., & Hillier, S. L. (2004). Antimicrobial resistance associated with the treatment of bacterial vaginosis. *American Journal of Obstetrics and Gynecology*, *191*(4), 1124–1129. https://doi.org/10.1016/j.ajog.2004.05.033

Belyakov, I. M., & Berzofsky, J. A. (2004). Immunobiology of Mucosal HIV Infection and the Basis for Development of a New Generation of Mucosal AIDS Vaccines. *Immunity*, *20*(3), 247–253. https://doi.org/10.1016/S1074-7613(04)00053-6

Betancur-Murillo, C. L., Aguilar-Marín, S. B., & Jovel, J. (2023). Prevotella: A Key Player in Ruminal Metabolism. *Microorganisms*, *11*(1). https://doi.org/10.3390/microorganisms11010001

Boskey, E. R., Cone, R. A., Whaley, K. J., & Moench, T. R. (2001). Origins of vaginal acidity: High D/L lactate ratio is consistent with bacteria being the primary source. *Human Reproduction (Oxford, England)*, *16*(9), 1809–1813. https://doi.org/10.1093/humrep/16.9.1809

Boskey, E. R., Telsch, K. M., Whaley, K. J., Moench, T. R., & Cone, R. A. (1999). Acid production by vaginal flora in vitro is consistent with the rate and extent of vaginal acidification. *Infection and Immunity*, *67*(10), 5170–5175. https://doi.org/10.1128/IAI.67.10.5170-5175.1999

Bouras, G., Nepal, R., Houtak, G., Psaltis, A. J., Wormald, P.-J., & Vreugde, S. (2023). Pharokka: A fast scalable bacteriophage annotation tool. *Bioinformatics*, *39*(1), btac776. https://doi.org/10.1093/bioinformatics/btac776

Bradshaw, C. S., Morton, A. N., Hocking, J., Garland, S. M., Morris, M. B., Moss, L. M., Horvath, L. B., Kuzevska, I., & Fairley, C. K. (2006). High recurrence rates of bacterial vaginosis over the course of 12 months after oral metronidazole therapy and factors associated with recurrence. *The Journal of Infectious Diseases*, *193*(11), 1478–1486. https://doi.org/10.1086/503780

Briselden, A. M., Moncla, B. J., Stevens, C. E., & Hillier, S. L. (1992). Sialidases (neuraminidases) in bacterial vaginosis and bacterial vaginosis-associated microflora. *Journal of Clinical Microbiology*, *30*(3), 663–666.

Brown-Jaque, M., Calero-Cáceres, W., & Muniesa, M. (2015). Transfer of antibiotic-resistance genes via phage-related mobile elements. *Plasmid*, *79*, 1–7. https://doi.org/10.1016/j.plasmid.2015.01.001

Brumfield, K. D., Huq, A., Colwell, R. R., Olds, J. L., & Leddy, M. B. (2020). Microbial resolution of whole genome shotgun and 16S amplicon metagenomic sequencing using publicly available NEON data. *PLOS ONE*, *15*(2), e0228899. https://doi.org/10.1371/journal.pone.0228899

Brynildsrud, O., Bohlin, J., Scheffer, L., & Eldholm, V. (2016). Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. *Genome Biology*, *17*(1), 238. https://doi.org/10.1186/s13059-016-1108-8

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*, 421. https://doi.org/10.1186/1471-2105-10-421

Carlier, J. P., Sellier, N., Rager, M. N., & Reysset, G. (1997). Metabolism of a 5-nitroimidazole in susceptible and resistant isogenic strains of Bacteroides fragilis. *Antimicrobial Agents and Chemotherapy*, *41*(7), 1495–1499. https://doi.org/10.1128/AAC.41.7.1495

Carter, K. A., Balkus, J. E., Anzala, O., Kimani, J., Hoffman, N. G., Fiedler, T. L., Mochache, V., Fredricks, D. N., McClelland, R. S., & Srinivasan, S. (2022). Associations Between Vaginal Bacteria and Bacterial Vaginosis Signs and Symptoms: A Comparative Study of Kenyan and American Women. *Frontiers in Cellular and Infection Microbiology*, *12*. https://www.frontiersin.org/articles/10.3389/fcimb.2022.801770

Castillo, Y., Delgadillo, N. A., Neuta, Y., Hernández, A., Acevedo, T., Cárdenas, E., Montaño, A., Lafaurie, G. I., & Castillo, D. M. (2022). Antibiotic Susceptibility and Resistance Genes in Oral Clinical Isolates of *Prevotella intermedia*, *Prevotella nigrescens*, and *Prevotella melaninogenica*. *Antibiotics*, *11*(7), 888. https://doi.org/10.3390/antibiotics11070888

Castro, J., Rosca, A. S., Muzny, C. A., & Cerca, N. (2021). Atopobium vaginae and *Prevotella bivia* Are Able to Incorporate and Influence Gene Expression in a Pre-Formed Gardnerella vaginalis Biofilm. *Pathogens*, *10*(2), 247. https://doi.org/10.3390/pathogens10020247

Cauci, S., & Culhane, J. F. (2011). High sialidase levels increase preterm birth risk among women who are bacterial vaginosis-positive in early gestation. *American Journal of Obstetrics and Gynecology*, *204*(2), 142.e1-9. https://doi.org/10.1016/j.ajog.2010.08.061

Cauci, S., Hitti, J., Noonan, C., Agnew, K., Quadrifoglio, F., Hillier, S. L., & Eschenbach, D. A. (2002). Vaginal hydrolytic enzymes, immunoglobulin A against *Gardnerella vaginalis* toxin, and risk of early preterm birth among women in preterm labor with bacterial vaginosis or intermediate flora. *American Journal of Obstetrics and Gynecology*, *187*(4), 877–881. https://doi.org/10.1067/mob.2002.127454

Ceri, H., Olson, M. E., Stremick, C., Read, R. R., Morck, D., & Buret, A. (1999). The Calgary Biofilm Device: New Technology for Rapid Determination of Antibiotic Susceptibilities of Bacterial Biofilms. *Journal of Clinical Microbiology*, *37*(6), 1771–1776.

Chalabaev, S., Chauhan, A., Novikov, A., Iyer, P., Szczesny, M., Beloin, C., Caroff, M., & Ghigo, J.-M. (2014). Biofilms formed by gram-negative bacteria undergo increased lipid a palmitoylation, enhancing in vivo survival. *mBio*, *5*(4), e01116-14. https://doi.org/10.1128/mBio.01116-14

Chee, W. J. Y., Chew, S. Y., & Than, L. T. L. (2020). Vaginal microbiota and the potential of Lactobacillus derivatives in maintaining vaginal health. *Microbial Cell Factories*, *19*(1), 203. https://doi.org/10.1186/s12934-020-01464-4

Chen, L., Yang, J., Yu, J., Yao, Z., Sun, L., Shen, Y., & Jin, Q. (2005). VFDB: A reference database for bacterial virulence factors. *Nucleic Acids Research*, *33*(Database issue), D325-328. https://doi.org/10.1093/nar/gki008

Cherpes, T. L., Meyn, L. A., Krohn, M. A., Lurie, J. G., & Hillier, S. L. (2003). Association between Acquisition of Herpes Simplex Virus Type 2 in Women and Bacterial Vaginosis. *Clinical Infectious Diseases*, *37*(3), 319–325. https://doi.org/10.1086/375819

Chetwin, E., Manhanzva, M. T., Abrahams, A. G., Froissart, R., Gamieldien, H., Jaspan, H., Jaumdally, S. Z., Barnabas, S. L., Dabee, S., Happel, A.-U., Bowers, D., Davids, L., Passmore, J.-A. S., & Masson, L. (2019). Antimicrobial and inflammatory properties of South African clinical Lactobacillus isolates and vaginal probiotics. *Scientific Reports*, *9*, 1917. https://doi.org/10.1038/s41598-018-38253-4

Chopra, I., & Roberts, M. (2001). Tetracycline Antibiotics: Mode of Action, Applications, Molecular Biology, and Epidemiology of Bacterial Resistance. *Microbiology and Molecular Biology Reviews*, *65*(2), 232–260. https://doi.org/10.1128/MMBR.65.2.232-260.2001

Chromas | Technelysium Pty Ltd. (n.d.). Retrieved May 21, 2024, from https://technelysium.com.au/wp/chromas/

Ciornei, C. D., Novikov, A., Beloin, C., Fitting, C., Caroff, M., Ghigo, J.-M., Cavaillon, J.-M., & Adib-Conquy, M. (2010). Biofilm-forming *Pseudomonas aeruginosa* bacteria undergo lipopolysaccharide structural modifications and induce enhanced inflammatory cytokine response in human monocytes. *Innate Immunity*, *16*(5), 288–301. https://doi.org/10.1177/1753425909341807

Clifton, Luke. A., Ciesielski, F., Skoda, M. W. A., Paracini, N., Holt, S. A., & Lakey, J. H. (2016). The Effect of Lipopolysaccharide Core Oligosaccharide Size on the Electrostatic Binding of Antimicrobial Proteins to Models of the Gram Negative Bacterial Outer Membrane. *Langmuir*, *32*(14), 3485–3494. https://doi.org/10.1021/acs.langmuir.6b00240

CLSI. (2020). *Performance Standards for Antimicrobial Susceptibility Testing* (M100 30).

Cobo, F., Rodríguez-Granger, J., Pérez-Zapata, I., Sampedro, A., Aliaga, L., & Navarro-Marí, J. M. (2019). Antimicrobial susceptibility and clinical findings of significant anaerobic bacteria in southern Spain. *Anaerobe*, *59*, 49–53. https://doi.org/10.1016/j.anaerobe.2019.05.007

Cohen, C. R., Wierzbicki, M. R., French, A. L., Morris, S., Newmann, S., Reno, H., Green, L., Miller, S., Powell, J., Parks, T., & Hemmerling, A. (2020). Randomized Trial of Lactin-V to Prevent Recurrence of Bacterial Vaginosis.

The New England Journal of Medicine, *382*(20), 1906–1915. https://doi.org/10.1056/NEJMoa1915254

Cornejo Ulloa, P., Krom, B. P., & van der Veen, M. H. (2021). Sex Steroid Hormones as a Balancing Factor in Oral Host Microbiome Interactions. *Frontiers in Cellular and Infection Microbiology*, *11*, 714229. https://doi.org/10.3389/fcimb.2021.714229

Council, S. E. (2013). *Prevotella melaninogenica,* an oral anaerobic bacterium, prevalent in cystic fibrosis chronic lung infection. University of North Carolina at Chapel Hill.

Dahlén, G. G. (1993). Black-pigmented Gram-negative anaerobes in periodontitis. *FEMS Immunology & Medical Microbiology*, *6*(2–3), 181–192. https://doi.org/10.1111/j.1574-695X.1993.tb00323.x

Dareng, E. O., Ma, B., Famooto, A. O., Akarolo-Anthony, S. N., Offiong, R. A., Olaniyan, O., Dakum, P. S., Wheeler, C. M., Fadrosh, D., Yang, H., Gajer, P., Brotman, R. M., Ravel, J., & Adebamowo, C. A. (2016). Prevalent high-risk HPV infection and vaginal microbiota in Nigerian women. *Epidemiology and Infection*, *144*(1), 123–137. https://doi.org/10.1017/S0950268815000965

Darriba, D., Posada, D., Kozlov, A. M., Stamatakis, A., Morel, B., & Flouri, T. (2020). ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. *Molecular Biology and Evolution*, *37*(1), 291–294. https://doi.org/10.1093/molbev/msz189

Davidov, E., & Kaufmann, G. (2008). RloC: A wobble nucleotide-excising and zinc-responsive bacterial tRNase. *Molecular Microbiology*, *69*(6), 1560–1574. https://doi.org/10.1111/j.1365-2958.2008.06387.x

De Filippis, F., Pasolli, E., Tett, A., Tarallo, S., Naccarati, A., De Angelis, M., Neviani, E., Cocolin, L., Gobbetti, M., Segata, N., & Ercolini, D. (2019). Distinct Genetic and Functional Traits of Human Intestinal *Prevotella copri* Strains Are Associated with Different Habitual Diets. *Cell Host & Microbe*, *25*(3), 444-453.e3. https://doi.org/10.1016/j.chom.2019.01.004

Deurenberg, R. H., Bathoorn, E., Chlebowicz, M. A., Couto, N., Ferdous, M., García-Cobos, S., Kooistra-Smid, A. M. D., Raangs, E. C., Rosema, S., Veloo, A. C. M., Zhou, K., Friedrich, A. W., & Rossen, J. W. A. (2017). Application of next generation sequencing in clinical microbiology and infection prevention. *Journal of Biotechnology*, *243*, 16–24. https://doi.org/10.1016/j.jbiotec.2016.12.022

Devoto, A. E., Santini, J. M., Olm, M. R., Anantharaman, K., Munk, P., Tung, J., Archie, E. A., Turnbaugh, P. J., Seed, K. D., Blekhman, R., Aarestrup, F. M., Thomas, B. C., & Banfield, J. F. (2019). Megaphages infect *Prevotella* and variants are widespread in gut microbiomes. *Nature Microbiology*, *4*(4), 693–700. https://doi.org/10.1038/s41564-018-0338-9

Dong, B., Huang, Y., Cai, H., Chen, Y., Li, Y., Zou, H., Lin, W., Xue, H., Feng, A., Zhao, H., Lu, Y., Gao, H., Mao, X., Wu, Z., Pan, D., & Sun, P. (2022). Prevotella as the hub of the cervicovaginal microbiota affects the occurrence of persistent human papillomavirus infection and cervical lesions in women of childbearing age via host NF-κB/C-myc. *Journal of Medical Virology*, *94*(11), 5519–5534. https://doi.org/10.1002/jmv.28001

Downes, J., & Wade, W. G. (2011). *Prevotella fusca* sp. Nov. And *Prevotella* scopos sp. Nov., isolated from the human oral cavity. *International Journal of Systematic and Evolutionary Microbiology*, *61*(4), 854–858. https://doi.org/10.1099/ijs.0.023861-0

DuBose, D. A., Lemaire, M., Basamania, K., & Rowlands, J. (1980). Comparison of plasma extraction techniques in preparation of samples for endotoxin testing by the Limulus amoebocyte lysate test. *Journal of Clinical Microbiology*, *11*(1), 68–72. https://doi.org/10.1128/jcm.11.1.68-72.1980

Duerden, B. I. (1993). Black-pigmented gram-negative anaerobes in genito-urinary tract and pelvic infections. *FEMS Immunology and Medical Microbiology*, *6*(2–3), 223–227. https://doi.org/10.1111/j.1574-695X.1993.tb00331.x

Dufty, J., Gkranias, N., & Donos, N. (2017). Necrotising Ulcerative Gingivitis: A Literature Review. *Oral Health & Preventive Dentistry*, *15*(4), 321–327. https://doi.org/10.3290/j.ohpd.a38766

Dy, R. L., Richter, C., Salmond, G. P. C., & Fineran, P. C. (2014). Remarkable Mechanisms in Microbes to Resist Phage Infections. *Annual Review of Virology*, *1*(1), 307–331. https://doi.org/10.1146/annurev-virology-031413-085500

Eade, C. R., Diaz, C., Wood, M. P., Anastos, K., Patterson, B. K., Gupta, P., Cole, A. L., & Cole, A. M. (2012). Identification and Characterization of Bacterial Vaginosis-Associated Pathogens Using a Comprehensive Cervical-Vaginal Epithelial Coculture Assay. *PLoS ONE*, *7*(11), e50106. https://doi.org/10.1371/journal.pone.0050106

Eastment, M. C., & McClelland, R. S. (2018). Vaginal Microbiota and Susceptibility to HIV. *AIDS (London, England)*, *32*(6), 687–698. https://doi.org/10.1097/QAD.0000000000001768

Egido, J. E., Costa, A. R., Aparicio-Maldonado, C., Haas, P.-J., & Brouns, S. J. J. (2021). Mechanisms and clinical importance of bacteriophage resistance. *FEMS Microbiology Reviews*, *46*(1), fuab048. https://doi.org/10.1093/femsre/fuab048

Eidhin, D. N., & Mouton, C. (1993). A rapid method for preparation of rough and smooth lipopolysaccharide from *Bacteroides, Porphyromonas* and *Prevotella*. *FEMS Microbiology Letters*, *110*(2), 133–138. https://doi.org/10.1111/j.1574-6968.1993.tb06309.x

Eley, A., Jones, B. M., & Al-Mushrif, S. (2000). Inhibition of chemotaxis by organic acids from anaerobes may prevent a purulent response in bacterial vaginosis. *Journal of Medical Microbiology*, *49*(11), 1023–1030. https://doi.org/10.1099/0022-1317-49-11-1023

Emiola, A., George, J., & Andrews, S. S. (2015). A Complete Pathway Model for Lipid A Biosynthesis in Escherichia coli. *PLoS ONE*, *10*(4), e0121216. https://doi.org/10.1371/journal.pone.0121216

Eschenbach, D. A. (2007). Bacterial Vaginosis: Resistance, Recurrence, and/or Reinfection? *Clinical Infectious Diseases*, *44*(2), 220–221. https://doi.org/10.1086/509584

Eschenbach, D. A., Davick, P. R., Williams, B. L., Klebanoff, S. J., Young-Smith, K., Critchlow, C. M., & Holmes, K. K. (1989). Prevalence of hydrogen peroxide-producing *Lactobacillus* species in normal women and women with bacterial vaginosis. *Journal of Clinical Microbiology*, *27*(2), 251–256. https://doi.org/10.1128/JCM.27.2.251-256.1989

Eschenbach, D. A., Thwin, S. S., Patton, D. L., Hooton, T. M., Stapleton, A. E., Agnew, K., Winter, C., Meier, A., & Stamm, W. E. (2000). Influence of the normal menstrual cycle on vaginal tissue, discharge, and microflora. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *30*(6), 901–907. https://doi.org/10.1086/313818

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics (Oxford, England)*, *32*(19), 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Fathy Mohamed, Y., Hamad, M., Ortega, X. P., & Valvano, M. A. (2017). The LpxL acyltransferase is required for normal growth and penta-acylation of lipid A in *Burkholderia cenocepacia*. *Molecular Microbiology*, *104*(1), 144–162. https://doi.org/10.1111/mmi.13618

Faught, B. M., & Reyes, S. (2019). Characterization and Treatment of Recurrent Bacterial Vaginosis. *Journal of Women's Health (2002)*, *28*(9), 1218–1226. https://doi.org/10.1089/jwh.2018.7383

Feldgarden, M., Brover, V., Gonzalez-Escalona, N., Frye, J. G., Haendiges, J., Haft, D. H., Hoffmann, M., Pettengill, J. B., Prasad, A. B., Tillman, G. E., Tyson, G. H., & Klimke, W. (2021). AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Scientific Reports*, *11*(1), 12728. https://doi.org/10.1038/s41598-021-91456-0

Fernández-Canigia, L., Cejas, D., Gutkind, G., & Radice, M. (2015). Detection and genetic characterization of β-lactamases in *Prevotella intermedia* and *Prevotella nigrescens* isolated from oral cavity infections and peritonsillar abscesses. *Anaerobe*, *33*, 8–13. https://doi.org/10.1016/j.anaerobe.2015.01.007

Ferreira, C. S. T., Marconi, C., Parada, C. M. G. L., Ravel, J., & da Silva, M. G. (2021). Sialidase Activity in the Cervicovaginal Fluid Is Associated With Changes in Bacterial Components of Lactobacillus-Deprived Microbiota. *Frontiers in Cellular and Infection Microbiology*, *11*, 813520. https://doi.org/10.3389/fcimb.2021.813520

Ferreira, L. Q., Avelar, K. E. S., Vieira, J. M. B. D., de Paula, G. R., Colombo, A. P. V., Domingues, R. M. C. P., & Ferreira, M. C. S. (2007). Association Between the cfxA Gene and Transposon Tn4555 in *Bacteroides distasonis* Strains and Other *Bacteroides* Species. *Current Microbiology*, *54*(5), 348–353. https://doi.org/10.1007/s00284-006-0411-0

Ferris, M. J., Masztal, A., Aldridge, K. E., Fortenberry, J. D., Fidel, P. L., & Martin, D. H. (2004). Association of *Atopobium vaginae*, a recently described metronidazole resistant anaerobe, with bacterial vaginosis. *BMC Infectious Diseases*, *4*, 5. https://doi.org/10.1186/1471-2334-4-5

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. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *47*(11), 1426–1435. https://doi.org/10.1086/592974

Fethers, K. A., Fairley, C. K., Morton, A., Hocking, J. S., Hopkins, C., Kennedy, L. J., Fehler, G., & Bradshaw, C. S. (2009). Early Sexual Experiences and Risk Factors for Bacterial Vaginosis. *The Journal of Infectious Diseases*, *200*(11), 1662–1670. https://doi.org/10.1086/648092

Fettweis, J. M., Brooks, J. P., Serrano, M. G., Sheth, N. U., Girerd, P. H., Edwards, D. J., Strauss, J. F., the Vaginal Microbiome Consortium, Jefferson, K. K., & Buck, G. A. (2014). Differences in vaginal microbiome in African American women versus women of European ancestry. *Microbiology*, *160*(10), 2272–2282. https://doi.org/10.1099/mic.0.081034-0

Fettweis, J. M., Serrano, M. G., Brooks, J. P., Edwards, D. J., Girerd, P. H., Parikh, H. I., Huang, B., Arodz, T. J., Edupuganti, L., Glascock, A. L., Xu, J., Jimenez, N. R., Vivadelli, S. C., Fong, S. S., Sheth, N. U., Jean, S., Lee, V., Bokhari, Y. A., Lara, A. M., … Buck, G. A. (2019). The vaginal microbiome and preterm birth. *Nature Medicine*, *25*(6), 1012–1021. https://doi.org/10.1038/s41591-019-0450-2

France, M. T., Fu, L., Rutt, L., Yang, H., Humphrys, M. S., Narina, S., Gajer, P. M., Ma, B., Forney, L. J., & Ravel, J. (2022). Insight into the ecology of vaginal bacteria through integrative analyses of metagenomic and metatranscriptomic data. *Genome Biology*, *23*(1), 66. https://doi.org/10.1186/s13059-022-02635-9

France, M. T., Ma, B., Gajer, P., Brown, S., Humphrys, M. S., Holm, J. B., Waetjen, L. E., Brotman, R. M., & Ravel, J. (2020). VALENCIA: A nearest centroid classification method for vaginal microbial communities based on composition. *Microbiome*, *8*(1), 166. https://doi.org/10.1186/s40168-020-00934-6

Fredricks, D. N., Fiedler, T. L., & Marrazzo, J. M. (2005). Molecular identification of bacteria associated with bacterial vaginosis. *The New England Journal of Medicine*, *353*(18), 1899–1911. https://doi.org/10.1056/NEJMoa043802

Freitas, A. C., Bocking, A., Hill, J. E., Money, D. M., Money, D., Bocking, A., Hemmingsen, S., Hill, J., Reid, G., Dumonceaux, T., Gloor, G., Links, M., O'Doherty, K., Tang, P., van Schalkwyk, J., Yudin, M., & the VOGUE Research Group. (2018). Increased richness and diversity of the vaginal microbiota and spontaneous preterm birth. *Microbiome*, *6*(1), 117. https://doi.org/10.1186/s40168-018-0502-8

Frias, J., Olle, E., & Alsina, M. (2001). Periodontal Pathogens Produce Quorum Sensing Signal Molecules. *Infection and Immunity*, *69*(5), 3431–3434. https://doi.org/10.1128/IAI.69.5.3431-3434.2001

Gajer, P., Brotman, R. M., Bai, G., Sakamoto, J., Schütte, U. M. E., Zhong, X., Koenig, S. S. K., Fu, L., Ma, Z., Zhou, X., Abdo, Z., Forney, L. J., & Ravel, J. (2012). Temporal Dynamics of the Human Vaginal Microbiota. *Science Translational Medicine*, *4*(132), 132ra52. https://doi.org/10.1126/scitranslmed.3003605

Galhardo, C. L., Soares, J. M., Simões, R. S., Haidar, M. A., Rodrigues de Lima, G., & Baracat, E. C. (2006). Estrogen effects on the vaginal pH, flora and cytology in late postmenopause after a long period without hormone therapy. *Clinical and Experimental Obstetrics & Gynecology*, *33*(2), 85–89.

García, N., Gutiérrez, G., Lorenzo, M., García, J. E., Píriz, S., & Quesada, A. (2008). Genetic determinants for cfxA expression in Bacteroides strains isolated from human infections. *The Journal of Antimicrobial Chemotherapy*, *62*(5), 942–947. https://doi.org/10.1093/jac/dkn347

Gazi, M. I., Cox, S. W., Clark, D. T., & Eley, B. M. (1995). Comparison of host tissue and bacterial dipeptidyl peptidases in human gingival crevicular fluid by analytical isoelectric focusing. *Archives of Oral Biology*, *40*(8), 731–736. https://doi.org/10.1016/0003-9969(95)00032-K

Geisinger, E., & Isberg, R. R. (2015). Antibiotic Modulation of Capsular Exopolysaccharide and Virulence in *Acinetobacter baumannii*. *PLoS Pathogens*, *11*(2), e1004691. https://doi.org/10.1371/journal.ppat.1004691

Georjon, H., & Bernheim, A. (2023). The highly diverse antiphage defence systems of bacteria. *Nature Reviews. Microbiology*, *21*(10), 686–700. https://doi.org/10.1038/s41579-023-00934-x

Gibb, B., Hyman, P., & Schneider, C. L. (2021). The Many Applications of Engineered Bacteriophages—An Overview. *Pharmaceuticals*, *14*(7), 634. https://doi.org/10.3390/ph14070634

Gibbons, R. J., & Macdonald, J. B. (1960). Hemin and vitamin K compounds as required factors for the cultivation of certain strains of *Bacteroides melaninogenicus*. *Journal of Bacteriology*, *80*(2), 164–170. https://doi.org/10.1128/jb.80.2.164-170.1960

Gilbert, N. M., Lewis, W. G., Li, G., Sojka, D. K., Lubin, J. B., & Lewis, A. L. (2019). *Gardnerella vaginalis* and *Prevotella bivia* Trigger Distinct and Overlapping Phenotypes in a Mouse Model of Bacterial Vaginosis. *The Journal of Infectious Diseases*, *220*(7), 1099–1108. https://doi.org/10.1093/infdis/jiy704

Gilchrist, C. L. M., & Chooi, Y.-H. (2021). clinker & clustermap.js: Automatic generation of gene cluster comparison figures. *Bioinformatics*, *37*(16), 2473–2475. https://doi.org/10.1093/bioinformatics/btab007

Gill, K., Happel, A., Pidwell, T., Mendelsohn, A., Duyver, M., Johnson, L., Meyer, L., Slack, C., Strode, A., Mendel, E., Fynn, L., Wallace, M., Spiegel, H., Jaspan, H., Passmore, J., Hosek, S., Smit, D., Rinehart, A., & Bekker, L. (2020). An open-label, randomized crossover study to evaluate the acceptability and preference for contraceptive options in female adolescents, 15 to 19 years of age in Cape Town, as a proxy for HIV prevention methods (UChoose). *Journal of the International AIDS Society*, *23*(10), e25626. https://doi.org/10.1002/jia2.25626

Gillet, E., Meys, J. F., Verstraelen, H., Bosire, C., De Sutter, P., Temmerman, M., & Broeck, D. V. (2011). Bacterial vaginosis is associated with uterine cervical human papillomavirus infection: A meta-analysis. *BMC Infectious Diseases*, *11*, 10. https://doi.org/10.1186/1471-2334-11-10

Goh, S. H., Potter, S., Wood, J. O., Hemmingsen, S. M., Reynolds, R. P., & Chow, A. W. (1996). HSP60 gene sequences as universal targets for microbial species identification: Studies with coagulase-negative staphylococci. *Journal of Clinical Microbiology*, *34*(4), 818–823.

Gorenc, G., Lukáš, F., & Avgustin, G. (2008). Examination of AI-2 quorum sensing System in *Prevotella bryantii* and *Prevotella ruminicola*-like strains by using bioluminiscence assay. *Acta Agriculturae Slovenica Institute of Animal Physiology and Genetics*, *90*. https://doi.org/10.14720/aas.2007.90.2.14960

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., & Tiedje, J. M. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *International Journal of Systematic and Evolutionary Microbiology*, *57*(Pt 1), 81–91. https://doi.org/10.1099/ijs.0.64483-0

Gosmann, C., Anahtar, M. N., Handley, S. A., Farcasanu, M., Abu-Ali, G., Bowman, B. A., Padavattan, N., Desai, C., Droit, L., Moodley, A., Dong, M., Chen, Y., Ismail, N., Ndung'u, T., Ghebremichael, M. S., Wesemann, D. R., Mitchell, C., Dong, K. L., Huttenhower, C., … Kwon, D. S. (2017a). Lactobacillus-Deficient Cervicovaginal Bacterial Communities Are Associated with Increased HIV Acquisition in Young South African Women. *Immunity*, *46*(1), 29–37. https://doi.org/10.1016/j.immuni.2016.12.013

Gosmann, C., Anahtar, M. N., Handley, S. A., Farcasanu, M., Abu-Ali, G., Bowman, B. A., Padavattan, N., Desai, C., Droit, L., Moodley, A., Dong, M., Chen, Y., Ismail, N., Ndung'u, T., Ghebremichael, M. S., Wesemann, D. R., Mitchell, C., Dong, K. L., Huttenhower, C., … Kwon, D. S. (2017b). *Lactobacillus*-Deficient

Cervicovaginal Bacterial Communities are Associated with Increased HIV Acquisition in Young South African Women. *Immunity*, *46*(1), 29–37. https://doi.org/10.1016/j.immuni.2016.12.013

Gottschick, C., Deng, Z.-L., Vital, M., Masur, C., Abels, C., Pieper, D. H., Rohde, M., Mendling, W., & Wagner-Döbler, I. (2017). Treatment of biofilms in bacterial vaginosis by an amphoteric tenside pessary-clinical study and microbiota analysis. *Microbiome*, *5*(1), 119. https://doi.org/10.1186/s40168-017-0326-y

Grant, J. S., Stafylis, C., Celum, C., Grennan, T., Haire, B., Kaldor, J., Luetkemeyer, A. F., Saunders, J. M., Molina, J.-M., & Klausner, J. D. (2020). Doxycycline Prophylaxis for Bacterial Sexually Transmitted Infections. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *70*(6), 1247–1253. https://doi.org/10.1093/cid/ciz866

Greenblum, S., Carr, R., & Borenstein, E. (2015). Extensive strain-level copy-number variation across human gut microbiome species. *Cell*, *160*(4), 583–594. https://doi.org/10.1016/j.cell.2014.12.038

Grossman, T. H. (2016). Tetracycline Antibiotics and Resistance. *Cold Spring Harbor Perspectives in Medicine*, *6*(4), a025387. https://doi.org/10.1101/cshperspect.a025387

Guédou, F. A., Van Damme, L., Deese, J., Crucitti, T., Becker, M., Mirembe, F., Solomon, S., & Alary, M. (2013). Behavioural and medical predictors of bacterial vaginosis recurrence among female sex workers: Longitudinal analysis from a randomized controlled trial. *BMC Infectious Diseases*, *13*, 208. https://doi.org/10.1186/1471-2334-13-208

Guidelines for the management of sexually transmitted infections. (2021). Geneva: World Health Organisation.

Guo, J., Bolduc, B., Zayed, A. A., Varsani, A., Dominguez-Huerta, G., Delmont, T. O., Pratama, A. A., Gazitúa, M. C., Vik, D., Sullivan, M. B., & Roux, S. (2021). VirSorter2: A multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome*, *9*(1), 37. https://doi.org/10.1186/s40168-020-00990-y

Guo, J., Vik, D., Pratama, A. A., Roux, S., & Sullivan, M. (2021). *Viral sequence identification SOP with VirSorter2*. https://www.protocols.io/view/viral-sequence-identification-sop-with-virsorter2-bwm5pc86

Guo, Y., You, K., Qiao, J., Zhao, Y., & Geng, L. (2012). Bacterial vaginosis is conducive to the persistence of HPV infection. *International Journal of STD & AIDS*, *23*(8), 581–584. https://doi.org/10.1258/ijsa.2012.011342

Gupta, S. K., Padmanabhan, B. R., Diene, S. M., Lopez-Rojas, R., Kempf, M., Landraud, L., & Rolain, J.-M. (2014). ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrobial Agents and Chemotherapy*, *58*(1), 212–220. https://doi.org/10.1128/AAC.01310-13

Gupta, V. K., Chaudhari, N. M., Iskepalli, S., & Dutta, C. (2015). Divergences in gene repertoire among the reference Prevotella genomes derived from distinct body sites of human. *BMC Genomics*, *16*(1), 153. https://doi.org/10.1186/s12864-015-1350-6

Hadfield, J., Croucher, N. J., Goater, R. J., Abudahab, K., Aanensen, D. M., & Harris, S. R. (2018). Phandango: An interactive viewer for bacterial population genomics. *Bioinformatics*, *34*(2), 292–293. https://doi.org/10.1093/bioinformatics/btx610

Haggerty, C. L., Hillier, S. L., Bass, D. C., Ness, R. B., & PID Evaluation and Clinical Health study investigators. (2004). Bacterial vaginosis and anaerobic bacteria are associated with endometritis. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *39*(7), 990–995. https://doi.org/10.1086/423963

Haggerty, C. L., Ness, R. B., Totten, P. A., Farooq, F., Tang, G., Ko, D., Hou, X., Fiedler, T. L., Srinivasan, S., Astete, S. G., & Fredricks, D. N. (2020). Presence and concentrations of select bacterial vaginosis-associated bacteria are associated with increased risk of pelvic inflammatory disease. *Sexually Transmitted Diseases*, *47*(5), 344–346. https://doi.org/10.1097/OLQ.0000000000001164

Haggerty, C. L., Totten, P. A., Tang, G., Astete, S. G., Ferris, M. J., Norori, J., Bass, D. C., Martin, D. H., Taylor, B. D., & Ness, R. B. (2016). Identification of novel microbes associated with pelvic inflammatory disease and infertility. *Sexually Transmitted Infections*, *92*(6), 441–446. https://doi.org/10.1136/sextrans-2015-052285

Hashimoto, M., Asai, Y., Tamai, R., Jinno, T., Umatani, K., & Ogawa, T. (2003). Chemical structure and immunobiological activity of lipid A from *Prevotella intermedia* ATCC 25611 lipopolysaccharide. *FEBS Letters*, *543*(1–3), 98–102. https://doi.org/10.1016/S0014-5793(03)00414-9

He, G.-X., Zhang, C., Crow, R. R., Thorpe, C., Chen, H., Kumar, S., Tsuchiya, T., & Varela, M. F. (2011). SugE, a new member of the SMR family of transporters, contributes to antimicrobial resistance in *Enterobacter cloacae*. *Antimicrobial Agents and Chemotherapy*, *55*(8), 3954–3957. https://doi.org/10.1128/AAC.00094-11

Hill, J. E., Goh, S. H., Money, D. M., Doyle, M., Li, A., Crosby, W. L., Links, M., Leung, A., Chan, D., & Hemmingsen, S. M. (2005). Characterization of vaginal microflora of healthy, nonpregnant women by chaperonin-60 sequence-based methods. *American Journal of Obstetrics and Gynecology*, *193*(3 Pt 1), 682–692. https://doi.org/10.1016/j.ajog.2005.02.094

Hillier, S. L., & Lau, R. J. (1997). Vaginal microflora in postmenopausal women who have not received estrogen replacement therapy. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *25 Suppl 2*, S123-126. https://doi.org/10.1086/516221

Hillier, S. L., Nugent, R. P., Eschenbach, D. A., Krohn, M. A., Gibbs, R. S., Martin, D. H., Cotch, M. F., Edelman, R., Pastorek, J. G., & Rao, A. V. (1995). Association between bacterial vaginosis and preterm delivery of a low-birth-weight infant. The Vaginal Infections and Prematurity Study Group. *The New England Journal of Medicine*, *333*(26), 1737–1742. https://doi.org/10.1056/NEJM199512283332604

Hladik, F., Sakchalathorn, P., Ballweber, L., Lentz, G., Fialkow, M., Eschenbach, D., & McElrath, M. J. (2007). Initial Events in Establishing Vaginal Entry and Infection by Human Immunodeficiency Virus Type-1. *Immunity*, *26*(2), 257–270. https://doi.org/10.1016/j.immuni.2007.01.007

Høiby, N., Ciofu, O., Johansen, H. K., Song, Z., Moser, C., Jensen, P. Ø., Molin, S., Givskov, M., Tolker-Nielsen, T., & Bjarnsholt, T. (2011). The clinical impact of bacterial biofilms. *International Journal of Oral Science*, *3*(2), 55–65. https://doi.org/10.4248/IJOS11026

Holm, J. B., France, M. T., Gajer, P., Ma, B., Brotman, R. M., Shardell, M., Forney, L., & Ravel, J. (2023). Integrating compositional and functional content to

describe vaginal microbiomes in health and disease. *Microbiome*, *11*(1), 259. https://doi.org/10.1186/s40168-023-01692-x

Ilhan, Z. E., Łaniewski, P., Tonachio, A., & Herbst-Kralovetz, M. M. (2020). Members of Prevotella Genus Distinctively Modulate Innate Immune and Barrier Functions in a Human Three-Dimensional Endometrial Epithelial Cell Model. *The Journal of Infectious Diseases*, *222*(12), 2082–2092. https://doi.org/10.1093/infdis/jiaa324

Jacobson, A. N., Choudhury, B. P., & Fischbach, M. A. (2018). The Biosynthesis of Lipooligosaccharide from *Bacteroides thetaiotaomicron*. *mBio*, *9*(2), 10.1128/mbio.02289-17. https://doi.org/10.1128/mbio.02289-17

Jacobsson, B., Pernevi, P., Chidekel, L., & Jörgen Platz-Christensen, J. (2002). Bacterial vaginosis in early pregnancy may predispose for preterm birth and postpartum endometritis. *Acta Obstetricia Et Gynecologica Scandinavica*, *81*(11), 1006–1010. https://doi.org/10.1034/j.1600-0412.2002.811103.x

Jespers, V., Kyongo, J., Joseph, S., Hardy, L., Cools, P., Crucitti, T., Mwaura, M., Ndayisaba, G., Delany-Moretlwe, S., Buyze, J., Vanham, G., & van de Wijgert, J. H. H. M. (2017). A longitudinal analysis of the vaginal microbiota and vaginal immune mediators in women from sub-Saharan Africa. *Scientific Reports*, *7*(1), 11974. https://doi.org/10.1038/s41598-017-12198-6

Jiao, Y., Hasegawa, M., & Inohara, N. (2014). The Role of Oral Pathobionts in Dysbiosis during Periodontitis Development. *Journal of Dental Research*, *93*(6), 539–546. https://doi.org/10.1177/0022034514528212

Johnston, C. D., Skeete, C. A., Fomenkov, A., Roberts, R. J., & Rittling, S. R. (2017). Restriction-modification mediated barriers to exogenous DNA uptake and incorporation employed by *Prevotella intermedia*. *PLOS ONE*, *12*(9), e0185234. https://doi.org/10.1371/journal.pone.0185234

Jung, Y.-G., Choi, J., Kim, S.-K., Lee, J.-H., & Kwon, S. (2015). Embedded biofilm, a new biofilm model based on the embedded growth of bacteria. *Applied and Environmental Microbiology*, *81*(1), 211–219. https://doi.org/10.1128/AEM.02311-14

Kahne, T., Lendeckel, U., Wrenger, S., Neubert, K., Ansorge, S., & Reinhold, D. (1999). Dipeptidyl peptidase IV: A cell surface peptidase involved in regulating T cell growth (review). *International Journal of Molecular Medicine*, *4*(1), 3–15. https://doi.org/10.3892/ijmm.4.1.3

Karaoglu, Ş., Aydín, F., Kiliç, S., & Kiliç, A. (2003). Antimicrobial Activity and Characteristics of Bacteriocins Produced by Vaginal Lactobacilli. *Turkish Journal of Medical Sciences*, *33*(1), 7–13. https://doi.org/-

Karched, M., Bhardwaj, R. G., Qudeimat, M., Al-Khabbaz, A., & Ellepola, A. (2022). Proteomic analysis of the periodontal pathogen *Prevotella intermedia* secretomes in biofilm and planktonic lifestyles. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-09085-0

Katsandri, A., Avlamis, A., Pantazatou, A., Houhoula, D. P., & Papaparaskevas, J. (2006). Dissemination of nim-class genes, encoding nitroimidazole resistance, among different species of Gram-negative anaerobic bacteria isolated in Athens, Greece. *The Journal of Antimicrobial Chemotherapy*, *58*(3), 705–706. https://doi.org/10.1093/jac/dkl285

Kellenberger, E., & Ryter, A. (1958). Cell Wall and Cytoplasmic Membrane of Escherichia coli. *The Journal of Biophysical and Biochemical Cytology*, *4*(3), 323–326.

Kenyon, C., Colebunders, R., & Crucitti, T. (2013). The global epidemiology of bacterial vaginosis: A systematic review. *American Journal of Obstetrics and Gynecology*, *209*(6), 505–523. https://doi.org/10.1016/j.ajog.2013.05.006

Kilár, A., Dörnyei, Á., & Kocsis, B. (2013). Structural characterization of bacterial lipopolysaccharides with mass spectrometry and on- and off-line separation techniques. *Mass Spectrometry Reviews*, *32*(2), 90–117. https://doi.org/10.1002/mas.21352

Kim, S.-J., Choi, E.-Y., Kim, E. G., Shin, S.-H., Lee, J.-Y., Choi, J.-I., & Choi, I.-S. (2007). *Prevotella intermedia* lipopolysaccharide stimulates release of tumor necrosis factor-α through mitogen-activated protein kinase signaling pathways in monocyte-derived macrophages. *FEMS Immunology & Medical Microbiology*, *51*(2), 407–413. https://doi.org/10.1111/j.1574-695X.2007.00318.x

Klatt, N. R., Klatt, N. R., Cheu, R., Birse, K., Zevin, A. S., Perner, M., Noël-Romas, L., Grobler, A., Westmacott, G., Xie, I. Y., Butler, J., Mansoor, L., McKinnon, L. R., Passmore, J.-A. S., Abdool Karim, Q., Abdool Karim, S. S., & Burgener, A. D. (2017). Vaginal bacteria modify HIV tenofovir microbicide efficacy in African women. *Science (New York, N.Y.)*, *356*(6341), 938–945. https://doi.org/10.1126/science.aai9383

Kondo, Y. (2021). Construction of a Mutant in *Prevotella melaninogenica* Using the Conjugation Transfer Method with *Escherichia coli*. *Methods in Molecular Biology (Clifton, N.J.)*, *2210*, 33–41. https://doi.org/10.1007/978-1-0716-0939-2_4

Kondo, Y., Sato, K., Nagano, K., Nishiguchi, M., Hoshino, T., Fujiwara, T., & Nakayama, K. (2018). Involvement of PorK, a component of the type IX secretion system, in *Prevotella melaninogenica* pathogenicity. *Microbiology and Immunology*, *62*(9), 554–566. https://doi.org/10.1111/1348-0421.12638

Könönen, E., Asikainen, S., Saarela, M., Karjalainen, J., & Jousimies-Somer, H. (1994). The oral gram-negative anaerobic microflora in young children: Longitudinal changes from edentulous to dentate mouth. *Oral Microbiology and Immunology*, *9*(3), 136–141. https://doi.org/10.1111/j.1399-302X.1994.tb00049.x

Könönen, E., Fteita, D., Gursoy, U. K., & Gursoy, M. (2022). *Prevotella* species as oral residents and infectious agents with potential impact on systemic conditions. *Journal of Oral Microbiology*, *14*(1), 2079814. https://doi.org/10.1080/20002297.2022.2079814

Könönen, E., Jousimies-Somer, H., & Asikainen, S. (1994). The most frequently isolated gram-negative anaerobes in saliva and subgingival samples taken from young women. *Oral Microbiology and Immunology*, *9*(2), 126–128. https://doi.org/10.1111/j.1399-302x.1994.tb00047.x

Könönen, E., Kanervo, A., Takala, A., Asikainen, S., & Jousimies-Somer, H. (1999). Establishment of Oral Anaerobes during the First Year of Life. *Journal of Dental Research*, *78*(10), 1634–1639. https://doi.org/10.1177/00220345990780100801

Konstantinidis, K. T., & Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(7), 2567–2572. https://doi.org/10.1073/pnas.0409727102

Koo, H., Falsetta, M. L., & Klein, M. I. (2013). The Exopolysaccharide Matrix. *Journal of Dental Research*, *92*(12), 1065–1073. https://doi.org/10.1177/0022034513504218

Kornman, K. S., & Loesche, W. J. (1982). Effects of estradiol and progesterone on Bacteroides melaninogenicus and *Bacteroides gingivalis*. *Infection and Immunity*, *35*(1), 256–263. https://doi.org/10.1128/iai.35.1.256-263.1982

Köser, C. U., Ellington, M. J., & Peacock, S. J. (2014). Whole-genome sequencing to control antimicrobial resistance. *Trends in Genetics*, *30*(9), 401–407. https://doi.org/10.1016/j.tig.2014.07.003

Koumans, E. H., Markowitz, L. E., Hogan, V., & CDC BV Working Group. (2002). Indications for Therapy and Treatment Recommendations for Bacterial Vaginosis in Nonpregnant and Pregnant Women: A Synthesis of Data. *Clinical Infectious Diseases*, *35*(Supplement_2), S152–S172. https://doi.org/10.1086/342103

Koumans, E. H., Sternberg, M., Bruce, C., McQuillan, G., Kendrick, J., Sutton, M., & Markowitz, L. E. (2007). The prevalence of bacterial vaginosis in the United States, 2001-2004; associations with symptoms, sexual behaviors, and reproductive health. *Sexually Transmitted Diseases*, *34*(11), 864–869. https://doi.org/10.1097/OLQ.0b013e318074e565

Kumar, M., Murugesan, S., Singh, P., Saadaoui, M., Elhag, D. A., Terranegra, A., Kabeer, B. S. A., Marr, A. K., Kino, T., Brummaier, T., McGready, R., Nosten, F., Chaussabel, D., & Al Khodor, S. (2021). Vaginal Microbiota and Cytokine Levels Predict Preterm Delivery in Asian Women. *Frontiers in Cellular and Infection Microbiology*, *11*, 639665. https://doi.org/10.3389/fcimb.2021.639665

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Lamoureux, C., Guilloux, C.-A., Courteboeuf, E., Gouriou, S., Beauruelle, C., & Héry-Arnaud, G. (2021). *Prevotella melaninogenica*, a Sentinel Species of Antibiotic Resistance in Cystic Fibrosis Respiratory Niche? *Microorganisms*, *9*(6), 1275. https://doi.org/10.3390/microorganisms9061275

Landlinger, C., Tisakova, L., Oberbauer, V., Schwebs, T., Muhammad, A., Latka, A., Van Simaey, L., Vaneechoutte, M., Guschin, A., Resch, G., Swidsinski, S., Swidsinski, A., & Corsini, L. (2021). Engineered Phage Endolysin Eliminates *Gardnerella* Biofilm without Damaging Beneficial Bacteria in Bacterial Vaginosis Ex Vivo. *Pathogens*, *10*(1), 54. https://doi.org/10.3390/pathogens10010054

Lane, D. J., Pace, B., Olsen, G. J., Stahl, D. A., Sogin, M. L., & Pace, N. R. (1985). Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proceedings of the National Academy of Sciences of the United States of America*, *82*(20), 6955–6959. https://doi.org/10.1073/pnas.82.20.6955

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)*, *23*(21), 2947–2948. https://doi.org/10.1093/bioinformatics/btm404

Larsen, J. M. (2017). The immune response to *Prevotella* bacteria in chronic inflammatory disease. *Immunology*, *151*(4), 363–374. https://doi.org/10.1111/imm.12760

Larsen, J. M., Musavian, H. S., Butt, T. M., Ingvorsen, C., Thysen, A. H., & Brix, S. (2015). Chronic obstructive pulmonary disease and asthma-associated *Proteobacteria*, but not commensal *Prevotella* spp., promote Toll-like receptor 2-independent lung inflammation and pathology. *Immunology*, *144*(2), 333–342. https://doi.org/10.1111/imm.12376

Leclercq, R. (2002). Mechanisms of Resistance to Macrolides and Lincosamides: Nature of the Resistance Elements and Their Clinical Implications. *Clinical Infectious Diseases*, *34*(4), 482–492. https://doi.org/10.1086/324626

Lee, S.-W., Okamoto, M., & Leung, K. (2021). Identification and characterization of a hemagglutinin gene from *Prevotella intermedia*. *Oral Biology Research*, *45*, 82–89. https://doi.org/10.21851/obr.45.02.202106.82

Leiros, H.-K. S., Kozielski-Stuhrmann, S., Kapp, U., Terradot, L., Leonard, G. A., & McSweeney, S. M. (2004). Structural Basis of 5-Nitroimidazole Antibiotic Resistance: THE CRYSTAL STRUCTURE OF Nima FROM DEINOCOCCUS RADIODURANS*. *Journal of Biological Chemistry*, *279*(53), 55840–55849. https://doi.org/10.1074/jbc.M408044200

Leitich, H., Bodner-Adler, B., Brunbauer, M., Kaider, A., Egarter, C., & Husslein, P. (2003). Bacterial vaginosis as a risk factor for preterm delivery: A meta-analysis. *American Journal of Obstetrics and Gynecology*, *189*(1), 139–147. https://doi.org/10.1067/mob.2003.339

Lennard, K., Dabee, S., Barnabas, S., Havyarimana, E., Blakney, A., Jaumdally, S., Botha, G., Zwane Nkuna, N., Bekker, L.-G., Lewis, D., Gray, G., Mulder, N., Passmore, J.-A., & Jaspan, H. (2019). Vaginal microbiota varies by geographical location in South African women.

Letunic, I., & Bork, P. (2007). Interactive Tree Of Life (iTOL): An online tool for phylogenetic tree display and annotation. *Bioinformatics (Oxford, England)*, *23*(1), 127–128. https://doi.org/10.1093/bioinformatics/btl529

Leung, K., Nesbitt, W. E., Okamoto, M., & Fukushima, H. (1999). Identification of a fimbriae-associated haemagglutinin from *Prevotella intermedia*. *Microbial Pathogenesis*, *26*(3), 139–148. https://doi.org/10.1006/mpat.1998.0258

Leung, K.-P., Subramaniam, P. S., Okamoto, M., Fukushima, H., & Lai, C.-H. (1998). The binding and utilization of hemoglobin by *Prevotella intermedia*. *FEMS Microbiology Letters*, *162*(2), 227–233. https://doi.org/10.1111/j.1574-6968.1998.tb13003.x

Lev-Sagie, A., Goldman-Wohl, D., Cohen, Y., Dori-Bachash, M., Leshem, A., Mor, U., Strahilevitz, J., Moses, A. E., Shapiro, H., Yagel, S., & Elinav, E. (2019). Vaginal microbiome transplantation in women with intractable bacterial vaginosis. *Nature Medicine*, *25*(10), 1500–1504. https://doi.org/10.1038/s41591-019-0600-6

Lewis, W. G., Robinson, L. S., Perry, J., Bick, J. L., Peipert, J. F., Allsworth, J. E., & Lewis, A. L. (2012). Hydrolysis of Secreted Sialoglycoprotein Immunoglobulin A (IgA) in ex Vivo and Biochemical Models of Bacterial Vaginosis *. *Journal of Biological Chemistry*, *287*(3), 2079–2089. https://doi.org/10.1074/jbc.M111.278135

Li, T., Zhang, Z., Wang, F., He, Y., Zong, X., Bai, H., & Liu, Z. (2020). Antimicrobial Susceptibility Testing of Metronidazole and Clindamycin against *Gardnerella vaginalis* in Planktonic and Biofilm Formation. *The Canadian Journal of*

Infectious Diseases & Medical Microbiology = Journal Canadien Des Maladies Infectieuses Et De La Microbiologie Medicale, *2020*, 1361825. https://doi.org/10.1155/2020/1361825

Li, Y., Wang, Z., Chen, J., Ernst, R. K., & Wang, X. (2013). Influence of Lipid A Acylation Pattern on Membrane Permeability and Innate Immune Stimulation. *Marine Drugs*, *11*(9), 3197–3208. https://doi.org/10.3390/md11093197

Lin, C.-Y., Lin, C.-Y., Yeh, Y.-M., Yang, L.-Y., Lee, Y.-S., Chao, A., Chin, C.-Y., Chao, A.-S., & Yang, C.-Y. (2020). Severe preeclampsia is associated with a higher relative abundance of *Prevotella bivia* in the vaginal microbiota. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-75534-3

Links, M. G., Dumonceaux, T. J., Hemmingsen, S. M., & Hill, J. E. (2012). The Chaperonin-60 Universal Target Is a Barcode for Bacteria That Enables De Novo Assembly of Metagenomic Sequence Data. *PLoS ONE*, *7*(11), e49755. https://doi.org/10.1371/journal.pone.0049755

Liu, L., Johnson, H. L., Cousens, S., Perin, J., Scott, S., Lawn, J. E., Rudan, I., Campbell, H., Cibulskis, R., Li, M., Mathers, C., Black, R. E., & Child Health Epidemiology Reference Group of WHO and UNICEF. (2012). Global, regional, and national causes of child mortality: An updated systematic analysis for 2010 with time trends since 2000. *Lancet (London, England)*, *379*(9832), 2151–2161. https://doi.org/10.1016/S0140-6736(12)60560-1

Liu, S. M., Brown, D. M., O'Donoghue, P., Upcroft, P., & Upcroft, J. A. (2000). Ferredoxin involvement in metronidazole resistance of *Giardia duodenalis. Molecular and Biochemical Parasitology*, *108*(1), 137–140. https://doi.org/10.1016/s0166-6851(00)00194-8

Lloyd-Price, J., Mahurkar, A., Rahnavard, G., Crabtree, J., Orvis, J., Hall, A. B., Brady, A., Creasy, H. H., McCracken, C., Giglio, M. G., McDonald, D., Franzosa, E. A., Knight, R., White, O., & Huttenhower, C. (2017). Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature*, *550*(7674), 61–66. https://doi.org/10.1038/nature23889

Löfmark, S., Edlund, C., & Nord, C. E. (2010). Metronidazole Is Still the Drug of Choice for Treatment of Anaerobic Infections. *Clinical Infectious Diseases*, *50*(Supplement_1), S16–S23. https://doi.org/10.1086/647939

Lopes dos Santos Santiago, G., Cools, P., Verstraelen, H., Trog, M., Missine, G., Aila, N. E., Verhelst, R., Tency, I., Claeys, G., Temmerman, M., & Vaneechoutte, M. (2011). Longitudinal Study of the Dynamics of Vaginal Microflora during Two Consecutive Menstrual Cycles. *PLoS ONE*, *6*(11), e28180. https://doi.org/10.1371/journal.pone.0028180

Lorenzo-Redondo, R., Ozer, E. A., Achenbach, C. J., D'Aquila, R. T., & Hultquist, J. F. (2021). Molecular epidemiology in the HIV and SARS-CoV-2 pandemics. *Current Opinion in HIV and AIDS*, *16*(1), 11–24. https://doi.org/10.1097/COH.0000000000000660

Löytynoja, A. (2014). Phylogeny-aware alignment with PRANK. In D. J. Russell (Ed.), *Multiple Sequence Alignment Methods* (pp. 155–170). Humana Press. https://doi.org/10.1007/978-1-62703-646-7_10

Lubbe, M. M., Stanley, K., & Chalkley, L. J. (1999). Prevalence of *nim* genes in anaerobic/facultative anaerobic bacteria isolated in South Africa. *FEMS Microbiology Letters*, *172*(1), 79–83. https://doi.org/10.1111/j.1574-6968.1999.tb13453.x

Luetkemeyer, A. F., Donnell, D., Dombrowski, J. C., Cohen, S., Grabow, C., Brown, C. E., Malinski, C., Perkins, R., Nasser, M., Lopez, C., Vittinghoff, E.,

Buchbinder, S. P., Scott, H., Charlebois, E. D., Havlir, D. V., Soge, O. O., Celum, C., & DoxyPEP Study Team. (2023). Postexposure Doxycycline to Prevent Bacterial Sexually Transmitted Infections. *The New England Journal of Medicine*, *388*(14), 1296–1306. https://doi.org/10.1056/NEJMoa2211934

Lugo-Miro, V. I., Green, M., & Mazur, L. (1992). Comparison of different metronidazole therapeutic regimens for bacterial vaginosis. A meta-analysis. *JAMA*, *268*(1), 92–95.

Ma, B., France, M. T., Crabtree, J., Holm, J. B., Humphrys, M. S., Brotman, R. M., & Ravel, J. (2020). A comprehensive non-redundant gene catalog reveals extensive within-community intraspecies diversity in the human vagina. *Nature Communications*, *11*, 940. https://doi.org/10.1038/s41467-020-14677-3

Machado, A., Jefferson, K. K., & Cerca, N. (2013). Interactions between Lactobacillus crispatus and Bacterial Vaginosis (BV)-Associated Bacterial Species in Initial Attachment and Biofilm Formation. *International Journal of Molecular Sciences*, *14*(6), 12004–12012. https://doi.org/10.3390/ijms140612004

Machado, A., Salgueiro, D., Harwich, M., Jefferson, K. K., & Cerca, N. (2013). Quantitative analysis of initial adhesion of bacterial vaginosis-associated anaerobes to ME-180 cells. *Anaerobe*, *23*, 1–4. https://doi.org/10.1016/j.anaerobe.2013.07.007

Machado, D., Castro, J., Palmeira-de-Oliveira, A., Martinez-de-Oliveira, J., & Cerca, N. (2016). Bacterial Vaginosis Biofilms: Challenges to Current Therapies and Emerging Solutions. *Frontiers in Microbiology*, *6*, 1528. https://doi.org/10.3389/fmicb.2015.01528

Mager, D. L., Ximenez-Fyvie, L. A., Haffajee, A. D., & Socransky, S. S. (2003). Distribution of selected bacterial species on intraoral surfaces. *Journal of Clinical Periodontology*, *30*(7), 644–654. https://doi.org/10.1034/j.1600-051x.2003.00376.x

Majigo, M. V., Kashindye, P., Mtulo, Z., & Joachim, A. (2021). Bacterial vaginosis, the leading cause of genital discharge among women presenting with vaginal infection in Dar es Salaam, Tanzania. *African Health Sciences*, *21*(2), 531–537. https://doi.org/10.4314/ahs.v21i2.7

Marconi, C., Donders, G. G. G., Bellen, G., Brown, D. R., Parada, C. M. G. L., & Silva, M. G. (2013). Sialidase activity in aerobic vaginitis is equal to levels during bacterial vaginosis. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, *167*(2), 205–209. https://doi.org/10.1016/j.ejogrb.2012.12.003

Marrazzo, J. M., Thomas, K. K., Agnew, K., & Ringwood, K. (2010). Prevalence and risks for bacterial vaginosis in women who have sex with women. *Sexually Transmitted Diseases*, *37*(5), 335–339.

Martin, M. (2011). Cutadapt removes adapter sequences from high-trhoughput sequencing reads. *EMBnet.Journal*, *17*(1), 10–12.

Masson, L., Passmore, J.-A. S., Liebenberg, L. J., Werner, L., Baxter, C., Arnold, K. B., Williamson, C., Little, F., Mansoor, L. E., Naranbhai, V., Lauffenburger, D. A., Ronacher, K., Walzl, G., Garrett, N. J., Williams, B. L., Couto-Rodriguez, M., Hornig, M., Lipkin, W. I., Grobler, A., … Abdool Karim, S. S. (2015). Genital inflammation and the risk of HIV acquisition in women. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *61*(2), 260–269. https://doi.org/10.1093/cid/civ298

Matsuki, T., Watanabe, K., Fujimoto, J., Takada, T., & Tanaka, R. (2004). Use of 16S rRNA gene-targeted group-specific primers for real-time PCR analysis of predominant bacteria in human feces. *Applied and Environmental Microbiology*, *70*(12), 7220–7228. https://doi.org/10.1128/AEM.70.12.7220-7228.2004

Mayer, B. T., Srinivasan, S., Fiedler, T. L., Marrazzo, J. M., Fredricks, D. N., & Schiffer, J. T. (2015). Rapid and Profound Shifts in the Vaginal Microbiota Following Antibiotic Treatment for Bacterial Vaginosis. *The Journal of Infectious Diseases*, *212*(5), 793–802. https://doi.org/10.1093/infdis/jiv079

McKinnon, L. R., Achilles, S. L., Bradshaw, C. S., Burgener, A., Crucitti, T., Fredricks, D. N., Jaspan, H. B., Kaul, R., Kaushic, C., Klatt, N., Kwon, D. S., Marrazzo, J. M., Masson, L., McClelland, R. S., Ravel, J., Van De Wijgert, J. H. H. M., Vodstrcil, L. A., & Tachedjian, G. (2019). The Evolving Facets of Bacterial Vaginosis: Implications for HIV Transmission. *AIDS Research and Human Retroviruses*, *35*(3), 219–228. https://doi.org/10.1089/aid.2018.0304

Michaelis, C., & Grohmann, E. (2023). Horizontal Gene Transfer of Antibiotic Resistance Genes in Biofilms. *Antibiotics (Basel, Switzerland)*, *12*(2), 328. https://doi.org/10.3390/antibiotics12020328

Mikamo, H., Kawazoe, K., Izumi, K., Sato, Y., & Tamaya, T. (1998). Studies on the clinical implications of anaerobes, especially *Prevotella bivia*, in obstetrics and gynecology. *Journal of Infection and Chemotherapy*, *4*(4), 177–187. https://doi.org/10.1007/BF02490164

Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D., & Gurevich, A. (2018). Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics (Oxford, England)*, *34*(13), i142–i150. https://doi.org/10.1093/bioinformatics/bty266

Millman, A., Bernheim, A., Stokar-Avihail, A., Fedorenko, T., Voichek, M., Leavitt, A., Oppenheimer-Shaanan, Y., & Sorek, R. (2020). Bacterial Retrons Function In Anti-Phage Defense. *Cell*, *183*(6), 1551-1561.e12. https://doi.org/10.1016/j.cell.2020.09.065

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, *37*(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Mir, A., Edraki, A., Lee, J., & Sontheimer, E. J. (2018). Type II-C CRISPR-Cas9 Biology, Mechanism and Application. *ACS Chemical Biology*, *13*(2), 357–365. https://doi.org/10.1021/acschembio.7b00855

Mirmonsef, P., Hotton, A. L., Gilbert, D., Burgad, D., Landay, A., Weber, K. M., Cohen, M., Ravel, J., & Spear, G. T. (2014). Free glycogen in vaginal fluids is associated with *Lactobacillus* colonization and low vaginal pH. *PloS One*, *9*(7), e102467. https://doi.org/10.1371/journal.pone.0102467

Mitra, A., MacIntyre, D. A., Ntritsos, G., Smith, A., Tsilidis, K. K., Marchesi, J. R., Bennett, P. R., Moscicki, A.-B., & Kyrgiou, M. (2020). The vaginal microbiota associates with the regression of untreated cervical intraepithelial neoplasia 2 lesions. *Nature Communications*, *11*, 1999. https://doi.org/10.1038/s41467-020-15856-y

Modena, B. D., Milam, R., Harrison, F., Cheeseman, J. A., Abecassis, M. M., Friedewald, J., Kirk, A. D., & Salomon, D. R. (2017). Changes in urinary microbiome populations correlate in kidney transplants with interstitial fibrosis and tubular atrophy documented in early surveillance biopsies. *American*

Journal of Transplantation : Official Journal of the American Society of Transplantation and the American Society of Transplant Surgeons, *17*(3), 712–723. https://doi.org/10.1111/ajt.14038

Moncla, B. J., Chappell, C. A., Debo, B. M., & Meyn, L. A. (2016). The Effects of Hormones and Vaginal Microflora on the Glycome of the Female Genital Tract: Cervical-Vaginal Fluid. *PLOS ONE*, *11*(7), e0158687. https://doi.org/10.1371/journal.pone.0158687

Morar, M., Bhullar, K., Hughes, D. W., Junop, M., & Wright, G. D. (2009). Structure and Mechanism of the Lincosamide Antibiotic Adenylyltransferase LinB. *Structure*, *17*(12), 1649–1659. https://doi.org/10.1016/j.str.2009.10.013

Morris, J. M., Roberts, C. L., Bowen, J. R., Patterson, J. A., Bond, D. M., Algert, C. S., Thornton, J. G., Crowther, C. A., & PPROMT Collaboration. (2016). Immediate delivery compared with expectant management after preterm pre-labour rupture of the membranes close to term (PPROMT trial): A randomised controlled trial. *Lancet (London, England)*, *387*(10017), 444–452. https://doi.org/10.1016/S0140-6736(15)00724-2

Morris, M., Rogers, P., & Kinghorn, G. (2001). Is bacterial vaginosis a sexually transmitted infection? *Sexually Transmitted Infections*, *77*(1), 63–68. https://doi.org/10.1136/sti.77.1.63

Mühlradt, P. F., & Golecki, J. R. (1975). Asymmetrical Distribution and Artifactual Reorientation of Lipopolysaccharide in the Outer Membrane Bilayer of Salmonella typhimurium. *European Journal of Biochemistry*, *51*(2), 343–352. https://doi.org/10.1111/j.1432-1033.1975.tb03934.x

Munoz, A., Hayward, M. R., Bloom, S. M., Rocafort, M., Ngcapu, S., Mafunda, N. A., Xu, J., Xulu, N., Dong, M., Dong, K. L., Ismail, N., Ndung'u, T., Ghebremichael, M. S., & Kwon, D. S. (2021). Correction to: Modeling the temporal dynamics of cervicovaginal microbiota identifies targets that may promote reproductive health. *Microbiome*, *9*, 206. https://doi.org/10.1186/s40168-021-01171-1

Muzny, C. A., Blanchard, E., Taylor, C. M., Aaron, K. J., Talluri, R., Griswold, M. E., Redden, D. T., Luo, M., Welsh, D. A., Van Der Pol, W. J., Lefkowitz, E. J., Martin, D. H., & Schwebke, J. R. (2018). Identification of Key Bacteria Involved in the Induction of Incident Bacterial Vaginosis: A Prospective Study. *The Journal of Infectious Diseases*, *218*(6), 966–978. https://doi.org/10.1093/infdis/jiy243

Muzny, C. A., & Schwebke, J. R. (2020). Asymptomatic Bacterial Vaginosis: To Treat or Not to Treat? *Current Infectious Disease Reports*, *22*(12), 32. https://doi.org/10.1007/s11908-020-00740-z

Muzny, C. A., & Sobel, J. D. (2022). The Role of Antimicrobial Resistance in Refractory and Recurrent Bacterial Vaginosis and Current Recommendations for Treatment. *Antibiotics*, *11*(4), 500. https://doi.org/10.3390/antibiotics11040500

Muzny, C. A., Taylor, C. M., Swords, W. E., Tamhane, A., Chattopadhyay, D., Cerca, N., & Schwebke, J. R. (2019a). An Updated Conceptual Model on the Pathogenesis of Bacterial Vaginosis. *The Journal of Infectious Diseases*, *220*(9), 1399–1405. https://doi.org/10.1093/infdis/jiz342

Muzny, C. A., Taylor, C. M., Swords, W. E., Tamhane, A., Chattopadhyay, D., Cerca, N., & Schwebke, J. R. (2019b). An Updated Conceptual Model on the Pathogenesis of Bacterial Vaginosis. *The Journal of Infectious Diseases*, *220*(9), 1399–1405. https://doi.org/10.1093/infdis/jiz342

Myer, L., Kuhn, L., Denny, L., & Wright, Jr., Thomas C. (2006). Recurrence of Symptomatic Bacterial Vaginosis 12 Months after Oral Metronidazole Therapy in HIV-Positive and -Negative Women. *The Journal of Infectious Diseases*, *194*(12), 1797–1799. https://doi.org/10.1086/509625

Myziuk, L., Romanowski, B., & Johnson, S. C. (2003). BVBlue Test for Diagnosis of Bacterial Vaginosis. *Journal of Clinical Microbiology*, *41*(5), 1925–1928. https://doi.org/10.1128/JCM.41.5.1925-1928.2003

Nagot, N., Ouedraogo, A., Defer, M., Vallo, R., Mayaud, P., & Van de Perre, P. (2007). Association between bacterial vaginosis and Herpes simplex virus type-2 infection: Implications for HIV acquisition studies. *Sexually Transmitted Infections*, *83*(5), 365–368. https://doi.org/10.1136/sti.2007.024794

Nakao, R., Ramstedt, M., Wai, S. N., & Uhlin, B. E. (2012). Enhanced Biofilm Formation by *Escherichia coli* LPS Mutants Defective in Hep Biosynthesis. *PLOS ONE*, *7*(12), e51241. https://doi.org/10.1371/journal.pone.0051241

Nasu, K., & Narahara, H. (2010). Pattern recognition via the toll-like receptor system in the human female genital tract. *Mediators of Inflammation*, *2010*, 976024. https://doi.org/10.1155/2010/976024

Nayfach, S., Camargo, A. P., Schulz, F., Eloe-Fadrosh, E., Roux, S., & Kyrpides, N. C. (2021). CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nature Biotechnology*, *39*(5), Article 5. https://doi.org/10.1038/s41587-020-00774-7

Nelson, D., Hanlon, A., Nachamkin, I., Haggerty, C., Mastrogiannis, D., Liu, C., & Fredricks, D. (2014). Early Pregnancy Changes in Bacterial Vaginosis-associated Bacteria and Preterm Delivery. *Paediatric and Perinatal Epidemiology*, *28*(2), 88–96. https://doi.org/10.1111/ppe.12106

Nelson, T. M., Borgogna, J. C., Michalek, R. D., Roberts, D. W., Rath, J. M., Glover, E. D., Ravel, J., Shardell, M. D., Yeoman, C. J., & Brotman, R. M. (2018). Cigarette smoking is associated with an altered vaginal tract metabolomic profile. *Scientific Reports*, *8*, 852. https://doi.org/10.1038/s41598-017-14943-3

Nelson, T. M., Borgogna, J.-L. C., Brotman, R. M., Ravel, J., Walk, S. T., & Yeoman, C. J. (2015). Vaginal biogenic amines: Biomarkers of bacterial vaginosis or precursors to vaginal dysbiosis? *Frontiers in Physiology*, *6*, 253. https://doi.org/10.3389/fphys.2015.00253

NIH HMP Working Group, Peterson, J., Garges, S., Giovanni, M., McInnes, P., Wang, L., Schloss, J. A., Bonazzi, V., McEwen, J. E., Wetterstrand, K. A., Deal, C., Baker, C. C., Di Francesco, V., Howcroft, T. K., Karp, R. W., Lunsford, R. D., Wellington, C. R., Belachew, T., Wright, M., … Guyer, M. (2009). The NIH Human Microbiome Project. *Genome Research*, *19*(12), 2317–2323. https://doi.org/10.1101/gr.096651.109

Nisha, K., Antony, B., & Udayalaxmi, J. (2019). Comparative analysis of virulence factors & biotypes of *Gardnerella vaginalis* isolated from the genital tract of women with & without bacterial vaginosis. *The Indian Journal of Medical Research*, *149*(1), 57–61. https://doi.org/10.4103/ijmr.IJMR_1674_16

Noyes, N., Cho, K.-C., Ravel, J., Forney, L. J., & Abdo, Z. (2018). Associations between sexual habits, menstrual hygiene practices, demographics and the vaginal microbiome as revealed by Bayesian network analysis. *PLoS ONE*, *13*(1), e0191625. https://doi.org/10.1371/journal.pone.0191625

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.

Journal of Clinical Microbiology, *29*(2), 297–301. https://doi.org/10.1128/JCM.29.2.297-301.1991
Nunn, K. L., Wang, Y.-Y., Harit, D., Humphrys, M. S., Ma, B., Cone, R., Ravel, J., & Lai, S. K. (2015). Enhanced Trapping of HIV-1 by Human Cervicovaginal Mucus Is Associated with *Lactobacillus crispatus*-Dominant Microbiota. *mBio*, *6*(5), 10.1128/mbio.01084-15. https://doi.org/10.1128/mbio.01084-15
Oduyebo, O. O., Anorlu, R. I., & Ogunsola, F. T. (2009). The effects of antimicrobial therapy on bacterial vaginosis in non-pregnant women. *The Cochrane Database of Systematic Reviews*, *3*, CD006055. https://doi.org/10.1002/14651858.CD006055.pub2
Oerlemans, E., Ahannach, S., Wittouck, S., Dehay, E., De Boeck, I., Ballet, N., Rodriguez, B., Tuyaerts, I., & Lebeer, S. (2022). Impacts of Menstruation, Community Type, and an Oral Yeast Probiotic on the Vaginal Microbiome. *mSphere*, *7*(5), e00239-22. https://doi.org/10.1128/msphere.00239-22
Oh, J., Byrd, A. L., Deming, C., Conlan, S., NISC Comparative Sequencing Program, Kong, H. H., & Segre, J. A. (2014). Biogeography and individuality shape function in the human skin metagenome. *Nature*, *514*(7520), 59–64. https://doi.org/10.1038/nature13786
Okamoto, M., Maeda, N., Kondo, K., & Leung, K.-P. (1999). Hemolytic and hemagglutinating activities of *Prevotella intermedia* and *Prevotella nigrescens*. *FEMS Microbiology Letters*, *178*(2), 299–304. https://doi.org/10.1111/j.1574-6968.1999.tb08691.x
O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., … Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*, *44*(D1), D733-745. https://doi.org/10.1093/nar/gkv1189
Oliver, W. W., & Wherry, W. B. (1921). Notes on some bacterial parasites of the human mucous membranes. *The Journal of Infectious Diseases*, *28*(4), 341–344. https://doi.org/10.1093/indfis/28.4.341
Onderdonk, A. B., Delaney, M. L., Hinkson, P. L., & DuBois, A. M. (1992). Quantitative and qualitative effects of douche preparations on vaginal microflora. *Obstetrics and Gynecology*, *80*(3 Pt 1), 333–338.
Onderdonk, A. B., Hecht, J. L., McElrath, T. F., Delaney, M. L., Allred, E. N., Leviton, A., & ELGAN Study Investigators. (2008). Colonization of second-trimester placenta parenchyma. *American Journal of Obstetrics and Gynecology*, *199*(1), 52.e1-52.e10. https://doi.org/10.1016/j.ajog.2007.11.068
Pabich, W. L., Fihn, S. D., Stamm, W. E., Scholes, D., Boyko, E. J., & Gupta, K. (2003). Prevalence and determinants of vaginal flora alterations in postmenopausal women. *The Journal of Infectious Diseases*, *188*(7), 1054–1058. https://doi.org/10.1086/378203
Palmer, M., Steenkamp, E. T., Blom, J., Hedlund, B. P., & Venter, S. N. (2020). All ANIs are not created equal: Implications for prokaryotic species boundaries and integration of ANIs into polyphasic taxonomy. *International Journal of Systematic and Evolutionary Microbiology*, *70*(4), 2937–2948. https://doi.org/10.1099/ijsem.0.004124
Paramel Jayaprakash, T., Wagner, E. C., van Schalkwyk, J., Albert, A. Y. K., Hill, J. E., & Money, D. M. (2016). High Diversity and Variability in the Vaginal

Microbiome in Women following Preterm Premature Rupture of Membranes (PPROM): A Prospective Cohort Study. *PLoS ONE*, *11*(11), e0166794. https://doi.org/10.1371/journal.pone.0166794

Park, B. S., Song, D. H., Kim, H. M., Choi, B.-S., Lee, H., & Lee, J.-O. (2009). The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature*, *458*(7242), 1191–1195. https://doi.org/10.1038/nature07830

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*, *25*(7), 1043–1055. https://doi.org/10.1101/gr.186072.114

Patterson, J. L., Girerd, P. H., Karjane, N. W., & Jefferson, K. K. (2007). Effect of biofilm phenotype on resistance of *Gardnerella vaginalis* to hydrogen peroxide and lactic acid. *American Journal of Obstetrics and Gynecology*, *197*(2), 170.e1-7. https://doi.org/10.1016/j.ajog.2007.02.027

Peeters, E., Nelis, H. J., & Coenye, T. (2008). Comparison of multiple methods for quantification of microbial biofilms grown in microtiter plates. *Journal of Microbiological Methods*, *72*(2), 157–165. https://doi.org/10.1016/j.mimet.2007.11.010

Penadés, J. R., Chen, J., Quiles-Puchalt, N., Carpena, N., & Novick, R. P. (2015). Bacteriophage-mediated spread of bacterial virulence genes. *Current Opinion in Microbiology*, *23*, 171–178. https://doi.org/10.1016/j.mib.2014.11.019

Peng, Y., Cai, X., Li, M., Deng, L., Wang, Y., Qiu, Y., Zhao, L., Xiao, Y., Xu, L., & Hou, Q. (2023). The first completed genome of species *Prevotella bivia*, assembled from a clinically derived strain PLW0727. *Journal of Global Antimicrobial Resistance*, *35*, 268–270. https://doi.org/10.1016/j.jgar.2023.10.009

Persson, R., Hitti, J., Verhelst, R., Vaneechoutte, M., Persson, R., Hirschi, R., Weibel, M., Rothen, M., Temmerman, M., Paul, K., & Eschenbach, D. (2009). The vaginal microflora in relation to gingivitis. *BMC Infectious Diseases*, *9*, 6. https://doi.org/10.1186/1471-2334-9-6

Peters, R. P. H., Garrett, N., Chandiwana, N., Brink, A. J., Cohen, K., Gill, K., Chidarikire, T., Wattrus, C., Nel, J. S., Moosa, M. Y. S., & Bekker, L.-G. (2022). South African HIV Clinicians Society 2022 guideline for the management of sexually transmitted infections: Moving towards best practice. *Southern African Journal of HIV Medicine*, *21*(1). https://doi.org/10.4102/sajhivmed.v23i1.1450

Peterson, J., Garges, S., Giovanni, M., McInnes, P., Wang, L., Schloss, J. A., Bonazzi, V., McEwen, J. E., Wetterstrand, K. A., Deal, C., Baker, C. C., Di Francesco, V., Howcroft, T. K., Karp, R. W., Lunsford, R. D., Wellington, C. R., Belachew, T., Wright, M., Giblin, C., … Guyer, M. (2009). The NIH Human Microbiome Project. *Genome Research*, *19*(12), 2317–2323. https://doi.org/10.1101/gr.096651.109

Petrina, M. A. B., Cosentino, L. A., Rabe, L. K., & Hillier, S. L. (2017). Susceptibility of bacterial vaginosis (BV)-associated bacteria to secnidazole compared to metronidazole, tinidazole and clindamycin. *Anaerobe*, *47*, 115–119. https://doi.org/10.1016/j.anaerobe.2017.05.005

Petrina, M. A. B., Cosentino, L. A., Wiesenfeld, H. C., Darville, T., & Hillier, S. L. (2019). Susceptibility of endometrial isolates recovered from women with clinical pelvic inflammatory disease or histological endometritis to antimicrobial agents. *Anaerobe*, *56*, 61–65. https://doi.org/10.1016/j.anaerobe.2019.02.005

Pritchard, L., Glover, R., Humphris, S., Elphinstone, J., & Toth, I. (2015). Genomics and Taxonomy in Diagnostics for Food Security: Soft-rotting enterobacterial plant pathogens. *Anal. Methods*, *8*. https://doi.org/10.1039/C5AY02550H

Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., & Korobeynikov, A. (2020). Using SPAdes De Novo Assembler. *Current Protocols in Bioinformatics*, *70*(1), e102. https://doi.org/10.1002/cpbi.102

Pumbwe, L., Chang, A., Smith, R. L., & Wexler, H. M. (2007). BmeRABC5 Is a Multidrug Efflux System That Can Confer Metronidazole Resistance in *Bacteroides fragilis*. *Microbial Drug Resistance*, *13*(2), 96–101. https://doi.org/10.1089/mdr.2007.719

Pybus, V., & Onderdonk, A. B. (1997). Evidence for a commensal, symbiotic relationship between *Gardnerella vaginalis* and *Prevotella bivia* involving ammonia: Potential significance for bacterial vaginosis. *The Journal of Infectious Diseases*, *175*(2), 406–413. https://doi.org/10.1093/infdis/175.2.406

Pybus, V., & Onderdonk, A. B. (1998). A commensal symbiosis between *Prevotella bivia* and *Peptostreptococcus anaerobius* involves amino acids: Potential significance to the pathogenesis of bacterial vaginosis. *FEMS Immunology and Medical Microbiology*, *22*(4), 317–327. https://doi.org/10.1111/j.1574-695X.1998.tb01221.x

Raetz, C. R. H., Guan, Z., Ingram, B. O., Six, D. A., Song, F., Wang, X., & Zhao, J. (2009). Discovery of new biosynthetic pathways: The lipid A story. *Journal of Lipid Research*, *50*, S103–S108. https://doi.org/10.1194/jlr.R800060-JLR200

Ranjit, E., Raghubanshi, B. R., Maskey, S., & Parajuli, P. (2018). Prevalence of Bacterial Vaginosis and Its Association with Risk Factors among Nonpregnant Women: A Hospital Based Study. *International Journal of Microbiology*, *2018*, 8349601. https://doi.org/10.1155/2018/8349601

Rather, M. A., Gupta, K., & Mandal, M. (2021). Microbial biofilm: Formation, architecture, antibiotic resistance, and control strategies. *Brazilian Journal of Microbiology: [Publication of the Brazilian Society for Microbiology]*, *52*(4), 1701–1718. https://doi.org/10.1007/s42770-021-00624-x

Ravel, J., Gajer, P., Abdo, Z., Schneider, G. M., Koenig, S. S. K., McCulle, S. L., Karlebach, S., Gorle, R., Russell, J., Tacket, C. O., Brotman, R. M., Davis, C. C., Ault, K., Peralta, L., & Forney, L. J. (2011). Vaginal microbiome of reproductive-age women. *Proceedings of the National Academy of Sciences of the United States of America*, *108 Suppl 1*(Suppl 1), 4680–4687. https://doi.org/10.1073/pnas.1002611107

Reissier, S., Penven, M., Guérin, F., & Cattoir, V. (2023). Recent Trends in Antimicrobial Resistance among Anaerobic Clinical Isolates. *Microorganisms*, *11*(6), 1474. https://doi.org/10.3390/microorganisms11061474

Richter, M., & Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences*, *106*(45), 19126–19131. https://doi.org/10.1073/pnas.0906412106

Robinson, L. S., Schwebke, J., Lewis, W. G., & Lewis, A. L. (2019). Identification and characterization of NanH2 and NanH3, enzymes responsible for sialidase activity in the vaginal bacterium *Gardnerella vaginalis*. *The Journal of Biological Chemistry*, *294*(14), 5230–5245. https://doi.org/10.1074/jbc.RA118.006221

Rossano, F., Rizzo, A., Sanges, M. R., Cipollaro de L'Ero, G., & Tufano, M. A. (1993). Human monocytes and gingival fibroblasts release tumor necrosis factor-alpha, interleukin-1 alpha and interleukin-6 in response to particulate and

soluble fractions of *Prevotella melaninogenica* and *Fusobacterium nucleatum*. *International Journal of Clinical & Laboratory Research*, *23*(3), 165–168. https://doi.org/10.1007/BF02592303

Rouli, L., Merhej, V., Fournier, P.-E., & Raoult, D. (2015). The bacterial pangenome as a new tool for analysing pathogenic bacteria. *New Microbes and New Infections*, *7*, 72–85. https://doi.org/10.1016/j.nmni.2015.06.005

Rousset, F., Depardieu, F., Miele, S., Dowding, J., Laval, A.-L., Lieberman, E., Garry, D., Rocha, E. P. C., Bernheim, A., & Bikard, D. (2022). Phages and their satellites encode hotspots of antiviral systems. *Cell Host & Microbe*, *30*(5), 740-753.e5. https://doi.org/10.1016/j.chom.2022.02.018

Roux, S., Adriaenssens, E. M., Dutilh, B. E., Koonin, E. V., Kropinski, A. M., Krupovic, M., Kuhn, J. H., Lavigne, R., Brister, J. R., Varsani, A., Amid, C., Aziz, R. K., Bordenstein, S. R., Bork, P., Breitbart, M., Cochrane, G. R., Daly, R. A., Desnues, C., Duhaime, M. B., … Eloe-Fadrosh, E. A. (2019). Minimum Information about an Uncultivated Virus Genome (MIUViG). *Nature Biotechnology*, *37*(1), Article 1. https://doi.org/10.1038/nbt.4306

Sakamoto, M., & Ohkuma, M. (2010). Usefulness of the *hsp60* gene for the identification and classification of Gram-negative anaerobic rods. *Journal of Medical Microbiology*, *59*(Pt 11), 1293–1302. https://doi.org/10.1099/jmm.0.020420-0

Sali, W., Patoli, D., Pais de Barros, J.-P., Labbé, J., Deckert, V., Duhéron, V., Le Guern, N., Blache, D., Chaumont, D., Lesniewska, E., Gasquet, B., Paul, C., Moreau, M., Denat, F., Masson, D., Lagrost, L., & Gautier, T. (2019). Polysaccharide Chain Length of Lipopolysaccharides From Salmonella Minnesota Is a Determinant of Aggregate Stability, Plasma Residence Time and Proinflammatory Propensity in vivo. *Frontiers in Microbiology*, *10*, 1774. https://doi.org/10.3389/fmicb.2019.01774

Sauter, C., & Wolfensberger, C. (1980). Interferon in human serum after injection of endotoxin. *Lancet (London, England)*, *2*(8199), 852–853. https://doi.org/10.1016/s0140-6736(80)90189-0

Schellenberg, J., Links, M. G., Hill, J. E., Dumonceaux, T. J., Peters, G. A., Tyler, S., Ball, T. B., Severini, A., & Plummer, F. A. (2009). Pyrosequencing of the chaperonin-60 universal target as a tool for determining microbial community composition. *Applied and Environmental Microbiology*, *75*(9), 2889–2898. https://doi.org/10.1128/AEM.01640-08

Schellenberg, J., Paramel Jayaprakash, T., Withana Gamage, N., Patterson, M. H., Vaneechoutte, M., & Hill, J. E. (2016). *Gardnerella vaginalis* Subgroups Defined by cpn60 Sequencing and Sialidase Activity in Isolates from Canada, Belgium and Kenya. *PLoS ONE*, *11*(1), e0146510. https://doi.org/10.1371/journal.pone.0146510

Schleicher, L., Herdan, S., Fritz, G., Trautmann, A., Seifert, J., & Steuber, J. (2021). Central Carbon Metabolism, Sodium-Motive Electron Transfer, and Ammonium Formation by the Vaginal Pathogen *Prevotella bivia*. *International Journal of Molecular Sciences*, *22*(21), Article 21. https://doi.org/10.3390/ijms222111925

Schloter, M., Lebuhn, M., Heulin, T., & Hartmann, A. (2000). Ecology and evolution of bacterial microdiversity. *FEMS Microbiology Reviews*, *24*(5), 647–660. https://doi.org/10.1111/j.1574-6976.2000.tb00564.x

Schuyler, J. A., Mordechai, E., Adelson, M. E., Sobel, J. D., Gygax, S. E., & Hilbert, D. W. (2016). Identification of intrinsically metronidazole-resistant clades of

Gardnerella vaginalis. *Diagnostic Microbiology and Infectious Disease*, *84*(1), 1–3. https://doi.org/10.1016/j.diagmicrobio.2015.10.006
Schwebke, J. R., Muzny, C. A., & Josey, W. E. (2014). Role of *Gardnerella vaginalis* in the pathogenesis of bacterial vaginosis: A conceptual model. *The Journal of Infectious Diseases*, *210*(3), 338–343. https://doi.org/10.1093/infdis/jiu089
Schwebke, J. R., Richey, C. M., & Weiss, H. L. (1999). Correlation of Behaviors with Microbiological Changes in Vaginal Flora. *The Journal of Infectious Diseases*, *180*(5), 1632–1636. https://doi.org/10.1086/315065
Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics (Oxford, England)*, *30*(14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153
Seemann, T. (2015). *ABRicate* [Computer software]. https://github.com/tseemann/abricate
Seña, A. C., Goldstein, L. A., Ramirez, G., Parish, A. J., & McClelland, R. S. (2021). Bacterial Vaginosis and Its Association With Incident *Trichomonas vaginalis* Infections: A Systematic Review and Meta-Analysis. *Sexually Transmitted Diseases*, *48*(12), e192. https://doi.org/10.1097/OLQ.0000000000001537
Sexually Transmitted Infections Management Guidelines 2018. (n.d.). Department of Health, Republic of South Africa.
Sexually transmitted infections treatment guidelines. (2021). Centers for Disease Control and Prevention.
Shaffer, M., Borton, M. A., McGivern, B. B., Zayed, A. A., La Rosa, S. L., Solden, L. M., Liu, P., Narrowe, A. B., Rodríguez-Ramos, J., Bolduc, B., Gazitúa, M. C., Daly, R. A., Smith, G. J., Vik, D. R., Pope, P. B., Sullivan, M. B., Roux, S., & Wrighton, K. C. (2020). DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Research*, *48*(16), 8883–8900. https://doi.org/10.1093/nar/gkaa621
Shah, H. N., & Collins, D. M. (1990). *Prevotella*, a new genus to include *Bacteroides melaninogenicus* and related species formerly classified in the genus *Bacteroides*. *International Journal of Systematic Bacteriology*, *40*(2), 205–208. https://doi.org/10.1099/00207713-40-2-205
Sharma, G., Garg, N., Hasan, S., & Shirodkar, S. (2022). *Prevotella*: An insight into its characteristics and associated virulence factors. *Microbial Pathogenesis*, *169*, 105673. https://doi.org/10.1016/j.micpath.2022.105673
Sherrard, L. J., Schaible, B., Graham, K. A., McGrath, S. J., McIlreavey, L., Hatch, J., Wolfgang, M. C., Muhlebach, M. S., Gilpin, D. F., Schneiders, T., Elborn, J. S., & Tunney, M. M. (2014). Mechanisms of reduced susceptibility and genotypic prediction of antibiotic resistance in *Prevotella* isolated from cystic fibrosis (CF) and non-CF patients. *The Journal of Antimicrobial Chemotherapy*, *69*(10), 2690–2698. https://doi.org/10.1093/jac/dku192
Shkoporov, A. N., O'Regan, O., Smith, L., Khokhlova, E. V., Draper, L. A., Ross, R. P., & Hill, C. (2024). Dynamic nature of viral and bacterial communities in human faeces. *iScience*, *27*(2), 108778. https://doi.org/10.1016/j.isci.2023.108778
Simhan, H. N., Caritis, S. N., Krohn, M. A., & Hillier, S. L. (2003). Elevated vaginal pH and neutrophils are associated strongly with early spontaneous preterm birth. *American Journal of Obstetrics and Gynecology*, *189*(4), 1150–1154. https://doi.org/10.1067/S0002-9378(03)00582-9

Socransky, S. S., Haffajee, A. D., Cugini, M. A., Smith, C., & Kent, R. L. (1998). Microbial complexes in subgingival plaque. *Journal of Clinical Periodontology*, *25*(2), 134–144. https://doi.org/10.1111/j.1600-051x.1998.tb02419.x

Sood, A., Ray, P., & Angrup, A. (2023). Anaerobic Gram-Negative Bacteria: Role as a Reservoir of Antibiotic Resistance. *Antibiotics (Basel, Switzerland)*, *12*(5), 942. https://doi.org/10.3390/antibiotics12050942

Sousa, L., Almeida, C., Muzny, C. A., & Cerca, N. (2023). Development of a *Prevotella bivia* PNA probe and a multiplex approach to detect three relevant species in bacterial vaginosis-associated biofilms. https://doi.org/10.1038/s41522-023-00411-6

Srinivasan, S., Hoffman, N. G., Morgan, M. T., Matsen, F. A., Fiedler, T. L., Hall, R. W., Ross, F. J., McCoy, C. O., Bumgarner, R., Marrazzo, J. M., & Fredricks, D. N. (2012). Bacterial Communities in Women with Bacterial Vaginosis: High Resolution Phylogenetic Analyses Reveal Relationships of Microbiota to Clinical Criteria. *PLOS ONE*, *7*(6), e37818. https://doi.org/10.1371/journal.pone.0037818

Srinivasan, S., Munch, M. M., Sizova, M. V., Fiedler, T. L., Kohler, C. M., Hoffman, N. G., Liu, C., Agnew, K. J., Marrazzo, J. M., Epstein, S. S., & Fredricks, D. N. (2016). More Easily Cultivated Than Identified: Classical Isolation With Molecular Identification of Vaginal Bacteria. *The Journal of Infectious Diseases*, *214*(Suppl 1), S21–S28. https://doi.org/10.1093/infdis/jiw192

Staden, R. (1996). The staden sequence analysis package. *Molecular Biotechnology*, *5*(3), 233–241. https://doi.org/10.1007/BF02900361

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Steffens, L. S., Nicholson, S., Paul, L. V., Nord, C. E., Patrick, S., & Abratt, V. R. (2010). Bacteroides fragilis RecA protein overexpression causes resistance to metronidazole. *Research in Microbiology*, *161*(5–3), 346–354. https://doi.org/10.1016/j.resmic.2010.04.003

Surette, M. G., Miller, M. B., & Bassler, B. L. (1999). Quorum sensing in *Escherichia coli*, *Salmonella typhimurium*, and *Vibrio harveyi*: A new family of genes responsible for autoinducer production. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(4), 1639–1644. https://doi.org/10.1073/pnas.96.4.1639

Suzuki, N., Fukamachi, H., Arimoto, T., Yamamoto, M., & Igarashi, T. (2012). Contribution of hly homologs to the hemolytic activity of *Prevotella intermedia*. *Anaerobe*, *18*(3), 350–356. https://doi.org/10.1016/j.anaerobe.2012.04.005

Swidsinski, A., Mendling, W., Loening-Baucke, V., Ladhoff, A., Swidsinski, S., Hale, L. P., & Lochs, H. (2005a). Adherent biofilms in bacterial vaginosis. *Obstetrics and Gynecology*, *106*(5 Pt 1), 1013–1023. https://doi.org/10.1097/01.AOG.0000183594.45524.d2

Swidsinski, A., Mendling, W., Loening-Baucke, V., Ladhoff, A., Swidsinski, S., Hale, L. P., & Lochs, H. (2005b). Adherent biofilms in bacterial vaginosis. *Obstetrics and Gynecology*, *106*(5 Pt 1), 1013–1023. https://doi.org/10.1097/01.AOG.0000183594.45524.d2

Swidsinski, A., Mendling, W., Loening-Baucke, V., Swidsinski, S., Dörffel, Y., Scholze, J., Lochs, H., & Verstraelen, H. (2008). An adherent *Gardnerella vaginalis* biofilm persists on the vaginal epithelium after standard therapy with oral

metronidazole. *American Journal of Obstetrics and Gynecology*, *198*(1), 97.e1-6. https://doi.org/10.1016/j.ajog.2007.06.039

Tajkarimi, M., & Wexler, H. M. (2017). CRISPR-Cas Systems in *Bacteroides fragilis,* an Important Pathobiont in the Human Gut Microbiome. *Frontiers in Microbiology*, *8*. https://www.frontiersin.org/articles/10.3389/fmicb.2017.02234

Tamarelle, J., Penaud, B., Tyssandier, B., Guichoux, E., de Barbeyrac, B., Peuchant, O., Baita, D., Ouziel-Duretz, C., Poudens, B., Brun, R., Jouvert, S., Tesson, A., Carrière, J., Diaz, M., Forget, C., Le Hen, I., Ahano-Ducourneau, F., Van, D. H., Robert, P., … Galet, J. (2023). Effects of azithromycin and doxycycline on the vaginal microbiota of women with urogenital *Chlamydia trachomatis* infection: A substudy of the Chlazidoxy randomized controlled trial. *Clinical Microbiology and Infection*, *29*(8), 1056–1062. https://doi.org/10.1016/j.cmi.2023.04.020

Tan, C. W., & Chlebicki, M. P. (2016). Urinary tract infections in adults. *Singapore Medical Journal*, *57*(9), 485–490. https://doi.org/10.11622/smedj.2016153

Taylor, B. D., Darville, T., & Haggerty, C. L. (2013). Does Bacterial Vaginosis Cause Pelvic Inflammatory Disease? *Sexually Transmitted Diseases*, *40*(2), 117. https://doi.org/10.1097/OLQ.0b013e31827c5a5b

Teng, W., Liao, B., Chen, M., & Shu, W. (2023). Genomic Legacies of Ancient Adaptation Illuminate GC-Content Evolution in Bacteria. *Microbiology Spectrum*, *11*(1), e0214522. https://doi.org/10.1128/spectrum.02145-22

Tesson, F., Hervé, A., Mordret, E., Touchon, M., d'Humières, C., Cury, J., & Bernheim, A. (2022). Systematic and quantitative view of the antiviral arsenal of prokaryotes. *Nature Communications*, *13*(1), 2561. https://doi.org/10.1038/s41467-022-30269-9

Tett, A., Pasolli, E., Masetti, G., Ercolini, D., & Segata, N. (2021). *Prevotella* diversity, niches and interactions with the human host. *Nature Reviews Microbiology*, *19*(9), 585–599. https://doi.org/10.1038/s41579-021-00559-y

Thomas, S. (1928). Döderlein's Bacillus: *Lactobacillus Acidophilus*. *The Journal of Infectious Diseases*, *43*(3), 218–227.

Tonkin-Hill, G., MacAlasdair, N., Ruis, C., Weimann, A., Horesh, G., Lees, J. A., Gladstone, R. A., Lo, S., Beaudoin, C., Floto, R. A., Frost, S. D. W., Corander, J., Bentley, S. D., & Parkhill, J. (2020). Producing polished prokaryotic pangenomes with the Panaroo pipeline. *Genome Biology*, *21*(1), 180. https://doi.org/10.1186/s13059-020-02090-4

Toprak, N. U., Akgul, O., Sóki, J., Soyletir, G., Nagy, E., Leitner, E., Wybo, I., Tripkovic, V., Justesen, U. S., Jean-Pierre, H., Stingu, S. C., Rodloff, A. C., Paparaskevas, J., Urban, E., Jamal, W., Rotimi, V. O., Veloo, A. C. M., Jeverica, S., Kulekci, G., … Morris, T. (2020). Detection of beta-lactamase production in clinical *Prevotella* species by MALDI-TOF MS method. *Anaerobe*, *65*, 102240. https://doi.org/10.1016/j.anaerobe.2020.102240

Torrone, E. A., Morrison, C. S., Chen, P.-L., Kwok, C., Francis, S. C., Hayes, R. J., Looker, K. J., McCormack, S., McGrath, N., van de Wijgert, J. H. H. M., Watson-Jones, D., Low, N., Gottlieb, S. L., & STIMA Working Group. (2018). Prevalence of sexually transmitted infections and bacterial vaginosis among women in sub-Saharan Africa: An individual participant data meta-analysis of 18 HIV prevention studies. *PLoS Medicine*, *15*(2), e1002511. https://doi.org/10.1371/journal.pmed.1002511

Tribble, G. D., Garza, J. J., Yeung, V. L., Rigney, T. W., Dao, D.-H. V., Rodrigues, P. H., Walker, C. B., & Smith, C. J. (2010). Genetic Analysis of Mobile tetQ

Elements in Oral *Prevotella* Species. *Anaerobe*, *16*(6), 604–609. https://doi.org/10.1016/j.anaerobe.2010.08.008

Ulger Toprak, N., Veloo, A. C. M., Urban, E., Wybo, I., Justesen, U. S., Jean-Pierre, H., Morris, T., Akgul, O., Kulekci, G., Soyletir, G., & Nagy, E. (2018). A multicenter survey of antimicrobial susceptibility of *Prevotella* species as determined by Etest methodology. *Anaerobe*, *52*, 9–15. https://doi.org/10.1016/j.anaerobe.2018.05.005

UNAIDS Fact Sheet. (2023).

van Teijlingen, N. H., Helgers, L. C., Sarrami-Forooshani, R., Zijlstra-Willems, E. M., van Hamme, J. L., Segui-Perez, C., van Smoorenburg, M. Y., Borgdorff, H., van de Wijgert, J. H., van Leeuwen, E., van der Post, J. A., Strijbis, K., Ribeiro, C. M., & Geijtenbeek, T. B. (2022). Vaginal bacterium *Prevotella timonensis* turns protective Langerhans cells into HIV-1 reservoirs for virus dissemination. *The EMBO Journal*, *41*(19), e110629. https://doi.org/10.15252/embj.2022110629

van Teijlingen, N. H., Helgers, L. C., Zijlstra - Willems, E. M., van Hamme, J. L., Ribeiro, C. M. S., Strijbis, K., & Geijtenbeek, T. B. H. (2020). Vaginal dysbiosis associated-bacteria *Megasphaera elsdenii* and *Prevotella timonensis* induce immune activation via dendritic cells. *Journal of Reproductive Immunology*, *138*, 103085. https://doi.org/10.1016/j.jri.2020.103085

Vancuren, S. J., & Hill, J. E. (2019). Update on cpnDB: A reference database of chaperonin sequences. *Database: The Journal of Biological Databases and Curation*, *2019*, baz033. https://doi.org/10.1093/database/baz033

Veloo, A. C. M., Baas, W. H., Haan, F. J., Coco, J., & Rossen, J. W. (2019). Prevalence of antimicrobial resistance genes in *Bacteroides* spp. And *Prevotella* spp. Dutch clinical isolates. *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, *25*(9), 1156.e9-1156.e13. https://doi.org/10.1016/j.cmi.2019.02.017

Veloo, A. C. M., Chlebowicz, M., Winter, H. L. J., Bathoorn, D., & Rossen, J. W. A. (2018). Three metronidazole-resistant *Prevotella bivia* strains harbour a mobile element, encoding a novel *nim* gene, *nimK*, and an efflux small MDR transporter. *Journal of Antimicrobial Chemotherapy*, *73*(10), 2687–2690. https://doi.org/10.1093/jac/dky236

Veloo, A. C. M., Tokman, H. B., Jean-Pierre, H., Dumont, Y., Jeverica, S., Lienhard, R., Novak, A., Rodloff, A., Rotimi, V., Wybo, I., & Nagy, E. (2020). Antimicrobial susceptibility profiles of anaerobic bacteria, isolated from human clinical specimens, within different European and surrounding countries. A joint ESGAI study. *Anaerobe*, *61*, 102111. https://doi.org/10.1016/j.anaerobe.2019.102111

Walboomers, J. M., Jacobs, M. V., Manos, M. M., Bosch, F. X., Kummer, J. A., Shah, K. V., Snijders, P. J., Peto, J., Meijer, C. J., & Muñoz, N. (1999). Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *The Journal of Pathology*, *189*(1), 12–19. https://doi.org/10.1002/(SICI)1096-9896(199909)189:1<12::AID-PATH431>3.0.CO;2-F

Wang, L., Hou, Y., Yuan, H., & Chen, H. (2022). The role of tryptophan in *Chlamydia trachomatis* persistence. *Frontiers in Cellular and Infection Microbiology*, *12*, 931653. https://doi.org/10.3389/fcimb.2022.931653

Wang, Y., Noël-Romas, L., Perner, M., Knodel, S., Molatlhegi, R., Hoger, S., Birse, K., Zuend, C. F., McKinnon, L. R., & Burgener, A. D. (2023). Non-

Lactobacillus-Dominant and Polymicrobial Vaginal Microbiomes Are More Common in Younger South African Women and Predictive of Increased Risk of Human Immunodeficiency Virus Acquisition. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *76*(8), 1372–1381. https://doi.org/10.1093/cid/ciac938

Waters, J. L., & Salyers, A. A. (2013). Regulation of CTnDOT Conjugative Transfer Is a Complex and Highly Coordinated Series of Events. *mBio*, *4*(6), 10.1128/mbio.00569-13. https://doi.org/10.1128/mbio.00569-13

Watts, D. H., Krohn, M. A., Hillier, S. L., & Eschenbach, D. A. (1992). The association of occult amniotic fluid infection with gestational age and neonatal outcome among women in preterm labor. *Obstetrics and Gynecology*, *79*(3), 351–357. https://doi.org/10.1097/00006250-199203000-00005

Wayne, L. G., Brenner, D. J., Colwell, R., Grimont, P., Krichevsky, M., Moore, L. H., Moore, W. E. C., Murray, R. G. E., Stackebrandt, E., Starr, M. P., & Truper, H. G. (1987). Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. *International Journal of Systematic Bacteriology*, *37*. https://doi.org/10.1099/00207713-37-4-463

Webb, K., Olagoke, O., Baird, T., Neill, J., Pham, A., Wells, T., Ramsay, K. A., Bell, S. C., Sarovich, D. S., & Price, E. P. (2022). Genomic diversity and antimicrobial resistance of *Prevotella* species isolated from chronic lung disease airways. *Microbial Genomics*, *8*(2). https://doi.org/10.1099/mgen.0.000754

Weisblum, B. (1995). Erythromycin resistance by ribosome modification. *Antimicrobial Agents and Chemotherapy*, *39*(3), 577–585. https://doi.org/10.1128/AAC.39.3.577

Weisburg, W. G., Barns, S. M., Pelletier, D. A., & Lane, D. J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *Journal of Bacteriology*, *173*(2), 697–703.

Westphal, O., & Jann, K. (1965). *Bacterial lipopolysaccharides*. *5*(83).

Whitfield, C., Williams, D. M., & Kelly, S. D. (2020). Lipopolysaccharide O-antigens—Bacterial glycans made to measure. *The Journal of Biological Chemistry*, *295*(31), 10593–10609. https://doi.org/10.1074/jbc.REV120.009402

Wiesenfeld, H. C., Hillier, S. L., Krohn, M. A., Landers, D. V., & Sweet, R. L. (2003). Bacterial vaginosis is a strong predictor of *Neisseria gonorrhoeae* and *Chlamydia trachomatis* infection. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *36*(5), 663–668. https://doi.org/10.1086/367658

Wiggins, R., Hicks, S., Soothill, P., Millar, M., & Corfield, A. (2001). Mucinases and sialidases: Their role in the pathogenesis of sexually transmitted infections in the female genital tract. *Sexually Transmitted Infections*, *77*(6), 402–408. https://doi.org/10.1136/sti.77.6.402

Wilson, C., Lukowicz, R., Merchant, S., Valquier-Flynn, H., Caballero, J., Sandoval, J., Okuom, M., Huber, C., Brooks, T. D., Wilson, E., Clement, B., Wentworth, C. D., & Holmes, A. E. (2017). Quantitative and Qualitative Assessment Methods for Biofilm Growth: A Mini-review. *Research & Reviews. Journal of Engineering and Technology*, *6*(4), http://www.rroij.com/open-access/quantitative-and-qualitative-assessment-methods-for-biofilm-growth-a-minireview-.pdf.

Wolf, L. J., & Stingu, C.-S. (2022). Antimicrobial Susceptibility Profile of Rare Anaerobic Bacteria. *Antibiotics*, *12*(1), 63. https://doi.org/10.3390/antibiotics12010063

Wu, G. D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.-Y., Keilbaugh, S. A., Bewtra, M., Knights, D., Walters, W. A., Knight, R., Sinha, R., Gilroy, E., Gupta, K., Baldassano, R., Nessel, L., Li, H., Bushman, F. D., & Lewis, J. D. (2011). Linking Long-Term Dietary Patterns with Gut Microbial Enterotypes. *Science (New York, N.y.)*, *334*(6052), 105–108. https://doi.org/10.1126/science.1208344

Xie, Y., Chen, J., He, J., Miao, X., Xu, M., Wu, X., Xu, B., Yu, L., & Zhang, W. (2014). Antimicrobial Resistance and Prevalence of Resistance Genes of Obligate Anaerobes Isolated From Periodontal Abscesses. *Journal of Periodontology*, *85*(2), 327–334. https://doi.org/10.1902/jop.2013.130081

Yamanaka, T., Yamane, K., Furukawa, T., Matsumoto-Mashimo, C., Sugimori, C., Nambu, T., Obata, N., Walker, C. B., Leung, K.-P., & Fukushima, H. (2011). Comparison of the virulence of exopolysaccharide-producing *Prevotella intermedia* to exopolysaccharide non-producing periodontopathic organisms. *BMC Infectious Diseases*, *11*, 228. https://doi.org/10.1186/1471-2334-11-228

Yamasaki, K., Kawanami, T., Yatera, K., Fukuda, K., Noguchi, S., Nagata, S., Nishida, C., Kido, T., Ishimoto, H., Taniguchi, H., & Mukae, H. (2013). Significance of Anaerobes and Oral Bacteria in Community-Acquired Pneumonia. *PLoS ONE*, *8*(5), e63103. https://doi.org/10.1371/journal.pone.0063103

Yan, C., Hong, F., Xin, G., Duan, S., Deng, X., & Xu, Y. (2022). Alterations in the vaginal microbiota of patients with preterm premature rupture of membranes. *Frontiers in Cellular and Infection Microbiology*, *12*. https://www.frontiersin.org/articles/10.3389/fcimb.2022.858732

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R. N., Anokhin, A. P., Heath, A. C., Warner, B., Reeder, J., Kuczynski, J., Caporaso, J. G., Lozupone, C. A., Lauber, C., Clemente, J. C., Knights, D., … Gordon, J. I. (2012). Human gut microbiome viewed across age and geography. *Nature*, *486*(7402), 222–227. https://doi.org/10.1038/nature11053

Yeoman, C. J., Yildirim, S., Thomas, S. M., Durkin, A. S., Torralba, M., Sutton, G., Buhay, C. J., Ding, Y., Dugan-Rocha, S. P., Muzny, D. M., Qin, X., Gibbs, R. A., Leigh, S. R., Stumpf, R., White, B. A., Highlander, S. K., Nelson, K. E., & Wilson, B. A. (2010). Comparative Genomics of *Gardnerella vaginalis* Strains Reveals Substantial Differences in Metabolic and Virulence Potential. *PLoS ONE*, *5*(8), e12411. https://doi.org/10.1371/journal.pone.0012411

Yin, W., Wang, Y., Liu, L., & He, J. (2019). Biofilms: The Microbial “Protective Clothing” in Extreme Environments. *International Journal of Molecular Sciences*, *20*(14), 3423. https://doi.org/10.3390/ijms20143423

Yokoyama, S., Hayashi, M., Goto, T., Muto, Y., & Tanaka, K. (2023). Identification of cfxA gene variants and susceptibility patterns in β-lactamase-producing *Prevotella* strains. *Anaerobe*, *79*, 102688. https://doi.org/10.1016/j.anaerobe.2022.102688

Zabor, E. C., Klebanoff, M., Yu, K., Zhang, J., Nansel, T., Andrews, W., Schwebke, J., & Jeffcoat, M. (2010). Association between periodontal disease, bacterial vaginosis, and sexual risk behaviors. *Journal of Clinical Periodontology*, *37*(10), 888–893. https://doi.org/10.1111/j.1600-051X.2010.01593.x

Zanoni, I., Bodio, C., Broggi, A., Ostuni, R., Caccia, M., Collini, M., Venkatesh, A., Spreafico, R., Capuano, G., & Granucci, F. (2012). Similarities and differences

of innate immune responses elicited by smooth and rough LPS. *Immunology Letters*, *142*(1), 41–47. https://doi.org/10.1016/j.imlet.2011.12.002

Zeigler, D. R. (2003). Gene sequences useful for predicting relatedness of whole genomes in bacteria. *International Journal of Systematic and Evolutionary Microbiology*, *53*(Pt 6), 1893–1900. https://doi.org/10.1099/ijs.0.02713-0

Zhang, Y., Zhen, M., Zhan, Y., Song, Y., Zhang, Q., & Wang, J. (2017). Population-Genomic Insights into Variation in *Prevotella intermedia* and *Prevotella nigrescens* Isolates and Its Association with Periodontal Disease. *Frontiers in Cellular and Infection Microbiology*, *7*. https://www.frontiersin.org/articles/10.3389/fcimb.2017.00409

Zhou, X., Brown, C. J., Abdo, Z., Davis, C. C., Hansmann, M. A., Joyce, P., Foster, J. A., & Forney, L. J. (2007). Differences in the composition of vaginal microbial communities found in healthy Caucasian and black women. *The ISME Journal*, *1*(2), Article 2. https://doi.org/10.1038/ismej.2007.12

Zhou, X., Hansmann, M. A., Davis, C. C., Suzuki, H., Brown, C. J., Schütte, U., Pierson, J. D., & Forney, L. J. (2010). The vaginal bacterial communities of Japanese women resemble those of women in other racial groups. *FEMS Immunology and Medical Microbiology*, *58*(2), 169–181. https://doi.org/10.1111/j.1574-695X.2009.00618.x

Ziklo, N., Huston, W. M., Hocking, J. S., & Timms, P. (2016). *Chlamydia trachomatis* Genital Tract Infections: When Host Immune Response and the Microbiome Collide. *Trends in Microbiology*, *24*(9), 750–765. https://doi.org/10.1016/j.tim.2016.05.007

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. *Journal of Clinical Microbiology*, *48*(5), 1812–1819. https://doi.org/10.1128/JCM.00851-09

Zuppi, M., Hendrickson, H. L., O'Sullivan, J. M., & Vatanen, T. (2022). Phages in the Gut Ecosystem. *Frontiers in Cellular and Infection Microbiology*, *11*. https://www.frontiersin.org/articles/10.3389/fcimb.2021.822562

Supplementary data

Table S 1: **Minimum inhibitory concentrations (MIC) of all strains to common genitourinary antibiotics, as determined by Etest methodology**

	MIC (µg/ml)[1]					
Strain[6]	**Metronidazole**	**Clindamycin**	**Azithromycin[4]**	**Amoxicillin[2]**	**Doxycycline[3]**	**Antibiotic treatment in individual?[5]**
UC016_V1_1	0.75	0.064	>256	>256	24	None
UC024_V2_1	0.75	0.032	8	0.023	8	None
UC033_V1_2	1.5	0.064	>256	96	12	None
UC037_V1_1	0.75	0.094	>256	1.5	12	None
UC040_V1_1	0.5	0.016	1.5	>256	16	None
UC040_V2_1	1	0.032	4	3	4	None
UC055_V2_3	0.25	0.16	1.5	6	6	None
UC055_V3_1	0.38	0.23	0.5	1.5	8	None
UC059_V1_1 ●	0.75	0.064	8	64	24	None
UC059_V1_2 ●	1	0.047	6	>256	24	None
UC064_V1_1	4	0.047	6	>256	0.19	None
UC065_V2_1	1.5	0.047	6	>256	16	None
UC071_V1_1 ●	1	0.032	12	>256	16	None
UC071_V1_1F ●	2	0.047	6	>256	24	None
UC074_V1_1	12	0.047	6	0.032	24	None
UC079_V2_1	1	0.032	3	64	6	None
UC095_V2_2	2	0.032	12	>256	24	None
UC096_V1_3	0.5	0.016	2	64	16	None
UC104_V2_1	1	0.032	6	2	4	None
UC107_V1_2	4	0.064	>256	64	48	None
UC107_V2_1	3	0.064	24	24	24	AZM
UC109_V3_2	3	>256	>256	3	16	AZM, CTX, MTZ
UC121_V1_3	12	>256	>256	24	32	None
UC126_V1_1 ●	1	0.094	>256	0.032	48	None
UC126_V1_2 ●	0.75	0.125	12	0.047	16	None
UC126_V1_3 ●	1.5	0.094	96	0.032	24	None
UC126_V1_3F ●	0.75	0.047	8	0.032	24	None
UC132_V1_1	0.75	0.016	0.75	>256	12	None
UC136_V1_1	0.75	0.032	3	0.25	4	None
UC143_V2_1 ●	1	0.032	>256	6	16	AZM, CTX
UC143_V2_2 ●	0.19	0.016	1.5	0.023	0.032	AZM, CTX
UC146_V1_5	0.25	0.032	3	>256	16	None
UC149_V3_1	>256	>256	>256	48	0.19	AZM
UC158_V3_1	1	0.032	3	>256	16	AZM, CTX, MTZ
UC162_V1_1	24	0.032	2	0.047	0.064	None
UC167_V3_1	6	0.032	3	>256	8	None
***B. fragilis* ATCC25285^T**	0.25	0.5	>256	1	0.032	N/A
***P. bivia* ATCC29303^T**	1.5	0.047	4	0.016	16	N/A

[1] Shading represents resistant ■, intermediate ■ and susceptible □
[2] Using CLSI breakpoint for amoxicillin/ clavulanic acid (CLSI, 2020)
[3] Using CLSI breakpoint for tetracycline (CLSI, 2020)
[4] No CLSI breakpoint. R ≥ 8µg/ml, S ≤ 2 µg/ml
[5] Antibiotic names abbreviated: MTZ (metronidazole), CTX (ceftriaxone) and AZM (azithromycin)
[6] Dots of the same colour represent those strains isolated from the same sample

Table S 2: **Full assembly metrics for all sequenced isolates and for reference *P. bivia and P. melaninogenica* sequences obtained from NCBI RefSeq database.**

Isolate	Genome size (bp)	N50	L50	%GC	No. contigs	No. CDS	No. tRNAs	Completeness	Contamination	%Coding density
Prevotella bivia										
UC008_V1_1	2457529	71083	13	39.41	94	2101	51	88.23	2.6	86.75
UC016_V1_1	2541583	57817	13	39.65	146	2191	53	92.64	4.95	86.47
UC024_V2_1	2489034	68716	11	39.32	104	2129	49	89.19	3.46	86.7
UC033_V1_2	2408000	113983	8	39.7	67	2042	50	89.45	3.12	86.89
UC037_V1_1	2506965	102549	7	39.51	63	2169	51	89.28	3.47	86.92
UC040_V1_3	2545855	77619	11	39.83	115	2148	50	88.77	3.79	86.5
UC040_V2_1	2361601	22779	31	39.45	464	1987	53	87.11	14.97	86.09
UC055_V2_3	2573297	55578	15	39.81	125	2229	48	91.05	5.8	86.72
UC055_V3_1	2552518	51990	16	39.86	115	2204	48	90.05	5.15	86.79
UC059_V1_1	2427339	117289	7	39.62	104	2049	50	88.9	3.14	86.61
UC059_V1_2	2419735	117285	7	39.63	95	2037	50	88.9	3.29	86.6
UC064_V1_1	2403854	64230	12	39.69	72	2038	49	88.75	2.07	86.56
UC065_V2_1	2414714	113480	8	39.83	70	2031	48	88.1	2.61	86.98
UC069_V1_1	2566899	69952	11	39.79	128	2209	50	88.78	4.62	86.98
UC071_V1_1	2543881	70372	11	39.75	90	2175	48	90.88	4.14	86.59
UC071_V1_1F	2593423	55582	14	39.48	130	2229	50	89.27	4.38	86.52
UC074_V1_1	2417062	50236	13	39.38	134	2052	49	89.54	2.42	86.19
UC079_V2_1	2422421	107802	7	39.76	68	2046	49	88.67	2.58	86.89
UC095_V2_2	2636482	68836	14	39.98	94	2257	49	88.31	4.89	86.73
UC096_V1_3	2679607	50119	18	39.8	161	2317	49	93.59	5.39	86.37
UC104_V2_1	2895896	1664	416	39.38	2131	NA	NA	89.38	57.61	85.26
UC107_V1_2	2567686	75731	11	39.66	97	2194	50	89.77	5.12	86.92
UC107_V2_1	2490372	89031	11	39.69	112	2129	50	88.77	4.21	86.91
UC109_V3_2	2543011	61695	11	39.96	108	2173	48	87.72	3.87	86.82
UC121_V1_3	2467808	82360	10	39.74	71	2077	48	89.25	5.82	86.71
UC126_V1_1	2497615	79389	11	39.7	128	2142	51	88.77	4.23	86.93
UC126_V1_2	2491510	79389	11	39.7	121	2137	50	88.77	4.15	86.86
UC126_V1_3	2698246	73513	13	39.78	331	2293	50	89.47	6.03	86.67
UC126_V1_3F	2491439	68755	13	39.69	120	2140	50	88.77	4.19	86.79
UC132_V1_1	2488560	103024	8	39.67	85	2122	50	91.98	3.93	86.55
UC136_V1_1	2376711	65684	12	39.2	120	2052	48	90.08	2.74	86.25
UC143_V2_1	2378842	72329	10	39.49	78	2012	49	87.56	2.31	87.04
UC146_V1_5	2580740	68280	11	39.82	116	2240	48	89.24	3.79	87.18
UC149_V3_1	2579807	67359	11	39.81	119	2236	48	89.17	3.69	87.22
UC158_V3_1	2891740	1381	497	39.38	2355	NA	NA	87.85	70.78	85.3
UC162_V1_1	2612014	54205	12	39.35	109	2333	50	90.54	4.66	86.65

UC167_V3_1	2458755	105690	7	39.91	71	2088	48	89.52	3.96	87.03
GCF_000177315.1	2424432	48627	16	39.77	121	2060	45	97.28	1.14	87.42
GCF_000262545.1	2521238	1982184	1	39.86	3	2138	50	97.28	1.14	86.48
GCF_000759045.1	2527024	50986	17	39.85	158	2144	39	88.53	5.55	87.69
GCF_000759165.1	2408181	34813	20	39.74	173	2061	41	87.67	3.75	87.65
GCF_000759245.1	2584224	38665	19	40	179	2204	40	89.82	5.88	87.72
GCF_001065995.1	2361895	17441	42	39.5	272	1993	36	90.49	4.62	87.4
GCF_001546565.2	2459295	48915	15	39.62	127	2129	44	90.48	4.68	87.21
GCF_001574405.1	2579172	68280	14	39.91	114	2277	45	91.09	5.33	86.81
P. melaninogenica										
UC143_V2_2	2904179	161629	6	40.67	51	2307	48	88.61	2.45	82.25
GCF_000144405.1	3168282	1796408	1	40.98	2	2513	52	98.52	0.48	82.44
GCF_000163035.1	3292341	217145	5	40.86	46	2599	47	97.06	1.12	79.78
GCF_000759305.1	2976254	8133	119	41.36	553	2338	31	84.79	8.32	83.51
GCF_002208725.2	3168230	1796407	1	40.98	2	2459	52	98.52	0.48	82.43
GCF_003609775.1	3090326	1737294	1	40.77	3	2407	52	89.31	2.31	82.27
GCF_013267595.1	3352987	1954177	1	40.87%	2	2700	51	96.68	1.12	82.26
GCF_018127885.1	3168321	1791969	1	40.92	2	2566	51	84.07	14.87	81.94
GCF_018127905.1	3169665	1793327	1	40.92	2	2609	51	83.75	14.87	81.86
GCF_018127925.1	3360806	1913733	1	41.17	2	2716	50	91.97	5.06	81.27
GCF_018127945.1	3334355	1967710	1	40.84	3	2748	53	90.37	3.68	82.21
GCF_018127965.1	3218481	1863200	1	40.86	2	2612	50	90.02	3.92	81.95
GCF_018128005.1	3619436	2182615	1	41.04	3	2983	51	92.76	5.86	82.51
GCF_018128045.1	3358299	1964469	1	41.1	2	2794	50	84.98	16.27	83.63
GCF_018128065.1	3342898	1860477	1	41.31	2	2833	50	85.53	27.39	83.83
GCF_019375485.1	3473651	154741	8	40.78	52	2855	49	92.66	4.16	82.89
GCF_019375525.1	3293085	163933	8	40.79	43	2688	53	92.79	3.83	82.68
GCF_019375575.1	3112350	191760	5	40.98	42	2448	45	91.7	1.47	82.58
GCF_019375605.1	3265850	164382	8	40.82	44	2660	47	92.79	3.83	82.7
GCF_019375715.1	3421699	88077	13	40.8	86	2837	44	85.51	13.38	82.09
GCF_019375745.1	3307101	87241	12	40.85	74	2681	49	84.75	13.07	81.81
GCF_019375775.1	3348634	75095	13	40.83	97	2725	46	84.72	13.53	81.71
GCF_019375815.1	3422092	100394	8	40.78	63	2814	48	91.36	4.59	83.04
GCF_019375935.1	3349000	100681	10	40.65	69	2663	50	94.11	5.52	82.44
GCF_019391735.1	3285599	92174	11	40.65	62	2621	50	93.16	4.05	81.93
GCF_019391775.1	3309181	164678	7	40.75	36	2709	47	84.81	20.73	82.56
GCF_019391815.1	3280014	165184	8	40.82	40	2690	49	92.56	4.49	82.76
GCF_019391945.1	3256297	21761	41	41.02	242	2548	63	91.63	5.99	82.1
GCF_020735905.1	3306754	1864514	1	40.71	2	2658	52	92.86	6.71	82.09
GCF_901875335.1	3192247	210629	6	40.52	57	2541	47	91.13	3.62	82.89
Prevotella scopos										
GCF_000614285.1	3184425	48387	19	40.70	113	2855	46	97.54	0.18	83.34
GCF_001683355.1	3303745	1859892	1	40.74	3	2685	50	99.29	0.18	82.74

Figure S 1: **Core genome alignment of *P. melaninogenica* sequences.** *P. melaninogenica* whole genome sequences were downloaded from the NCBI RefSeq database and aligned with UC143_V2_2. Bootstrap values over 60 are displayed on the branches. The tree scale represents a phylogenetic distance of 0.1 nucleotide substitutions per site.

Table S 3: **Jaccard similarity results for all predicted complete prophage genomes**

		Closest match in the NCBI RefSeq *Caudoviricetes* database			
Bacterial host	**Jaccard Similarity**	**Genome accession**	**Taxonomic name**	**Genus**	**Morphotype***
UC016_V1_1	0.0118352	GCF_028515065.1	*Pseudomonas* phage UF_RH1	*Septimatrevirus*	*Siphoviridae*
UC024_V2_1	0.0163379	GCF_002593485.1	*Lactococcus* phage 936 group phage Phi4.2	*Skunavirus*	*Siphoviridae*
UC040_V1_3	0.018491	GCF_002619845.2	*Geobacillus* phage TP-84	*Saundersvirus*	*Siphoviridae*
UC059_V1_1	0.0191408	GCF_000875405.1	*Mycobacterium* phage DD5	*Fromanvirus*	*Siphoviridae*
UC059_V1_2	0.0191408	GCF_000875405.1	*Mycobacterium* phage DD5	*Fromanvirus*	*Siphoviridae*
UC065_V2_1	0.0167778	GCF_001470175.1	*Mycobacterium* phage Badfish	*Pegunavirus*	*Siphoviridae*
UC069_V1_1	0.014402	GCF_001470175.1	*Mycobacterium* phage Badfish	*Pegunavirus*	*Siphoviridae*
UC079_V2_1	0.015506	GCF_003423085.1	*Leptospira* phage LE4	*Nylescharonvirus*	*Myoviridae*
UC107_V1_2	0.0139024	GCF_000860225.1	*Staphylococcus* phage 47	*Triavirus*	*Siphoviridae*
UC109_V3_2	0.0214559	GCF_003441895.1	*Mycobacterium* phage Aggie	*Charlievirus*	*Siphoviridae*
UC143_V2_1	0.0168051	GCF_000954995.1	*Pseudomonas* phage vB_PaeP_C2-10_Ab22	*Bruynoghevirus*	*Podoviridae*
UC162_V1_1	0.0138516	GCF_000859445.1	*Staphylococcus* phage X2	*Phietavirus*	*Siphoviridae*
UC167_V3_1	0.0208078	GCF_002621805.1	*Streptococcus* phage P7632	*Moineauvirus*	*Siphoviridae*

*Morphotype based on previous phage taxonomy – *Siphoviridae*-type: long, non-contractile tails, *Myoviridae*-type: short, contractile tails, *Podoviridae*-type: very short, non-contractile tails.

Table S 4: **Number of each identified anti-phage mechanism per isolate genome**

Isolate	Paris	RM	Cas_CI	AbiD	AbiE	AbiH	Avs	DarTG	Gabija	GAPS4	Hachim	Lamass	NLR	PD-T7-	PrrC	Retron	RloC	SanaTA	SoFIC	SspBC	Theoris
UC008_V1_1	1	1			1	1											1			1	
UC016_V1_1	1	7																1			
UC024_V2_1	1	6			1																
UC033_V1_2	1	2			1						1						1				
UC037_V1_1	1	2			1						1						1				
UC040_V1_3	1	2		1	1	1	1														
UC040_V2_1	1																1			1	
UC055_V2_3	1	8		2														1			
UC055_V3_1	1	8		2														1			
UC059_V1_1	1	8			1																
UC059_V1_2	1	8			1																
UC064_V1_1	1	8		1	1				1	1								1			
UC065_V2_1	1	4															1				
UC069_V1_1	1	6		1														1			
UC071_V1_1	1	3			1	1					1										
UC071_V1_1F	1	5			1																
UC074_V1_1	1	2		1								1									
UC079_V2_1	1	4															1				
UC095_V2_2	1	2			1	1					1										
UC096_V1_3	1	3		1	1										1			1			
UC107_V1_2	1	9		2	1																
UC107_V2_1	1	8		1														1			
UC109_V3_2	2	4			2								1					1	1		
UC121_V1_3	1	2		1																1	
UC126_V1_1	1	8		1														1			
UC126_V1_2	1	8		1														1			
UC126_V1_3	1	8		1														1			
UC126_V1_3F	1	8		1														1			
UC132_V1_1	1	4			1					1	1			1			1	1			
UC136_V1_1	1	3												1			1				
UC143_V2_1	1	1			1	1											1			1	
UC143_V2_2	1	5		2					1							1		1			
UC146_V1_5	1	9		1	2			1						1				1			
UC149_V3_1	1	9			2			1						1				1			
UC162_V1_1	1	4		1	1	1															1
UC167_V3_1	1	5		1													1	1			

Appendix

Media composition

Brain heart infusion (BHI) (1L)

37g	BHI powder
5g	Yeast extract
0.5g	Cysteine
5mL	Haemin
200µl	Vitamin K_3
995ml	Water

Laked Brucella blood agar with vancomycin and kanamycin (LKV)

28.1g	Brucella broth base
5ml	Haemin
2ml	Vitamin K_3
2ml	Cysteine
15g	Agar
100mg	Kanamycin
7.5mg	Vancomycin
50mg	Laked sheep blood
1L	Water

Antibiotic assay medium (AAM)

5g	Tryptone
5g	Meat extract
7g	Yeast extract
1ml	Tween 80
2.5g	Sodium acetate
200mg	$MgSO_4 \times 7\ H_2O$
50mg	$MnSO_4 \times H_2O$

7g	Casamino acids
1L	Water

NYCIII medium (1L)

2.4g	HEPES
7.5g	Protease peptone
3.8g	Yeast extract
900ml	Water
5g	NaCl
5g	Glucose
100ml	10% (v/v) inactivated horse serum
12g	Agar (for solid media)

Ethics documentation

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

Room 45 E-52-E-Floor- Old Main Building
Groote Schuur Hospital
Observatory 7925
Telephone [021] 406 6492
Email: hrec-submissions@uct.ac.za
Website: www.health.uct.ac.za/fhs/research/humanethics/forms

26 August 2022

HREC REF:475/2022

A/Prof J Passmore
Division of Medical Virology
S3.33 Level 3, Entrance 5 -FHS
Email: jo-ann.passmore@uct.ac.za
Student: wipkir001@myuct.ac.za

Dear A/Prof Passmore

PROJECT TITLE : CHARACTERISATION OF VAGINAL PREVOTELLA STRAINS FROM A COHORT OF ADOLESCENT SOUTH AFRICAN INDIVIDUALS WITH AND WITHOUT BACTERIAL VAGINOSIS (SUB-STUDY - 801/2014)
(MSC CANDIDATE - MISS KIRSTEN WELP)

Thank you for your response letter, addressing the issues raised by the Faculty of Health Sciences Human Research Ethics Committee (HREC).

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study.

Approval is granted for one year until the 30 August 2023.

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.
(Forms can be found on our website: www.health.uct.ac.za/fhs/research/humanethics/forms)

The HREC acknowledge that the student: Miss Kirsten Welp will also be involved in this study.

Please quote the HREC REF 475/2022 in all your correspondence.

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please note that for all studies approved by the HREC, the principal investigator **must** obtain appropriate institutional approval, where necessary, before the research may occur.

Yours sincerely

PROFESSOR M BLOCKMAN
CHAIRPERSON, FACULTY OF HEALTH SCIENCES HUMAN RESEARCH ETHICS COMMITTEE

HREC.REF475.2022

FHS016: Annual Progress Report / Renewal

HREC office use only (FWA00001637; IRB00001938) This serves as notification of annual approval, including any documentation described below.			
☑ Approved	Annual progress report	Approved until/next renewal date	30.8.2024
☐ Not approved	See attached comments		
Signature Chairperson of the HREC/ Designee	Signed by candidate	Date Signed	7/1/23

Note: Please email this form and supporting documents (if applicable) in a combined pdf-file to hrec-enquiries@uct.ac.za.
Please clarify your plan for research-related activities during COVID-19 lockdown.
Please use the latest form found on our website:
http://www.health.uct.ac.za/fhs/research/humanethics/forms

Comments to PI from the HREC

Principal Investigator to complete the following:

1. Protocol information

Date (when submitting this form)			
HREC REF Number	475/2022	Current Ethics Approval was granted until	30 August 2023
Protocol title	Characterisation of vaginal *Prevotella* strains from a cohort of adolescent South African individuals with and without bacterial vaginosis		
Protocol number (if applicable)			
Are there any sub-studies linked to this study?	☐ Yes	☒ No	
If yes, could you please provide the HREC Reference number for all sub-studies? **Note:** A separate FHS016 must be submitted for each sub-study.			
Principal Investigator	Professor Jo-Ann Passmore		

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)