Department of
Metabolism, Digestion and Reproduction

Imperial College London

Metabolic profiling of the lower reproductive tract for preterm birth prediction and stratification

A thesis submitted by

Katia Capuccini
CID 01445001

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy to Imperial College London

Abstract

Preterm birth (PTB) and its associated complications remain the leading cause of death in children under 5 years of age, worldwide. Despite extensive research, our understanding of the underlying causes of PTB and our ability to predict women at risk remain limited. However, 40% of all PTB cases are linked with an infectious aetiology. Recent studies have established an association between vaginal bacterial composition and risk of PTB. Depletion of *Lactobacillus* species and increased bacterial diversity associate with increased PTB risk. Current techniques used to study vaginal microbiota are costly, time-consuming, and do not provide information on host:microbiota interactions, which are crucial for understanding PTB phenotypes. The biochemical composition of cervicovaginal fluid (CVF) has been shown to reflect the microbial composition of the vagina. The work presented in this thesis set out to test the hypothesis that CVF metabolic profiles rapidly captured by a novel platform, direct on swab analysis by Desorption Electrospray Ionisation-Mass Spectrometry (DESI-MS), can be used for PTB risk stratification and outcome prediction. To test this, women at high risk of PTB were recruited and vaginal swabs analysed using DESI-MS. Integration of the resulting metabolic profiles with metataxonomics and immunological data enabled the identification of metabolic signatures capable of simultaneously predicting bacterial composition and host inflammatory status. Results also indicated that the CVF metabolome as assessed by DESI-MS is partly influenced by prior pregnancy history and can moderately predict cervical shortening. Analysis of samples with cervical sutures *in situ* showed that DESI-MS -based prediction of bacterial composition is not affected by the procedure. A sub analysis of vaginal swabs highlighted a limitation of DESI-MS to detect progesterone and other steroid hormones in CVF. However, robust detection and quantitation of progesterone, androstenedione, and cortisol could be achieved in a pilot study of CVF samples collected using menstrual cups analysed via an adopted Ultra-performance liquid chromatography (UPLC-MS) assay. No differences in the levels of these compounds were observed between women subsequently delivering at term or preterm. Finally DESI-MS swab profiling performed on a small format, portable mass spectrometer was shown to provide strong capacity to monitor vaginal microbial colonisation in women receiving a live biotherapeutic intervention during pregnancy. These results represent an important step toward clinical translation of DESI-MS and highlight its potential as a novel tool for preterm birth risk phenotyping, patient stratification and treatment response monitoring.

Statement of Originality and Contributions

All the work herein described has been performed by the Author, unless otherwise stated.

DESI-MS profiling for the VMET validation data set used in chapter III and IV was generated by Drs Pamela Pruski and Holly Lewis. DNA extraction and 16S rRNA gene sequencing for all cervicovaginal swab samples in VMET2 cohort were conducted by Dr Denise Chan. The UPLC-MS/MS steroids assay described in chapter V was adapted for cervicovaginal fluid by Dr Lucia Olmo Garcia and performed by the National Phenome Centre at Imperial College London. DNA extraction, 16S rRNA gene sequencing and immune profiling for FLIP cervicovaginal swab samples used in chapter VI were performed by Dr Erna Bayar. Some of the figures in this thesis have been adapted from publications of which I was a main author or co-author.

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Publications and conference presentations

Articles published during these studies

Pamela Pruski, Gonçalo D. S. Correia, Holly V. Lewis., **Katia Capuccini**, et al. "Direct on-swab metabolic profiling of vaginal microbiome host interactions during pregnancy and preterm birth." Nat Commun 12, 5967 (2021).

Katia Capuccini, Mark Sullivan, and David A. MacIntyre. "The reproductive tract microbiome and health", The International Microbiology Literacy Initiative (2023), In press

Conference presentations

Katia Capuccini, Pamela Pruski, Gonçalo D. S. Correia, Holly V. Lewis, Denise Chan, Richard G. Brown, Yun S. Lee, Anna L. David, Sarah J. Stock, Jane E. Norman, Vasso Terzidou, T. G. Teoh, Lynne Sykes, Phillip R. Bennett, Zoltan Takats and David A. MacIntyre. "Rapid prediction of vaginal microbiota composition in pregnancy by direct on-swab metabolic profiling"
Oral presentation at the UK Preterm Birth annual conference, 15th - 16th November 2021

Katia Capuccini, Gonçalo D. S. Correia, Pamela Pruski, Holly Lewis, Lynne Sykes, Phillip R. Bennett, Zoltán Takáts, and David A. MacIntyre. "Direct Swab Analysis By Desorption Electrospray Ionisation Mass Spectrometry For Preterm Birth Risk Stratification In Patients With Cervical Shortening"
Podium presentation at BMSS Ambient Ionisation SIG Meeting 2023, 1st February, 2023 (AstraZeneca, Macclesfield, UK)

Katia Capuccini, Gonçalo D. S. Correia, Pamela Pruski, Holly Lewis, Lynne Sykes, Phillip R. Bennett, Zoltán Takáts, and David A. MacIntyre. "Direct Swab Analysis By DESI-MS For Preterm Birth Risk Stratification In Patients With Cervical Shortening"
Podium presentation in the scientific session of ''Emerging Technologies for Point of Care Diagnostics'' at MSACL 2023 13th Annual Conference & Exhibits, Monterey, CA, April 2 - 7, 2023*. Recipient of the Young Investigator Educational Grant*

Katia Capuccini, Gonçalo D. S. Correia, Pamela Pruski, Holly Lewis, Lynne Sykes, Phillip R. Bennett, Zoltán Takáts, and David A. MacIntyre. “Direct swab analysis by desorption electrospray ionisation mass spectrometry (DESIMS) for preterm birth risk stratification in patients with cervical shortening” Poster presentation in the session "Microbes and their metabolites underpinning host-microbe interactions" Microbiology Society Annual Conference 2023, Birmingham April 17-20, 2023. *Recipient of the Society conference travel grant*

Fabienne D. Hanton, **Katia Capuccini**, Aanya Hirdaramani, Aylin Hanyaloglou, Sherrianne Ng, Yun S Lee, Phillip R Bennett, and David A. MacIntyre “Mechanistic Insights into the Role of Lactate Acidification on Microbial-Induced Inflammation in Vaginal Epithelial Cells”. Poster presentation at the 70th Annual Scientific meeting of Society for Reproductive Investigation, Brisbane, Australia March 21-25, 2023

Acknowledgements

I would like to express my deepest gratitude to Professor David MacIntyre for his precious supervision that made this work possible. Your contagious enthusiasm and passion for science was a constant source of encouragement and inspiration. Was a great opportunity to learn from you. Thank you for always showing me the positive side with a kind sense of humour.

I would also like to thank Professor Zoltan Takats for his technical expertise and guidance in the Mass Spectrometry field, and Professor Phillip Bennett for his crucial clinical perspective and experienced insights on this work.

I would like to extend my gratitude to Drs Pamela Pruski and Eftychios Manoli for sharing with me their deep knowledge of mass spectrometry but also their friendship and many fun memories in the long hours of data acquisition. Also, many thanks to Dr Gonçalo Correia for his precious help in data modelling and patience in explaining intricate bioinformatics concepts to me.

My gratitude goes also to Dr Paschalia Pantazi, Maria Arianoglou, Pav Karia, Dr Sherrianne Ng and all the fantastic members of our team at IRDB for creating a positive, stimulating, and collaborative environment where was really a privilege to be. Particularly, thanks to DrTanweer Beleil for her caring words in difficult moments.

A special heartfelt thank you to Dr Belen Gimeno-Molina who walked close to me this path from the start, becoming on the way a cherished friend always with a coffee and a chocolate at hand. My gratitude extends to the many friends, both close and far away, that showed me love and support in countless ways. A special place in my heart for Tishya, Ilaria and Elena.

I express my profound gratitude to my mother, whose sacrifices, love, and unwavering support have made me feel her presence even when distant.

Lastly, but certainly not least, thank you Alex. No words can adequately convey the gratitude I have for you. Your support and love have been instrumental in my journey, and I simply could not have accomplished this without you. Thank you for pushing me forward and believing in me, even during times when I doubted myself.

Table of Contents

Abstract 2

Statement of Originality and Contributions 3

Copyright Declaration 4

Publications and conference presentations 5

Articles published during these studies 5

Conference presentations 5

Acknowledgements 7

Table of Contents 8

List of Figures 12

List of Tables 15

Abbreviations 16

Chapter I 19

1 INTRODUCTION 19

1.1 Preterm birth 19

1.2 Regulation of parturition in term and preterm pregnancies 21

1.2.1 Term parturition 21

1.2.2 Preterm parturition 28

1.3 PTB Prediction strategies 30

1.4 PTB Prevention and treatment strategies 32

1.4.1 Progesterone 32

1.4.2 Cerclage 33

1.4.3 Antibiotic therapy 34

1.5 Host:microbiota interactions 36

1.5.1 Vaginal microbiota in health and disease 37

1.5.2 Vaginal microbiota in pregnancy and PTB 41

1.6 Therapeutic strategies for modification of microbiota communities 44

1.7 Identification of vaginal microbiota 47

1.7.1 Conventional methods 47

1.7.2 Molecular-based methods for bacterial identification and characterisation of community composition....48
1.8 Constituents of the cervicovaginal mucosa....50
1.8.1 Carbohydrates....51
1.8.2 Short Chain Fatty Acids (SCFAs)....52
1.8.3 Lipids....53
1.8.4 Amino acids and Biogenic Amines....53
1.9 Assessment of microbiota-host interactions using metabolic profiling....54
1.9.1 Mass Spectrometry (MS)....55
1.9.2 MS ion sources....56
1.10 Data modelling approaches for metabolic profiling data....64
1.10.1 Multivariate modelling approaches....64
1.11 Study rationale....66
1.12 Hypothesis and Aims....67
Chapter II....68
2 MATERIALS AND METHODS....68
2.1 Studies subjects and samples collection....68
2.1.1 Patient cohorts and sampling....68
2.2 Characterisation of vaginal microbiota community composition....71
2.2.1 16S rRNA gene sequencing-based metataxonomic profiling....71
2.3 Cytokine/chemokine profiling in cervicovaginal swabs samples....73
2.4 DESI-MS based metabolite profiling of clinical swab samples....74
2.4.1 Reagents and materials....74
2.4.2 Direct swab analysis by DESI-MS using an Orbitrap Mass Spectrometer....74
2.4.3 DESI-MS using a small-format RDa/BioAccord TOF Mass Spectrometer....78
2.4.4 Liquid extraction methods of cervicovaginal soft cup samples....82
2.4.5 LC-MS steroid assay....84
2.5 Data processing and Statistics....87
2.5.1 Statistical analysis chapter III....87
2.5.2 Statistical analysis chapter IV....90
2.5.3 Statistical analysis chapter V....92
2.5.4 Statistical analysis chapter VI....92
Chapter III....96

3 RAPID MICROBIOTA AND INFLAMMATION PREDICTION USING DIRECT SWAB ANALYSIS BY DESI-MS 96
3.1 Introduction 96
3.2 Hypothesis and aims 98
3.3 Results 98
3.3.1 Sample cohort 98
3.3.2 Analysis of swab shape, geometry, and physical characteristics 102
3.3.3 Data acquisition and preliminary confounders analyses 102
3.3.4 Characterisation of vaginal microbiota composition using metataxonomics 104
3.3.5 Integration of DESI-MS profiles and vaginal microbiota compositional data 106
3.3.6 DESI-MS prediction of host immune response 109
3.3.7 Levels of key metabolites in relation to inflammation markers 112
3.3.8 Assessment of cervical vaginal inflammation and microbiota composition in term and preterm birth outcomes 116
3.3.9 Assessment of cervical vaginal inflammation in response to cervical cerclage 117
3.4 Discussion 118
Chapter IV 123
4 THE CERVICOVAGINAL METABOLOME AND PRETERM BIRTH RISK STRATIFICATION 123
4.1 Introduction 123
4.2 Hypothesis and aims 125
4.3 Results 126
4.3.1 Analysis of previous pregnancy history on PTB risk in the studied patient cohorts 126
4.3.2 The CVF metabolome and previous pregnancy history 129
4.3.3 Assessment of the ability of DESI-MS derived CVF metabolic profiles to predict PTB 133
4.3.4 Prediction of cervical shortening and intervention response using DESI-MS swab profiling 134
4.3.5 Metabolic profiling to monitor cerclage efficacy 136
4.3.6 Prediction of microbial composition after cerclage intervention 137
4.4 Discussion 140
Chapter V 146

5 MASS SPECTROMETRY-BASED MEASUREMENT OF PROGESTERONE AND STEROIDS IN THE CERVICOVAGINAL NICHE DURING PREGNANCY 146

5.1 Introduction 146

5.2 Hypothesis and aims 149

5.3 Results 150

5.3.1 Progesterone measurement with DESI-MS 150

5.3.2 Measurement of progesterone and related steroids in cervicovaginal fluid by UPLC-MS/MS steroids assay 151

5.4 Discussion 157

Chapter VI 160

6 ASSESSMENT OF CVF MICROBIOTA-HOST INTERACTIONS IN PREGNANT WOMEN RECEIVING LIVE BIOTHERAPEUTIC INTERVENTION 160

6.1 Introduction 160

6.2 Hypothesis and aims 163

6.3 Results 163

6.3.1 Sample cohort 163

6.3.2 Assessment of DESI-MS metabolic profiles acquired by RDa 165

6.3.3 Microbial composition prediction using DESI-MS metabolic profiling 166

6.3.4 Prediction of LACTIN-V colonisation by DESI-MS 171

6.3.5 Inflammatory status prediction using DESI-MS metabolic profiling 172

6.4 Discussion 175

Chapter VII 180

7 DISCUSSION 180

7.1 Background 180

7.2 Summary of study findings 181

7.3 Future Work 185

7.4 Conclusion 186

References 187

Appendix 229

List of Figures

Figure 1.1 - Graphic summary of possible PTB aetiologies and of the main events triggered in preterm labour. 20
Figure 1.2 Steroid hormone biosynthesis pathway. 24
Figure 1.3 - Chart describing serum progesterone levels in menstrual cycle and in pregnancy. 29
Figure 1.4 - Photo of vaginal ultrasound for cervical measurement. 31
Figure 1.5 – Schematic representation of estrogens levels variation through different life phases and impact on microbial composition 41
Figure 1.6 – Schematic representation of vaginal mucosa in a scenario of Lactobacillus dominance vs Lactobacillus depletion 44
Figure 1.7 - Schematic representation of main soft ionisation sources 58
Figure 1.8 – Schematic representation of DESI-MS source. 62
Figure 2.1 - Schematic of study design for FLIP cohort. 70
Figure 2.2 - Schematic representation of soft cup samples processing. 71
Figure 2.3 - Schematic representation of DESI-MS source on LTQ-Orbitrap platform. 75
Figure 2.4 - Schematic representation of DESI-MS source on RDa ACQUITY platform. 79
Figure 3.1 – Schematic representation of VMET2 study design. 99
Figure 3.3 – Representative images of different swab shapes and characteristics 102
Figure 3.4 - PCA coloured according to different possible confounders 103
Figure 3.5 - PLS-DA coloured according to time points 103
Figure 3.6 Hierarchical clustering analysis of the 16S rRNA gene sequencing data from the VMET2 study cohort. 105
Figure 3.7 - Heatmap of DESI-MS derived discriminatory metabolic features and their relationship with Lactobacillus dominance or depletion. 106
Figure 3.8 - Examples of metabolic features significantly differing between Lactobacillus dominant and Lactobacillus depleted samples. 107
Figure 3.9 - ROC curves comparing DESI-MS classification performance between different vaginal CSTs. 108
Figure 3.10 - Assessment of correlation between cervicovaginal metabolic features and immune mediators. 110
Figure 3.11 - Association between predicted log-transformed value of immune markers by DESI-MS and measured log-transformed values by multiplexed immunoassay. 111
Figure 3.12 – Box and whisker plots showing levels of succinate (m/z 117.0197366), propionate (m/z 73.02988371) and lactate (m/z 89.02485758) in relation with microbial composition. 113

Figure 3.13 - Linear model showing correlation between succinate/lactate ratio and main immune markers (IL-1b, IgM, C3b, C5,) *114*
Figure 3.14 – Box and whiskers plot showing different levels of succinate/lactate ratio in high or low classes of immune markers (IL-1β, IL-8, TNF-α, IL-6) *115*
Figure 3.15 - Box and whiskers plot showing immunoassay measured or DESI-MS predicted levels of IL-1b in patient receiving cerclage with different materials. *116*
Figure 3.16 - Box and whiskers plot showing immunoassay measured or DESI-MS predicted levels of IL-1b in patient receiving cerclage with different materials. *117*
Figure 4.1 - Bar plot showing percentages of different outcomes for each risk group. *126*
Figure 4.2 – Prevalence of CST composition based on PTB birth risk category. *127*
Figure 4.3 – Bar plot showing percentages of different Ethnicity prevalence for each risk group. *128*
Figure 4.4 - Bar plot showing percentages of different CST composition for each risk group divided by ethnicity. *128*
Figure 4.5 - PCA of DESI-MS metabolic profiles of CVF collected from white pregnant women in the VMET and VMET2 cohorts n = 233. ..... *129*
Figure 4.6 - PLS-DA model including all white patients from VMET and VMET2 cohorts n = 233 and two classes. *131*
Figure 4.7 - Boxplots of the 6 shared significant features in VMET and VMET2 study. *132*
Figure 4.8 - PLS-DA model including all patients from VMET2 and VMET respectively. Trying to discriminate for outcome. *133*
Figure 4.9 - VMET samples subset selected because of their intervention and availability of longitudinal data before and after intervention *134*
Figure 4.10 - ROC curve negative and positive ion mode obtained using only significant features for prediction. *135*
Figure 4.11 – Stacked bar plot representing microbial composition in patients with long or short cervix after cerclage intervention. *136*
Figure 4.12 – PLSDA showing the difference between patients with long or short cervix after cerclage intervention. *137*
Figure 4.13 - ROC curve negative and positive ion mode predicting microbial composition at genera level for people who received a cerclage *138*
Figure 4.14 - ROC curve negative and positive ion mode showing prediction ability at different taxonomic levels *139*
Figure 5.1- Example of a positive ion mode DESI-MS spectrum where some progesterone adducts are visible. *150*
Figure 5.2- Histograms showing the distribution of the intensities of the $[M+H]^+$ adducts detected by DESI-MS. *151*
Figure 5.3 - Gestational age at time of soft cup sampling of a subset of VMET2 patients. *152*

Figure 5.4 – Schematic representation of the soft cup sample preparation procedure from collection to analysis by UPLC-MS/MS. 153
Figure 5.5 – Box plots representing values of progesterone, cortisol, androstenedione in the cervicovaginal mucosa..... 154
Figure 5.6 - Scatterplot with values of progesterone, cortisol, androstenedione in the cervicovaginal mucosa during gestation…… ……..155
Figure 5.7 - Individuals with longitudinal sampling values of progesterone, cortisol, androstenedione, and difference between timepoint A and B..... 156
Figure 5.8 - Boxplot comparing values of progesterone, cortisol, androstenedione between term and preterm cases. 156
Figure 6.1 - PCA plot of all FLIP samples coloured by date of acquisition 165
Figure 6.2 – PCA plots of all FLIP samples coloured by two main confounders. 166
Figure 6.3 – Heatmap representing relative concentrations of DESI-MS derived metabolic features and their relationship with Lactobacillus dominance or depletion. 167
Figure 6.4 – Example of two metabolic features higher in LDOM or LDEPL and their change across gestation..... 168
Figure 6.5– ROC curve showing prediction of Lactobacillus dominance or depletion status in FLIP samples and comparison with VMET2 data performance 169
Figure 6.6 - ROC curve showing contrast of predictions of CSTs in FLIP samples and comparison with VMET2 data performance..... 170
Figure 6.7 - ROC curve showing prediction of responder status in FLIP samples 171
Figure 6.8 - Boxplots showing differences between timepoint 0 and timepoint 1 of metabolites in responders and non responders groups. 172
Figure 6.9 - Scatterplot showing relation between predicted and measured levels of main inflammation markers. 173
Figure 6.10 - Boxplots showing levels of main metabolites correlated with immune markers between LDOM and LDEPL samples. 174

List of Tables

Table 2.1 - Instrumental parameters used with the LTQ- Orbitrap Discovery mass spectrometer..... 76
Table 2.2 - DESI-MS geometric and electro-solvent parameters used for the medical swabs analysis. 77
Table 2.3 - Dimensional differences between LTQ-Orbitrap and RDa ACQUITY platforms. 78
Table 2.4 - RDa ACQUITY platform operating parameters. 80
Table 2.5 - RDa ACQUITY platform DESI swab stage set up parameters. 81
Table 2.6 - Multiple reaction monitoring (MRM) parameters for the targeted analytes and their isotopically labelled versions. 85
Table 3.1 - Demographic characteristics of VMET 2 cohort 100
Table 4.1 - Random forest results for CSTs comparison 139
Table 5.1 Demographic and pregnancy outcome information of patients providing soft cup samples 152
Table 6.2 AUC, Accuracy and Sensitivity for each CST contrast in FLIP study and VMET2 study. . 170

Abbreviations

12-HETE	12-Hydroxyeicosatetraenoic acid
17-OHPC	17-Hydroxyprogesterone caproate
20αHSD	20alpha-hydroxysteroid dehydrogenase
20α-OHP	20α-Hydroxyprogesterone
3β-HDS	3β-Hydroxysteroid Dehydrogenase
AKR1C1	Aldo-keto reductase family 1 member C1
ANOVA	Analysis Of Variance
AP-1	Activator Protein 1
APC	Antigen Presenting Cell
ASV	Amplicon Sequence Variant
AUC	Area Under the Curve
BMI	Body Mass Index
BV	Bacterial Vaginosis
BVAB	Bacterial Vaginosis-Associated Bacteria
CCL-2	Chemokine (C-C motif) ligand 2
CFU	Colony-Forming Unit
COX-2	Cyclooxygenase-2
CRH	Corticotrophin-releasing hormone
CRH-BP	CRH-binding protein
CST	Community State Type
CV	Cross Validation
CVF	Cervical-Vaginal Fluid
CX43	Connexin 43
DC	Direct Current
DESI-MS	Desorption Electrospray Ionisation Mass Spectrometry
DHEA	Dehydroepiandrosterone
DHEA-S	Dehydroepiandrosterone sulphate
DNA	Deoxyribonucleic Acid
EDTA	Ethylenediaminetetraacetic Acid
EI	Electron Impact
EMMPRIN	Extracellular Matrix Metalloproteinase Inducer
ER	Estrogen Receptor
ESI	Electrospray ionisation
FDA	Food and Drug Administration
FDCS	Caesarean Section at Full cervical Dilatation
FDR	False Discovery Rate
fFN	Fetal Fibronectin
FIGO	International Federation of Gynecology and Obstetrics
FUT2	Fucosyltransferase 2
GA	Gestational Age
GC	Gas Chromatography
GM-CSF	Granulocyte-macrophage colony-stimulating factor
GPR30	G protein-coupled estrogen receptor
GPX	Glutathione peroxidase
HIV	Human Immunodeficiency Virus

HMDB	Human Metabolome Database
HPV	Human Papillomavirus
IFN-γ	Interferon gamma
IL	Interleukin
IL-RA	Interleukin Receptor Antagonist
IQR	Interquartile Range
IRDB	Institute of Reproductive and Developmental Biology
IS	Internal Standard
IVF	In Vitro Fertilization
LAG-3	Lymphocyte Activation Gene 3
LC	Liquid Chromatography
LDL	Low-Density Lipoprotein
LDEPL	*Lactobacillus* depleted
LDOM	*Lactobacillus* dominant
LLETZ	Large Loop Excision of the Transformation Zone
LLOQ	Lower Limit of Quantification
LPS	Lipopolysaccharide
LTQ	Linear Trap Quadrupole
MALDI	Matrix-Assisted Laser Desorption/Ionization
MBL	Mannose-Binding Lectin
MCT-1	Monocarboxylate Transporter
MMP	Matrix Metalloproteinase
MRM	Multiple Reaction Monitoring
MS	Mass Spectrometry
MTL	Mid-Trimester Loss
m/z	Mass to charge ratio
NF-kB	Nuclear Factor-kappa B
NGS	Next-Generation Sequencing
NHS	National Health Service
NK	Natural Killer
NMR	Nuclear Magnetic Resonance
NPV	Negative Predictive Value
NRES	National Research Ethics Service
OTU	Operational Taxonomic Unit
P4	Progesterone
PAMG	Placental alpha macroglobulin
PAMP	Pathogen-Associated Molecular Pattern
PBS	Phosphate-Buffered Saline
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PD-1	Programmed Cell Death Protein 1
PERMANOVA	Permutational Multivariate Analysis of Variance
PG	Prostaglandin
PGE-2	Prostaglandin E
phIGFBP -1	Phosphorylated Insulin-Like Growth Factor Binding Protein 1
PLS-DA	Partial Least Squares Discriminant Analysis
POC	Point of Care
ppm	Part per million

PPROM	Preterm Prelabour Rupture of Membranes
PPV	Positive Predictive Value
PR	Progesterone Receptor
PTB	Preterm Birth
QC	Quality Control
QTOF	Quadrupole Time-of-Flight
RDP	Ribosomal Database Project
REC	Research Ethics Committee
REIMS	Rapid Evaporative Ionization Mass Spectrometry
REML	Restricted Maximum Likelihood
RF	Random Forest
RNA	Ribonucleic Acid
ROC	Receiver Operating Characteristic curve
rRNA	Ribosomal RNA
RU486	Mifepristone
SCFA	Short-Chain Fatty Acid
SD	Standard Deviation
sPTB	Spontaneous Preterm Birth
STD	Sexually Transmitted Disease
STI	Sexually Transmitted Infection
TE	Tris-EDTA
TIC	Total Ion Current
TLR	Toll-Like Receptor
TM	Targeted Metabolomics
TNF-α	Tumour Necrosis Factor Alpha
TOF	Time-of-Flight
TQ	Triple Quadrupole
UK	United Kingdom
ULOQ	Upper Limit of Quantification
UPLC	Ultra-High Performance Liquid Chromatography
USA	United States of America
UTI	Urinary Tract Infection
VIP	Variable Importance Projection
VMET	Vaginal Microbiome and Metabolome in Pregnancy
WHO	World Health Organization
ZEB-1	Zinc-Finger E-Box-Binding Homeobox-1

Chapter I

1 INTRODUCTION

1.1 Preterm birth

Preterm birth (PTB) describes delivery before the 37th week of gestational age (GA)[1]. It remains the primary cause of death of children under 5 years of age worldwide [2], and survivors often suffer from long-term health complications [3]. Globally, PTB is estimated to occur in around 11% of all pregnancies, however PTB rates vary widely across countries with the highest rates in South-eastern Asia and sub-Saharan Africa [4], [5]. PTB is estimated to affect 10.09% of all pregnancies in the USA in 2020 and 7.6% in the UK in 2021 [6]. Both within the USA and the UK race is a major determinant of PTB, with babies from the black ethnic group being the most affected. In the UK, PTB rates in the black ethnic group were reported at 8.7% in 2021 [4], [7]. In many countries, including high-income countries, rates of preterm birth have been steadily increasing over several decades [1]. This increase is only partially attributable to better management of fetal viability and reductions in stillbirth rates [8].

PTB is often sub-classified based on GA at delivery and on underlying aetiology [1]. The WHO defines delivery before 28 weeks as ‘extreme preterm birth’, delivery between 28 and 32 weeks as ‘very preterm’ and delivery between 32 and 37 weeks ‘moderate or late preterm’[9]. Approximately two-thirds of all PTBs occur spontaneously with the remaining one-third being clinically induced, often due to maternal or fetal complications [10]. Approximately 30-40% of all preterm birth are preceded by premature rupture of the fetal membranes (PPROM), which is often associated with an infectious aetiology [3]. The third trimester of pregnancy represents a crucial period for organ growth and development and, although many organs are immature at birth, brain and lungs are the most susceptible to adverse consequences [11]. Therefore, the risk of mortality and adverse health outcomes increases with lower gestational ages at birth. PTB is often associated with impaired pulmonary functions, higher risks of cerebral palsy, cognitive disability, and seizure disorder in prematurely born children [12]. Additional PTB consequences include renal, cardiovascular, and endocrine disfunctions that could manifest as late as in adult life [12], highlighting the significant long-term health and economic consequences of being born prematurely [2].

Addressing PTB remains a major priority in medical research. Yet despite decades of investigations, the ability to predict and treat preterm birth remains highly challenging. This is likely due to the fact that PTB is "a syndrome attributable to multiple pathologic processes" yet it is treated as a single disease (Figure 1.1) [3], [10]. There are several well-established risk factors of PTB including a history of PTB, maternal age, race, BMI, maternal drug and alcohol use, smoking status and socio-economic status [13]–[15]. The risk of PTB in women with previous history of preterm delivery is estimated to be between 10% and 57% depending on the GA at the previous delivery and the number of previous pregnancies [16]. Increased risk of PTB is associated with both young [17] and advanced maternal age [18], [19]. Epidemiological studies have also highlighted that ethnicity is an important determinant of PTB risk [13], [20]. Kistka et al. reported that Black women experience preterm birth at increased rates and at earlier gestations compared to White women [21]. Overall, the PTB rate of Black, African American and Afro-Caribbean is 16-18% compared to 5-9% in White women [3].

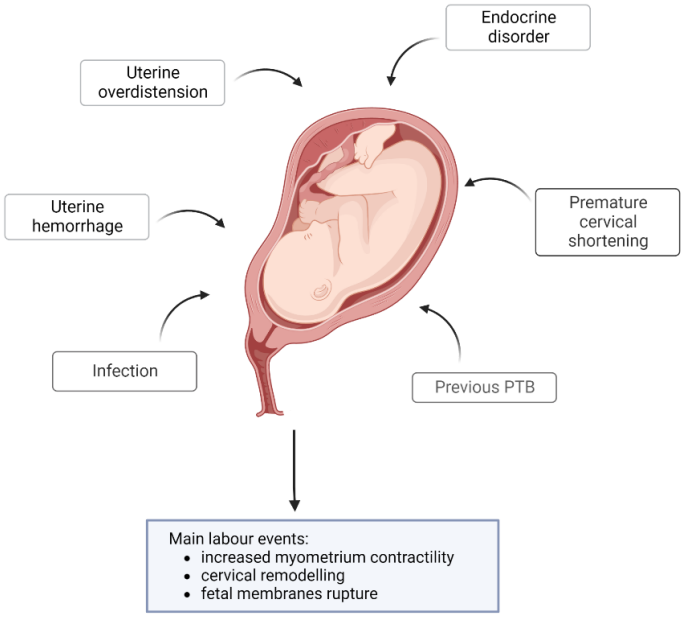


Figure 1.1 - Graphic summary of possible PTB aetiologies and of the main events triggered in preterm labour.

Despite the identification of these epidemiological risk factors, the largest proportion of PTB cases occur in women with low-risk pregnancies and/or those women experiencing their first pregnancy [22]. Moreover, an infectious aetiology is associated with up to 40% of all PTB

cases [23] and there is strong mechanistic evidence that pathogen colonisation of the reproductive tract can lead to the development of a preterm birth phenotype. To understand how this occurs, it is important to first consider underlying mechanisms of healthy term parturition.

1.2 Regulation of parturition in term and preterm pregnancies

1.2.1 Term parturition

The normal length of a healthy term pregnancy is considered to be between the 37th and the 42nd week of gestation [24]. Although the precise mechanism underpinning the onset of human parturition at term is yet to be elucidated, it is clear that it involves carefully orchestrated interaction between mechanical, endocrine and particularly, pro-inflammatory signals that prepare the uterus and other gestational tissues for the expulsion of the fetus [25]. In the 1980s Liggins observed that structural changes in the cervix at the time of labour resembled an inflammatory process, and more recent evidence indicates activation of pro-inflammatory pathways in gestational tissues prior to the onset of labour [26]–[30]. In addition to parturition, modulation of pro- or anti-inflammatory pathways plays a key role in other essential physiological stages of pregnancy [31]. At the beginning of pregnancy, the maternal immune system establishes a pro-inflammatory environment [31] that helps the myometrium to proliferate and become thicker to facilitate implantation of the embryo [25]. Throughout the rest of gestation, the uterus is maintained in a quiescent state which is coordinated via the action of hormones such as progesterone as well as adaption of the maternal immune system that remains tolerant towards the fetus [32]. Around the time of labour onset, it is hypothesised that changes in progesterone signalling in addition to signals coming both from the mother [33] and the fetus [34], [35], promote a cascade of pro-inflammatory signals that drives activation of pro-labour and pro-tissue remodelling pathways [24].

The main mechanical changes occurring at the time of labour include activation of myometrial contractility and the softening and dilatation of the cervix. The transition of the myometrium from a quiescent to an active state is linked to the increased number of gap junctions forming between adjacent myocytes. This allows for electrical potential to be propagated in a synchronised fashion [24]. The myometrium can then start developing regular contractions,

with higher intensity and frequency [24]. Throughout most of pregnancy, the cervix provides a mechanical and chemical barrier isolating the fetus from the external environment. Just prior to labour onset, it experiences drastic structural changes that involve major changes to key components of the cervical tissue including fibroblasts and extracellular matrix proteins (e.g. collagen, extracellular proteins, elastin, glycosaminoglycans) that provide the structural integrity, strength and stability of the cervix [36]. Biochemical changes in the cervix in the weeks leading up to labour onset include changes in the water content of the cervix and breakdown of the collagen structure, resulting in an increased elasticity and plasticity of the entire cervix [24].

Together with these changes, an influx of leucocytes, natural killer cells (NK) and neutrophils has been observed in the reproductive tract close to the onset of labour [37]. Immune cells infiltrates in the cervix, myometrium and chorio-amniotic membranes are associated with the production of pro-inflammatory cytokines such as IL-8, IL-6, TNF-α, IL1-β, CCL2 [38], [39]. Additionally, leukocyte infiltration in the cervix and fetal membranes drives production of metalloproteinases (MMPs) involved in fetal membrane rupture and cervical ripening [40]. As this inflammatory pathway activation occurs in the absence of infection, it is often referred to as “sterile inflammation” [41]. Hereafter are discussed the main hormonal mediators of the inflammation mechanism.

1.2.1.1 Hormonal mediators of pregnancy and parturition

Corticotrophin-releasing hormone

Corticotrophin-releasing hormone (CRH) is a hormone produced by the hypothalamus and, during pregnancy, by the trophoblasts in the placenta and chorion [42]. It is a peptide hormone composed of 41 amino-acids. Its structure is highly-conserved among vertebrates and it is processed from a precursor molecule called preproCRH, consisting of 196 amino-acids. The precursor molecule contains a hydrophobic sequence required for secretion at its amino-terminal end and the 41 peptides of the mature hormone at the carboxylic end [43]. CRH has been suggested to have a role in human parturition timing through a “placental clock” mechanism [44]. Plasma levels of CRH increase during the second and third trimester of pregnancy and peak during labour before quickly dropping in the postpartum phase, reaching pre-pregnancy levels 48h after delivery [45]. During most of the pregnancy the presence of a high-affinity CRH-binding protein (CRH-BP) renders circulating CRH biologically inactive,

but during the last weeks of pregnancy, CRH-BP levels decrease significantly which results in an increase of free and biologically active CRH in the blood stream [45]. Challis and Hopper suggested that CRH is responsible for a series of positive-feedback hormonal loops to induce labour [46]. This begins in the placenta and the fetus where CRH promotes the production of fetal cortisol and dehydroepiandrosterone sulphate (DHEA-S). Fetal cortisol is known to induce further CRH production, while DHEA-S is converted in estrogen that indirectly promotes the formation of gap junctions in the myometrium and enhances the prostaglandins activity [42]. In the amniotic cavity CRH promotes the synthesis and release of prostaglandins from the amnion, chorion, and decidua that, in turn, stimulate further secretion of CRH in the placenta. Some studies have also implicated CRH in the regulation of myometrial contractility. [47]. However, the precise mechanism of action of CRH remains to be determined.

Estrogens

Estrogens are a family of steroid hormones secreted by the adrenal cortex, ovary, and, during pregnancy, produced mainly by the placenta [48]. This group includes estradiol, estrone, and estriol [48]. As all the other steroid hormones, estrogens are synthetised from cholesterol and therefore have the characteristic lipidic cyclopentanophenanthrene four rings structure. Estradiol, estrone and estriol differ for the number of hydroxyl and keto groups around the steroid backbone. Details of their synthesis from cholesterol are shown in Figure 1.2. Estrogens are not directly responsible for uterine contractions but are crucial for the passage of the myometrium from a quiescent to an active state via promotion of transcriptional pathways leading to the production of prostaglandins, increased expression of oxytocin receptors, the formation of gap junctions between myometrial cells and the upregulation of enzymes involved in muscle contraction [49]. Estrogens are also involved in the cervical ripening process through the activation of collagenase and elastase that remodel and restructure the collagen fibres of the cervix [49]. In other mammals such as rats, mice and sheep circulating estrogens levels decrease just at the onset of parturition [50], [51]. However, this does not occur in humans where instead, levels remain high throughout the whole gestation, including the parturition phase [45]. It is hypothesised that the effects of estrogens at the time of labour are altered by changes in the expression and transcriptional activity of estrogen receptors (ERs). In humans 3 types of ERs have been described: two nuclear receptors (ERα and ERβ) and one membrane G protein coupled receptor, GPR30 [45]. While ERα promotes estrogen transcriptional activity, ERβ inhibits it. Thereby the ratio between ERα and ERβ may impact gene expression by estrogen [45].

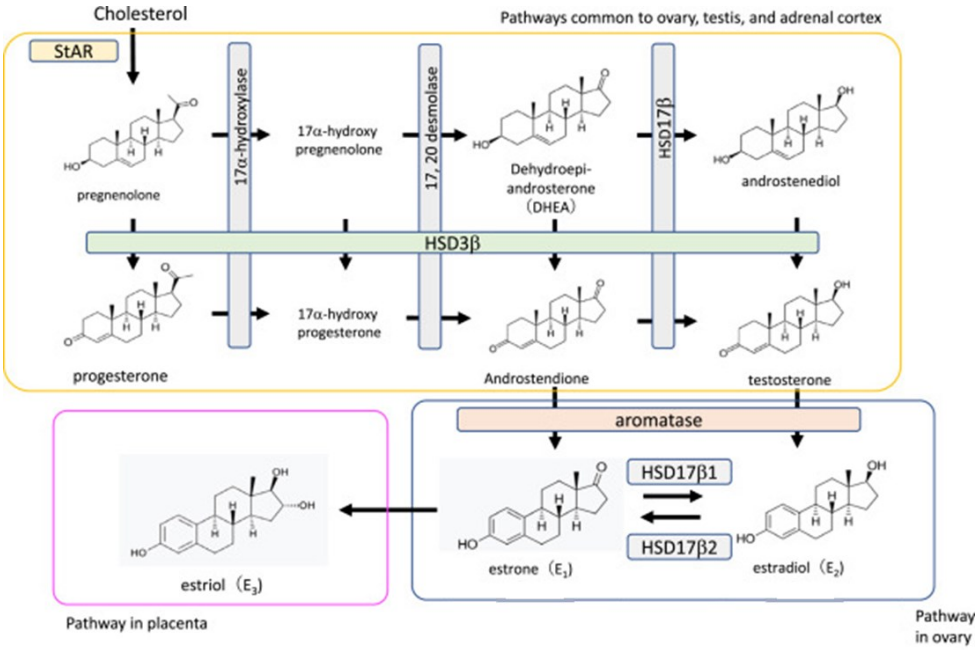


Figure 1.2 Steroid hormone biosynthesis pathway. Schematic representation of steroidogenesis from Cholesterol. Cholesterol is transferred from the cytoplasm to the mitochondrion by the steroidogenic acute regulatory protein (StAR). Cholesterol is then converted in pregnenolone by cytochrome p450scc. Pregnenolone is converted to progesterone by 3b-hydroxysteroid dehydrogenase (3b-HSD). 3bHSD is also responsible for converting 17-hydroxyprognenolone to 17-hydroxyprogesterone, dehydroepiandrosterone (DHEA) to androstenedione. Aromatase is the enzyme that converts testosterone and androstenedione to estradiol (E2) and estrone (E1), respectively. Figure modified from [52].

Oxytocin

Oxytocin is a peptide hormone produced by hypothalamus and was once considered to be the initiator of parturition due to the observation that clinical administration of oxytocin can induce labour in late stages of pregnancy [45], [53]. Oxytocin is a small cyclic peptide consisting of nine amino acids. Six amino-acids are structured in a ring through a disulphide bridge, and the remaining three are forming a tail on the amino-terminal end [54]. In many animals, including humans, oxytocin levels increase during labour, as does expression of oxytocin receptors [55]–[57]. The increased expression of oxytocin receptors at term increases the sensitivity of the myometrium to oxytocin [42]. Placental oxytocin acts directly on the myometrium to cause contractions and indirectly by up-regulating prostaglandin production. Prostaglandins, in turn, act on the myometrium to up-regulate oxytocin receptors and gap junctions, thereby promoting uterine contractions [49].

Prostaglandins

Prostaglandins (PGs) are lipids with a hormone-like function. They are unsaturated carboxylic acids containing 20 carbons. The major precursor for PGs is arachidonic acid, which is stored in glycerophospholipids, and gives PGs their characteristic five atom ring [49]. For most of pregnancy, levels of prostaglandins are suppressed by progesterone but increasing estrogen levels stimulate storage of glycerophospholipids in fetal membranes[49]. A series of fetal membrane lipases controls the release of arachidonic acid from the membranes that is then free to be transformed into prostaglandins. There is strong evidence that PGs are involved in the initiation of labour in humans [58], [59]. Just prior to and during labour, levels of prostaglandins increase and stimulate myometrial contractility and ripening of the cervix [45]. PGs can stimulate uterine contractions by altering the cellular membrane permeability and therefore increasing the intracellular calcium concentration [60]. They promote the formation of gap junctions facilitating the transmission of contraction signals horizontally throughout the myometrium[61]. Secondarily, they upregulate the expression of oxytocin receptors [62]. Regarding their role in cervical ripening, PGs induce increased elastase and collagenase activity. PGE2 has been shown to stimulate IL-8 production, with consequent recruitment of neutrophils and release of collagenase upon degranulation [63]. Also, PGE2 induces functional progesterone withdrawal (mentioned in the following section) by increasing the PR-A/PR-B expression ratio [63].

Progesterone

Progesterone is a steroid hormone that, as the name suggests, is “pro-gestation”, meaning it promotes maintenance of pregnancy [64]. The mechanism of action of progesterone has been hypothesised as early as in 1950s when Csapo proposed the “progesterone block” theory [65]. According to this theory, progesterone maintains pregnancy by blocking myometrial activity; removal of progesterone’s functionality initiates labour [66]. There are now several lines of evidence to support [66]–[68].

In non-pregnant women progesterone is mainly produced by the corpus luteum in the ovary during the second half of the menstrual cycle. During pregnancy the corpus luteum remains the main source of progesterone for the first 7 to 9 weeks until the trophoblasts of the placenta take over production [69].

The synthesis of all steroid hormones, including progesterone, starts with the conversion of two molecules of Acetyl-CoA in squalene and lanosterol, two precursors of cholesterol [70]. Cholesterol is present in cell membranes and also the blood stream in the form of low-density

lipoproteins (LDLs). LDLs interact with cellular membranes and are shuttled inside the cell through lysosomes. Once in cytoplasm, free cholesterol is transported to the mitochondria where it is converted in pregnenolone [71]. Pregnenolone is then released from the mitochondria and once in the cytoplasm is converted in progesterone by 3β-Hydroxysteroid Dehydrogenase (3β-HDS) [72].

Progesterone circulates in the bloodstream bound to cortisol-binding globulin and serum albumin and exerts its function by binding to progesterone receptors (PRs). Several classes of PRs have been identified, including nuclear (nuclear PR – nPR) and membrane bound (membrane PR – mPR). Binding of progesterone to nuclear receptors mediates cellular transcriptional activity. There are two main isoforms of nuclear PR: PR-A and PR-B. They are expressed in variable amount and ratios in all the tissue targeted by progesterone but during pregnancy there is a prevalence of PR-B expression. Binding of progesterone to PR-B, induces a conformational change that removes chaperone proteins enabling the formation of dimers that translocate inside the nucleus where they bind to specific progesterone-response elements within promoter regions of target genes. Many of these target genes are involved in inflammatory response. Progesterone exerts significant immunomodulatory effects by the inhibition of pro-inflammatory transcription factors such as NF-kB and AP-1[73], [74]. This occurs via upregulation of the NF-kB inhibitor IkBa [73] and increased expression of ZEB-1, which suppresses the expression of contractile genes such as CX43 and oxytocin receptor. In this way, for most of pregnancy, progesterone maintains uterine quiescence by inhibiting the activation of pro-inflammatory and pro-contraction genes.

Non-genomic effects of progesterone have also been observed and linked both to the action of membrane and nuclear receptors. The action of membrane receptors could affect the integrity of fetal membranes. Luo et al. recently reported that progesterone blocks TNF-α induced apoptosis in the fetal membrane preserving their integrity and preventing rupture [75]. Nevertheless, the role of mPR and non-genomic effects of progesterone in parturition timing remain unclear.

In many mammals, progesterone withdrawal initiates onset of labour. For example, in sheep at the end of gestation there is a stimulation of placental enzymes converting progesterone to estrogen. Additionally, increased production of heat-shock proteins near term, prevent the dimerization of the PRs necessary for their genomic activity [51]. However, in humans and some primates, systemic progesterone levels remain high at the end of gestation. Despite this, the use of a PR antagonist (e.g. RU486) initiates parturition highlighting the critical role of progesterone in mediating labour onset [76]–[78]. One way in which the action of progesterone

is thought to be inhibited at the end human pregnancy is through a “functional” progesterone withdrawal. There are several molecular mechanisms proposed to achieve this [58], [66]. The first involves a change in the expression of the PR isoforms. Towards the end of pregnancy, the expression of PR-A, a functional antagonist of progesterone, is increased. This leads to an increase in the PR-A:PR-B ratio and consequently a decrease in progesterone signalling and withdrawal of its action. This in turn releases the so-called ‘progesterone break’ leading to increased expression of pro-inflammatory and pro-contractile genes. There is evidence that increased stretch of the uterine walls at term may be a mechanical mechanism that promotes a switch in the ratio of the PR isoforms [79]. Another proposed mechanism for progesterone withdrawal involves its metabolism by the enzyme 20α-hydroxysteroid dehydrogenase (20αHSD) [68]. 20αHSD catalyses the conversion of progesterone into 20α-hydroxyprogesterone (20α- OHP), an inactive metabolite [80]. It has been previously reported an increase of 20α- OHP in human myometrium at term compared to progesterone, suggesting an increased metabolism towards the end of gestation [80]. In mice, 20αHSD has been robustly associated with labour: in 20α- HSD knockout mice it has been observed an increase of progesterone levels and delays in parturition[81]. Additionally, the activity of 20αHSD it is associated with increased inflammation. Nadeem et al. showed that inflammatory mediators associated with term and preterm labour induce localized progesterone withdrawal in myometrial cells [68]. This mechanism involves the increased expression of *AKR1C1*, the gene encoding for 20α- HSD. The study found up-regulation of *AKR1C1* in human myometrium in term parturition and the up-regulation of the homologous murine gene in myometrium in both term and preterm mice models [68]. The consequent conversion of progesterone in its inactive metabolite would therefore remove progesterone inhibition effect and lead to labour.

1.2.2 Preterm parturition

There are multiple different underlying mechanisms that could potentially cause the onset of preterm labour and delivery.

1.2.2.1 The role of infection in preterm labour

Infection is linked to approximately 40% of all cases of preterm birth [82]. There are several potential routes of infection including haematogenous spread or the introduction of infection during invasive procedures (e.g. amniocentesis) [83]. However, it is widely considered that the most common route for infection is ascension from the vaginal canal [3], [84]. This mechanism is readily demonstrable in animal models. For example, Suff et al. showed using a bioluminescent strain of the pathogen *E. coli* which permitted non-invasive tracking of colonisation, that vaginal administration leads to ascending infection via the cervix followed by the uterus and placenta [73]. In humans, evidence for ascending vaginal infection is supported by several observations. Firstly, bacteria found in the placenta and fetal membranes (amnion and chorion) of PTB cases are similar to the ones present in the vagina [85], [86]. Additionally, chorioamnionitis (the inflammation of fetal membranes) is usually observed near the site of membrane rupture, near the cervix [87]. Finally, in twin pregnancies the microbial invasion of the amniotic cavity is observed in the first born twin, closest to the cervix [88].

In the event of a pathogen colonisation of the reproductive tract, the immune system responds via the innate and the adaptive immune systems. The innate response is mediated by the complement system and by antigen presenting cells (APCs), which include macrophages, neutrophils, and dendritic cells. The adaptive immune system relies on lymphocytes and on their ability to produce cytokines and antibodies [89]. Detection and recognition of pathogens by APCs occurs via surface specific pattern recognition receptors, called toll-like receptors (TLRs), capable of recognising and binding “pathogen associated molecular patterns” (PAMPs) [90]. Examples of PAMPs include components of bacterial outer structures such as lipopolysaccharides (LPS), peptidoglycans, lipoproteins but also bacterial nucleic acids [91]. Activated APCs are then responsible for the maturation of T and B lymphocytes that produce pro-inflammatory cytokines and immunoglobulins, respectively. This leads to the activation of pro-inflammatory transcription factors such as NF-kB and AP-1 in the gestational tissues and increased production of pro-labour and pro-tissue remodelling proteins such as prostaglandins

and matrix metalloproteinases and eventually untimely onset of parturition and labour [92], [93].

1.2.2.2 Progesterone in preterm parturition: the deficiency hypothesis

Given the importance of progesterone in maintaining a healthy gestation until term, its role in preterm deliveries has been widely investigated. Blood progesterone levels start to rise midway through the menstrual cycle and continues to rise in pregnancy (see Figure 1.3).

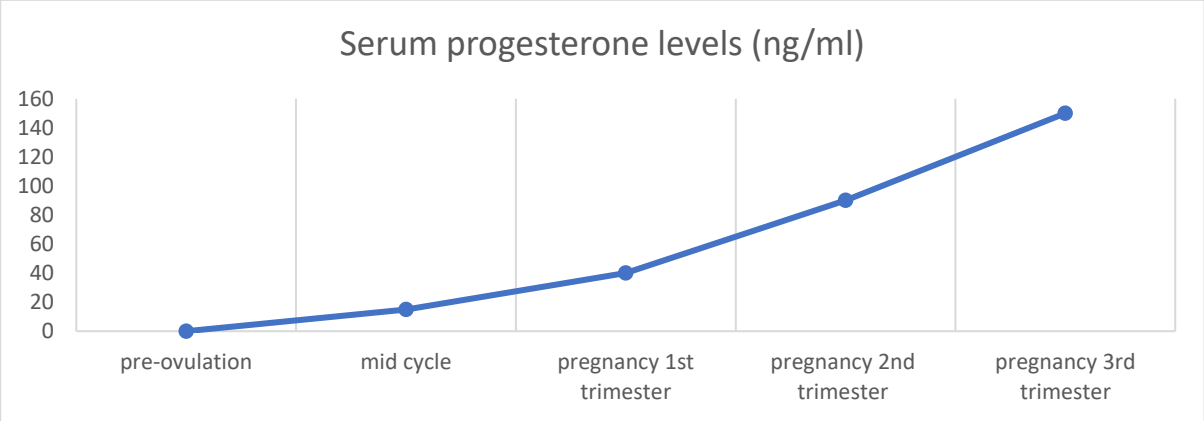


Figure 1.3 - Chart describing serum progesterone levels in menstrual cycle and in pregnancy.
During the follicular phase serum levels of progesterone are less than 1.5ng/ml increasing progressively until ovulation (15ng/ml). In pregnancy progesterone levels increase steadily throughout gestation from a level of 40 ng/ml at the end of the first trimester to a level of 150ng/ml at the end of the third trimester [94].

It has been hypothesised that some preterm deliveries could be caused by a local progesterone deficiency. This theory, formulated in 2011 by Roberto Romero, is largely based on the fact that administering progesterone receptor antagonists any time during pregnancy leads to cervical ripening and often, the onset of labour [95]. However, only a few studies have demonstrated a correlation between low serum progesterone levels in early stages of gestation and poor pregnancy outcome (more specifically miscarriage and threatened miscarriage) [96], [97]. To date, levels of progesterone in cervicovaginal fluid during pregnancy have not been determined, nor a potential relationship with gestational age at birth assessed.

1.3 PTB Prediction strategies

Detecting women at high risk of PTB during early stages of pregnancy is currently challenged by the absence of a reliable test that can confidently predict cervical shortening, preterm labour or PPROM. The identification of PTB-associated phenotypes early in pregnancy would facilitate appropriate clinical decision on risk modification or preventative interventions.

The main indicator for prediction of PTB is the analysis of known risk factors. Risk factors for PTB include demographic characteristics like the socio-economic status, advanced maternal age, malnutrition, behavioural factors like smoking, alcohol consumption, drug abuse, stress and obstetric history [3]. Among obstetric history factors, the most significant risk factor for PTB is a previous history of PTB, which confers an estimated 2.5-fold increased relative risk of PTB compared to women with no previous PTB history [98]–[100]. Obviously, this specific predictor is only useful for women that had previous pregnancies and it is not useful for primigravid women.

Another significant risk factor for PTB associated with the patient's obstetric history, is a previous treatment for cervical precancerous lesions. Precancerous lesions of the cervix are changes in cervical cells that make them more likely to develop into cancer. Treatment for this condition is usually the excision of the affected portion of the cervix through a procedure called large loop excision of the transformation zone (LLETZ). Women with previous LLETZ have increased risk of both early and late sPTB compared to women who did not have a diagnosis of precancerous changes in the cervix [101], [102].

In general, using risk factor assessment alone as a predictor of PTB is unreliable as over 50% of pregnancies that deliver preterm have none of these risk factors present [103]. The prediction tool currently most used in clinical practice, is the measurement of mid-trimester (16-24 weeks) cervical length. Cervical length measurement is obtained through transvaginal ultrasound scan, quantifying the distance between the internal and external opening of the cervix. Measurements above 25mm are considered an indication of a normal cervix and term delivery, while measurements below 25mm are considered to double the risk of PTB in women with a previous history of PTB (Figure 1.4). Currently, transvaginal ultrasound is not recommended for low risk population despite its safety, reliability and reproducibility [104]–[106].

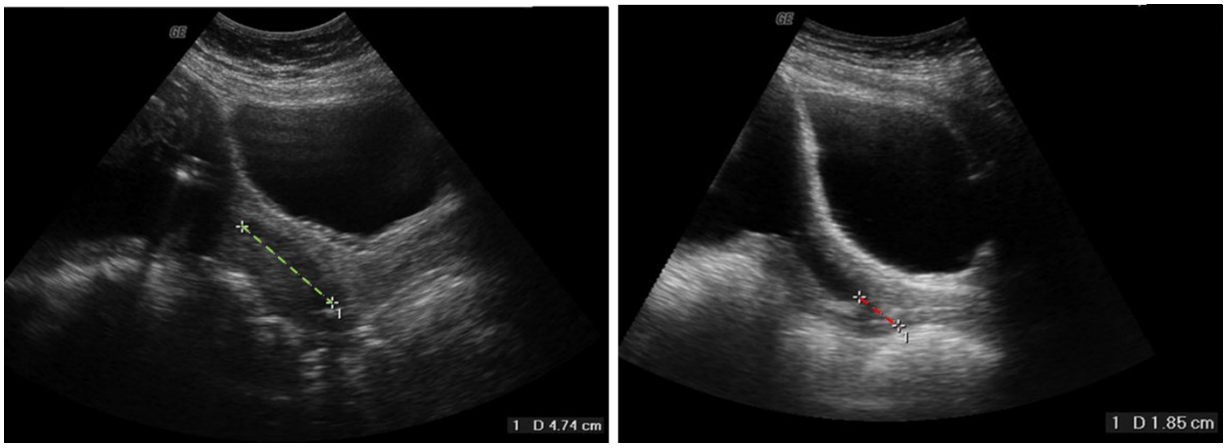


Figure 1.4 - Photo of vaginal ultrasound for cervical measurement. Vaginal ultrasound example of normal cervical length (left) and short cervical length (right). Image adapted from [107].

Another tool in predicting early PTB-associated phenotypes is the use of biochemical cervicovaginal fluid markers. More specifically fetal fibronectin (fFN), Phosphorylated insulin-like growth factor binding protein-1 (phIGFBP1) and Placental alpha macroglobulin (PAMG-1).

fFN is an extracellular matrix glycoprotein produced by fetal cells located between the chorion and decidua [108]. Around the 18th week of gestation these tissues fuse together and fFN concentrations become undetectable. After this threshold, the presence of fFN into the CVF is presumed to be due to an inflammatory, infectious or mechanical event that disturbs the extracellular matrix. Its presence at elevated concentrations has been associated with an increased likelihood of preterm delivery in high risk women [109], [110]. The test is qualitative with a positive/negative response and is used in women at high risk around 22-34 weeks of gestational age. The test has a low positive predictive value and it is mainly used for its negative predictive value as a reassurance of future unlikely intervention [111], [112]. One of the main limitations of this test is that can be prejudiced by recent vaginal digital examination, unprotected sexual intercourse, vaginal bleeding or amniotic fluid contamination from ruptured fetal membranes [113].

PhIGFBP1 is secreted by placental decidual cells. Similarly to fFN, the presence of this factor in the cervicovaginal fluid is the consequence of some disruptive event [114]. The test involving PhIGFBP1 has small positive predictive ability and, like the fFN equivalent, it is used as a reassurance tool for its negative predictive value [115]. However this test is cheaper and not affected by recent examinations, sexual intercourse etc. [115].

Finally, PAMG-1 test has similar predictive characteristics to the previous two. PAMG-1 is a glycoprotein synthesized by the decidua. It is present in the amniotic fluid in high concentrations but high vaginal concentrations are find in cases of threatened preterm delivery due to transudation through chorioamniotic pores in fetal membranes during uterine contractions, or by the degradation of extra cellular matrix in fetal membranes due to inflammatory process of labour or infection [116].

Collectively, these cervicovaginal biomarkers tests are useful for their negative predictive value in reassuring the patient at high risk of PTB but are not capable of reliably predicting PTB.

1.4 PTB Prevention and treatment strategies

The strategies adopted in the prevention of preterm birth consist of both primary and secondary prevention measures. Primary prevention measures are directed towards all women before or during pregnancy and include weight optimisation, smoking cessation, nutritional supplementation [117]. Low maternal BMI values and smoker status are considered strong predictors for adverse pregnancy outcome [118], [119]. Smoking cessation in pregnancy has been reported to reduce the proportion of women who experience PTB and reduce the incidence of low birth weight [120]. On the other hand, nutrient supplementation strategies are considered controversial. Although low serum levels of iron, folate and zinc are associated with preterm birth and stillbirth [121], results of nutritional supplementation studies and preterm birth prevention have been conflicting [122].

Secondary prevention measures are directed toward women already know to be at heightened risk of PTB [117]. These interventions include cerclage, antibiotic treatment, progesterone supplementation and the use of pre- and probiotics.

1.4.1 Progesterone

Progesterone is the most widely used preventative measure against PTB. Its use is largely based upon previously discussed evidence. It promotes uterine quiescence both through interactions between nuclear and membrane receptors and by lowering levels of pro-inflammatory cytokines, oxytocin, and intracellular calcium [123], [124]. It has also been suggested that a

subset of PTB cases are caused by a local deficiency of progesterone in the cervicovaginal environment yet further investigation is needed to confirm this claim [95]. According to recent recommendations, women with a singleton pregnancy and a previous PTB history should be offered progesterone supplementation to prevent recurrent preterm labour [117]. Two different forms of progesterone have been used to treat PTB, natural progesterone (P4) and 17α-hydroxyprogesterone caproate (17-OHPC) [125]. However, these should be treated as two different compounds. Natural P4 used therapeutically is typically manufactured from plant extracts and can be administered orally, vaginally, rectally or by intramuscular injection [125]. However, due to "first pass effect" the best administration route appears to be through a vaginal pessary. This would also avoid the unwanted side effects such as nausea, headache, and sleepiness occurring after oral administration [126]–[128]. 17-OHPC is a synthetical version of a naturally occurring progesterone metabolite called 17α-hydroxyprogesterone (17-OHP). It is usually administered through a intramuscular injection [125].

Many trials have been conducted to investigate the efficacy of P4 to prevent preterm birth, yet findings have proven inconsistent. For example, Da Fonseca *et al.,* published a randomised placebo-controlled trial of natural vaginal progesterone in 142 women with prior history of PTB [129]. Also, two large multicentre trials investigated vaginal progesterone in women with a short cervix [130], [131]. All these studies demonstrated significant reduction in the rates of PTB. However, a larger study by Crowther *et al*., conducted in New Zealand and Australia failed to demonstrate any effect using vaginal progesterone [129]. A randomised trial of 463 women with a singleton pregnancy and a prior history of PTB compared the effects of placebo with weekly 17OHPC. The results showed a significant reduction in the rate of preterm birth at less than 37, 35 and 32weeks [132]. In multiple pregnancies, P4 does not appear to prevent PTB [133].

Thus, while there is support for biological plausibility in the use of progesterone for preventing PTB, evidence of the mechanisms involved is lacking. Further investigation is needed in the field to clarify the effectiveness of this treatment.

1.4.2 Cerclage

Cervical cerclage is a widely used surgical procedure that consists of placing a stitch around the cervix, with the aim of providing mechanical support [134]. In women with a history of spontaneous preterm birth or a short cervical length, cervical cerclage is estimated to reduce

the risk of preterm birth by approximately 20% [135]. Cerclage is typically given to women who have been found to have a cervical length <25mm, which is considered a risk factor for preterm birth. A short cervix is most commonly diagnosed using a transvaginal ultrasound [136]. In women with a history of cervical shortening, a cerclage is sometimes placed prior to diagnosis of a short cervix or in some cases, even before pregnancy. However, cerclage does not work in a large portion of women and meta-analysis on four randomised trials demonstrated an increased risk of PTB in women with twin gestations after a cerclage [137]. Furthermore, cerclage has been linked with increased risk of rupture of membranes, chorioamnionitis, and an increased risk of infection [134], [138]. The increased risk of vaginal infections following cerclage appears to be related to the material chosen for performing the procedure. A study published in 2016 reported that cervical cerclage performed with braided filaments induces vaginal dysbiosis which is often associated with pathogen enrichment, increased production of pro-inflammatory cytokines and early cervical remodelling. In contrast, women receiving a monofilament suture has not impact on their vaginal microbiome and inflammatory pathways [139].

1.4.3 Antibiotic therapy

Given the substantial evidence linking infection to the pathophysiology of PTB, is not surprising that there has been extensive investigations on the use of antibiotics for the prevention of preterm birth [140]. Many studies have tried to investigate the use of different formulations, mainly in the context of BV, in pregnancy. However, the results on whether antibiotics are effective in decreasing PTB incidence have been inconsistent [141], [142].

This is due to the heterogeneity of the trials conducted, that present often different timing of the treatment, types of antibiotic used, administration routes and kinds of patient cohorts [143].

For example the PREMEVA trial, the largest trial to date, is a double blinded trial for oral clindamycin in early pregnancy. In this trial women with BV diagnosis and low background risk for PTB were included. No difference in the rate of miscarriages or very preterm birth has been observed [144]. This study presents a number of limitations to consider including the low number of primary outcomes reported (1.2% of the overall pregnancies) and the fact that almost 20% of women did not complete their assigned treatment course. The trial has also faced criticism for the method by which BV was diagnosed [143] and for not assessing the effectiveness of the treatment to resolve BV [145]. However, if these data hold true, they

collectively highlight the ineffectiveness of oral clindamycin to prevent BV-associated preterm birth in low-risk women.

Some aspects need to be carefully considered when analysing this and other trials. The treatment of women with an existing BV diagnosis, implies that the infection is already happening and the inflammatory pathway activation may have already started. This suggests that a prompt diagnosis of infection could potentially increase the effectiveness of the therapy [146]. Also, the action of specific antibiotics (e.g. metronidazole) relies on bacterial lysis with consequent possible release of endotoxins [147]. These molecules are known activators of inflammation in gestational tissues and therefore might accentuate the already present inflammatory phenotype [148]–[150]. Additionally, some antibiotics have been reported to be particularly effective against *Lactobacillus* species but not necessarily as effective against pathogens, causing an involuntary enrichment of undesirable species in the vaginal mucosa [151]–[153]. Finally, there is also evidence that BV-associated bacteria have high rates of antibiotic resistance [154], [155].

In the current clinical practice, for women presenting with PPROM, prophylactic administration of oral erythromycin at a dose of 250 mg for 10 days is used as first line treatment in many countries [156]. This is based upon the results of the ORACLE I trial. This trial reported short-term neonatal benefits, prolongation of pregnancy for 48 hours, reduced need for supplemental oxygen, reduction in neonatal morbidity (neonatal death, chronic lung disease, or major cerebral abnormality) in women randomized to erythromycin [157].

The positive effects of erythromycin are attributed to the inhibition of ascending vaginal infection. However, considering the low levels of erythromycin in the vaginal mucosa following oral administration, this seems quite unlikely. The concentrations of erythromycin measured in the vaginal mucosa appear to be effective against *Lactobacillus* species but not against most vaginal pathogens [158], [159]. Another suggested mechanism is the erythromycin's inhibitory activity of neutrophil migration, oxidative burst, cytokine and metalloproteinases release that has been already demonstrated in bronchial epithelial cells [160].

In summary, it is assumed that erythromycin treatment helps resolve vaginal and reproductive tract infection as well as associated inflammation reducing risk of neonatal infection and subsequent morbidity and mortality. However, a study published by Brown *et al*. investigated the effects of erythromycin treatment on vaginal microbiota and found that the treatment

worked differently for women with different vaginal microbial composition profiles. For women that presented with a *Lactobacillus spp.* dominated vaginal microbiome before the treatment, erythromycin was associated with a strong shift towards colonisation by highly diverse and dysbiotic communities and a reduction of *Lactobacillus* levels. For women that were *Lactobacillus* depleted before the treatment, a reduced diversity was reported with unchanged bacterial load. Babies born through the birth canal of women with vaginal dysbiosis in this cohort had higher rates of sepsis and chorioamnionitis. These findings highlight how treatment can have a potentially detrimental effect on a subgroup of women experiencing PPROM [153].

1.5 Host:microbiota interactions

Despite the apparent ineffectiveness of antibiotic to prevent preterm birth there is growing evidence that microorganisms in the reproductive tract during pregnancy are an important modifier of preterm birth risk. Microbiota is defined as the collection of microorganisms (i.e. bacteria, fungi, viruses and archaea) within a specific environment [161]. In contrast, the microbiome defines a microbial community with distinct properties and functions that include its interactions with its environment which can be described as its "theatre of activity" [161], [162]. Microbiota colonise external and internal mucosal surfaces in a number comparable to human host cells [163].

Like other body niches, the vagina harbours a dynamic and complex microbiota. Many of these microorganisms have developed a commensal relationship with the host. The importance of commensal bacteria in maintaining the host health has been extensively demonstrated with the use of "germ-free" animals [164]. In these studies the elimination of the microbiota had important consequences at different levels for the host's health [165], [166]. In return the host provides suitable growth conditions (e.g. organic nutrients, pH, osmotic pressure and temperature) [167].

Different body environments have different chemical and physical properties that leads to the development of exclusive, niche-specific host-microbiome relationships [168]. For example. the mucous membrane generally presents a surface layer of epithelial cells over a deeper layer of loose connective tissue. The epithelial cells in mucous membranes are covered in mucus, a thick liquid secretion containing mucin, a water-soluble glycoprotein, and numerous complex sugars. Degradation of mucins and complex sugars requires the action of multiple enzymes

such as sulphatase, esterase and sialidase [169], which in turn frees carbon sources for microbial metabolism and provides adhesion and attachment sites. The structure of sugars presented on mucins is determined by the action of an enzyme called α (1,2) fucosyltransferase. This enzyme is encoded by a gene called *FUT2* and is responsible of the expression of histo-blood group antigens (AB0) in mucosal tissues [170]. About 20% of the population has a non-functional FUT2 gene and therefore presents a "non-secretor" phenotype [171]. The secretory status can influence the microbiota composition by directly hampering the interaction between the cell and the bacterial surface as well as modifying carbon source availability [172]. Recent studies have shown how secretor status can mediate host susceptibility to specific pathologies (e.g., Chron's disease, HIV infection etc.) [173], [174]. There is also some evidence to indicate secretory status can also influence the composition of vaginal microbiota and its relationship with the host's health [175].

1.5.1 Vaginal microbiota in health and disease

Unlike other mucosal surfaces, the vaginal mucosa harbours a microbiota that is often dominated by single genus of bacteria. The predominance of *Lactobacillus* species in the human vaginal environment has been known since 1892 when Albert S. Döderlein discovered Gram-positive bacillus that at the time he named Döderlein's bacilli [176]. *Lactobacilli* are Gram-positive, rod-shaped bacteria that produce lactic acid as the major end product. They have been traditionally classified as oxygen-tolerant anaerobes [177], [178]. The prevalence of *Lactobacillus* species in the reproductive tract seems to be unique to humans and not shared by other mammals. Even in non-human primates, *Lactobacilli* account for <5% of the total microbiota [179]. The presence of *Lactobacilli* in the human reproductive tract is typically associated with states of optimal host health, which is achieved via several mechanisms. The main mechanism of action of *Lactobacillus* spp. is the production of lactic acid in both D- and L- isomers [180], [181]. The accumulation of lactic acid in the cervicovaginal environment acidifies the vaginal mucosa to a pH between 3.8 – 4.5. While *Lactobacillus* species have evolved to grow in such acidic conditions, numerous potentially pathogenic vaginal microbes are unable to survive. More specifically, at a pH value < pKa 3.9, the protonated or uncharged form of lactic acid predominates and, in these conditions, lactic acid is membrane-permeant [182]. Lactic acid presents a selective capacity to lyse bacteria other than *Lactobacillus spp.* [182], [183] and the mechanisms through which it causes bacterial death include the

acidification of the cytosol, the disruption of intracellular functions [184], the increase permeability of the membrane to other antimicrobial molecules [185]. Lactic acid also exerts immunomodulatory actions on the cervicovaginal epithelium cells [186]. It promotes an anti-inflammatory state by stimulating the production of anti-inflammatory cytokines and inhibiting the production of pro-inflammatory ones [186], [187].

Apart from lactic acid, *Lactobacilli* produce hydrogen peroxide (H_2O_2), which inhibits the growth of catalase-negative anaerobic organisms [188], and other antimicrobial peptides such as bacteriocins, bacteriocin-like substances and biosurfactants. Furthermore, the lactic acid physiological concentrations and the consequent acidification of vaginal secretions, enhance the protective activities of these antibacterial compounds [189].

Lactobacillus species have been reported to also promote autophagy of intracellular bacteria, viruses and protozoa [190], and inhibit the growth of pathogen by competitively bind the vaginal epithelium surface [180].

Historically, studies of vaginal microbiota were conducted using culture-dependent methods, which are intrinsically biased towards the growth of more "easily cultivable" species. This provides only a partial perspective on the vaginal microbiota composition [191]. The recent development of culture-independent approaches, particularly high-throughput 'next-generation' sequencing technologies, allowed a deeper and less-biased characterisation of the vaginal microbiota. The first large scale application of these types of approaches to the vagina was performed by Ravel et al. in 2011. Trying to define a "core vaginal microbiome" Ravel et al., instead showed that vaginal bacterial profiles of health reproductive age women could be grouped into five different so called "community state types" or CSTs [192]. A CST is defined as "a group of community states with similar microbial phylotype composition and abundance" [193]. Among the five original CSTs that were identified, four were dominated by *Lactobacillus spp.* (CST I - *L. crispatus*, CST II – *L. gasseri*, CST III – *L. iners*, CST V – *L. jensenii*), while CST IV was depleted in these species and had increased bacterial diversity. Species commonly identified as part of CST IV include various anaerobic bacteria such as *Prevotella* spp., *Megasphaera* spp., *Gardnerella vaginalis*, *Sneathia* spp., *Atopobium vaginae*, *Eggerthella* spp, *Mobiluncus* spp., *Lachnospiraceae* spp. (including BVAB 1-3), *Dialister* spp., *Parvimonas* spp., *Veillonella* spp., *Streptococcus* spp., *Staphylococcus* spp., *Gemella* spp., *Porphyromonas* spp., *Bacteroides* spp., *Leptrochia* spp., *Mycoplasma* spp., *Ureaplasma* spp., *Escherchia/Shigella* spp, and *Ruminococcaceae* spp. [193]–[197]. More recent larger scale studies suggest the likely existence of >5 CSTs. For example, a study of >13000 vaginal

samples by France et al. identified 13 different CSTs using a centroid classification algorithm called VALENCIA [198]. The additional CST subclassifications usually associate with the presence of *Bifidobacterium* or rarer *Lactobacillus* species [199], [200].

Not all the CSTs are thought to have the same impact on host's health and not all of them are found in women with equal prevalence. While CST I, III and IV are common, CST II and CST V are rare, making their significance for health less clear [201]. Interspecies functional differences could explain the different impact that different *Lactobacillus* dominated CSTs have on women's health. For example, *L. iners* is often found to be a community member in women suffering Bacterial Vaginosis (BV), a condition of bacterial overgrowth of the vagina that causes malodorous discharge, representing a risk factor for STDs or adverse pregnancy outcomes [202]–[204] [205]. In contrast, *L. crispatus* is rarely found in women with BV. Despite being related, there exists major functional differences between these species. For example, some strains of *L. crispatus* are able to directly metabolise glycogen, which is present in high concentrations in the vaginal epithelium. Its primary metabolic output is the D-isomer of lactic acid, which inhibits inflammatory pathways and the production of metalloproteinases in the niche. As a result of a much smaller genome, *L. iners* can only produce L-lactic acid and often requires co-colonising bacteria to produce cofactors and enzymes to support survival and growth [181] [206]. Consequently, this species is particularly susceptible to environmental changes and partially explains why it is often found to be part of a less stable transitional microbiota.

CST IV microbiota are characterised by *Lactobacillus* depletion, high bacterial diversity and load, and a higher vaginal pH (>4.5). The composition of a CST IV microbiota can be very variable. *Lactobacillus* depletion and the general presence of *Gardnerella, Prevotella* and *Atopobium* species are common features, but CST IV can also harbour high proportion of *Sneathia, Mobilincus* and even *L. iners* [194]. CST IV associated species often produce sialidases and fucosidases which degrade superficial mucus glycan and cytolysin to exploit nutrients from epithelial cells. The consequent damage of the mucus layer and damage of the vaginal epithelium activate pro-inflammatory pathways that link CST IV to symptomatic states and adverse health outcomes [207]. Numerous studies have associated CST IV with increased risk of contracting sexually transmitted infections (STIs) including HIV [208]–[214]. Similarly, a high-diversity microbiota has been linked to higher risk for human papilloma virus (HPV) infections and development and progression of consequent neoplastic lesions [215]–[217]. For these reasons, *L. iners* (CST III) and CST IV microbiota are often considered to be

"suboptimal microbiota" compositions. Nevertheless, it is important to note that not all women presenting a CST IV microbiota have BV symptoms and there is a significantly higher prevalence of these community compositions in cohorts of healthy, asymptomatic Black and Hispanic women compared to Asians and White ones [192], [218]. This indicates that host genetic factors may be partly responsible for influencing bacterial composition across populations [180], [219], [220].

More specifically studies conducted in the UK and Canada described a higher prevalence of BV related microbiotas in Afro-Caribbean and Aboriginal communities respectively while studies conducted in Spain and China found the same pattern in Gipsy and Tibetan communities [220]. The African population, especially in the sub-Saharan countries, present very low prevalence of *L. crispatus* and their microbiota are mainly dominated by *L. iners* or by a mixture of anaerobes [221].

Hygiene practices, cultural norms, socioeconomical status and quality of the health system are all contributing factors that should be considered when discussing the variability of vaginal microbial communities[15], [222], [223].

An important factor in promoting *Lactobacillus* spp. dominance of the vaginal niche is estrogen. Increased levels of estrogen leads to accumulation of glycogen in vaginal epithelial cells which can then be broken down into complex sugars that can be preferentially utilised by *Lactobacillus* species [180]. Consistent with this, both premenarchal and post-menopausal women, who have lower levels of circulating estrogens have reduced prevalence of *Lactobacillus* dominated microbiota [224], [225]. Hormonal variation of the menstrual cycle can also lead to differences in the vaginal microbiota composition (Figure 1.5). In the secretory phase of the menstrual cycle (higher estrogen and progesterone concentrations) there is more stability in terms of microbiota composition which correlates with high levels of steroid sex hormones [226], [227]. Similarly has also been observed that hormonal contraception is associated with a decrease in the risk of presenting BV, because it generates greater microbiota stability [220], [228].

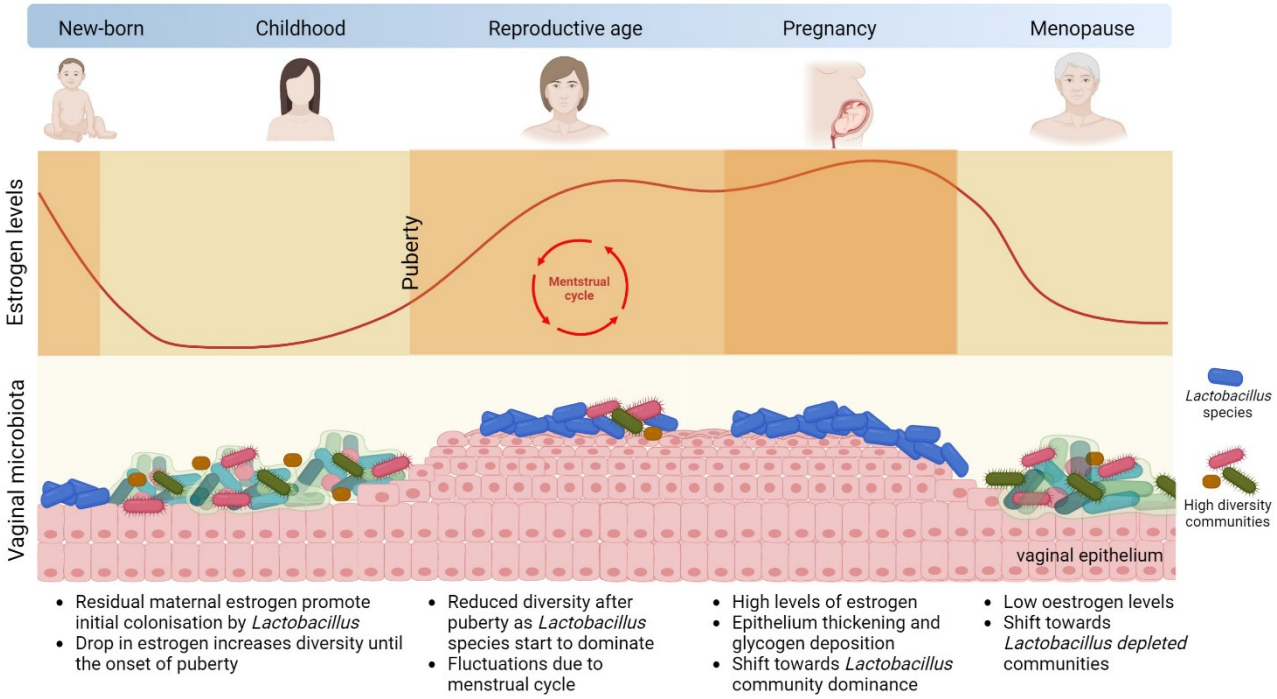


Figure 1.5 – Schematic representation of estrogens levels variation through different life phases and impact on microbial composition

1.5.2 Vaginal microbiota in pregnancy and PTB

Consistent with a role for promoting *Lactobacillus* species dominance in the vaginal niche, the estrogenic state of heathy pregnancy is associated with increased relative abundance of *Lactobacillus* species. In healthy pregnancy, bacterial communities have been reported to shift between *Lactobacillus* dominant CSTs but rarely towards high diverse communities [229]. In the post-partum period the rapid drop in circulating estrogen levels is coupled with a drastic increase in bacterial diversity in the vaginal environment [230]. These effect can last up to one year, offering a plausible explanation for the higher risk of preterm birth associated with short interpregnancy intervals [115], [230].

The association between vaginal ascending infection and PTB, has led to a deeper characterisation of the vaginal microbiota in high-risk pregnancies. Initial studies reported inconsistent results that were likely reflective of design limitations. Some of the studies lacked of geographic and ethnic variation in the study populations, had limited cohort sizes and

statistical power, or considered on the same level different PTB cases, including late PTBs (36–37 weeks of gestation) that are less likely to be caused by an infection [139], [230]–[232].

Ethnic variation of the study population is particularly relevant. Hyman et al. found that bacterial diversity was greater in women of Caucasian origin with PTB but that species diversity was generally higher among African Americans [232]. On the other hand, the predominantly African American cohort described in the work of Romero et al., found no differences in bacterial taxa abundance or CST composition between women delivering preterm or term [229]. Di Giulio et al. found a correlation between *Lactobacillus*-depleted vaginal CSTs and reduced gestational age at delivery in white women. They also showed that risk for PTB was greater in women with high abundances of *Gardnerella* or *Ureaplasma* [230]. The same group later conducted a study comparing White women in California and Black women in Alabama. Their result showed that whilst *Lactobacillus* depletion and greater abundance of Gardnerella were more common in Black women, it represented a risk factor for PTB only in White women. *Lactobacillus crispatus* however was found to be protective against PTB in both groups [233]. Finally, Stout et al. in a cohort of Black women, found that abundance of specific Lactobacillus species did not correlate to risk or protection from PTB, but a decrease in vaginal diversity was associated with PTB [234].

Generally, has been observed that on predominantly Caucasian populations exists a positive association between increased vaginal microbial diversity and risk of PTB with many studies identifying *G. vaginalis*, *Sneathia spp.*, *Prevotella spp.*, and *Mollicutes spp.* as microbial markers of risk [153], [233], [235]–[237]. Conversely, US studies performed in black women cohorts, have associated decreased vaginal diversity with increased risk of PTB [234], [238] or have shown no correlation between the vaginal microbiome and PTB risk [229].

Additionally, European studies have also shown how different *Lactobacillus spp.* are associated with different degrees of risk. In a study including women of European, Middle Eastern or Asiatic origin, an increased risk of PTB was associated with *Lactobacillus iners* and a protective effect was shown for *Lactobacillus crispatus* dominance [239]. Similarly, our group previously showed that in a UK population, *Lactobacillus iners* dominance at 16 weeks was significantly associated with both premature cervical shortening and PTB before 34 weeks [240]. In contrast, *Lactobacillus crispatus* dominance was highly predictive of term birth [240].

Additionally, an *L. iners* dominated microbiota in the same period is more likely to transition towards CST IV communities increasing the potential risk for an adverse pregnancy outcome

[241]. This demonstrates how vaginal microbiota composition in the early weeks of pregnancy might have important repercussions further ahead in the gestation.

There are several ways in which a highly diverse vaginal bacterial community can drive a preterm birth phenotype. Firstly, these communities are often enriched with potentially pathogenic bacteria. Recognition and response to these pathogens by the host innate and adaptive immune response leads to increased local inflammation [242]. This is mediated by pro-inflammatory transcriptions factors such as Nuclear factor kappa B (NF-κB) and activator protein 1 (AP-1). NF-κB and AP-1 are tissue-specific pro-inflammatory transcription factors that are responsible for the inflammatory activation in gestational tissues (myometrium and fetal membranes) before the onset of labour [27], [30], [92]. Upon activation, NF-kB and AP-1 induce the expression of pro-labour genes such as cyclooxygenase-2 (COX2), which promote prostaglandin synthesis, cervical ripening, and myometrial contractility [243]. Also they promote the synthesis of connexin-43 (CX43), a major gap-junction protein that promotes connectivity and contractility of uterine cells [244], metalloproteinases (e.g., MMPs-1, -3, -8, and -9) involved in fetal membrane and cervical remodelling [92], [245] and a number of other pro-inflammatory cytokines and chemokines such as IL8, IL1β, and TNFα [246]–[248].

An alternative mechanism through which vaginal microbiome may modulate upper reproductive tract inflammatory pathways, does not involve the direct colonization of the upper gestational tissues. It has been shown in fact how in mice, vaginal administration of bacterial lipopolysaccharide is sufficient to induce cervical remodelling, macrophage infiltration, and subsequent preterm birth [249].

Pathogen colonisation of the lower reproductive tract also often involves displacement of *Lactobacillus spp.* leading to reduced production of anti-inflammatory lactic acid, with a consequent removal of its immunomodulator effect. It has been shown for example how treating human cervicovaginal epithelial cells cultured in transwells with lactic acid under physiological conditions, causes a great increase in the production of the anti-inflammatory cytokine IL-1RA [250], [251]. A study also showed how D-lactic acid has a role in modulating the ability of L-lactic acid to induce the production of the extracellular matrix metalloproteinase inducer (EMMPRIN) from vaginal epithelial cells [252]. EMMPRIN has two functions. Firstly, it is a crucial cofactor for monocarboxylate transporter-1 (MCT-1) facilitating the transport of lactate anions out of the cell. This prevents the cell death from increased cytosolic acidity [23]. Secondly, EMMPRIN also induces the production of matrix metalloproteinase-8

(MMP-8) that is implicated in degrading the cervical mucus plug. Therefore EMMPRIN upregulation due to a change in the ratio of D- and L- lactate isomers, might eventually lead to passage of bacteria from the lower to the upper reproductive tract, with consequent infection and PTB in pregnant women [253].

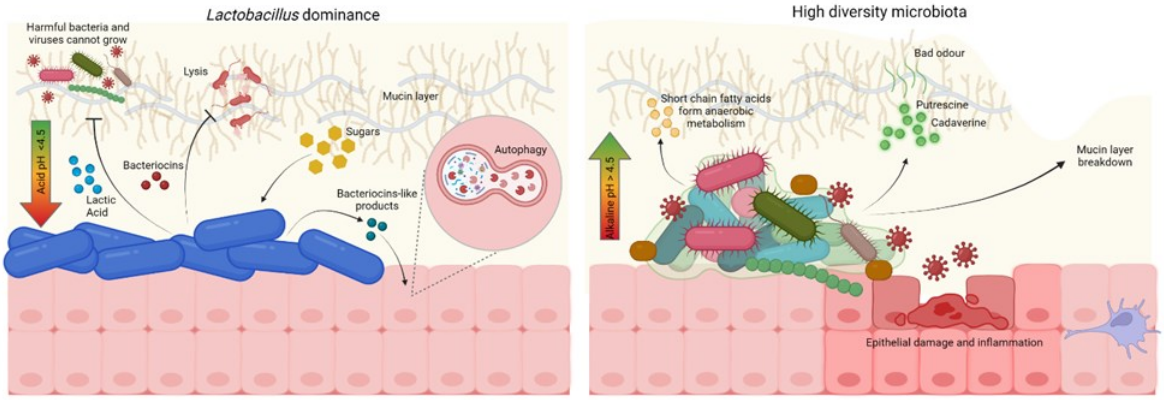


Figure 1.6 – Schematic representation of vaginal mucosa in a scenario of *Lactobacillus* dominance vs *Lactobacillus* depletion

1.6 Therapeutic strategies for modification of microbiota communities

As mentioned in section 1.4, a common intervention offered to women to lower inflammation and reduce the risk of PTB is progesterone. Despite its anti-inflammatory properties have been demonstrated *in-vitro*, progesterone does not seem to have an effect in modulating the vaginal microbial composition [213], [240]. Similarly, the attempt of modifying the vaginal microbiota composition using antibiotics has often proved ineffective if not detrimental. [235], [254]. Given the unsuccess of these interventions in regulating high risk vaginal microbial communities, the focus of research has now moved towards the concept of "positive modulation" [242]. Recent studies have investigated the use of live bio-therapeutics and probiotics in non-pregnant women demonstrating effective colonisation of *Lactobacillus* species, and reduction of recurrent BV [242], [255], [256].

The rationale behind the use of prebiotics and probiotics in the prevention of PTB relies on the assumption that by modulating the microbial composition of the vaginal environment it would be possible to consequently modulate the immunological response of the host and reduce the risk of preterm labour. Both prebiotics and probiotics have been investigated in this context.

The term prebiotic was first defined during 1990s [257] in the context of gut microbiota studies as a ‘‘non-digestible food ingredient that beneficially affects the host by selectively stimulating the growth and/or activity of one or a limited number of bacteria already resident in the colon’’. Can be considered prebiotics dietary fibres, resistant starch, and oligosaccharides [258]. The mechanisms of action of these prebiotics depend on the preferential metabolism of carbohydrate from *Lactobacillus spp.* compared to pathogenic ones. *Lactobacillus spp.* in fact metabolise oligosaccharides more readily than pathogens and consequently they grow faster on this substrate. This leads to an increase concentration of lactic acid and bacteriocins in the vaginal environment that further hampers the growth of pathogenic species [259]. Several oligosaccharides have been used in the form of pessaries to stimulate the growth of specific vaginal bacteria species in the vagina [260], [261].

Probiotics are defined as “live microorganisms that, when administered in adequate amounts, confer a health benefit on the host” [262]. From a regulatory point of view this definition covers both food supplements and products used as drugs. Since in commonly used terms probiotic is generally referred to food supplements, a specific term for a medicinal product was deemed necessary by FDA and WHO. For this reason in 2010 FDA proposed to consider a drug status for products containing live microorganisms that were aimed at the cure of prevention of diseases or conditions in human beings [263]. In 2012 an official FDA guideline was published creating for the first time the Live Biotherapeutic category [264]. Therefore, while probiotics are products that aim at maintaining or enhance a healthy state in the population that uses them, live biotherapeutics are considered drugs and are aimed at curing or preventing a disease.

Live biotherapeutics have been extensively studied in relation to their benefits in treatment or prevention of numerous diseases including enterocolitis, diarrhoea, and recurrent urinary tract infections [265]–[268]. More recently the interest in the use of live biotherapeutics grew in the PTB field [258], [269], [270]. The species most widely used as live biotherapeutics are *Lactobacillus spp.* and *Bifidobacterium spp*. Similarly, to what has been observed for the prebiotics, the mechanism of action of live biotherapeutics also involves a preferential growth of *Lactobacillus* species. An increased amount of *Lactobacilli* in the area determines an increased concentration of antimicrobial compounds, a competition for resources and an increased production of lactic acid that maintains the pH of the vagina at an inhospitable level for pathogens [182].

The main modes of administration of live biotherapeutics investigated in reproductive health studies have been the oral and the vaginal route. Several authors consider that oral

administration might be more effective, reduce the risk of pathogen ascension from the perineum area, and also considered less invasive by patients during pregnancy [258]. For oral administration, 10^9–10^{11} colony-forming units (cfu) are believed to be the necessary minimal dose of *Lactobacilli* for passage through the intestine and arrival to the vaginal environment [271], [272]. Once in the vaginal environment, the colonization can be influenced by many factor such as glycogen level, vaginal washing or use of antibiotics [273].
Studies conducted in humans reported that the use of *Lactobacilli* products reduce the recurrence of BV and increase the presence of *Lactobacillus spp.* in the vagina when assumed orally [271], [274]. Oral live biotherapeutics have also been used during pregnancy. However, a recent randomised clinical trial reported no effect of *Lactobacillus* and *Bifidobacterium* use in extending the gestational age [275].

One of the main disadvantages of using oral administration for live biotherapeutics products, is that bacteria need to survive the low pH levels of the upper gastrointestinal tract before they can reach the vaginal environment [276]. To treat vulvovaginal infections and BV, live biotherapeutics have been administered vaginally in several studies, in order to control the recolonization of *Lactobacillus spp.* without concerns about the bacterial survival rate [277]–[281]. However, to date, no study has yet investigated the use of vaginally administered live biotherapeutics in pregnant women. Our group recently submitted a clinical trial called FLIP-1 [282]. In this study women at high risk of PTB are offered a supplementation of LACTIN-V administered vaginally with a specifically designed applicator. LACTIN-V is the commercial name of *L. crispatus* CTV-05. The main objective of this study is to determine whether the vaginal live biotherapeutics is associated with any change in the colonisation pattern of the women involved [282].

1.7 Identification of vaginal microbiota

1.7.1 Conventional methods

Historically, identification and classification of vaginal bacteria have relied upon a combination of microscopy and culture-based techniques. Clinically, vaginal diagnostics still relies on assessment of this type of information combined with symptomatology. For example, the Amsel criteria is widely used in clinical settings as a diagnostic tool for BV [283]. Originally published in 1983 it consists of four parameters: i) presence of white-coloured discharge, ii) a pH higher than 4.5, iii) the presence of clue cells (epithelial cells covered in bacteria) detected using wet-mount microscopy, and iv) a fishy odour when 10% potassium hydroxide solution is added to the sample (the so-called "whiff test") [284] [285]. When three out of four criteria are present, the diagnosis of BV is made.

In laboratory-based research settings, Nugent scoring is considered gold standard due to superior reproducibility [283]. The Nugent score system was first described in 1991 and is based upon Gram staining and microscopy assessment of vaginal smears [284]. Gram staining is a technique that uses a dye commonly known as crystal violet to distinguish between "gram-positive" and "gram-negative" bacteria. Different bacteria have different cell wall structures that retain the dye differently giving the cell a darker or lighter appearance when observed through the microscope. Gram-positive bacteria usually have thick cell walls dense with proteoglycans that trap the dye and appear dark blue under the microscope. Gram-negative bacteria have more complex membrane systems that can release the dye more easily and therefore appear pink [286]. Additionally, Nugent scoring involves assessment of the shape and size of the bacteria including the presence or absence of large gram-positive rods (*Lactobacillus* type), small gram-variable rods (*Gardnerella* type) and curved gram-negative rods (*Mobiluncus* type). A score for each morphotype is assigned and finally a comparative scoring system is used to determine diagnosis [287]. Comparative studies showed that Amsel's criteria have higher specificity, but lower sensitivity compared to Nugent scoring [288], [289]. Regardless, both diagnostic approaches are high subjective and provide a very limited insight into host response, which is a critical consideration of pathology and disease risk.

Culture and microscopy are still key methods used in clinical microbiology settings given their affordability and wide availability. Further, culture provides effective and cheap means for studying therapeutical resistance or pathogenetic characteristics [290]. However, it is estimated

that 80% of the microbes that colonise the human body are unable to be cultured *ex vivo* [271]. Moreover, the highly specific nutritional requirements of many species means that culture ultimately biases the positive culture of some limiting the broad assessment of complex microbial communities. This has led to the development of culture-independent, molecular-based methods that often use direct assessment of bacterial DNA and other molecular markers as discussed below.

1.7.2 Molecular-based methods for bacterial identification and characterisation of community composition

Over the last two decades, next-generation sequencing (NGS) has become the major molecular-based approach used for microbiota assessment. After extracting DNA from a sample there are two major NGS-based approaches that can be taken. The first of these is called 'metataxonomics' and involves the amplification and sequencing of specific genes that act as phylogenetic markers [291]. For the study of bacteria, the major target is the hypervariable regions of the 16S rRNA gene. This gene has been widely used for several reasons. The 16S rRNA gene is present in all bacteria and its function has remained unaltered for a long period of time, indicating that changes in the sequence are attributable to evolutionary differences between species [292], [293]. An important characteristic of the 16S rRNA encoding bacterial gene is the presence of nine "hypervariable regions" (V1-V9). As the name implies, these sequences show high diversity among diverse bacteria, thus enabling their use for species identification [294]. However, current high-throughput sequencing platforms are limited to the analysis of short, partials sequence amplicons, and this sequence diversity also means that no single hypervariable region can alone differentiate among all bacteria species [295], [296]. For this reason, different hypervariable regions are often amplified together for sequencing. In the context of the vaginal microbiota, the V1 and V3 regions are often targeted for their ability to robustly classify of the most common vaginal *Lactobacillus* species (*L. crispatus, L. iners, L. gasseri,* and *L. jensenii*) [297]. Despite this advantage, the commonly used universal forward (27F) primer has mismatches against the *G. vaginalis* 16S rRNA gene sequence. Given the importance of *G. vaginalis* as a key bacterium associated with high diversity communities, it is essential to prevent potential underrepresentation of its sequences in the sampled data. A proposed solution for this issue is the use of a mixed formulation of the 27F forward primer (27F-YM), which retains the rRNA gene ratio of *Lactobacillus spp*. to *Gardnerella spp.* [295].

Primers targeting the V6 region have also been used to characterised vaginal microbiota however this region has known bias against *Sneathia*, *Leptotrichia*, *Ureaplasma* and *Mycoplasma* species [298].

One of the main limitations of metataxonomics is that heavily relies in PCR. PCR can suffer from bias introduced by differing amplification efficiencies between templates. This can often compromise the correct evaluation of a community structure [299].

The second major NGS-based approach to characterisation of microbiota is whole genome sequencing or 'metagenomics'. Metagenomics consists in the shotgun sequencing of all DNA extracted from a sample. It requires the bioinformatic assembly of microbial genomes, mapped against reference genomes [300]. Metagenomics avoids PCR bias, and it is not restricted to only bacterial sequences. In addition, the coverage of the whole genome beyond the 16S rRNA gene region means that specific, strain level discrimination is achievable [301]. Additionally, the reads can be accurately mapped to multiple known and annotated reference genomes providing further information about genes, proteins and putative function [302]. However, there are several drawbacks of this approach. Although reducing year-on-year, metagenomics is still comparatively expensive compared to metataxonomics. It also requires substantially more informatic expertise and computational expense and when used for assessment of human vaginal samples, a large amount (around 90%) of the generated sequence data are from the human host [283].

The recent advancements in molecular-based and culture-independent techniques, have greatly improved our understanding of the vaginal microbial composition. However, both metataxonomics and metagenomics approaches focus primarily on the identification and characterization of bacterial communities and fail to fully capture the functional interactions between the microbiota and the host, that ultimately define the health or disease phenotype. To achieve a comprehensive understanding of the vaginal microbiota and its implications for the female reproductive health is critical to integrate information from molecular-based techniques with biochemical and immunological data.

1.8 Constituents of the cervicovaginal mucosa

The bacterial composition of the vaginal environment has a fundamental role in shaping the chemical composition of the cervicovaginal fluid both in health and disease [303]–[305]. The cervicovaginal fluid consists of products derived both from the host (e.g., glandular secretions, mucus produced by cervix) and resident microbes. Consequently, a different microbial composition results in a different chemical fingerprint on the vaginal fluid. Commensal microbes produce metabolites that interact with host cells to help establish a mutual relationship. These metabolites contribute to the highly complex mucosal environment of the CVF. CVF consists of several components: secretions from various glands located in the vulvar area including sweat, sebaceous, Bartolini's glands and Skene's glands; plasma transudate from the vaginal wall; exfoliated cells; bacterial products; cervical mucus; endometrial and oviducts fluids and secretion from vaginal immune cells [306], [307]. From a biochemical perspective CVF is composed for 92–95% of water, ions and 5–8% solid matter [308]. The solid fraction includes mucins (high-molecular weight glycoproteins), proteoglycans and lipids. Mucins are constituted by 20% of protein and 80% of sugars and are responsible for the viscosity and elasticity of mucus [308].

CVF possesses several characteristics that make it a good candidate for the identification diagnostic or prognostic biomarker, outperforming other bodily fluids [309]–[311]. For example, the collection of CVF is usually simple and painless. CVF is usually collected through cervicovaginal washing, swabs, tampons, sponges or diaphragm-like devices (e.g. soft cups) [311]. Compared to blood, that is present in the entire body, CVF is proximal and limited to the reproductive tract offering the advantage of biomarkers more specific for gynaecological conditions [312]. Additionally, due to the relatively small volume/organ ratio of CVF, potential biomarkers are less diluted, thus increasing their sensitivity[312], [313].

However, secretions of the genital tract are greatly influenced by levels of estrogens and progesterone resulting in a substantial variability in concentration of a biomarker. For this reason, is practically impossible to determine a "normal" amount of a specific component. This makes sometimes difficult to distinguish between real biomarkers and natural occurring variations [310]. Apart from hormonal variation also several external factors can influence the vaginal metabolome. For example has been proven how breakdown products of several drugs such as painkillers, cocaine and antidepressants but also nicotine degradation metabolites can

be found in vaginal fluids [314]. In this chapter, however, the focus will be mainly on the bacterial metabolism and its role in shaping the metabolome.

1.8.1 Carbohydrates

In the vaginal environment carbohydrates can be found both in the mucus produced by the cervix and in the form of glycogen accumulated in the epithelial cells [315]–[317]. As described earlier, the accumulation of glycogen in the epithelial cells is estrogen-dependent [318]. The mechanism through which glycogen transfers in the vaginal canal is still unclear but it is hypothesised that this happens through the squamous epithelium turnover mechanism [319]. The vaginal epithelium is composed of four layers: the basal layer, the suprabasal layer, the upper intermediate layer, and the superficial layer. The upper intermediate layer contains large, rounded cells with cytoplasmic stores of glycogen [320], [321]. It is believed that the shedding of glycogen-containing epithelial cells results in the release of free glycogen in the cervicovaginal fluid [322].

Once released into the vaginal canal, most species of *Lactobacilli* are not capable of directly metabolising it and instead rely on host production of α-amylase, which breaks down glycogen in smaller polymers: maltose, maltotriose, maltopentaose and maltodextrins [323]. In the anaerobic condition of the vaginal environment, *Lactobacilli* preferentially convert these smaller polymers into pyruvate via glycolysis that is later converted in L or D-lactate by L- or D-lactate dehydrogenase [324].

Differences exist between sugar metabolism among different species of *Lactobacillus*. Both *L. crispatus* and *L. iners* are able to metabolise glucose, mannose, maltose, and trehalose, but only *L. crispatus* has the necessary genetic apparatus to ferment lactose, galactose, sucrose and fructose [202]. On the other hand, *L. iners* is able to use glycerol as a carbon source for glycolysis [325].

Mucins within the cervicovaginal mucus are also an important source of carbohydrates due to complex sugar structures that “decorate” their surface. Mucins are a family of glycoproteins that combine a long protein backbone with O-linked glycans terminated by a variety of sugar units [326]. The removal of these terminal residues, typically sialic acid or fucose, is the first step of mucin degradation and metabolism. The enzymes responsible for these processes are sialidase and fucosidase, respectively. A study by France et al., recently investigated sialidase

and fucosidase enzymes expression in different CST types. They showed that the expression of both enzymes is higher in CST IV communities and lower in *Lactobacillus* dominated ones. This highlights mucin metabolism as one mechanism by which BV-associated taxa can thrive within the niche with sialidases and fucosidases expressed by species such as *G.vaginalis, Prevotella spp., B.breve, P.amnii, P.bivia, P.timonensis and Propionibacterium* [259]. Metabolic profiling studies of women with BV and healthy controls have also shown that concentrations of maltose, maltotriose, maltopentaose and maltodextrins are positively correlated to relative abundances of *L. crispatus* and *L. jensenii* and negatively correlated in women diagnosed with BV [327]. Women affected by BV are also characterised by lower levels of simple sugars, alcohols and metabolites derived from the glycerol metabolism [328].

Differences in carbohydrate metabolism result in the production of different metabolic end products. Highly diverse communities are usually associated with lower levels of lactate but some CST IV associated bacteria, facultatively or strictly anaerobic, can produce lactate. These include *Atopobium, Streptococcus, Staphylococcus, Megashara and Leptotrichia* [329]. This may explain how, despite differences in microbial composition, some non-*Lactobacillus* vaginal communities maintain their ecological function and are well tolerated by the host [329]. The antimicrobial function of lactate appears to be strongly related to vaginal pH. Lactic acid is a weak acid (pKa = 3.89), indicating that it only partially dissociates in water to form lactate and a proton (H+) [186]. This means that when at pH above 4, lactate exists predominantly in its anion form [324]. However, when at a pH <3.9, most lactate will be in the protonated form which is permeable to microbial membranes and can therefore enter the microbial cytosol. The consequent acidification of the microbial cytosol leads to impaired cell functioning and cell death [185]. The equilibrium between lactate producing species in the vaginal mucosal environment is therefore important in the maintenance of vaginal health.

1.8.2 Short Chain Fatty Acids (SCFAs)

Short-chain fatty acids (SCFAs) are fatty acids with less than 6 carbon atoms and are the main metabolic products of anaerobic bacterial fermentation [330]. Accordingly, their detection in the vaginal environment is mainly linked with BV-associated species [331]. The SCFAs measured in the cervicovaginal fluids of BV patients include acetate, propionate, butyrate, succinate, formate, valerate and caproate [306], [332]–[336]. A few examples of BV-associated

species that produce organic acids include: *Peptococcus*, *Dialister*, *Gardnerella vaginalis*, *Bacteroides*, gram-positive cocci, and *Clostridium* [334], [337]–[339].

In comparison with lactic acid, acetic, propionic, butyric, and succinic acid are weaker, with pKa values of 4.76, 4.82, 4.87, and 4.16–5.61, respectively. When in vaginal conditions with a pH < 4.5, they exist in their protonated form and exhibit similar permeability properties to lactic acid. However, the amounts in which they are usually observed in these conditions are negligible and therefore they do not contribute to lowering the pH in a significative way [186]. Nevertheless, SCFAs have immunomodulatory functions. The treatment of vaginal epithelial cells with SCFAs in conditions that simulate dysbiosis (higher pH), has shown an upregulation of pro-inflammatory cytokines such as TNFα and IL-1β but a downregulation of IL-6 and IL-8 [340]. The net effect on the immune system is still unknown [324].

1.8.3 Lipids

Lipids have an important role in the maintenance of vaginal homeostasis and their profile can be significantly altered during pathological processes. A study conducted on BV diagnosed women found higher levels of 12-HETE, a signalling eicosanoid and known biomarker for inflammation, as well as lower levels of its precursor, arachidonate suggesting a conversion of arachidonate to 12-HETE by BV-associated bacteria [327]. Carnitine, and acyl-carnitines such as acetylcarnitine, propionylcarnitine, and butyrylcarnitine are found in lower levels in BV [327]. Carnitine is a product of lysine or methionine degradation and is involved in transport of long-chain fatty acids.

A study on vulvovaginal candidiasis and cytolytic vaginosis also distinguished different lipidomic profile related to the pathological status. In women with vulvovaginal candidiasis, higher concentrations of lipids related to inflammation and oxidative stress were reported. In women suffering from cytolytic vaginosis the lipids observed were linked to cellular apoptosis and in control women the lipids detected are known to be involved in the maintenance of epithelial integrity [341].

1.8.4 Amino acids and Biogenic Amines

Biogenic amines are organic compounds with one or more amine (NH2) group(s), and they are the products of amino acid metabolism in the vagina. More specifically their synthesis involves

the decarboxylation of precursor amino acids [342]. The biogenic amines most commonly found in the cervicovaginal fluids are putrescine, cadaverine, spermidine and spermine. Putrescine and cadaverine have two amino groups each, while spermidine and spermine contain three and four amino groups respectively [343]. Even though host cells can produce biogenic amines which contribute to immune regulation, lipid metabolism, nucleic acid stabilisation and cell division, certain biogenic amines are exclusively of microbial origin [344]. These compounds are responsible for the fishy odour associated with BV [327], [331], [345], [346] and their presence suggests that BV-associated bacteria use amino acids as their preferred carbon and nitrogen sources [327]. Consistent with this, BV metabolic profiles are characterised by lower levels of arginine, which is a precursor for the synthesis of putrescine [327]. BV is also characterised by reduced levels of amino acids in vaginal secretions including leucine, isoleucine, alanine, valine tyrosine and tryptophan [331].[340].

1.9 Assessment of microbiota-host interactions using metabolic profiling

Consistent with the above described impact of metabolism, several recent metabolomics studies have demonstrated a clear relationship between microbial composition and the metabolic signature of cervicovaginal fluids [303], [304], [327], [331], [346]–[348]. Metabolomics is the study of small molecules (typically with molecular masses less than 20’000 Da), within cells, biofluids or tissues and includes not only metabolites but also lipids, sugars, amino acids, fatty acids, phenolic compounds, and alkaloids. Metabolomics is often considered the “end point of the ‘omics’ cascade” as it provides the closest representation to the phenotype of the studied system [349]. One of the biggest challenges of this discipline is dealing with the enormous complexity and variability of metabolites. Metabolites belong to a wide variety of compound classes and that differ in their physical and chemical properties. For this reason, no single analytical instrument can simultaneously detect all metabolites present in a given sample. Currently, combinations of the most thorough analytical methods can identify and quantify around 8,000 of the 114,100 metabolites estimated to be present in humans [350]. The most widely used analytical platforms include nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) [351]–[353]. This thesis will focus on Mass Spectrometry methods. Mass spectrometry has the possibility to be coupled with separation techniques such as gas chromatography (GC), or liquid chromatography (LC) to enhance metabolic resolution. A description of the advantages and disadvantages of these analytical techniques is provided below.

1.9.1 Mass Spectrometry (MS)

Mass spectrometry is an analytical technique that studies ionised molecules in the gas phase with the aim of molecular weight determination. The ions are measured as their mass-to-charge ratio (m/z) and relative abundance [356]. Using this technique is possible that multiple molecular formulas match the measured molecular weight. In these cases, tandem MS (MS/MS) can be applied in order to confirm molecular identification. MS/MS is based on a fragmentation principle, where the molecular ions are dissociated in smaller pieces through the impact with an inert gas (usually nitrogen or helium). The first stage of MS/MS (MS1) allows the separation of molecular ions based on their m/z ratio. In the second stage (MS2) ions of a particular m/z ratio are selected and fragmented through collision induced dissociation (CID). During the collision, some of the kinetic energy released is then converted into internal energy that causes the fracture of chemical bonds. These fragmentations depend on the chemical structure of the parental molecule and result in interpretable and identifiable patterns. The smaller product ions are finally detected by the mass spectrometer [357].

Both targeted profiling (where metabolites to be assayed are selected *a priori*) and untargeted profiling can be carried out using MS-based metabolite screening. One advantage of MS techniques is that they provide excellent quantitative performance at low concentrations for many biomarkers.

In biological samples, sample preparation is a crucial step for the success of the analysis. The theoretical aim of sample preparation is to enrich the sample of metabolites of interest and remove interfering substances. Sample preparation in metabolomics generally consists of sampling a representative portion of the sample, stabilizing the sample (usually freeze-drying), and extracting it [358]. Depending on the analysis done, filtration, centrifugation, pre-concentration and derivatization steps may also be required. It is important to note how all sample preparation procedure, independently of how accurate, will cause some analyte loss [349]. Sample preparation is often a complex step in the analysis that requires a lot of expertise and can be costly both economically and in terms of time. Additionally, for complex mixtures as biological samples are, it is often necessary to separate molecules prior to the introduction into the mass spectrometer [356]. For example, if direct injection of the sample is used, chemical isomers cannot be distinguished because they have the same exact mass. A chromatographic separation, usually liquid or gas chromatography (LC-MS or GC-MS), added before the sample injection, delivers additional information on the physical and chemical

properties of the metabolite separating them, and reduces the problem of ion suppression [349]. Ion suppression is an effect of competition for ionisation efficiency between two or more analytes or one analyte and another endogenous or exogenous compound [359]. This can interfere with a meaningful data analysis [360]. Liquid chromatography can reduce ion suppression caused by coeluting compounds, isobaric interferences and often can separate isomers. The separation obtained will also result in better detection limits and MS data quality due to reduced background noise. Likewise, the use of gas chromatography provides good resolution, analyte-specific detection, and quantification of metabolites, but requires a sufficient vapor pressure and thermal stability of the analytes. Hence, the analysis of polar metabolites usually requires derivatization at the functional group to reduce polarity and increase thermal stability and volatility, increasing the overall complexity of sample preparation [349].

Apart from the inlet system (LC, GC, direct injection, etc.), the mass spectrometer consists of an ion source, mass analyser and detector. The choice of different components can alter the result of the analysis: depending on the sample inlet system or ionisation technique used, for example, a specific metabolite class may be enriched or depleted leading to biased representations [361]. Choosing the right combination of inlet system, ion sources and analyser is important for the desired result. For clinical applications, faster and simpler methodologies are preferred, especially in terms of simpler sample preparation procedures and rapid results [362].

1.9.2 MS ion sources

To analyse molecules by mass spectrometry, they must be ionised first. This can be achieved by variety of different ionisation methods, which can be classified as “hard” or “soft” according to the level of energy applied in ion generation [363]. Hard ionisation processes, which employ high levels of energy, produce characteristic mass spectra made up of smaller fragments of the parent molecule. Softer ion sources are more likely to produce ions of the molecular species, with simpler fragmentation spectra [357]. Examples of ion sources in the metabolomic field include electron impact ionisation (EI), electrospray ionisation (ESI), and matrix-assisted laser desorption ionisation (MALDI).

EI is a hard ionisation technique where ions are produced by collision of sample molecules in gaseous form with high energy electrons emitted by an electron source [364]. EI data is highly

reproducible between different instruments, however, the analytes most suited for EI analysis are volatile and thermally stable ones. In addition, being usually coupled with GC, EI requires complex sample preparation procedures and the mass spectra produced are dominated by fragment ion peaks while the precursor ion is present in low intensity or completely absent [365].

ESI is a soft ionisation technique, suitable to study intact metabolites, proteins, or even macro-molecular complexes. For this, an electric field is applied to the solution being analysed as it passes through a capillary. The electric field causes a charge accumulation on the liquid surface at the end of the capillary. This charging produces a Taylor cone, which breaks down to release highly charged droplets [366]. The solvent in these droplets is evaporated with heated inert gas and the charges get concentrated on the droplet surface. When surface charge repulsion exceeds surface tension, droplets break up into smaller droplets. This process, known also as Coulomb explosion, produces ionised form of the molecule [367]. ESI produces protonated and deprotonated molecules in positive and negative mode. The technique allows ionisation of a wide range of molecules directly from the liquid phase including small molecules and lipids as well as large molecules such as peptides and proteins [365].

MALDI is another popular soft ionisation technique used in metabolomics. For this technique, the sample needs to be mixed with an organic matrix and then irradiated with laser pulses. After the laser pulse, the irradiated spot is rapidly heated, and molecules are released from the surface. The matrix molecules ablated from the surface of the sample, absorb the laser energy and carry the analyte molecules into the gas phase as well. During the ablation process, the analyte molecules are usually ionized by being protonated or deprotonated with the nearby matrix molecules. MALDI most commonly produces ionised molecules that a single positive charge [365]. One of the drawbacks of MALDI, especially when analysing small molecules, is the occurrence of signal suppression effect. Signal suppression is usually caused by the presence of ion suppressing molecules in the matrix used such as salts, alkali-metals and lipids that compete with the analyte for ionization [368].

A schematic representation of ESI and MALDI ion sources is depicted in Figure 1.7.

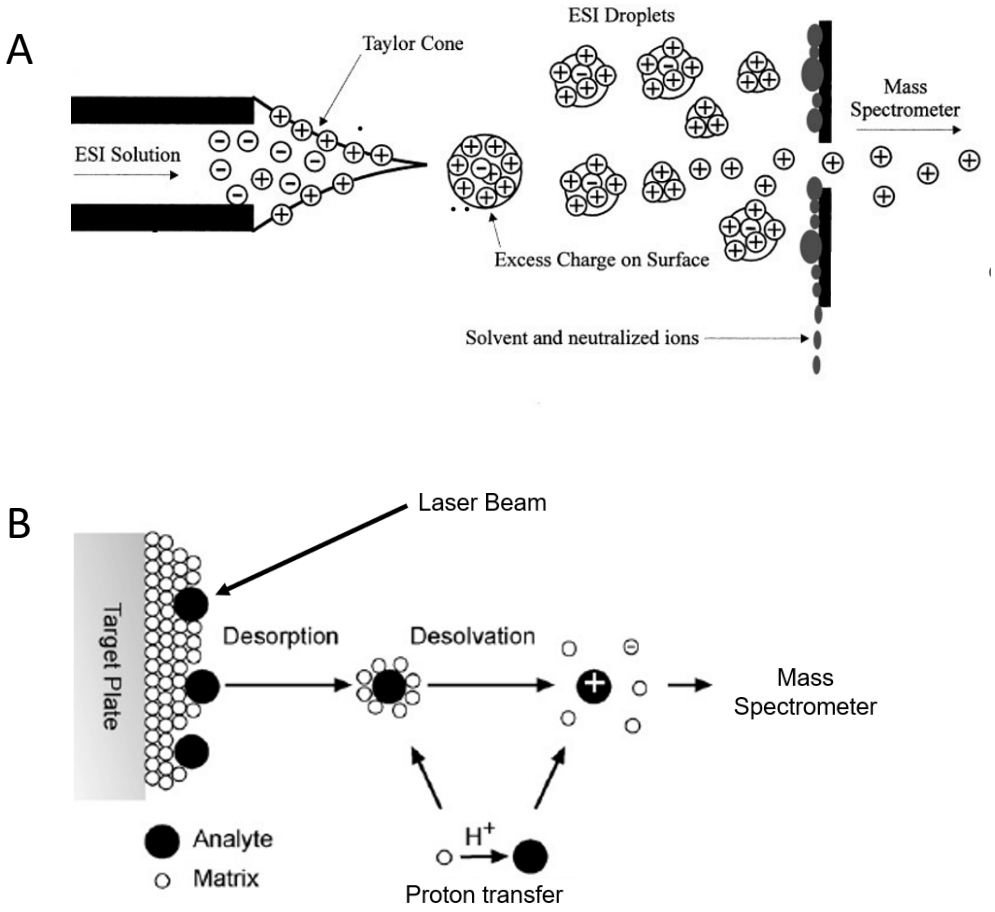


Figure 1.7 - Schematic representation of main soft ionisation sources. (A) Electrospray (ESI) figure adapted from [369]. (B) Matrix-assisted laser desorption ionisation (MALDI) figure adapted from [370].

1.9.2.1 Mass analysers

The function of the mass analyser is to resolve ions based on their mass to charge ratio (*m/z*) before they hit a detector to produce a signal recorded in the spectrum. An accurate measurement of the *m/z* ratio is crucial for a robust metabolic profiling. The accuracy of a measurement is the accordance between the measured *m/z* value and the true theoretical *m/z* value of the exact mass of the detected ion [371]. Mass accuracy in mass spectrometry is usually expressed in parts per million (ppm).

Other important features of mass analysers are their resolution power, appropriate mass range for the aim, good sensitivity, and lower/upper linear dynamic ranges. Resolving power in mass spectrometry is defined as the ability of the instrument to distinguish between two peaks at m/z values differing by a small amount [372]. The mass range is the range of *m/z* values for which the analyser is capable to record a mass spectrum [373]. If quantitative measurements are required, it is important to consider the lower and upper linear dynamic range limit. The linear dynamic range is a range in which the ion signal is directly proportional to the analyte concentration. Inside these limits mass spectrometry can provide an indication of the relative abundances of analyte concentrations within the sample [373].

The most common types of analysers in use include quadrupole, ion trap, time-of-flight (TOF), and Fourier transform analysers (ion cyclotron and Orbitrap). The choice of one over another largely depends on the objectives of the metabolomic profiling project since all of them have their positives and negatives. Quadrupole mass analysers work as mass filters. They consist of four cylindrical rods disposed in parallel with each other. The opposite rods are electrically connected and both a radio frequency (RF) potential and a direct current (DC) are applied through them. The combination of RF and DC potential causes the ions entering the analyser to oscillate in trajectories that are dependent on their *m/z* ratio. Depending on combination of RF and DC applied, only ions of a particular m/z have stable trajectories and pass through the quadrupole. The remaining ions collide with the quadrupole rods and are filtered out. Ions of different *m/z* values can be analysed tuning the RF potential and the DC accordingly [374]. The main advantages of quadrupole analysers are in their dimension – suitable for a benchtop – their low cost and their reliability. However, they have limited mass range and resolution, especially when singly charged ions are present or complex mixtures of compounds with similar masses are analysed [375].

TOF analysers rely on the particularly simple concept that ions with different *m/z* ratios take different times to reach the detector. Ions enter a TOF analyser in pulses and are accelerated through an electrical potential on the “acceleration grid”. The accelerated ions have equal kinetic energy and are left to drift in a field free region. In this part of the TOF analyser ions are separated according to their masses and the time it takes for them to reach the detector can be readily converted in a *m/z* value [376]. TOF analysers have several strengths compared to other analysers. They have a virtually infinite mass range, allowing separation of ions with masses of few Daltons to well over 100 kDa, an accuracy of 5ppm, and fast scanning capabilities. These characteristics make TOF the analyser of choice when performing targeted

and non-targeted metabolomic analyses or when observing high mass biomolecules such as proteins [377]. Although new configurations have allowed TOF analysers to be interfaced ion sources that provide a continuous flux of ions, they have initially limited to the use with only pulsed ion sources such as MALDI [375].

Orbitrap analysers use a Fourier transform to convert a signal, produced by ions oscillating in a "trap", from the time domain to a frequency domain [378]. The trap is composed of three main parts, an inner spindle electrode covered by two concave electrodes facing each other. A voltage potential is applied between the inner and outer electrodes, creating a linear electric field. Ions are introduced tangentially into the orbitrap as a "packet" between the inner and outer electrodes. The electric field between the electrodes, makes the ions move towards the inner electrode while their tangential velocity creates an opposing centrifugal force. At a specific potential between the inner and outer electrodes, the ions remain in a spiral path around the inner electrode. However, due to the conical shape of the electrodes, a harmonic axial oscillation in the ions is induced. The axial harmonic frequencies are proportional to the m/z of the ions. The outer electrodes also act as receiver plates that detect the back-and-forth axial harmonic motion of the ions. This signal is then converted from the time domain to the frequency domain [365], [378]. One of the major advantages of the orbitrap analyser is its high mass resolution, excellent mass accuracy (0.5-5 ppm), a wide dynamic range and a mass range of up to 2000 Da thus making it suitable for metabolomic applications [379].

1.9.2.2 Ambient ionisation

Ambient ionisation refers to a group of techniques where desorption of the sample and ionisation are coupled. The invention of ambient ionisation allowed mass spectra to be recorded on samples in their original environment without sample preparation procedures or the need to create gas-phase ions outside the mass spectrometer [380], [381]. The elimination of extraction and separation procedure prior the ionisation increased greatly the simplicity and speed of the method. Ambient ionisation works for a wide range of analytes both highly polar/non-volatile and nonpolar/volatile from various surfaces and matrices. In addition, newer methods have enabled ionisation of the analytes directly from a sampling device or even the simultaneous sampling and analysis on a single device. These characteristics opened the application of ambient ionisation in a variety of different setting including forensics and diagnostics. For example spraying directly from medical swabs [382], is ideal for on-site emergency

toxicological screens, roadside drug testing or point-of-care (POC) diagnostic assays, and the non-invasive nature of swab sampling makes it easy for patients to comply in clinical settings.

Analytical techniques relying on ambient ionisation include: touch spray (TS) [382]–[384], paper spray (PS) [385], rapid evaporative ionisation mass spectrometry (REIMS) [386], probe electrospray ionisation (PESI) [387], coated blade spray (CBS) [388], atmospheric pressure solids analysis probe [389], wooden-tip electrospray [390] and desorption electrospray ionisation- mass spectrometry (DESI-MS). In recent years there have been many successful applications of different ambient ionisation techniques. TS uses a sampling device that consists in a metallic object with a sharp and rugged tip that also represents the ionisation substrate. Both a solvent and a high voltage are applied resulting in the field emission of analyte-containing charged droplets. The droplets undergo a similar process to ESI-MS and ions are formed. Touch spray has been proven particularly effective in screening for presence or absence of particular compounds with the potential for semiquantitative analyses [383], [384]. A study from Jarmusch et al. used TS for the detection of strep throat causing bacterium directly from medical swabs, observing lipids metabolic profile of samples taken from culture and from simulated throat swabs [384]. Pirro et al, similarly reported that TS can be used for direct drug analysis from oral fluid medical swabs [382].
PS ionization generates ions directly from samples absorbed on cellulose filter paper surfaces. PS could be used for direct analysis of blood, plasma, saliva, and urine making it particularly suitable for POC diagnostics. Similarly to TS, both a solvent and electrical potential are applied on the paper surface and generate a ESI-like ionisation process [391]. The most popular application of PS-MS is on the drug detection field where numerous studies highlighted the potential for quantitative use against the abuse or for therapeutic monitoring purposes [392]–[396]. Other applications range from agricultural [397] to forensic [398] to microbiological [399] proving the versatility of the technique.

Over the past few years different solutions were proposed to facilitate the transition of MS analysis towards *in vivo* conditions for clinical applications in surgery settings. The most important example of this transition is the iKnife [400].
The iKnife (Intelligent Knife) was invented in 2013 by Prof. Takats et co-workers [401]. It consists in the combination of an electrosurgical tool (the knife) and an MS system based on REIMS. The smoke produced during the tissues dissection in a surgical operation is considered to be extremely rich of biological information and in this system, it is collected by aspiration

and transmitted to a QTOF-MS instrument. The iKnife allows for direct monitoring in vivo from patient tissues guiding for example cancer surgery in real-time [400], [401].

1.9.2.2.1 Desorption Electrospray Ionisation (DESI)-MS

Desorption electrospray ionisation (DESI-MS) was the first form of ambient ionisation described and it remains the most widely used [381], [402]. DESI-MS is a soft ionisation technique and consists of a pneumatically assisted electrospray directing charged solvent onto the surface of interest. The impacting solvent droplets dissolve the analytes in the sample and prompt the formation of secondary charged droplets. The secondary droplets are subject to evaporation in the atmospheric pressure space of the mass spectrometer inlet capillary. The formation of the ions happens through an ESI-like process that frequently generates multiply charged ions and solvent adducts [403]. A schematic representation of the DESI-MS source is shown in Figure 1.8.

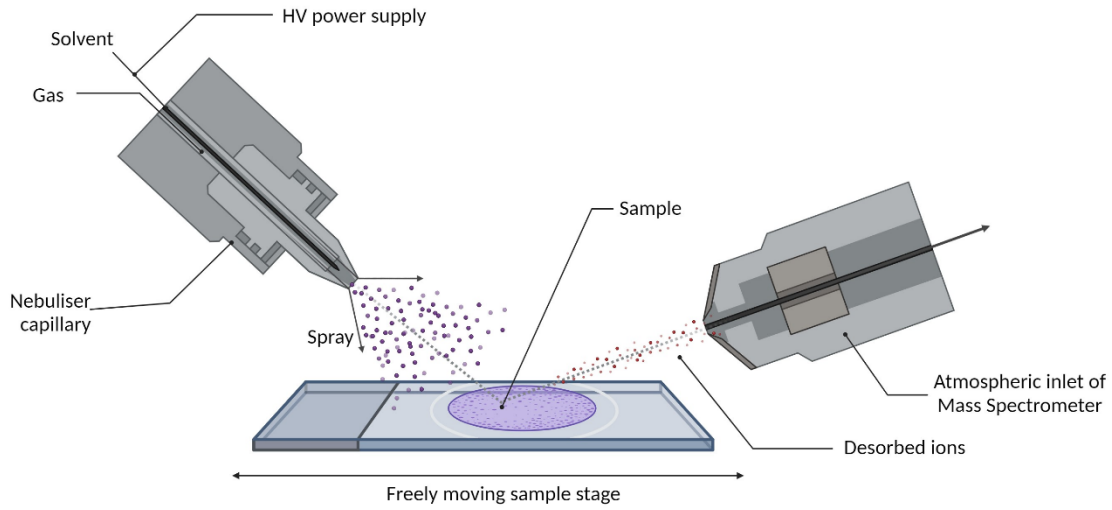


Figure 1.8 – Schematic representation of DESI-MS source. Figure adapted from [404]

DESI experiments can be performed on a wide range of compounds (polar and nonpolar, low and high molecular weight) present in a wide range of matrices and in various physical states. DESI-MS does not require complex sample preparation procedures and the acquisition is extremely rapid (< 3mins) [405]. However, the geometry of the DESI-MS setup is particularly important for a successful detection of the molecules of interest. Small molecules for example are preferably observed in positive ion mode spectra when an inlet-sample distance of 5mm is

maintained and the electrospray is positioned at 40° from the sample surface. Larger molecules like oligosaccharides or proteins require smaller distance from the inlet capillary and sharper impact angles of the solvent [380].

DESI-MS presents several advantageous characteristics. It is a versatile technique that has been applied to a wide range of surfaces including plastic, glass, paper and polymers [406], [407]. It is non-destructive and presents high throughput, sensitivity, specificity, discrimination power and a wide mass range from simple amino acids through drug molecules, to peptides and proteins [406]. For these reasons the fields where DESI found an application are various and include high throughput analysis, clinical diagnostics, food analysis and forensics. For example, DESI-MS has been successfully applied to dried blood spot analysis and has shown a good quantitative evaluation for the level of different [408]. Extensive work has been done using DESI-MS in the identification of cancerous tissues. Studies have been reported for breast cancer [409], [410], prostate cancer [383], [411]–[413], and brain cancer among others [414]–[417]. Most of these works rely on the detection of lipid profiles to discriminate tumour cells from healthy tissue. This is because tumour lesions cause a remarkable alteration of the lipid profile in the cancerous cells, but also because lipids are easily ionised by DESI-MS and become predominant in the spectrum [418]. Additionally, several studies have shown the effectiveness of DESI-MS in discriminating different purified bacterial sample at species level directly from glass slide or agar plates[405], [419], [420].

Nonetheless DESI-MS also presents several drawbacks. The main limitation is represented by the poor quantitative accuracy induced by matrix effect, that also compromises the efficiency of ionisation and droplet formation. Plus, the strict dependency of DESI-MS on a precise geometry impacts measurement automation and reproducibility. To overcome these setbacks modification to the original DESI-MS setup have been suggested such as a geometry independent DESI or the reactive DESI. With reactive DESI ionisation efficiency is improved by the use of derivatization agents in the sprayed solvent [406]. From the point of view of clinical application, however, another limitation of this technique relates to its structure. DESI-MS is usually coupled with bulky mass analysers which limit the potential as point of care diagnostic tool.

1.10 Data modelling approaches for metabolic profiling data

When the field of metabolomics began to emerge in the late 1990s, statistical analysis and data modelling methods were relatively straight forward and appropriate for the limited number of metabolites or mass spectral features that could be resolved using the available instrumentation [421]–[423]. However, the development of new techniques and instrumentation enabled the detection and measurement of thousands of different metabolites in a single experiment. High-throughput generation of high-dimensional datasets have led to a situation where the number of datapoints measured in an experiment far outweighs the number of observations. Consequently, the interpretation of the acquired data has become more complicated, necessitating the use of robust statistical tools to analyse the complex interactions between thousands of analytes [424].

After data processing, metabolomics measurements usually result in signal intensities across a wide dynamic range. Univariate and multivariate statistics are often used in combination to extract relevant information from these complex datasets [425]. Univariate models in metabolomics analyse one variable at a time, focusing on a small number of selected features out of the many measured. Examples include t-tests, Wilcoxon rank-sum test, Kruskal-Wallis tests for comparing means between two groups, and ANOVA for multiple groups [426], [427]. However, the limitation of univariate analysis is the inability to build predictive models using multiple metabolites simultaneously [428]. In contrast, multivariate approaches are more commonly used in metabolomics as they capture changes in single metabolites between groups and explore relationships within the entire dataset. Principal component analysis (PCA), partial least square regression (PLS) and Random Forest analysis (RF) are widely utilized multivariate techniques [421].

1.10.1 Multivariate modelling approaches

A common approach used in multivariate methods is the generation of new sets of variables that retain the largest possible amount of variance of the original dataset, while reducing the number of variables to consider [429]. Methods used to achieve this can be supervised and unsupervised, depending on whether previous knowledge of the dataset structure (e.g., different groups or class membership) is used to build a model.

PCA is an unsupervised method commonly used as the initial step in metabolomics data analysis [425]. It simplifies the description of the biological system by identifying linear combinations of the original variables known as "principal components." These components maximize the variance in the data while being uncorrelated (orthogonal) with each other. The first principal component explains the highest amount of variance, followed by subsequent components that explain additional variance while maintaining independence from the previous ones [430]. PCA helps reveal potential structure within metabolic profile data without prior knowledge or assumptions about the groups or classes present in the data. Additionally, the new dataset produced by a PCA is divided in a score matrix and a loading matrix. Imagining each sample as a point in a chemical space defined by the principal components, the score matrix contains the coordinates of each sample in the newly defined space. The position of each point in space can be interpreted as a relationship existing between samples characteristics. The closer two points are the more similarities they share in the metabolic profile. Likewise, loading matrices contain information about the linear combination of the original variables. Each variable is transformed in a score expressing its weight on the principal components. The use of PCA in the first stages of a study gives an informative overview of the general structure of the dataset and highlights obvious relationship among groups of samples. For this reason, PCA can be viewed as an observational tool rather than a dedicated model for investigating potential biomarkers in a metabolomics study[431].

Analysis of variance specific to groupings or classes of data (e.g., disease versus control), can often be achieved using supervised modelling approaches. The most widely used supervised model in metabolomics analysis is Partial Least Square regression (PLS). PLS is an extension of PCA where the components are calculated not only trying to maximise the general variance but also maximising the separation between different classes [432]. Similar to PCA, PLS creates a new subspace from the initial number of variables and aims at predicting the class membership of each sample based on a reduced number of factors [425]. PLS is often coupled with discriminant analysis (PLS-DA) in the case of categorical classification of the groups. The main advantage in using PLS-DA is its ability to handle high-dimensional metabolomic datasets that are characterised by inherently high noise levels and collinearity [433]. Additionally PLS-DA provides useful tools for result interpretation such as variable importance projection (VIP) and regression coefficients that can be used to identify the most relevant features responsible for group separation [434]. One drawback of PLS is that a large number of variables is usually necessary to achieve a good prediction accuracy and, in the cases

where number of variables significantly exceed the number of observations, the model is likely to be prone to "over-fitting" [435]. However, the use of appropriate cross-validation strategies can mitigate the overfitting problems of PLS-DA [436]. In cross validation the dataset is randomly split up in several groups. One of the groups is used as the test set and the rest are used as the training set. The model is trained on the training set and the performance evaluated on the model run with the test set. The stability of model parameters is then assessed by repeating this procedure until all the subsets of samples have been used at least once as the testing set [437].

In clinical metabolomics, large individual variability due to a variety of different factors such as demographic background, pathological conditions, or treatments can often be observed [438]. Random Forest (RF) is a versatile machine learning algorithm that can handle such complex data structures and nonlinear relationships. It presents several advantages such as simple theory, fast speed, stability with noisy data and little or no overfitting [439]. The RF algorithm was firstly introduced by Breiman et al. in 2001 [440]. It consists of a combination of decision trees classifiers fitted into distinct subset of samples using the bootstrap technique. Bootstrap is a resampling method where the entire dataset is randomly subsampled in smaller sets with a balanced number of samples. During the classification procedure each tree is fitted to a distinct subset leaving out the rest of the dataset that is subsequently used for prediction purposes. An additional level of randomness is added in the choice of the variable used to split nodes in a tree. Only a smaller subgroup of the available variables is used each time. These two procedures assure an internal validation.

In summary, the choice of a specific multivariate model depends on the specific objectives and characteristics of the data, and the interpretability requirements of the analysis.

1.11 Study rationale

PTB is the primary cause of death for children under 5 years old globally. Despite extensive research and many known risk factors, there is still a lack for reliable prediction, prevention and treatment strategies.

However, substantial evidence exists of a link between vaginal microbiome composition in pregnancy and PTB risk phenotypes. Moreover, changes in the vaginal microbiota composition and the corresponding host response, are reflected in the metabolic and chemical profile of the cervicovaginal mucosa.

Current approaches used to characterize the vaginal microbiota composition, are costly, time consuming, and lack insight into host immune and inflammatory response, which is a key aspect of the mechanisms driving PTB phenotype.

Recent advancements in mass spectrometry and in metabolic profiling methodologies, particularly through ambient ionisation mass spectrometry, made possible to obtain detailed metabolic profiling of clinical samples within minutes.

The information derived from cervicovaginal fluid metabolic profiling can potentially give new insights into both microbial composition and host response. This could have diagnostic application for PTB, help monitoring therapeutic strategies but also provide rapid mechanistic insight underlying the aetiology of preterm birth risk.

1.12 Hypothesis and Aims

Complex interactions between vaginal microbiota and host during pregnancy are reflected in the cervicovaginal metabolome. This information can be rapidly captured by direct swab analyses using DESI-MS to facilitate preterm birth risk stratification and prediction of patient outcome.

The specific aims of the project are as follows:

1. To test the ability of DESI-MS to predict underlying microbial composition and host inflammatory status in vaginal swab samples collected during pregnancy.

2. Investigate the ability of DESI-MS to predict PTB risk based on known preterm birth risk factors (e.g., microbiome composition, PTB history, ethnicity, BMI)

3. Characterise the levels of naturally occurring progesterone in cervicovaginal fluid and test the hypothesis that sees a local progesterone deficiency as a possible cause for PTB.

4. Use DESI-MS to monitor microbial colonisation and characterise host response in a cohort of women receiving vaginal live biotherapeutics during pregnancy.

Chapter II

2 MATERIALS AND METHODS

2.1 Studies subjects and samples collection

2.1.1 Patient cohorts and sampling

2.1.1.1 VMET and VMET 2 study cohorts

The majority of samples used while undertaking this thesis were from the VMET and VMET2 (Vaginal Microbiome and Metabolome in Pregnancy 1 and 2) pregnancy studies. These studies were conducted with approval of the NHS National Research Ethics Service (NRES) Committees London-City and East (REC 12/LO/2003) and London–Stanmore (REC 14/LO/0328), and by the North of Scotland Research Ethics Service (REC 14/NS/1078). Experiments were performed in accordance with the approved institutional guidelines after written informed consent was provided by all participants. Pregnant women were recruited at four major hospitals; Queen Charlotte's and Chelsea and St Mary's Hospitals (London, UK), Chelsea & Westminster Hospital (NHS Trust, London, UK), University College London Hospital (NHS Foundation Trust, London, UK) and the Royal Infirmary of Edinburgh, Scotland, UK.

For VMET and VMET2, recruitment criteria included pregnant women with a singleton pregnancy with and without risk factors for PTB. Exclusion criteria included women under 18 years of age, those who had sexual intercourse in the previous 72h, vaginal bleeding during the previous week, antibiotic use during the previous 2 weeks, multiple pregnancies, HIV infection or hepatitis-C infection. For all participants, detailed maternal clinical metadata and pregnancy outcome data was collected. This included ethnicity self-reported as White (European ancestry), Black (African or African-Caribbean ancestry) or Asian (Pakistani, Indian, Bangladeshi or Sri Lankan ancestry), maternal age, BMI, gestational age, and birth outcome. Deliveries prior to the 37th week of gestation were annotated as preterm births. An average of five high-vaginal swabs were collected from pregnant women at different gestational ages: 8-16, 16-19, 19-26, 26-30, and 32-38 weeks. At each time point, cervicovaginal fluid was sampled with three rayon Transwabs which were used for 16S rRNA amplicon-based

metataxonomic profiling, bacterial cultivation, immune profiling and metabolite profiling. Samples were taken by trained clinicians by rotating the swabs in the vaginal cavity both in a clockwise and anti-clockwise direction for 5 seconds. The swab used for metabolic profiling was then transferred into a sterile 2 ml microcentrifuge tube without any buffer or culture solution for storage at −80°C.

2.1.1.2 FLIP1 study cohort

Samples were also collected from patients enrolled in the FLIP1 study, which aimed to evaluate the safety, tolerance, and acceptability of a live biotherapeutic (LACTIN-V) use in a cohort of pregnant women at risk of preterm birth. Patients were recruited between September 2019 and April 2021 at Queen Charlotte's and Chelsea Hospital, Imperial Healthcare NHS Trust. The study was conducted in accordance with the Declaration of Helsinki and ethical approval was granted by the London Surrey Borders Research Ethics Committee (IRAS 262987). Health Research Authority and Health and Care Research Wales Approval was given (REC reference 19/LO/1018). Study participants were selected from patients referred to the Prematurity Prevention Service at Imperial Healthcare NHS Trust because of their increased risk of PTB. Inclusion criteria for the study included a series of well-established risk factors for PTB including a previous history of PTB, PPROM, mid-trimester loss (MTL), caesarean section at full cervical dilatation (FDCS), trachelectomy, previous cervical tear or congenital uterine anomaly, a short cervix in a previous pregnancy or women who had previously undergone a large loop excision of the transformation zone of the cervix (LLETZ). To be included in the study women had to be over 18 years old, between 11 and 15+6 weeks of gestational age and able to provide informed consent. Exclusion criteria include HIV positivity and use of antibiotic therapy up to 1-week prior to recruitment.

Women recruited for the study were instructed on how to use LACTIN-V during their first consultation. LACTIN-V is commercially available in pre-filled applicators containing *Lactobacillus crispatus* CTV-05 powder at a dose of 2 x 10^9 colony-forming units (CFU), preserved with inactive ingredients. Study participants were required to take a total of 11 doses of LACTIN-V. The first phase of the intervention included 5 daily doses followed by a maintenance phase of 6 weekly doses. Participants were also advised to store the product in a refrigerator and abstain from sexual intercourse for the first 14 days of the study and for 48 hours before each follow-up visit. Women were sampled at 1 (T1), 4 (T2), 6 (T3), 14 (T4) and

22 (T5) weeks after recruitment as well as at delivery (T6) corresponding to approximately 15, 18, 20, 28, 36 and 40-weeks gestation (see Figure 2.1). Additionally, neonatal stool samples were collected 14 days after delivery. At each visit, vaginal swab samples were collected using two rayon Transwabs for later 16S rRNA amplicon-based metataxonomic, immune and metabolite profiling. Samples were taken and stored by trained clinicians as described in 2.1.1.1. In addition, a sterile disposable menstrual disc (Softdisc^{TM}, Instead Inc.) was used to collect cervicovaginal fluid at T4. For this, the soft cup was inserted in the vaginal canal by a trained clinician during the time of the visit. The disc was kept in place for 20 minutes before being removed and then placed in a labelled 50ml Falcon tube. The weight of the empty soft cup was also recorded before and after collection to allow for determination of the wet weight of the cervicovaginal fluid collected. Samples were transported to the Institute of Reproductive and Developmental Biology (IRDB) with 15min of collected for processing and storage.

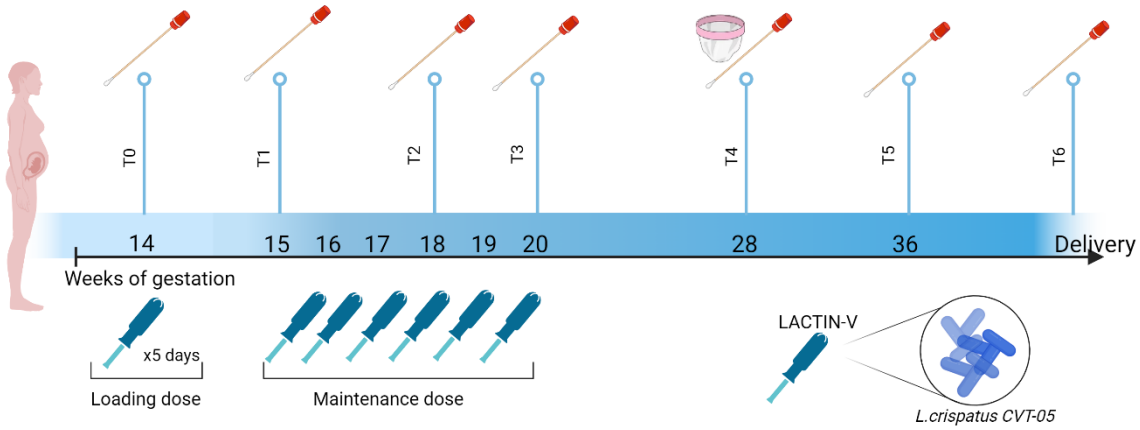


Figure 2.1 - Schematic of study design for FLIP cohort.

2.1.1.3 Collection of cervicovaginal fluid using soft cups

A schematic representation of the soft cup samples processing procedure is shown in Figure 2.2. Once the samples are brought into the lab at IRDB, they are logged in the sample logbook with the corresponding relevant information. The soft cup is then weighted, and the empty soft cup weight previously recorded is used to calculate the cervicovaginal fluid amount. To ensure full recovery of the cervicovaginal fluids from the cup, a spatula is used to thoroughly move the fluid towards the centre of the disc. The fluid is then diluted with a corresponding volume

of PBS that will achieve a dilution of 1:5 (v/v). PBS and vaginal fluids are mixed through pipetting and transferred in a clean 5ml tube. The total volume of the diluted liquid is then aliquoted in different Eppendorf tubes to be used for different downstream analyses. Some of the aliquots are centrifuged at 500g for 10mins to form pellet and supernatant aliquots. For this thesis however only uncentrifuged raw material has been used. After aliquoting the samples are stored in the freezer at -80°C.

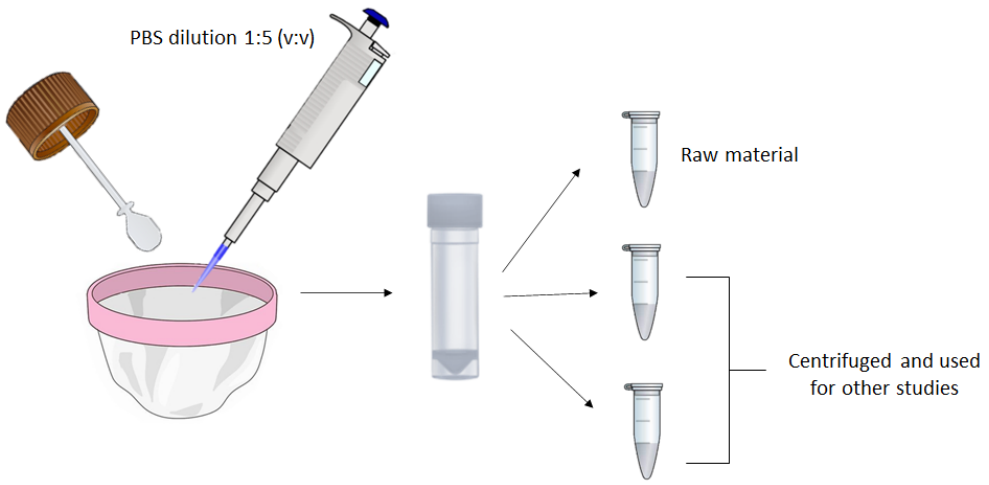


Figure 2.2 - Schematic representation of soft cup samples processing.

2.2 Characterisation of vaginal microbiota community composition

2.2.1 16S rRNA gene sequencing-based metataxonomic profiling

DNA extraction and 16S rRNA gene amplicon sequencing for all matching cervicovaginal swab samples in the VMET 1 cohort were conducted by the members of the MacIntyre laboratory as previously described [139], [235], [441]. For this, cervicovaginal fluid was extracted from the samples by vortexing the swab re-suspended in Amies Liquid Medium transport buffer. The released cell extract was then transferred to a sterile DNase/RNase-free 2 ml microcentrifuge tube where an enzymatic lysis step was performed for 1h at 37 ˚C. The enzyme mix consists of 50 µL of lyzosyme (10 mg/mL), 6 µL of mutanolysin (25,000 U/mL; Sigma- Aldrich), 3 µL of lysostaphin (4,000 U/mL in sodium acetate; Sigma- Aldrich), and 41 µL of TE buffer (10 mM Tris·HCL and 50 mM EDTA, pH 8.0). To ensure the mechanical

disruption of the sample, an additional lysis step was conducted using Mikro-Dismembrator (Sartorius UK Ltd, Surrey, United Kingdom) with acid-washed glass beads for 1 min at 1000 x *g*. The lysate obtained was then purified with the QIAamp DNA Mini kit (Qiagen, Manchester, UK) and the DNA was eluted in 100 ml AE buffer. In order to check the integrity of the extracted bacterial DNA, a PCR amplification was performed using the universal forward and reverse primers:
27F-5'-AGAGTT-TGATCCTGGCTCAG-3' and 338R-5'-'GCTGCCTCCCGTAGGAGT-3' [442]. For this study cohort, the V1-V3 hypervariable regions of 16S rRNA genes were amplified for sequencing using a forward and reverse fusion primer. The forward primer was constructed with the Illumina i5 adapter (5′-AATGATACGGCGACCACCGAGATCTACAC-3′), an 8–base pair (bp) bar code, a primer pad (forward, 5′-TATGGTAATT-3′), and the 28F primer (5′-GAGTTTGATCNTGGCTCAG-3′). The reverse fusion primer consisted of the Illumina i7 adapter (5′-CAAGCAGAAGACGGCATACGAGAT-3′), an 8-bp bar code, a primer pad (reverse, 5′-AGTCAGTCAG-3′), and the 519R primer (5′-GTNTTACNGCGGCKGCTG-3′). Sequencing was conducted on an Illumina MiSeq platform (Illumina Inc.) at Research and Testing Laboratory (Lubbock, TX, USA). The generated sequence reads were analysed using the MiSeq SOP Pipeline of the Mothur package [443]. The SILVA (www.arb-silva.de/) and STIRRUPS ribosomal RNA databases were used to perform taxonomic assignments, , and the subsequent classification was undertaken using the RDP (Ribosomal Database Project) database reference sequence files and the Wang method [444]–[446]. Operational taxonomic unit taxonomies (phylum to genus) were determined using the RDP MultiClassifier script. Species-level taxonomies were determined using USEARCH with 16S rRNA gene sequences from the cultured representatives from the RDP database (http://rdp.cme.msu.edu/) [447].

Sample preparation, DNA extraction and PCR amplification for VMET 2 samples were conducted following the same workflow as previously described in 2.2.1. However, for these samples V1-V2 hypervariable regions were amplified for sequencing using a mixed primer set that has been shown to maintain the rRNA gene ratio of *Lactobacillus* spp. to *Gardnerella* [295]. More specifically, the forward primer set (28f-/YM) consists of a mixture of the following primers in a ratio of 4:1:1:1; 28F-Borrellia GAGTTTGATCCTGGCTTAG; 28F-Chlorlex GAATTTGATCTTGGTTCAG; 28F- Bifido GGGTTCGATTCTGGCTCAG; 28FYM GAGTTTGATCNTGGCTCAG.

The reverse primer was 388 R GCTGCCTCCCGTAGGAGT. Sequencing was performed on an Illumina MiSeq platform (Illumina Inc.) at Research and Testing Laboratory, (RTL Genomics) in Lubbock, Texas, USA.

To calculate microbial abundance the 16S amplification primers were trimmed using Cutadapt [233] followed by quality control using FastQC [448]. Amplicon sequence variant (ASV) counts per sample were calculated using the Qiime2 pipeline, where we used DADA2 for denoising [233], [449]. ASVs were taxonomically classified to species level using the STIRRUPS reference database [446]. Counts were normalised with relative log-expression using DESeq2 [450]. Measures of microbial diversity: inverse Simpson and Shannon indexes and richness (species observed) were calculated using the Phyloseq package in R [451]. For analyses of the vaginal microbial composition, at genus level, samples were grouped into those that were *Lactobacillus* dominant (≥75% Lactobacillus species), or *Lactobacillus* deplete, (<75% Lactobacillus species). At species level, samples were assigned to CSTs using the VALENCIA centroid classification algorithm [198].

2.3 Cytokine/chemokine profiling in cervicovaginal swabs samples

Cytokine and chemokine profiling for all matching cervicovaginal swab samples in the VMET 2 cohort were performed by Dr Denise Chan [452]. For this, samples were thawed on ice and cervicovaginal fluid was recovered from the swab sponge using a sterile syringe and applying a pressure to aspirate a volume of ~350µl. The recovered liquid was then centrifuged at 3000 × *g* for 10 min. The supernatant after centrifugation was transferred to a clean Eppendorf tube and the resulting solution was used for Luminex® immunoassays to quantify the targeted cytokines, chemokines, complement components and immunoglobulins. Analytes were picked according to literature evidence of involvement in inflammatory change related to PTB, BV and microbial dysbiosis [453]–[455]. The selected analytes were IL-8, IL-1β, IL-2, IL-4, IL-5, IL-6, IL-10, IL-18, IFN-γ, GM-CSF, TNF-α, C5, C5a, MBL, C3b/iC3b, IgM, IgG1, IgG2, IgG3, IgG4, IgA and IgE. IL-8 was measured on a single plex Human Premixed Analyte Kit (R&D Systems/Bio-Techne), following a tenfold dilution using Calibrator Diluent RD6-52. The remaining cytokines, IL-1β, IL-2, IL-4, IL-5, IL-6, IL-10, IL-18, IFN-γ, GM-CSF and TNF-α, were measured neat on a multiplex plate, Human Premixed Multi-Analyte Kit (R&D Systems/Bio-Techne). Complement analytes, C5, C5a and MBL were detected using the Human Complement Magnetic Bead panel 1 (Milliplex® Merck/Millipore), while for

C3b/iC3b detection, Human Complement Magnetic Bead panel 2 (Milliplex® Merck/Millipore) was used. Dilution was not required for complement assays. Among the immunoglobulins, only IgG1 required a 50-fold dilution. This was performed with the Universal Assay Buffer included in the assay kit. IgM, IgG1, IgG2, IgG3, IgG4, IgA and IgE were measured with the ProcartaPlex Human Antibody Isotyping panel 7-plex (Thermofisher). Immunoassays were used together with the Bio-Plex 200 system (Bio-Rad Laboratories Ltd). All samples were analysed on 96-well plates in duplicates. A common pooled sample was analysed across all plates to assess plate-to-plate variation, and the coefficient of variation was calculated across all plates for standards. Bio-Plex Manager software was then used to determine concentrations of assayed samples (in pg/mL) via interpolation of values from the standard curves.

2.4 DESI-MS based metabolite profiling of clinical swab samples

2.4.1 Reagents and materials

All chemicals used in the experiments described in this section were of analytical reagent grade. UPLC grade methanol and water for DESI-MS analysis were purchased from Sigma Aldrich (St Louis, MO).

2.4.2 Direct swab analysis by DESI-MS using an Orbitrap Mass Spectrometer

DESI-MS experiments for VMET and VMET2 cohorts were conducted on an LTQ-Orbitrap Discovery mass spectrometer (Thermo Scientific, Bremen, Germany). A schematic representation of the hardware used for DESI-MS experiments is shown in Figure 2.3. The DESI-MS source used was built in-house and consisted of an electrospray emitter mounted on a movable stage to allow a customisable positioning of the DESI-MS sprayer and the swab surface in the x-y-z direction. The DESI sprayer described below was built and assembled by Doctor Pamela Pruski as part of her PhD completed in the laboratories of Profs Zoltan Takats and David MacIntyre [456]. The central piece of the sprayer is constituted by a Swagelok T element (1/16”). On both sides of the T element two sleeves and ferrules of the same measure (1/16”) hold in place two fused silica capillaries. Perpendicularly to this, an additional sleeve holds the gas tubing for the nitrogen line. The solvent capillary has an internal diameter of

50 µm, an outer diameter of 150 µm. This is inserted in the gas capillary 1.5 cm long with 250 µm internal diameter, and 363 µm outer diameter. A DC voltage source (part of the LTQ-Orbitrap MS outsource) is connected to the solvent line prior to entry into the Swagelok T element. The gas and solvent capillaries, and the nitrogen inlet are connected to the central Swagelok T-element, together with a metal cap spray emitter (labelled as "Electrospray" in Figure 2.3). The combination of pressured nitrogen flow, with the flow of the charged solvent results in the generation of the electrospray, which is directed to the sample surface via the emitter piece. In this swab analysis setup, the electrospray is targeted directly at the surface of the swab, which is held in place on fixed stage element connected to a 9V rotation motor ("Rotating swab support" in Figure 2.3). The maximum rotation speed achieved is of 405 rotation per minute (rpm), with a remote controllable speed dimmer which allows easy adjustment of rotation speed.

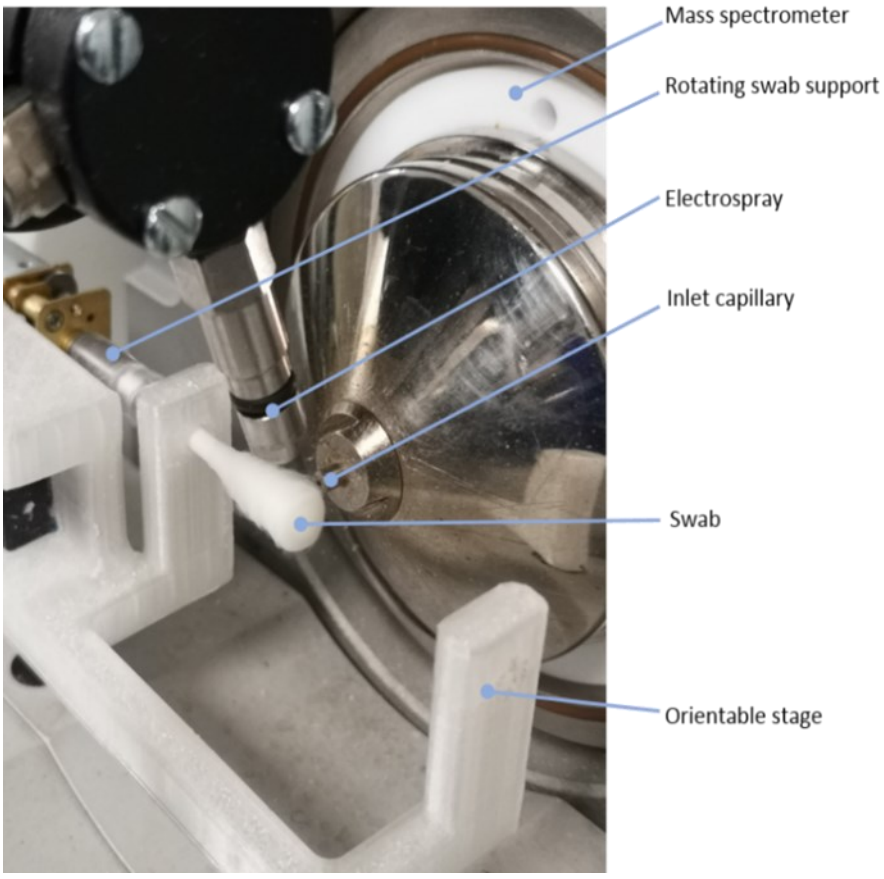


Figure 2.3 - Schematic representation of DESI-MS source on LTQ-Orbitrap platform.

This configuration allows the surface of the swab tip to be targeted by the DESI-MS electrospray in close proximity to the MS inlet capillary. The solvent dissolves the analytes present on the swab, secondary droplets are formed and molecules are ionised in a process

described in 1.9.3.2.1. The ions are then transferred to the mass analyser through the inlet capillary and the interface containing ion optical elements that direct ions through the sequential gas conductance limits.

2.4.2.1 MS instrumental settings

The instrumental settings used for the DESI-MS analysis of medical swabs on the LTQ-Orbitrap Discovery mass spectrometer are summarised in Table 2.1.

Table 2.1 - Instrumental parameters used with the LTQ- Orbitrap Discovery mass spectrometer.

Parameters	***MS settings***
Mass range	50 – 1000 *m/z*
Mass resolution	30.000
Max. injection time	500 ms
Microscans	1
Capillary temperature	250 ºC
Capillary Voltage	50 V
Tube Lens Voltage	120 V

To limit batch effect and ensure good quality of the data acquired, a regular cleaning of the inlet capillary, the optics element and of the whole instrument was performed weekly. Additionally, the inlet capillary cleaning was performed before each day of analysis. Also, a mass accuracy and mass resolution calibration of the electrospray ion source were performed daily before use following the manufacture's manual using a 50:50 mixture of negative/positive ion calibration solution (Thermo Scientific).

2.4.2.2 DESI-MS swab setup and experimental parameters

Once the swab is mounted onto the swab holder it was gently manoeuvred to ensure the distance between the swab and the capillary was maintained at less than 2mm and good signal intensity (TIC larger than 10e6) was observed. The DESI sprayer was kept at 1.5-2mm of distance directed towards the centre of the medical swab. The swab was analysed through a clockwise rotation while being sprayed with a methanol/water solution (95:5, v/v) nebulised with nitrogen gas with 7 bar flow pressure. The sprayer was also connected to a DC potential of 4.3 kV. A fixed solvent flow of 10 µl per minute was delivered by the integrated LTQ-Orbitrap syringe pump operating with a 0.5 mL volume Hamilton syringe. A summary of all used parameters for the DESI-MS analysis of swabs is given in Table 2.2.

Table 2.2 - DESI-MS geometric and electro-solvent parameters used for the medical swabs analysis.

Parameters	**Settings**
Nebulising gas pressure	7 bar
Spray voltage	4.3 kV
Solvent flow rate	10 µL min-1
Spray solvent	methanol/water (95:5, v/v)
Tip-to-surface height	<< 2 mm
Distance inlet-sample	2 mm
Angle tip-sample	50°
Angle inlet	10°
Distance tip-inlet	4.5 mm

2.4.2.3 Data pre-processing of DESI-MS .raw data

The DESI mass spectral data were converted from .raw files into .mzXML format via the ProteoWizard msConverterGUI (Vanderbilt University, Nashville, TN, USA) [457]. R was used to average the first 30 spectra per sample and save them as .csv files [458]. Mass spectral data were then imported into an in-house customised "MALDIquant"- based graphic user interface (GUI) platform with a resolution of 0.01 Δ*m/z* and the peaks detected in each sample were matched sequentially (window of 0.01 *m/z*), thus obtaining a matrix of mass spectral profiles.

2.4.3 DESI-MS using a small-format RDa/BioAccord TOF Mass Spectrometer

DESI-MS experiments for the FLIP-1 cohort were conducted using ACQUITY RDa/BioAcccord ToF mass spectrometer (Waters Corporation, Wimslow, UK). The RDA platform has several differences from the LTQ-Orbitrap and the decision to transition to this platform was driven by the objective to have direct on-swab metabolic profiling made available in clinical environments. The RDa platform has a substantially smaller footprint than the LTQ-Orbitrap (see Table 2.3), is portable and is amenable to bedside testing.

Table 2.3 - Dimensional differences between LTQ-Orbitrap and RDa ACQUITY platforms.

	RDa ACQUITY	***LTQ-Orbitrap***
Height	75 cm	143 cm
Width	40 cm	87 cm
Depth	80 cm	160 cm
Weight	80 kg	205 kg

Another important difference between the two platforms, is the mass analyser used. The RDa ACQUITY is a TOF mass spectrometer and has generally lower mass resolution compared to the orbitrap analyser used in the LTQ-Orbitrap machine. An image of the hardware used for

DESI-MS experiments is presented in Figure 2.4. The DESI stage was designed by WATERS and Dr. Emrys Jones. The roboticized rotation mechanism was designed by Dr. Eftychios Manoli and Dr. Mark Runciman. The ACQUITY RDa uses a DESI XS high-performance sprayer mounted on a custom 3D printed stage that allows adjusting the sprayer distance and angle from the sample. As part of the DESI stage, a roboticized sampling device was used, that provides both adjustable rotation and movement along the linear plane. The device included a magnetic, adjustable swab grip that was used to facilitate the analysis of different swab types with differing diameters and lengths. All functionalities including rotation, linear speed, and pause functions were controlled by the software (or manually) and seamlessly integrated with the analysis workflow. The entire surface of the clinical swabs was analysed in an automated way by DESI-MS through clockwise rotation of the swab toward the MS capillary.

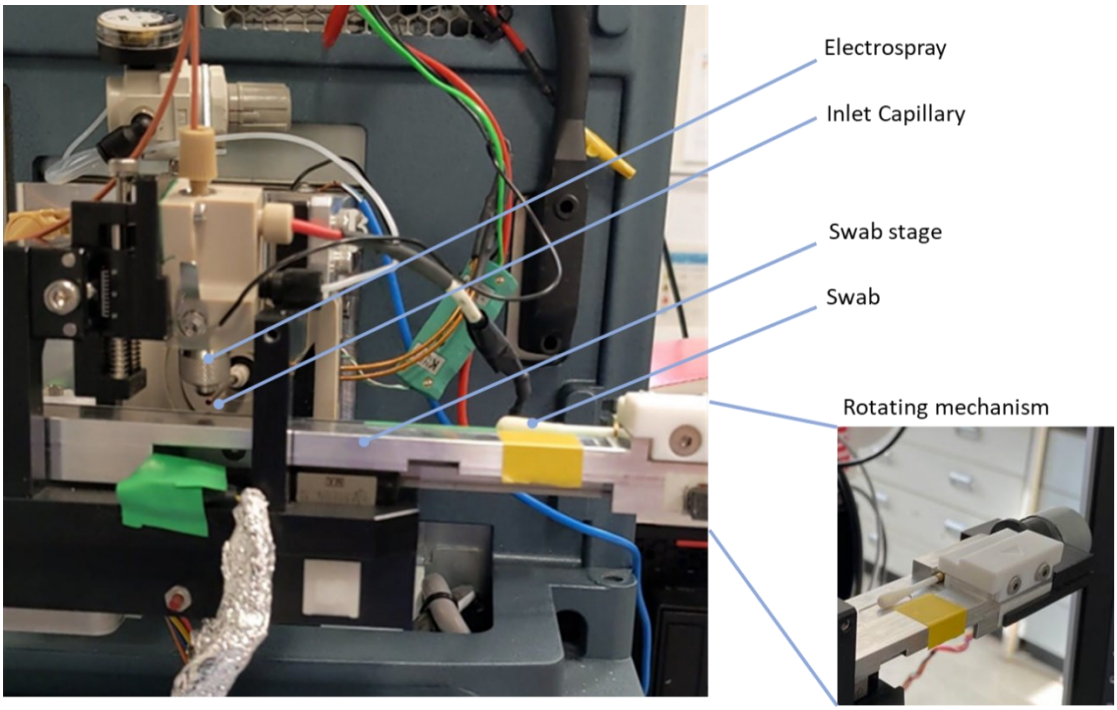


Figure 2.4 - Schematic representation of DESI-MS source on RDa ACQUITY platform.

Randomization of samples was done in blocks and mass spectra were acquired for each individual sample. This means that each patient's samples were run sequentially but were randomised in order of gestational age. Randomization of patient order was performed using a Python script and including criteria of the metadata such as microbial composition, parity and gestational age.

2.4.3.1 DESI-MS instrument set-up

The instrumental settings used for the DESI-MS analysis of medical swabs on the ACQUITY RDa mass spectrometer are summarised in Table 2.4.

The inlet capillary was cleaned before and at the end of each analytical run and additional cleaning was performed during acquisition when material from the swabs would clog the inlet opening or the backing pressure would decrease. After the analysis the instrument, DESI interface, capillary and stage were removed and cleaned using a mixture of methanol and water.

All the parts were submerged in a solvent mixture of methanol and water and placed in an ultrasonic bath and sonicated for 15 minutes. After that, parts were dried on air before use.

Mass accuracy and mass resolution calibration of the electrospray ion source were performed twice following the manufacturer's manual using a 50:50 mixture of negative/positive ion calibration solution.

Table 2.4 - RDa ACQUITY platform operating parameters.

Parameters	*MS settings*
Mass range	50 – 1000 m/z
Mass resolution	≥10.000
Max. injection time	500 ms
Scan Time	1 sec
Capillary temperature	250 ºC
Capillary Voltage	50 V
Tube Lens Voltage	120 V

2.4.3.2 DESI-MS swab set up and experimental parameters.

A summary of DESI-MS parameters on the RDa used for the analysis of swabs is provided in Table 2.5. Swabs were placed into a rotating holder positioned orthogonally in front of the MS inlet capillary with a swab–capillary distance of approximately 2 mm. The DESI sprayer tip was pointed to the swab centre with a tip-sample distance of 1.5–2 mm and a distance between the tip and the inlet capillary of 2 mm. The entire surface of the medical swabs was analysed by DESI-MS through clockwise rotation of the swab toward the MS capillary using a mixture of methanol/water solution (95:5, v/v) nebulised with nitrogen gas with 3.5 to 4.5 bar flow pressure. The sprayer was also connected to a DC potential of 0.8kV and 1kV for negative and positive mode respectively. The solvent flow rate was 30µl per minute.

Table 2.5 - RDa ACQUITY platform DESI swab stage set up parameters.

Parameters	Settings
Nebulising gas pressure	3.5 – 4.5 bar
Spray voltage	1 kV (positive mode) 0.8 kV (negative mode)
Solvent flow rate	30 µL min-1
Spray solvent	methanol/water (95:5, v/v)
Tip-to-surface height	<< 2 mm
Distance inlet-sample	2 mm
Angle tip-sample	35°-40°
Angle inlet	10°
Distance tip-inlet	2 mm

2.4.3.3 Data pre-processing of .raw data

The DESI mass spectral data were converted from the proprietary Waters .raw format into the open .mzML format with the ProteoWizard msconvert tool (Vanderbilt University, Nashville, TN, USA) [457]. The vendor centroid peak detection algorithm was applied during conversion. The spectra were then further processed in Python (v3.9.12), with the pymzML (v2.5.1) library used to parse mzML files [459]. For each DESI-MS experiment, individual scans with a TIC of less than 75,000 counts were discarded and the centroid peaks present in the remaining scans were grouped into MS features using a Python implementation of the MALDIquant *binPeaks* function, with the matching tolerance parameter set to 150 ppm, and the 'strict' grouping method [458]. Peaks present in less than 25% of the scans were removed. The *m/z* values and intensities of each feature were then averaged to generate a single spectrum per DESI-MS experiment. The *m/z* axis of each spectrum was re-calibrated using a linear calibration function (y=mx +b) fitted to the observed *m/z* deviation values for 3 endogenous reference ion signals: lactic acid $[M-H]^-$, cholesterol sulphate $[M-H]^-$ and PG 36:1 $[M-H]^-$, with a theoretical *m/z* value of 89.02442, 465.3044, and 775.5489, respectively. After this procedure, peaks were grouped across samples to generate the final MS data matrix with the MALDIquant *binPeaks* algorithm, with tolerance=150 ppm and the 'relaxed' grouper setting.

2.4.4 Liquid extraction methods of cervicovaginal soft cup samples

For liquid extraction, CVF samples recovered from soft cups were thawed on ice and then 200 µl transferred to a clean Eppendorf tube. To this, 800 µl methanol:water (2:1, v/v) was added. To enable the detection and quantification of progesterone and other hormones in the cervicovaginal fluid 100µl of a mixture of internal standards was added to each sample. Analytical standards of testosterone (TES), progesterone (PROG), 21-Deoxycortisol (21-DEOXY-HC), androstenedione (AND), 11-Deoxycortisol-2,2,4,6,6-d5 (d5-11-DEOXY-HC), corticosterone (COR), estrone (E1), β-estradiol (E2), estriol (E3), aldosterone (AL), were purchased from Sigma Aldrich (Gillingham, UK). Progesterone-2,2,4,6,6,17α,21,21,21-d9 (d9-PROG), cortisol (HC), 17α-hydroxyprogesterone (17-OH-PROG), and 11-deoxycorticosterone (11-DEOXY-COR) were acquired from Steraloids (Newport, RI, USA). The progesterone C^{13} standard was purchased from Merck® (PROGESTERONE-2,3,4-13C3,

99atom % 13C). After the addition of the internal standards the samples were vortexed for 2 mins and then centrifuged at 4 ºC for 10 mins at 16 000 x *g*. The supernatant was then collected in a clean Eppendorf tube and the pellet was saved for additional extraction steps. To improve the recovery of lipids, the pellet was reextracted using 400 µl of isopropanol:water (4:1, v/v). After vortexing for 2mins, the sample was centrifuged at 4 ºC for 10 mins at 16 000 x *g*. The supernatant from the isopropanol extraction is then added to the supernatant obtained from the methanol extraction. The pellet was discarded. Samples extracted were then placed in Savant DNA 120 SpeedVac® Concentrator for ~4h or until the solvent was completely evaporated. Blank samples with only internal standards were prepared using the exact procedure described above. A schematic workflow for the liquid/solid extraction workflow used for metabolite-type analytes from vaginal mucosa swab samples is illustrated in Figure 2.5.

The lyophilised sample was then redissolved with 100 µL of water/methanol/acetonitrile (6.4:1.2:2.4, v/v/v) containing isotopically labelled steroids used as internal standards (IS). Finally, after reconstitution (5 min shaking and 5 min sonication), the extract was transferred to the analysis plate, spined down and placed in the sample manager (kept at 4°C).

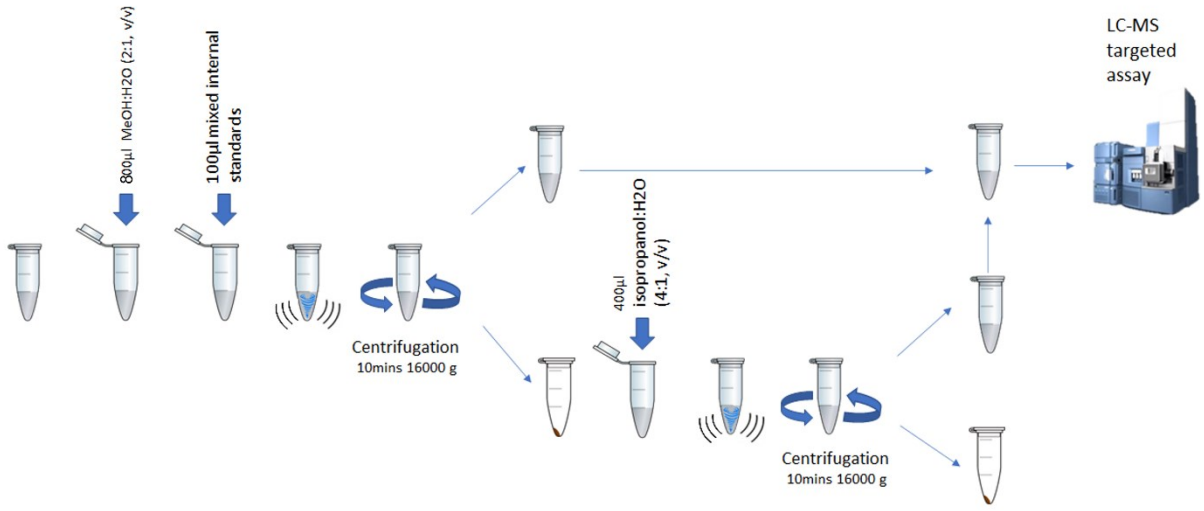


Figure 2.5 - Workflow for metabolite analysis of soft cup samples by liquid extraction and then LC-MS.

2.4.5 LC-MS steroid assay

Assaying of steroids was performed by the National Phenome Centre at Imperial College London. The assay was adapted for use on CVF by Dr. Lucia Olmo Garcia. The steroid assay was performed on an ACQUITY UPLC system coupled to a Waters Xevo TQ-S tandem quadrupole instrument through an electrospray ionization source (Waters Corp., Milford, MA, USA). Chromatographic separation was carried out with an Acquity HSS T3 column (1.0 × 100 mm, 1.8 µm particle size, 100 Å pore size) (Waters, Wilmslow, U.K.) operating at 45°C and with an injection volume of 20 µl. Mobile phase composition was 5 mM ammonium acetate in water adjusted to pH=4.5 with acetic acid (Phase A) and 50 mM ammonium acetate in water adjusted to pH=4.5 with acetic acid/methanol/acetonitrile (1:3:6, v/v/v) (Phase B). Steroids were eluted with a 0.2 mL/min mobile phase gradient: 30-40% Phase B from 0 to 2.5 min, kept isocratic for 2 min; 40-50% Phase B from 4.5 to 8 min; 50-90% Phase B from 8 to 8.2 min; 90-99% Phase B from 8.2 to 9.5 min. A further washing step was added to avoid carryover and column clogging increasing the overall run time to 13 min per sample. To that end, 20 µl of acetonitrile/isopropanol (1:1, v/v) were injected while keeping 99% Phase B at 0.3 mL/min. After 2.3 min of isocratic elution, initial conditions (30% Phase B, 0.2 mL/min) were reached in 0.2 min and the column was re-equilibrated for 1 min prior to the next sample injection. Estrogens were detected in negative ionisation mode with 1.5 kV capillary voltage; the rest of the steroids were determined in positive polarity with 1 kV capillary voltage. The other source parameters were similar in both ionisation modes: 80 V source offset, 150°C source temperature, 600°C desolvation temperature, 1200 L/h desolvation gas flow, 150 L/h cone gas flow, 7.0 bar nebulizer gas, and 0.15 mL/min collision gas flow. Multiple reaction monitoring (MRM) parameters for all the targeted analytes and their isotopically labelled versions are presented in Table 2.6.

Table 2.6 - Multiple reaction monitoring (MRM) parameters for the targeted analytes and their isotopically labelled versions.

Precursor Name	Label type	Use	Precursor m/z	Precursor Charge	Product m/z	Product Charge	Cone Voltage	Explicit Collision Energy	Retention Time
E3	light	Quantifier	287.2	-1	171.1	-1	52	36	1.7
		Qualifier	287.2	-1	145.1	-1	52	38	
d_3-E3	heavy	IS	290.2	-1	147.0	-1	8	44	
AL	light	Quantifier	361.2	1	343.2	1	38	16	2.1
		Qualifier	361.2	1	315.2	1	38	18	
d_4-AL	heavy	IS	365.2	1	346.2	1	38	16	
HC	light	Quantifier	363.2	1	327.2	1	10	14	2.7
		Qualifier	363.2	1	121.1	1	40	24	
d_4-HC	heavy	IS	367.2	1	121.1	1	40	24	
21-DEOXY-HC	light	Quantifier	347.2	1	311.2	1	36	16	4.0
		Qualifier	347.2	1	121.1	1	36	26	
d_8-21-DEOXY-HC	heavy	IS	355.3	1	125.1	1	36	16	
COR	light	Quantifier	347.2	1	121.1	1	42	26	4.2
		Qualifier	347.2	1	329.2	1	42	14	
d_4-COR	heavy	IS	351.2	1	333.2	1	42	14	
11-DEOXY-HC	light	Quantifier	347.2	1	97.1	1	38	22	4.5
		Qualifier	347.2	1	79.0	1	38	50	
d_5-11-DEOXY-HC	heavy	IS	352.3	1	100.1	1	38	22	
E2	light	Quantifier	271.2	-1	145.2	-1	10	32	5.6
		Qualifier	271.2	-1	256.0	-1	10	18	
d_2-E2	heavy	IS	273.2	-1	185.1	-1	10	36	
TES	light	Quantifier	289.2	1	97.1	1	38	22	6.5
		Qualifier	289.2	1	253.2	1	38	16	
$^{13}C_3$-TES	heavy	IS	292.2	1	100.1	1	38	22	
E1	light	Quantifier	269.2	-1	145.0	-1	2	40	6.7
		Qualifier	269.2	-1	159.1	-1	2	40	
d_2-E1	heavy	IS	271.2	-1	145.1	-1	2	32	

AND	light	Quantifier	287.2	1	97.1	1	42	20	7.2
		Qualifier	287.2	1	109.0	1	42	26	
$^{13}C_3$-AND	heavy	IS	290.2	1	100.1	1	42	20	
11-DEOXY-COR	light	Quantifier	331.2	1	97.1	1	54	20	7.2
		Qualifier	331.2	1	79.0	1	54	48	
$^{13}C_3$-11-DEOXY-COR	heavy	IS	334.2	1	100.1	1	54	20	
17-OH-PROG	light	Quantifier	331.2	1	97.1	1	44	24	8.2
		Qualifier	331.2	1	81.4	1	44	44	
d_8-17-OH-PROG	heavy	IS	339.3	1	100.1	1	44	24	
PROG	light	Quantifier	315.2	1	97.1	1	34	22	9.2
		Qualifier	315.2	1	297.3	1	34	16	
^{13}C -PROG	heavy	IS	318.27	1	100.1	1	10	20	

2.5 Data processing and Statistics

2.5.1 Statistical analysis chapter III

2.5.1.1 Multivariate exploratory analysis of the DESI-MS metabolic profiles

PCA analysis was performed to examine the main sources of variation and screen for potential confounders and batch effects such as spectral acquisition date, analytical batch, or recruitment centre. PCA models were created in the R programming language (R version 4.2.2) with the 'pcaMethods' package and plotted using the R-packages 'ggplot2' (version 3.4.0).

Linear mixed-effects modelling analyses were performed in R (R version 4.2.2) using the 'lme4' package (version 1.1.31). For all each metabolic feature, a linear mixed-effects model with the following formula was fitted:

$$Metabolite \sim GestationalAge + CST + MaternalAge + BMI + Ethnicity + (GestationalAge||SubjectID)$$

This model includes fixed effects for gestational age, CST composition, maternal age, BMI, and ethnicity, and a random intercept and slope with gestation for each pregnancy to account for repeated measures ((GestationalAge||SubjectID) term). All models were fitted using a restricted maximum likelihood (REML) estimator ('REML = TRUE' in lmer function). The 'emmeans' package (version 1.8.2) was used to perform contrast coding and obtain effect size estimates and P-values for contrasts and trends of interest. For the comparisons between different CSTs and LDOM vs LDEPL, assumed values for maternal age, BMI and gestational age were used (age = 30 years, BMI = 25, and gestational age = 20 weeks). The mean levels and contrasts obtained were averaged over all levels of ethnicity and CSTs.

The comparison between Lactobacillus dominant vs depleted was encoded as a comparison between the grand mean of CST I, II III, V and VII levels vs the mean of CST IV. The p-values for each contrast tested were calculated using the Kenward–Roger approximation, implemented in the 'pbkrtest' package (v0.4.8.6).

For each of the main contrasts of interest (CST I vs III, CST I vs IV, CST III vs IV and LDOM vs LDEPL), all p-values obtained across all metabolic variables were combined and corrected

for false discovery rate (FDR) with the Benjamini–Hochberg method, with a signature cut-off value of cut-off value of 5%.

The heatmaps were created in Python (version 3.8.5), 'matptlotlib' (version 3.3.2) and 'seaborn' (version 0.11.0). Only features which were statistically significant after false discovery rate correction and replicated in both VMET and VMET 2 datasets were plotted. To identify *m/z* features present in both VMET and VMET 2 datasets, for each feature on one study was found a marker with an m/z error of less than 5 ppm in the final FDR corrected signature from the other study. The linear mixed effect model, semi-partial r2 and conditional R2 measures were calculated using the R packages 'r2glmm' (version 0.1.2) and 'MuMIn' (version 1.47.1), respectively.

PERMANOVA (Permutational Multivariate Analysis of Variance) analysis was conducted with the 'vegan' package (version 2.6.4). The adonis2 function was used, with method = 'euclidean', default value of permutations = 999, and by adding terms sequentially following their order in the model formula:

$$log(MetaboliteMatrix + 1) \sim CST + GestationalAge + Ethnicity + BMI + Age + SubjectID$$

2.5.1.2 Prediction of microbial composition

Random forest classifiers were used to predict CST composition and *Lactobacillus* dominant or depleted status. For the *Lactobacillus* dominant vs depleted comparison, samples corresponding to CST I, II, III, V or VII were classified as "dominant" and samples corresponding to CST IV or VI were classified as "depleted". Random forest classifiers were trained using solely the metabolic profile variables as predictors, without any other clinical or demographical variable.

All random forest (RF) models were performed in R, using the 'randomForest' package (version 4.7.1.1) and the 'caret' package (version 6.0.93), for model fitting, performance metric calculation and cross-validation. RF classifiers were trained using constant default parameters: number of trees 'ntree' = 1000, and number of variables equal to 1/3 to the total metabolic variables ('mtry = number of metabolic variables/3'). A stratified (to ensure the training and testing datasets have a similar distribution of samples across all classes) seven-fold cross-validation procedure with 15 repeats was used for all models.

The performance of the RF models was assessed by calculating performance metrics using

only data left out from each CV round test sets (performed by default by 'caret'). For the classifiers, ROC curves, precision–recall curves, accuracy, sensitivity, specificity, positive-predictive value and negative predictive values were calculated and plotted with the R packages 'caret', 'pROC' (version 1.18.0), 'plotROC' (version 2.3.0), 'precrec' (version 0.12.9) and 'ggplot2'. For the regression analysis, the R2 was calculated with 'caret' and predictions plotted with 'ggplot2'.

2.5.1.3 Immune-metabolite association analysis

Linear regression models were used to investigate the association between immune markers and metabolites. Specifically, the formula used was:

$$lm(log(ImmuneMarker\ +\ 1) \sim Metabolite)$$

All the immune marker - cytokine combinations were investigated, and the corresponding F-test P-values were calculated. Benjamini–Hochberg false discovery rate correction was used for each analysis, with a 5% FDR cut-off.

The heatmaps were generated using Python (version 3.8.5), 'matptlotlib' (version 3.3.2) and 'seaborn' (version 0.11.0), using the results of the linear model analysis and the random forest regression models.

For the predictions of immune markers levels from metabolic features, random forest classifiers were used. Initially, the immune marker concentrations were log-transformed after the addition of a constant offset log(x + 1). All random forest models were trained in R, using the 'randomForest' (version 4.7.1.1) package in combination with the 'caret' package (v6.0-86), for model fitting, performance metric calculation and cross-validation. RF classifiers were trained as described in the previous paragraph (2.6.1.1).

Immune markers were also associated with specific metabolic features such as lactate (m/z 89.0248) succinate (m/z 117.0197), and propionate (m/z 73.0298). To investigate this relationship a linear model was fit following the formula:

$$lm(metabolite \sim immune\ marker\ +\ CST\ +\ GestationalAge\ +\ BMI\ +\ Ethnicity\ +\ Age)$$

To more accurately represent the phenotype of the cervicovaginal mucosa, the ratios between lactate and propionate were calculated, as well as lactate and succinate. Linear models were fitted using these ratios instead of the individual metabolites included in the previous formula.

2.5.2 Statistical analysis chapter IV

2.5.2.1 Multivariate analysis of the DESI-MS metabolic profiles in relation to previous pregnancy history

A preliminary exploratory analysis was undertaken with PCA to observe metabolic differences for the different risk groups. R (R version 4.2.2) was used, and plots were produced with the R-packages 'ggplot2' (version 3.4.0). Patients were divided based on their previous pregnancy history in "cervical" group including people that had previous cervical intervention, and "PTB" group including people with a previous history of preterm delivery.

Supervised analyses were then performed to assess the relationship between metabolic profile and risk factors. Only one sample per patient was included, with a cut off gestational age of 20 weeks to include only the first timepoint available for each patient. Prior to fitting the multivariate models, features were filtered based on a univariate analysis. Significant features (p-value <0.05) were selected using a linear model fit using the following formula:

$$lm(log(Metabolite) + 1) \sim risk_class + cst + bmi + gestational_age + ethnicity)$$

PLSDA models were created with the 'caret' package (version 6.0.93) introducing in the formula the parameter method = "pls". PLSDA model was performed using only significant features identified through the linear model and was cross validated. The plot was generated with 'ggplot2' package (version 3.4.0).

2.5.2.2 Prediction of future cervical cerclage intervention requirement

To investigate the ability of DESI-MS profiling to predict the need of a patient to receive a cerclage intervention later in gestation, a subset of patients from VMET and VMET2 cohorts meeting the following criteria were selected: availability of longitudinal samples and, for patients who received a cerclage, availability of a sample before intervention. The metabolic

profiles of patients who received the intervention were compared with samples from a control group of women who did not receive any intervention during their current pregnancy. The control group was matched with the cases in terms of gestational age and microbial composition at the genus level. Out of the 174 patients with longitudinal data, based on these criteria 49 were defined as “cases” and 125 as “controls”.

A linear model with CV was used to select significant features between cases and controls. The formula used was the following:

$$lm(log(Metabolite) + 1) \sim case\ VS\ controls\ class + cst + bmi + gestational_age + ethnicity)$$

This formula was applied to all the metabolic features in the dataset.

Using only significant features (p-value <0.05) a random forest classifier was applied to the selected subset of data. Random forest model was performed using the ‘randomForest’ package (version 4.7.1.1) and the ‘caret’ package (version 6.0.93) with the same parameters described in 2.6.1.1.

2.5.2.3 Multivariate analysis of the DESI-MS metabolic profiles in relation to cerclage efficacy

Supervised multivariate analysis using PLSDA was performed to examine differences in metabolic profiles between individuals who had a long cervix after receiving a cerclage intervention versus those who still experienced shortening despite the intervention.
The samples used in this analysis were selected according to the following criteria: one sample per patient taken after cerclage intervention but before 28th week of gestation, taken as a cut-off for a physiological cervical shortening period in pregnancy.

We selected the metabolic features that were significantly different between the group with a short cervix and the group with a long cervix (p-value < 0.05). To identify these features, we fitted a linear model with CV using the following formula:

$$lm(log(Metabolite) + 1) \sim cervical\ lenght\ class + cst + bmi + gestational_age + ethnicity)$$

As previously described, PLSDA models were created with the 'caret' package (version 6.0.93) introducing in the formula the parameter method = "pls". PLSDA model was performed using only significant features identified through the linear model and was cross validated. The plot was generated with 'ggplot2' package (version 3.4.0).

2.5.2.4 Microbial composition prediction after cerclage intervention

Using the same set of samples selected according to the criteria explained in 2.6.2.4, we performed RF analysis to predict the microbial composition after the cerclage intervention. Random forest classifiers were used to predict CST composition both at genera level for *Lactobacillus* dominant or depleted status and for CST comparisons. For the *Lactobacillus* dominant vs depleted comparison, samples corresponding to CST I, II, III, V or VII were classified as "dominant" and samples corresponding to CST IV or VI were classified as "depleted". Random forest classifiers were trained and plotted in the same way described in 2.6.1.1.

2.5.3 Statistical analysis chapter V

To investigate the relationship between progesterone and other steroid hormone concentration and gestational age the following linear model was used:

$$lm(steroid\ level\ \sim gestational_age)$$

Scatterplots of the steroids levels across gestational age were produced with the R package 'ggplot2' (version 3.4.0). Comparisons between single steroids levels in patients experiencing PTB versus patients reaching term were performed with two-tailed Wilcoxon-Mann-Whitney tests, using the default R base function "wilcox.test".

2.5.4 Statistical analysis chapter VI

2.5.4.1 Multivariate exploratory analysis of the DESI-MS metabolic profiles

PCA models were created in R, with the pcaMethods function, and results visualized graphically with 'ggplot2' (version 3.4.0).

Score plots were coloured according to different factors, including spectral data acquisition date, patient ethnicity and visible presence of blood on the sample.

Linear mixed modelling was performed in R using the 'lme4' package (version 1.1.31). For each metabolic feature, a linear mixed-effects model was applied utilizing the formula outlined below:

$$Log(Metabolite + 1) \sim GestationalAge + CST + MaternalAge + BMI + Ethnicity + age + (GestationalAge||SubjectID)$$

This model includes fixed effects for gestational age, CST composition, maternal age, BMI, and ethnicity, and the random effect structure (GestationalAge||SubjectID) is similar to the one described in section 6.2.1.2 The methodology for model fitting, contrast coding, effect size estimate and p-values was the same as described in section 6.2.1.2.

Random forest classifiers were used to predict CST composition and *Lactobacillus* dominant or depleted status. For the *Lactobacillus* dominant vs depleted comparison, samples were classified as "dominant" or "depleted" following the same criteria outlined above. Random forest classifiers were trained using solely the metabolic profile variables, and no other clinical or demographical variable. Additionally, RF models were employed to compare the prediction of different combinations of CSTs.

All random forest (RF) models were performed in R, using the 'randomForest' package (version 4.7.1.1) and the 'caret' package (version 6.0.93), as described above.

The performance of the RF models was assessed by calculating the R2 using only data left out from each CV round test sets (performed by default by 'caret'). For the classifiers, ROC curves, precision–recall curves, accuracy, sensitivity, specificity, positive-predictive value and negative predictive values were calculated and plotted with the R packages 'caret', 'pROC' (version 1.18.0), 'plotROC' (version 2.3.0), 'precrec' (version 0.12.9) and 'ggplot2'. For the regression analysis, the R2 was calculated with 'caret' (version 6.0.93).

2.5.4.2 Responder status prediction

In this study, metataxonomic and PCR data were collected from patients immediately after the loading dose of the live biotherapeutic intervention (LACTIN-V). Patients were classified as "responders" or "non-responders" based on the presence or absence of a specific primer for LACTIN-V in these samples. Based on these criteria were identified 47 "responders" and 13 "non-responders". The aim of the following analysis is to explore whether metabolic profiling data can be used to predict the response to therapy before it starts.

A linear model was used to select significant features between responders and non-responders. The formula used was the following:

$$lm(log(Metabolite) + 1) \sim responders\ VS\ NonResponders\ class + cst + bmi + gestational_age + ethnicity)$$

This formula was applied to all the metabolic features in the dataset.

Using only significant features (p-value <0.05) a random forest classifier was applied to the selected subset of data. RF model was performed using the 'randomForest' package (version 4.7.1.1) and the 'caret' package (version 6.0.93) with the same parameters described in 2.6.1.1. In the attempt to improve the model, the RF classifier was also applied to a subset of patients that were *Lactobacillus* dominant at timepoint 0 (responders = 40, non-responders = 10). The same was not possible for *Lactobacillus* depleted group due to low numbers (responders = 7, non-responders = 3).

2.5.4.3 Immune status prediction

Linear models were used to investigate the association between immune markers and metabolites measured in FLIP cohort samples. Specifically, the formula used was:

$$lm(log(ImmuneMarker + 1) \sim Metabolite)$$

All the immune marker - cytokine combinations were investigated, and the corresponding F-test P-values calculated. Benjamini–Hochberg false discovery rate correction was used for each analysis, with a 5% FDR cut-off.

For the predictions of immune markers levels from metabolic features, random forest classifiers were used. Initially, the immune marker concentrations were log-transformed after addition of

a constant offset log(x + 1). All random forest models were trained in R, using the 'randomForest' in combination with the 'caret' package (version 6.0-86), for model fitting, performance metric calculation and cross-validation. RF classifiers were trained as described in paragraph 2.6.1.1.

The relation between predicted and measured levels of immune markers was plotted as a scatterplot using 'ggplot2' (version 3.4.0).

Chapter III

Results presented in this chapter have been published in Pruski, P., Correia, G.D.S., Lewis, H.V., Capuccini K et al. Direct on-swab metabolic profiling of vaginal microbiome host interactions during pregnancy and preterm birth. Nat Commun 12, 5967 (2021).

3 RAPID MICROBIOTA AND INFLAMMATION PREDICTION USING DIRECT SWAB ANALYSIS BY DESI-MS

3.1 Introduction

The key role of microbiota in shaping the reproductive tract pathophysiology is now firmly established. While a high diversity microbiota is generally considered a hallmark of health in the gut, in the vagina, optimal health is typically associated with low bacterial diversity and predominance of *Lactobacillus* species. Increased vaginal bacterial diversity and depletion of *Lactobacillus* species are characteristic of bacterial vaginosis (BV) and are associated with higher incidence of sexually transmitted infections including HIV and HPV [212], [460], [461], lower efficacy of treatments, higher rates of IVF failure [462], miscarriage and PTB [153], [240], [463].

Bacteria contribute significantly to the chemical and metabolic milieu of the CVF. Bacteria use human-derived nutrients for their metabolism and their by-products shape the vaginal metabolome (i.e., the complete set of small molecules in the environment). Consistent with this, several studies have shown that BV is associated with a CVF metabolic signature that is distinct from women with optimal *Lactobacillus* dominated community compositions [327], [347]. For example, women with BV have lower vaginal levels of amino acids and higher levels of amino acid catabolites, higher levels of amines such as putrescine, cadaverine, and tyramine and increased concentrations of SCFAs such as succinate, acetate, and propionate. *Lactobacillus* dominance of the vagina is associated with high levels of intact amino acids, dipeptides and lactate [327].

There is also substantial evidence that bacterial metabolites play an important role in mediating inflammatory responses in the cervicovaginal niche. For example, Mirmonsef *et al*. showed that SCFAs produced by BV-associated bacteria can induce pro-inflammatory cytokine production [464]. High levels of both acetate and butyrate induce IL-8, IL-6, and IL-1β

cytokine production by vaginal epithelial cells and acetate and propionate are activators of neutrophil chemotactic response and stimulation of IFN-γ [465]. Moreover, Al-Mushrif *et al*. showed that succinate inhibits chemotactic responses, which is proposed to be one of the mechanisms by which BV-associated bacteria avoid activation of neutrophils in the vaginal niche. In contrast, production of lactate by *Lactobacillus* species is associated with low levels of inflammation [466].

During pregnancy, untimely activation of inflammation in the reproductive tract is associated with PTB. As demonstrated in several animal models, this can occur when pathogenic bacteria ascending from the vaginal canal are recognised by antigen presenting cells causing a pro-inflammatory cascade involving the activation of pro-inflammatory transcription factors that subsequently upregulate pro-labour and pro-tissue remodelling pathways [149]. In humans this is supported by data in patients receiving cervical cerclage using a braided suture material, which has recently been shown to promote vaginal bacterial dysbiosis. Concurrent with a change toward high diversity vaginal bacterial communities enriched by pathogens, these patients experience increased local production of immune mediators like MBL, C3b, C5a, IL-8, IL-6, and IL-1β followed by increased cervical vascularisation, which is an indicator of cervical ripening [452].

Given the importance of microbiota-host interactions in the context of preterm birth risk, there exists an unmet clinical need to provide rapid bedside diagnostic assessment of vaginal microbiota composition and host inflammatory status. However, current diagnostic methods rely on a combination of clinical observations, culture and microscopy techniques [285]. These are highly subjective, time consuming and fail to capture information regarding local inflammatory activation. While molecular-based tools like NGS provide insight into microbial community composition, their associated cost and time, as well as the fact that they cannot provide information about host response. limits their application in clinical settings.

Recently, our laboratory has developed Direct Swab Analysis by DESI-MS, which provides a detailed chemical signature directly from a routine clinical swab within minutes [467]. This method uses a form of ambient mass spectrometry [20],[21] whereby a pneumatically assisted charged electrospray of solvent is directed onto the surface of a rotating medical swab that dissolves analytes in the sample. New drops impacting the swab surface cause the release of

secondary droplets containing analytes, which are ionised and directed towards the spectrometer where they are characterised and quantified in less than a few minutes.

3.2 Hypothesis and aims

The work presented in this chapter primarily set out to test the hypothesis that vaginal microbiota-host interactions are reflected in the metabolic profile of the cervicovaginal fluid, which can be rapidly and accurately assessed using Direct Swab Analysis by DESI-MS. The secondary objective of the work was to investigate if the CVF metabolome captured by DESI-MS can provide information useful for monitoring microbiota and host response to treatments received during pregnancy (e.g. cerclage) which could modify their risk phenotype.

To test this, Direct Swab Analysis by DESI-MS was performed on a large cohort of women sampled throughout pregnancy (VMET 2, n = 205; 573 swabs). Metataxonomic profiling of the samples was performed using16S rRNA sequencing data and immuno-profiling data was performed using multiplexed Luminex immuno-assays. For each sample, metabolic profiles, metataxonomic and immunological data were integrated to identify metabolic signatures predictive of microbial composition and local inflammatory status.

3.3 Results

3.3.1 Sample cohort

A total of 205 pregnant women at high risk of PTB recruited for the VMET2 cohort were included in this study. Swab samples were collected longitudinally during pregnancy at approximately three different timepoints. Each of the 573 swabs collected was analysed by DESI-MS in both negative and positive ion modes. Additionally, the same number of swab samples collected from the same cohort of women was used for metataxonomic profiling of vaginal bacteria using 16S rRNA sequencing and a subset of these (n= 391) also underwent immune profiling (see Figure 3.1). Further information regarding the cohort demographic can be found in Table 3.1. The PTB rate in this high-risk population was of 21% (n = 44).

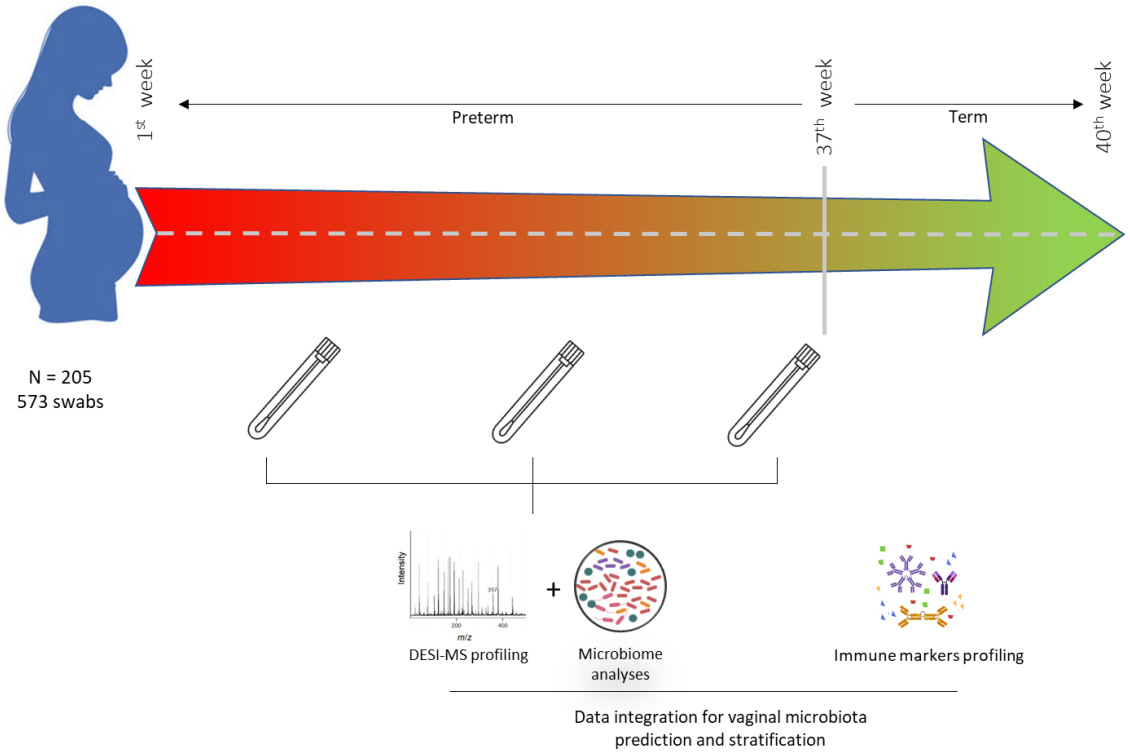


Figure 3.1 – Schematic representation of VMET2 study design.

A total of 205 women at high risk of PTB were sampled longitudinally during their pregnancy. The swab samples collected were used for metabolic profiling through DESI-MS, microbial composition analysis through 16s rRNA sequencing and a subset of them (n= 391) for immune markers profiling. Integration of these data was used to build a robust prediction of vaginal microbial composition.

To evaluate the sampling consistency at each of the three time points across gestation a plot of the distribution of sampling as a function of weeks of gestation was generated (Figure 3.2). The median gestational age of sampling was 14.6 (min 10.5 – max 18.5), 21.3 (19 – 26.3) and 30.5 (27.3 – 36) weeks for timepoints 1, 2 and 3 respectively.

In addition to these samples, an independent sample set derived from 160 pregnant women previously recruited as part of the VMET study were used for validation. A total of 455 swabs were collected from this group and were used for metabolic profiling both with DESI-MS and LC-MS. This enabled cross-platform performance comparison of the predictive metabolomic models generated from DESI-MS. Metataxonomic profiling was also performed on these samples. Data was generated from the VMET validation cohort by Drs Pamela Pruski and Holly Lewis.

Table 3.1 - Demographic characteristics of VMET 2 cohort

	IPTB	SPTB	TERM	TOTAL
Subject, n	3 (1%)	41 (20%)	161 (79%)	205
Samples, n	9	105	459	573
Mean maternal age, years ± sd (range)	38.0 ± 3.1 (34.0–41.0)	32.7 ± 5.9 (20.0–44.0)	33.2 ± 4.7 (20–43)	33.3 ± 4.9 (20–44)
Mean sample gestation, weeks (±sd range)	23.4 ± 7.2 (15.3–34.5)	21.5 ± 6.4 (12.2–32.6)	22.4 ± 6.8 (11.0–35.6)	22.3 ± 6.8 (10.5–36)
Mean bmi, kg/m2 (±sd, range)	28.9 ± 3.5 (24.0–32.0)	25.5 ± 5.6 (20.0–46.1)	25.3 ± 5.1 (18.1–39.9)	25.3 ± 5.1 (18.1–46.1)
Ethnicity, n/n (%)				
Caucasian	2/130 (2%)	20/130 (15%)	108/130 (73%)	130/205 (63%)
Asian	0/28 (0%)	8/28 (29%)	20/28 (71%)	28/205 (14%)
Black	1/35 (3%)	10/35 (29%)	24/35 (69%)	35/205 (17%)
Others	0/12 (0%)	3/12 (25%)	9/12 (75%)	12/205 (6%)
Not recorded*	0/0 (0%)	0/0 (0%)	0/0 (0%)	0/205 (0%)
Gravida				
0	0/1 (0%)	1/1 (100%)	0/1 (0%)	1/205 (0%)
1	0/45 (0%)	5/45 (11%)	40/45 (89%)	45/205 (22%)
2	1/59 (2%)	15/59 (25%)	43/59 (73%)	59/205 (29%)
>3	2/100 (2%)	20/100 (20%)	78/100 (78%)	100/205 (49%)
Previous mid trimester loss/ptb				
No	1/86 (1%)	7/86 (8%)	78/86 (91%)	86/205 (42%)
Yes	2/119 (2%)	34/119 (29%)	83/119 (70%)	119/205 (58%)
Previous excisional cervical treatment				
No	0/105 (0%)	15/105 (14%)	90/105 (86%)	105/205 (51%)
Yes	1/100 (1%)	28/100 (28%)	71/100 (71%)	100//205 (49%)
Cervical cerclage material				
Braided	1/26 (4%)	13/26 (50%)	12/26 (46%)	26/205 (13%)
Monofilament	0/30 (0%)	8/30 (27%)	22/30 (73%)	30/205 (15%)
Abdominal	0/2 (0%)	0/2 (0%)	2/2 (100%)	2/205 (1%)
No intervention	2/142 (1%)	18/142 (13%)	122/142 (86%)	142/205 (69%)
Not recorded*	0/5 (0%)	2/5 (40%)	3/5 (60%)	5/205 (2%)

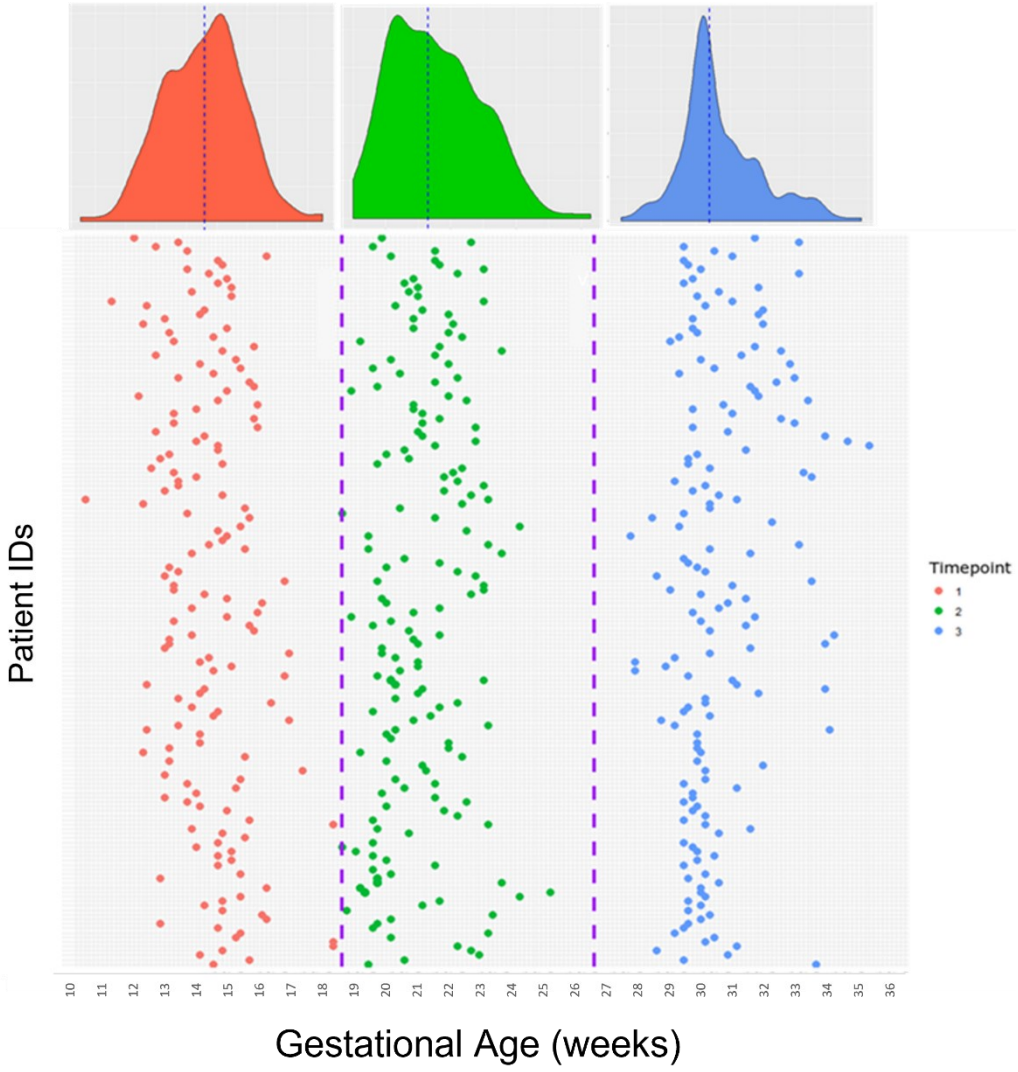


Figure 3.2 – Scatterplot describing the distribution of gestational age across three sampling timepoints. Each row of the plot represents one patient sampled longitudinally across three different timepoints. Timepoint 1 (shown in red) has a minimum gestational age of 10.5 weeks, a maximum of 18.5 weeks and a median of 14.6 weeks. Timepoint 2 (shown in green) has a minimum gestational age of 19 weeks, a maximum of 26.3 weeks and a median of 21.3 weeks. Timepoint 3 (shown in blue) has a minimum gestational age of 27.3 weeks, a maximum of 36 weeks and a median of 30.5 weeks.

3.3.2 Analysis of swab shape, geometry, and physical characteristics

To monitor the potential impact of swab characteristics (e.g. deformation of shape) on the quality of spectral data, photos of each swab sample were taken and visual inspection performed. This included assessment of the shape and colour of the swab compared to an unused swab. Three examples of different swab characteristics are shown in Figure 3.3. Most vaginal swabs were light-yellow in colour with rounded and tightly woven tips (Figure 3.3A). Occasionally, evidence of swab tip deformation (Figure 3.3B) or blood presence (Figure 3.3C) was noted. In total, 88 swabs had signs of deformation (~15%) and 16 swabs presented visible traces of blood (~2.5%).



Figure 3.3 – Representative images of different swab shapes and characteristics. (A) The vast majority of swabs analysed in this project had a tightly woven and consistent tip morphology. Some swabs were found to have (B) deformed tip morphology or (C) contain blood.

3.3.3 Data acquisition and preliminary confounders analyses

Acquisition of DESI-MS data was performed for approximately 2 minutes/swab consistent with previously published methods. Following data processing, this generated an average of 1925 peaks in negative mode and 1716 peaks in positive mode per sample. Assessment of the resulting spectral profiles was first performed using unsupervised multivariate modelling approaches to evaluate the potential presence of batch effects. PCA of unscaled spectral data showed that the majority of variation in the metabolic profiling data was not attributable to the hospital of origin (Figure 3.4 A). Similar results were obtained when data was modelled for maternal age, BMI, ethnicity, and gestational age at sampling (data not shown). A proportion of variance could however be explained by the date of spectral acquisition. This is mostly observable along the second component and represents a minor fraction (4.9 %) of the total variability of the dataset (Figure 3.4 B). Models produced using unit variance scaling (to provide even weighting to all variables in the model - data not shown), produced similar findings.

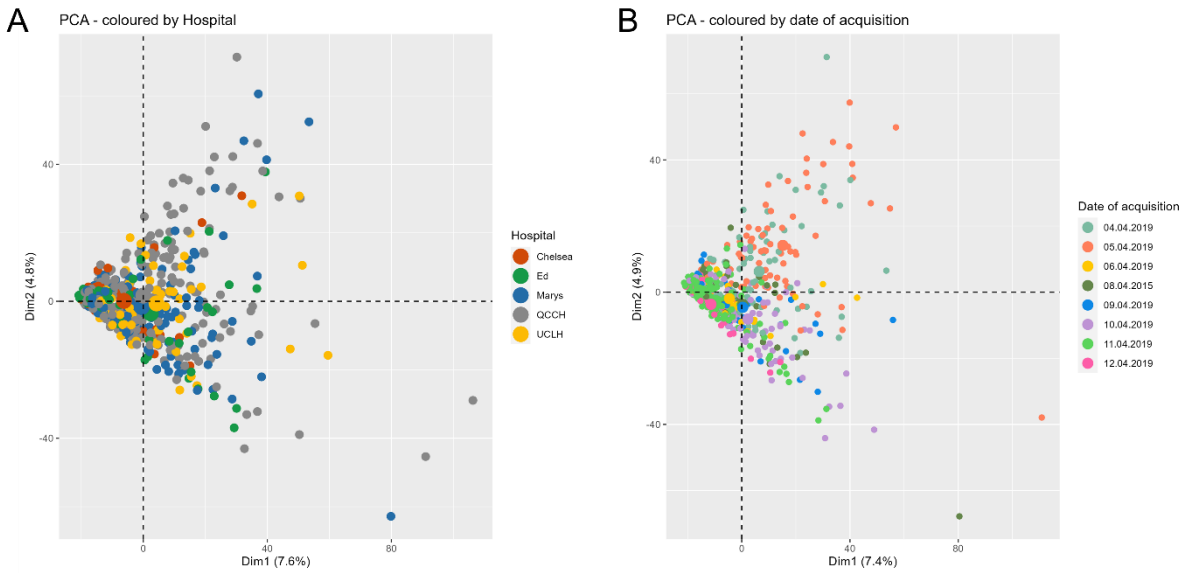


Figure 3.4 - PCA coloured according to different possible confounders. PCA of the non-scaled data coloured according to date of acquisition of the spectrum (A), hospital where the sample was taken (B). The cumulative variance explained by the first two components of this PCA is 12%.

Supervised analysis of the DESI-MS profiling data using a supervised modelling approach in the form of PLS was also performed to determine if any changes in the CVF metabolome associated with advancing gestation were detectable (Figure 3.5). The resulting PLS score plot showed some trend of separation of samples according to time point of sampling, however this accounted for only approximately 8.4% of the observed variability in the dataset.

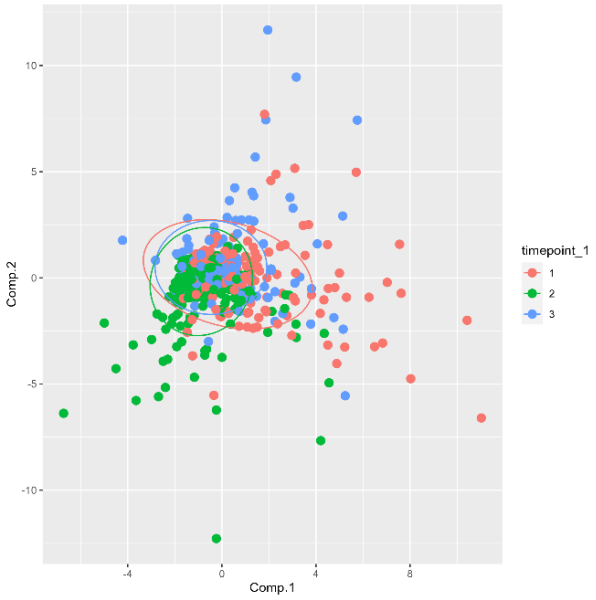


Figure 3.5 - PLS-DA coloured according to time points. Supervised PLS analysis shows that advancing gestation has a relatively minor impact on metabolome accounting for approximately 8.4% of explained variance in the model (cumulative percentage).Accuracy = 0.44, Sensitivity = 0.49, Specificity = 0.74, PPV = 0.49, NPV = 0.75.

3.3.4 Characterisation of vaginal microbiota composition using metataxonomics

The microbiota composition of the vaginal samples in the patient cohort was first grouped on the basis of relative abundances using Ward's linkage hierarchical clustering analysis with Jensen-Shannon distance metric of bacterial species (Figure 3.6). The hierarchical clustering was based on the operational taxonomic unit (OTU) counts aggregated by summing at species level. The optimal number of groups or clusters was determined using Silhouette scores which lead to a total of 11 groups being identified. To facilitate comparison of results with the validation cohort VMET and previously published literature, the number of clusters were reduced to 6 for the VMET2 samples. CST groupings including in the subsequent analyses were as follows: CST I (*L. crispatus*); CST II (*L. gasseri*); CST III (*L. iners*); CST IV (*Gardnerella vaginalis*); CST V *(L. jensenii*) and CST VI (*Bifidobacterium breve*). Samples were also grouped at genus level into *Lactobacillus*-dominant (LDOM) or *Lactobacillus*-depleted (LDEPL) categories. LDOM group included samples classified as CST I, II, III and V (n=427 samples), and *Lactobacillus* depleted (LDEPL) included samples from CST IV and CST VI (n=112 samples). Additionally, scores of fit for each CST were generated for each sample with the lowest mean scores observed for the CST IV cluster.

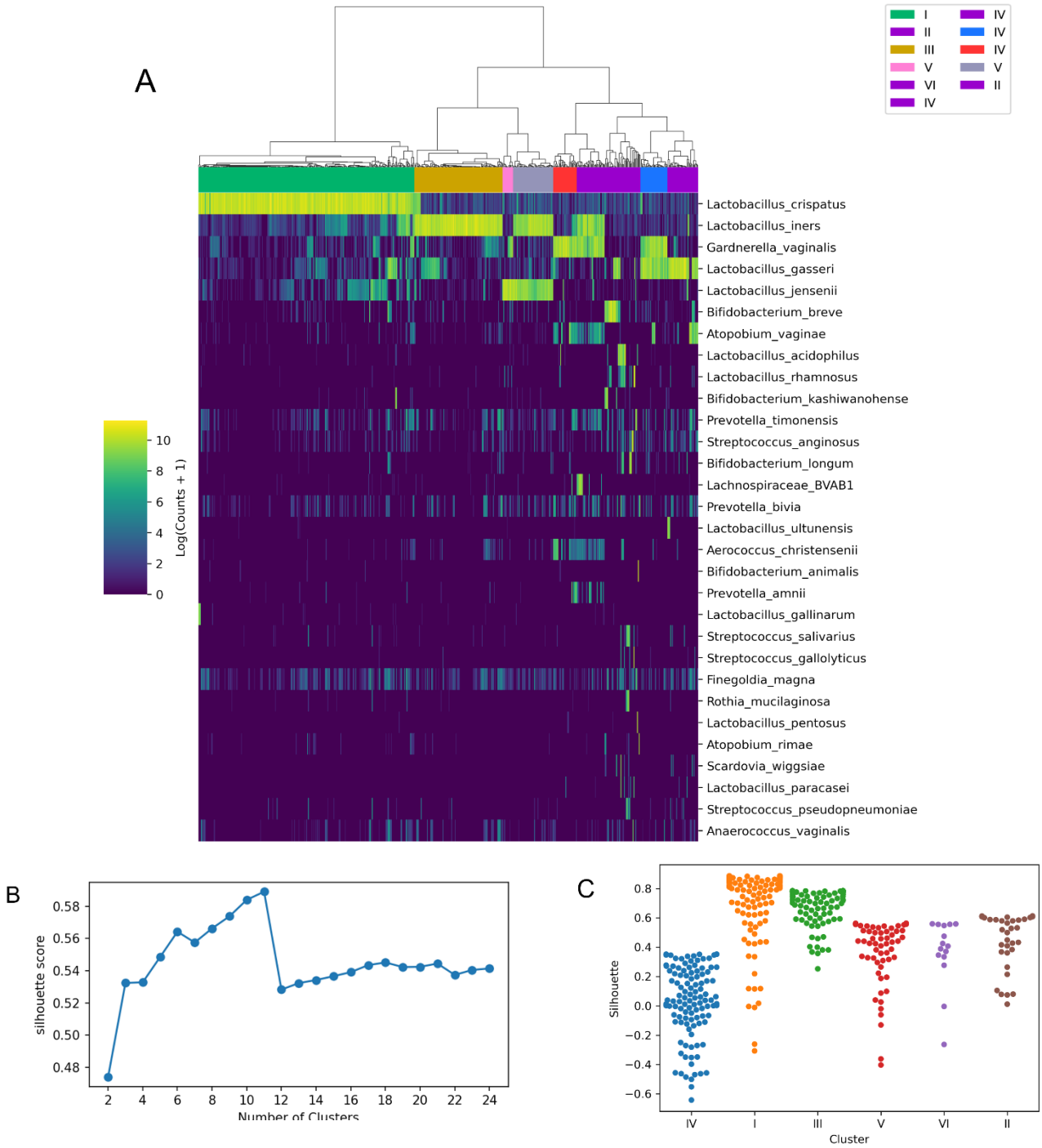


Figure 3.6 Hierarchical clustering analysis of the 16S rRNA gene sequencing data from the VMET2 study cohort. A) Heatmap of log-transformed counts per microbial taxa found in the vaginal bacterial communities. Hierarchical clustering of samples was based on their raw OTU counts aggregated at species level, using Ward's method and the square-root of the Jensen-Shannon divergence as distance metric. B) Analysis of the average silhouette scores suggests an optimal (maximum average silhouette score) number of 11 clusters. Clusters were further grouped into a smaller set of 6 clusters consistent with previously described community state types. C) Sample specific silhouette score values representing how each sample fits into the final cluster.

3.3.5 Integration of DESI-MS profiles and vaginal microbiota compositional data

The metabolic features measured through DESI-MS, both in positive and negative ion modes, were input into a linear mixed effect model to identify features significantly discriminatory between LDOM and LDEPL samples (Figure 3.7). A total of 113 metabolite features were found to be significantly different between the two groups. A complete list of the 113 identified metabolites is presented in the Appendix Table 1. The majority of these metabolites were found to not change concentration across gestation (data not shown). A permutational multivariate analysis of variance (PERMANOVA) was also performed to estimate the proportion of total metabolome variance explained by the microbial CST, gestational age, BMI and maternal age (Appendix Table 2). This analysis showed that inter-individual variability explained the majority of variance in the DESI-MS metabolic profiles (35 – 45.2%) followed by CST (2.2–25.2%). Ethnicity (0.5–2.4%) and gestational age (0.4–1.8%) accounted for a relatively small proportion of variance in the metabolic profiles in the dataset.

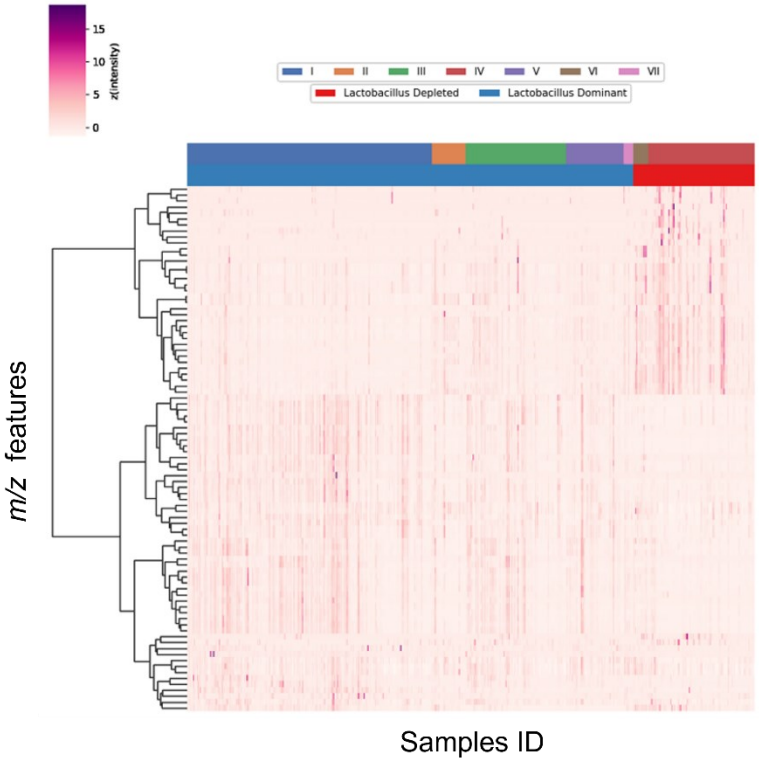


Figure 3.7 - Heatmap of DESI-MS derived discriminatory metabolic features and their relationship with Lactobacillus dominance or depletion. Heatmap representing relative concentrations of negative ion mode DESI-MS metabolic features (shown in rows, n = 113) significantly differing between *Lactobacillus* spp. dominated (in blue) and *Lactobacillus* spp. depleted (in red) samples (shown in columns).

Annotation and identification of differentially abundant metabolites was performed using MS/MS experiments with selected metabolites presented in Figure 3.8. The features that were more abundant in the LDEPL group included hydroxycaproic acid, isoarachidic acid, phytenic acid, nevronic acid, oleic acid, phenyllactic acid and phospopantotenic acid. Features found to be higher in the LDOM group included isoleucyl-serine, Valyl-Alanine, N-acetylputrescine, 5-aminopentanamide, N-palmitoyl-cysteine and Thiomalic acid.

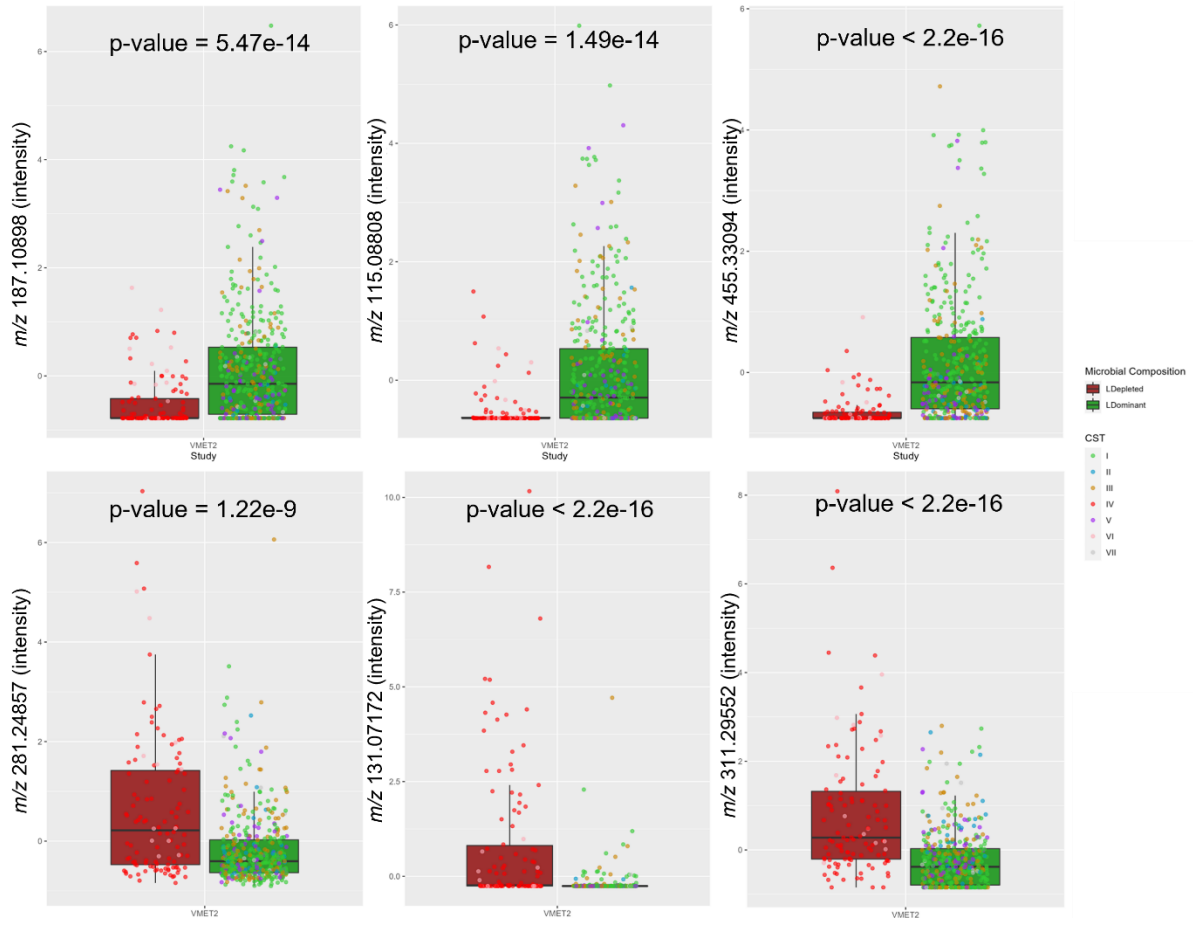


Figure 3.8 - Examples of metabolic features significantly differing between Lactobacillus dominant and Lactobacillus depleted samples. P-values reported in the figure were calculated using a Wilcoxon rank sum test. The interior horizontal line represents the median, with the upper and lower bounds of the box representing the 25th and 75th percentile values, respectively. Whiskers are drawn from the corresponding box boundary to the most extreme data point located within the box bound ± 1.5 × IQR (interquartile range).

We next examined if MS features identified in both positive and negative DESI-MS mode, could be used to predict the microbial composition of the vaginal mucosa. For this, a random forest classifier model was built using the intensity values of each feature and the performance of the model was assessed with ROC curve analysis. The resulting ROC curves are shown in

Figure 3.9. Robust predictions (AUC>90) could be obtained for the classification of the samples at genera level (LDOM vs LDEPL) particularly using features identified in negative ion mode (AUC = 90.6, sensitivity = 54.5) compared to positive ion mode (AUC = 83.1, sensitivity = 44.1). Similar results were obtained for models discriminating between the major vaginal CSTs. More specifically CSTI vs CSTIII AUC_{neg} = 97.2, AUC_{pos} = 70.8 (sensitivities 96.4 and 93.5 respectively), CSTI vs CSTIV AUC_{neg} = 95.6, AUC_{pos} = 87.7 (sensitivities 92.4 and 93.7 respectively), CSTI vs CSTV AUCneg= 95.9, AUCpos = 77.4 (sensitivities 74.4 and 19.7 respectively), CSTIII vs CSTIV AUCneg = 92.2, AUCpos= 86.8 (sensitivities 85.7 and 83.5 respectively), CSTIII vs CSTV AUCneg = 86.5, AUCpos= 75.5 (sensitivities 91.6 and 92.3 respectively).

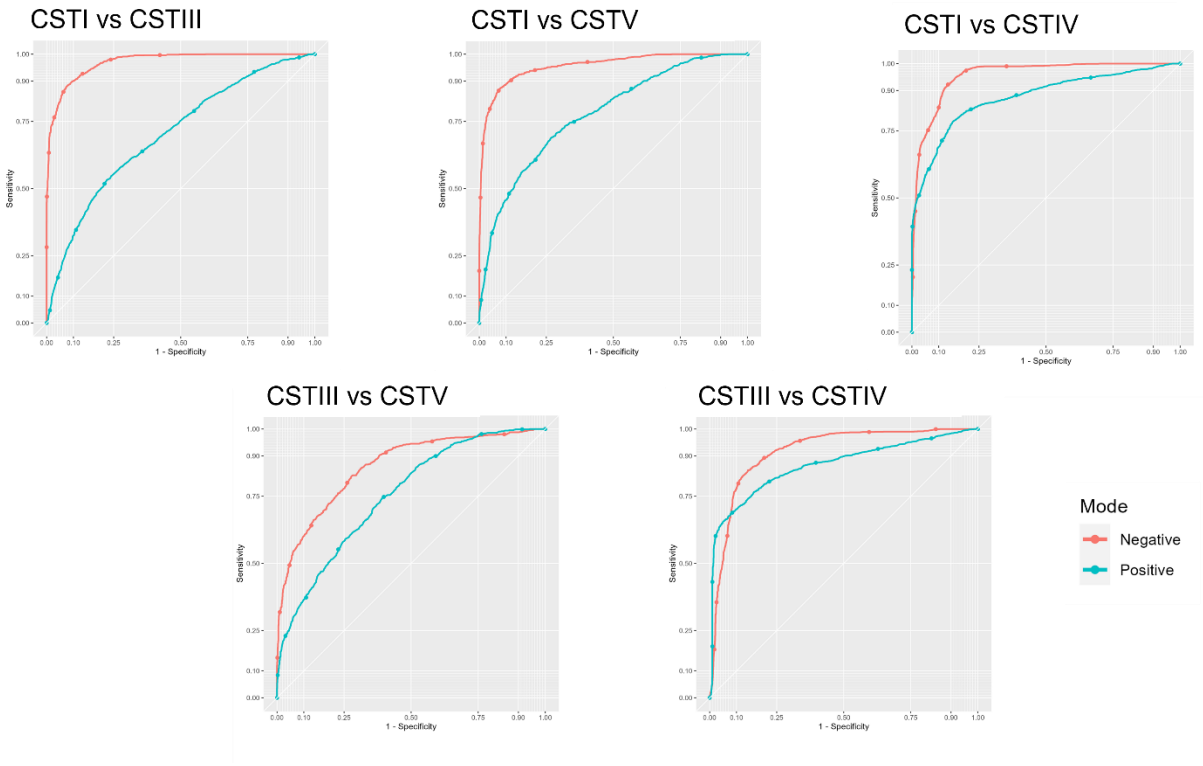


Figure 3.9 - ROC curves comparing DESI-MS classification performance between different vaginal CSTs. ROC-curve analysis showing performance of direct swab analysis by DESI-MS operating in both negative and positive ion polarity modes to predict various CSTs contrasts in VMET2. CSTI vs CSTIII AUCneg = 97.2, AUCpos = 70.8 (sensitivities 96.4 and 93.5 respectively), CSTI vs CSTIV AUCneg = 95.6, AUCpos = 87.7 (sensitivities 92.4 and 93.7 respectively), CSTI vs CSTV AUCneg= 95.9, AUCpos = 77.4 (sensitivities 74.4 and 19.7 respectively), CSTIII vs CSTIV AUCneg = 92.2, AUCpos= 86.8 (sensitivities 85.7 and 83.5 respectively), CSTIII vs CSTV AUCneg = 86.5, AUCpos= 75.5 (sensitivities 91.6 and 92.3 respectively).

3.3.6 DESI-MS prediction of host immune response

To investigate the relationship between DESI-MS metabolic profiles and CVF immune markers, a subset of samples from VMET2 cohort (n=391) was used to measure a panel of 22 soluble markers including cytokines, chemokines, immunoglobulins, and component system molecules. Specifically, levels of the following mediators where determined: IFN-y, IL-10, IL-2, IL-3, IL-5, TNF-a, GM-CSF, IL-1b, IL-18, IL-4, IL-6, IL-8, C5, C5a, MBL, C3b/iC3b, IgA, IgE, IgG1, IgG2, IgG3, IgG4, and IgM. DESI-MS features were then used to predict the log transformed concentrations of each immune marker using random forest regression. Cross validation was performed, and robust prediction was defined as cross-validated R^2 values> 0.25. Strong prediction of the log concentration values could be demonstrated for IL-1b, (CV R^2= 0.51), IL-8 (CV R^2= 0.37), C3b/iC3b (CV R^2= 0.31), IgG3 (CV R^2= 0.31) and MBL (CV R^2= 0.26) (Figure 3.10 A).

Immune markers (n=10) and DESI-MS features (n=23) with CV $R^2 > 0.1$ for the linear regression were selected and presented in a heatmap (Figure 3.10 B). A strong relationship between the metabolic signature and the immune phenotype could be observed that was characterised by higher levels of long chain fatty acids, glycerophospholipids and ceramides associated with higher local inflammation.

To further evaluate the ability of DESI-MS to predict the inflammation markers levels in the cervicovaginal fluid, both the log transformed value of the measured immune-assay measurements and the predicted value obtained using the significant DESI-MS features previously identified through RF analysis were plotted (Figure 3.11). A linear regression line was fitted between the immune marker concentrations and the correspondent predicted values. For each of the examples shown, a robust correlation between the values was observed indicating the capacity of DESI-MS to predict immune status. Additionally, each point of the plot representing a sample was coloured according to the classification in LDOM or LDEPL obtained from sequencing data and a boxplot of the predicted value for each of the two groups was also produced. This highlighted a strong relationship between increased levels of inflammatory markers in *Lactobacillus* depleted communities.

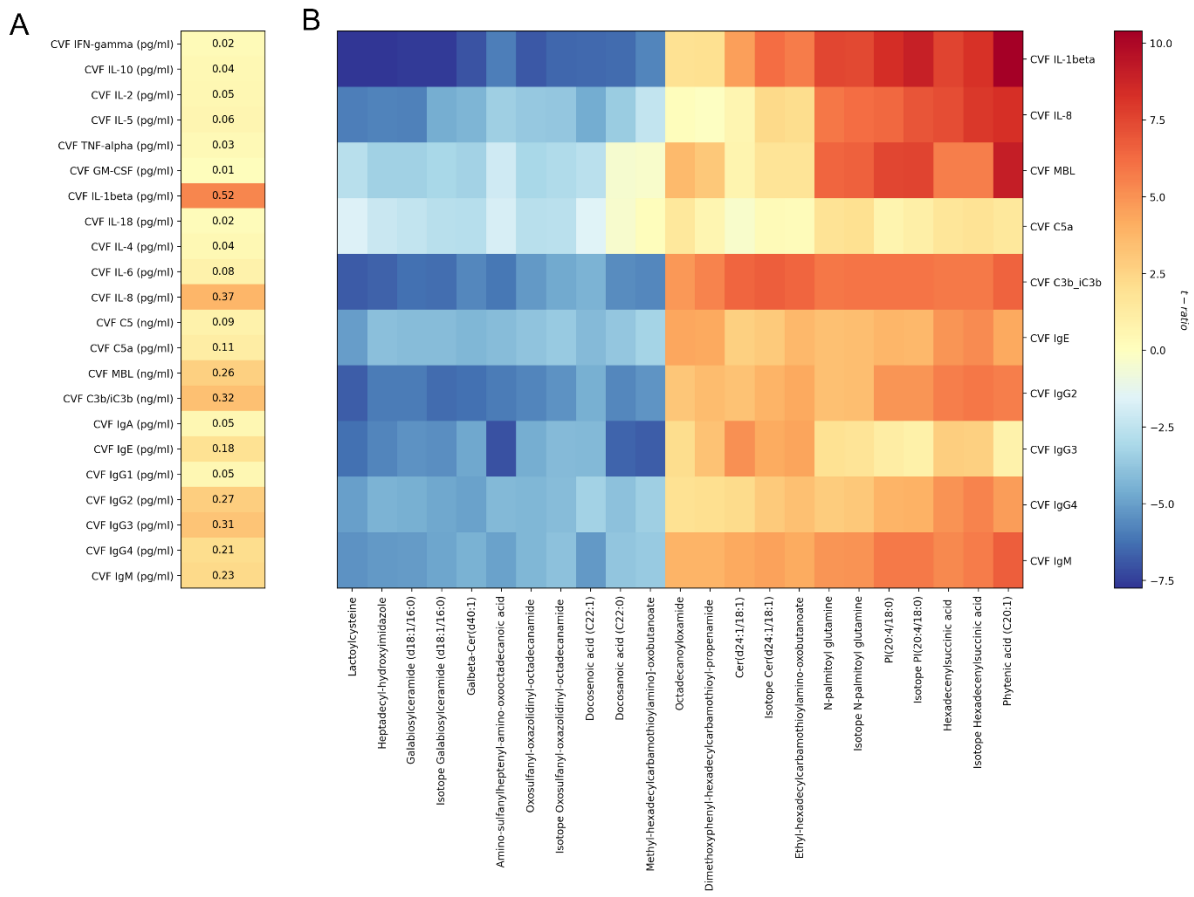


Figure 3.10 - Assessment of correlation between cervicovaginal metabolic features and immune mediators. A) Cross-validated R2 value for all 22 corresponding measured immune mediator concentrations. B) Heatmap of top 23 significantly correlated metabolite features with top 10 immune mediators (IL-1β, IL-8, MBL, C5, C3b/C3bi, IgE, IgG2, IgG3, IgG4, IgM). t-ratio ranges from +10 (red) to −7.5 (blue).

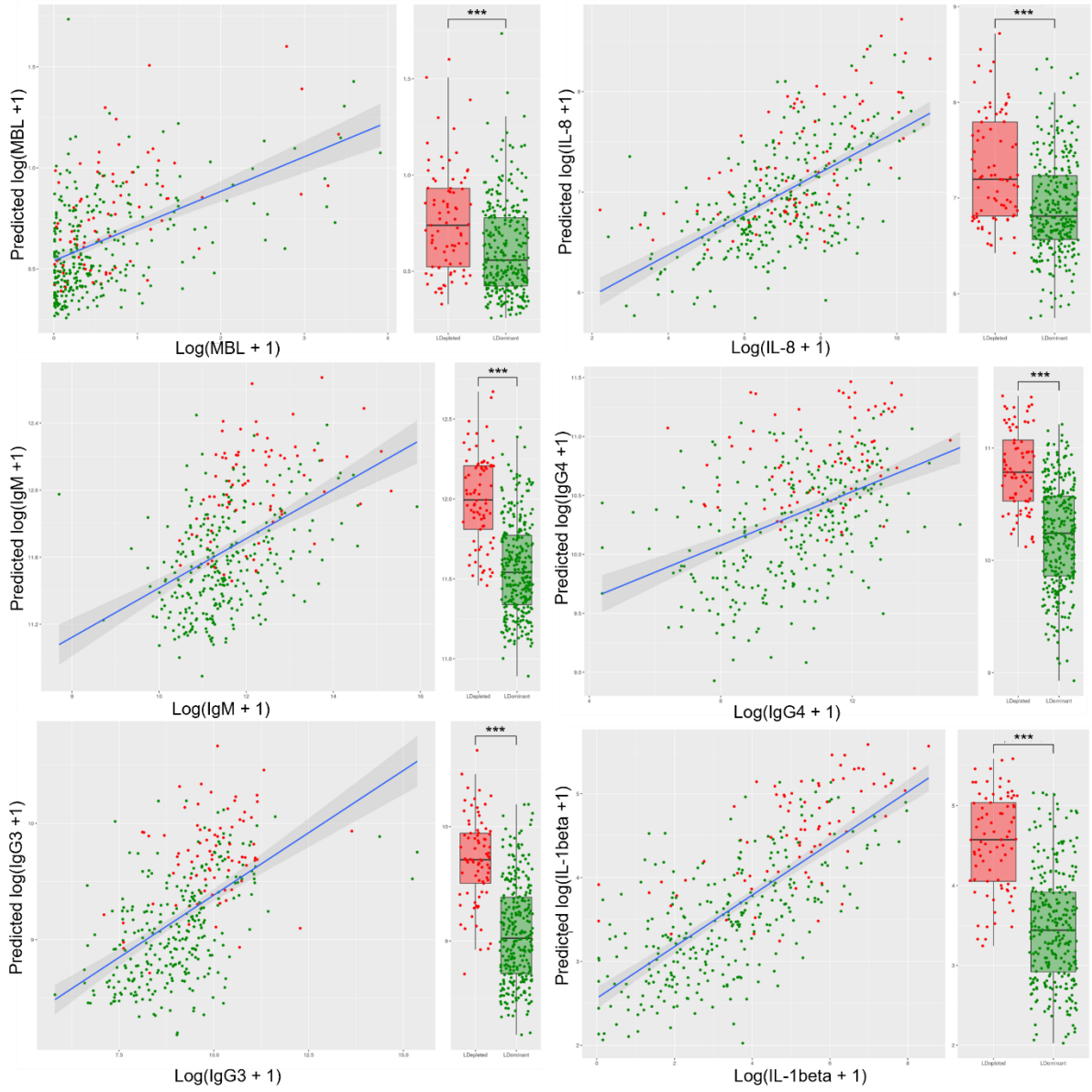


Figure 3.11 - Association between predicted log-transformed value of immune markers by DESI-MS and measured log-transformed values by multiplexed immunoassay. IL-1β (R2 = 0.51), IL-8 (R2 = 0.37), C3b/iC3b (R2 = 0.33), IgG3 (R2 = 0.31), IgG2 (R2 = 0.27), MBL (R2 = 0.26). A linear regression line was fitted to the log-transformed values and their corresponding prediction DESI-MS features derived from a random forest regression were used to predict the log transformed concentrations of each immune marker. A box plot of predicted immune marker levels for LDEP (red) and LDOM (green) samples is also presented. The lower, interior horizontal line, and upper bounds of the box represent the 25th, median and 75th percentile values, respectively. Whiskers are drawn from the corresponding box boundary to the most extreme data point located within the box bound ± 1.5 × IQR (interquartile range). P values are reported for a two-tailed Welch t-test for the difference in mean predicted immune markers between LDEP and LDOM (*** = p-value less than 0.001).

3.3.7 Levels of key metabolites in relation to inflammation markers

To test the plausibility that metabolite features identified by DESI-MS could correlate with local inflammatory status, the relationship between immune markers and selected metabolites with previously described microbiota-associated immune modifying functions in the cervicovaginal niche were examined. For this, means intensity values of features corresponding to lactate (*m/z* 89.02485758), succinate (*m/z* 117.0197366) and propionate (m/z 73.02988371) were first calculated and compared against vaginal CSTs (Figure 3.12). Propionate and succinate levels were found to significantly differ among CSTs (p-values 3.34e-7 and 1.84e-6 respectively). Specifically, succinate was significantly higher in CST IV vs CST I (p-value 0.0001), CST IV vs CST III (p-value <0.0001), and in CST VI vs CST V (p-value 0.0001). Similarly, propionate was significantly higher in CST IV vs CST I (p-value 0.04), CST II vs CST III (p-value 0.0052), CST II vs CST V (p-value 0.006), CST IV vs CST III (p-value 0.0001). and CST IV vs CST V (p-value 0.0003). Moreover, significant differences in the ratio of succinate to lactate and propionate to lactate in LDOM and LDEPL samples were observed (succinate/lactate p-value = 3.16e-7, propionate/lactate p-value = 3.648e-07). Specifically, a high succinate/lactate ratio was observed in CST IV vs CST I (p-value <0.0001), CST IV vs CST III (p-value <0.0001), and CST IV vs CST V (p-value 0.0003). A positive correlation between the succinate/lactate ratio and the concentration of several immune mediators was also observed including higher levels of succinate correlating with higher levels of inflammatory markers like IL1-β, IgM and some components of the complement cascade including C3b and C5 (Figure 3.13). See Appendix Figures 1, 2 and 3 and Appendix Table 3, 4, 5 for a full list of the linear models and corresponding R^2 and p-values.

.

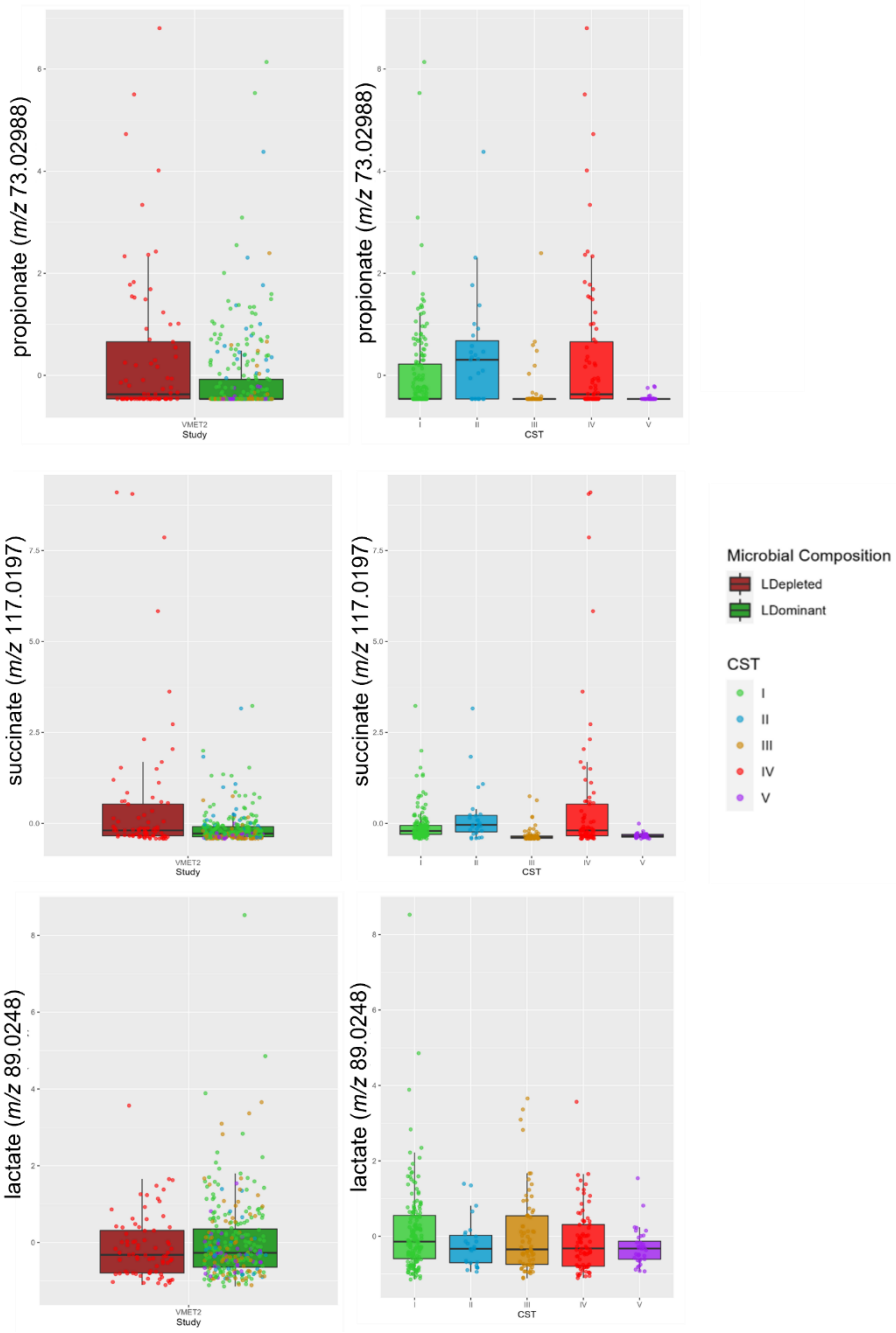


Figure 3.12 – Box and whisker plots showing levels of succinate (m/z 117.0197366), propionate (m/z 73.02988371) and lactate (m/z 89.02485758) in relation with microbial composition. Succinate was significantly higher in CST IV vs CST I (p-value 0.0001), CST IV vs CST III (p-value <0.0001), and in CST VI vs CST V (p-value 0.0001). Propionate was significantly higher in CST IV vs CST I (p-value 0.04), CST II vs CST III (p-value 0.0052), CST II vs CST V (p-value 0.006), CST IV vs CST III (p-value 0.0001). and CST IV vs CST V (p-value 0.0003). Lactate was not found to be significantly different among CSTs. All box and whisker plots are drawn with the lower, horizontal interior line, and upper bounds of the box representing the 25th percentile, median and 75th percentile values, and whiskers extending from the lower or upper box bonds to the position of the most extreme data point within ± 1.5 × IQR (interquartile range).

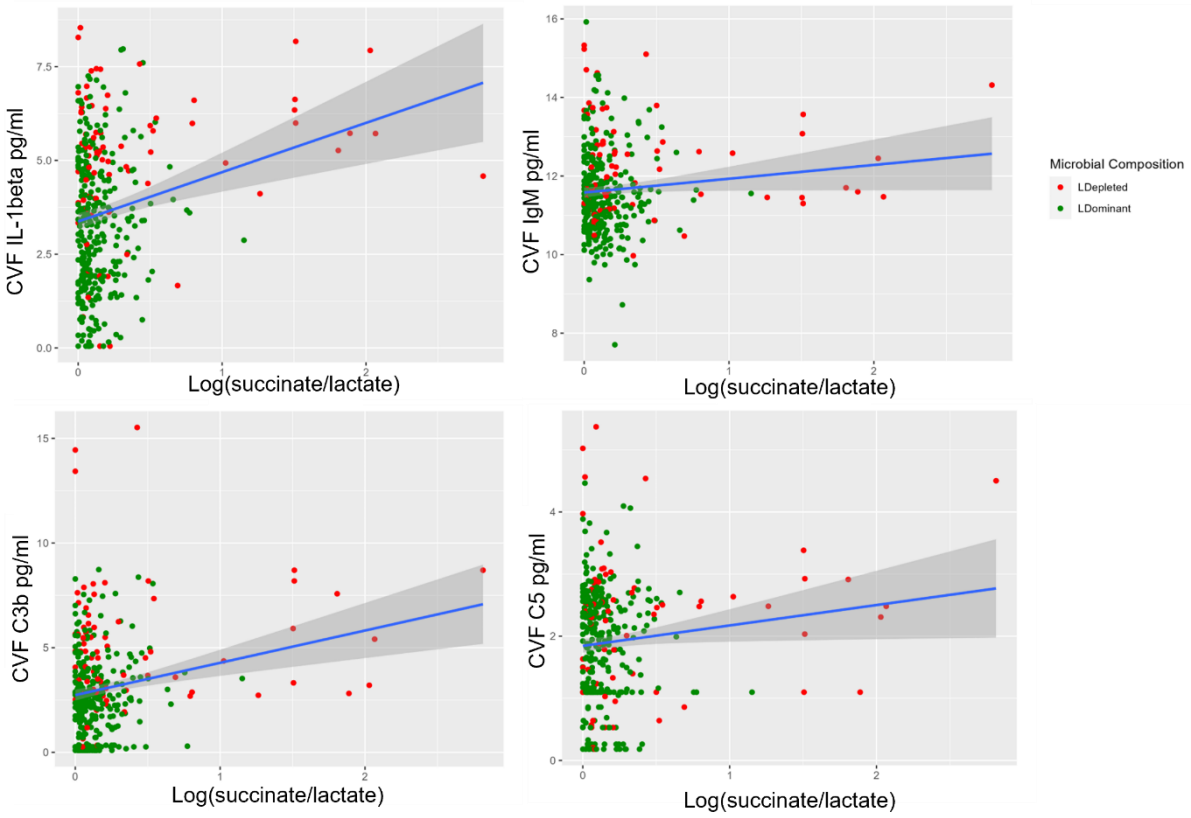


Figure 3.13 - Linear model showing correlation between succinate/lactate ratio and main immune markers (IL-1b, IgM, C3b, C5,) IL-1b R2 = 3.04e-2, p-value = 6.81e-4, IgM R2=1.63e-2, p-value 1.34e-2, C3b R2 = 4.18e-2, p-value = 6.61e-5, C5 R2 = 2.06e-2, p-value = 5.33 e-3 . The dot representing the samples in the plot were also coloured based on Lactobacillus dominance or depletion.

A similar relationship between the succinate:lactate ratio and inflammation was observed in samples from *Lactobacillus* depleted vaginal microbial communities. For this, samples were stratified into two different classes according to their immune marker levels (Figure 3.14). The "low" group was defined as samples that have an immune marker level equal or lower than the median value with the "high" group including samples with immune marker levels above the median value. A significantly higher mean succinate:lactate ratio value was observed in those samples categorised as "high" based on almost all of the assessed pro-inflammatory markers. However, the opposite relationship was observed in the case of the anti-inflammatory cytokine IL-6. Similar results are obtained when using propionate/lactate ratio. Results shown in Appendix Figure 5.

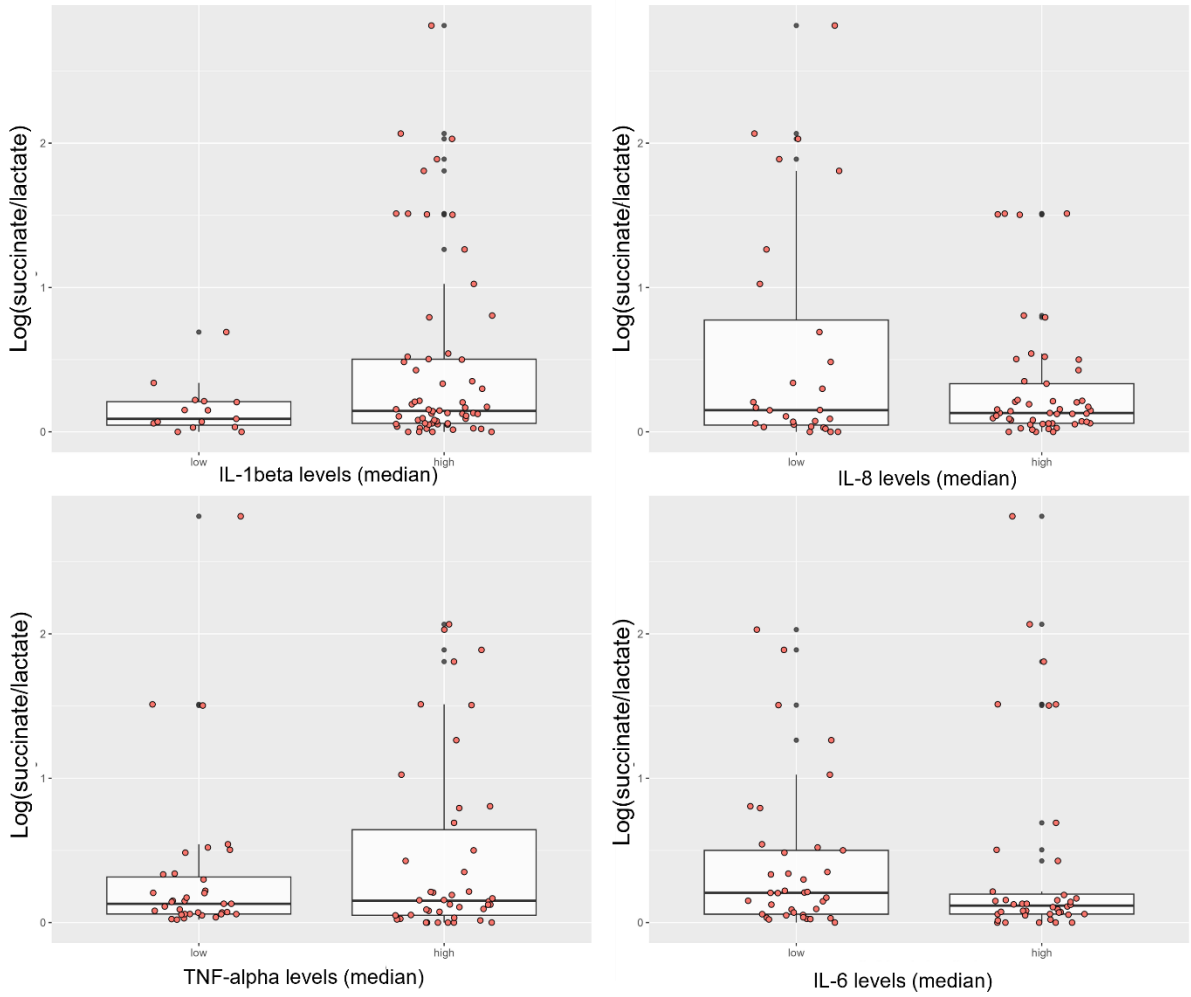


Figure 3.14 – Box and whiskers plot showing different levels of succinate/lactate ratio in high or low classes of immune markers (IL-1β, IL-8, TNF-α, IL-6) All box and whisker plots are drawn with the lower, horizontal interior line, and upper bounds of the box representing the 25th percentile, median and 75th percentile values, and whiskers extending from the lower or upper box bonds to the position of the most extreme data point within ± 1.5 × IQR (interquartile range).

3.3.8 Assessment of cervical vaginal inflammation and microbiota composition in term and preterm birth outcomes

We next examined the relationship between pregnancy outcome, microbiota composition and inflammatory status. We observe that both women that will deliver at term and women experiencing preterm birth have higher levels of IL-1β when they present a *Lactobacillus* depleted microbiota composition. More importantly, this relationship is true both when comparing levels of IL-1β, measured directly by immunoassay or levels predicted by DESI-MS. However, levels were significantly higher in preterm birth cases (preterm 59.63 pg/ml vs term 1.94 pg/ml, $p = 0.006$, Welch two sample t-test, Figure 3.15 A). The same relationship was observed for MBL, where increased levels, either measured directly by immunoassay or predicted by DESI-MS, were also observed in high-diversity cervicovaginal fluid of women subsequently delivering preterm (term 0.042 ng/ml vs preterm 0.692 ng/ml, $p = 0.007$, Welch two-sample t-test, Figure 3.15 B).

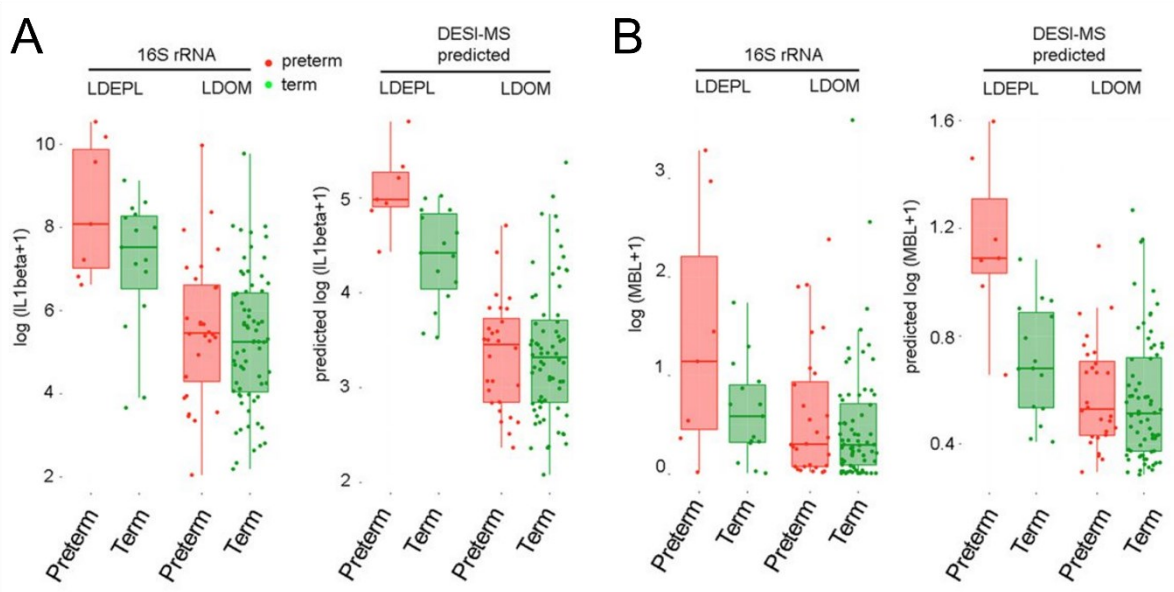


Figure 3.15 - Box and whiskers plot showing immunoassay measured or DESI-MS predicted levels of IL-1b in patient receiving cerclage with different materials. LDEPL vaginal composition was associated with increased IL-1β and MBL levels compared to LDOM. The highest levels were observed in LDEPL women subsequently experiencing preterm delivery. The same relationship for both IL-1β and MBL was observed also when levels and vaginal microbiota composition were predicted using DESI-MS ($n = 103$ pregnancies, 103 swabs).

3.3.9 Assessment of cervical vaginal inflammation in response to cervical cerclage

Finally we investigated the relationship between inflammation and pregnancy outcomes in women receiving a cerclage intervention. In particular we analysed the differences in suture material and we observed that IL-1β is higher in women delivering preterm after a braided cerclage compared to women delivering at term (n= 13 pregnancies). On the contrary there is no relationship between IL-1β and pregnancy outcomes following a monofilament cerclage (n = 21 pregnancies). When observing the levels of IL-1b predicted through DESI-MS was possible to observe the same relationship (Figure 3.16).

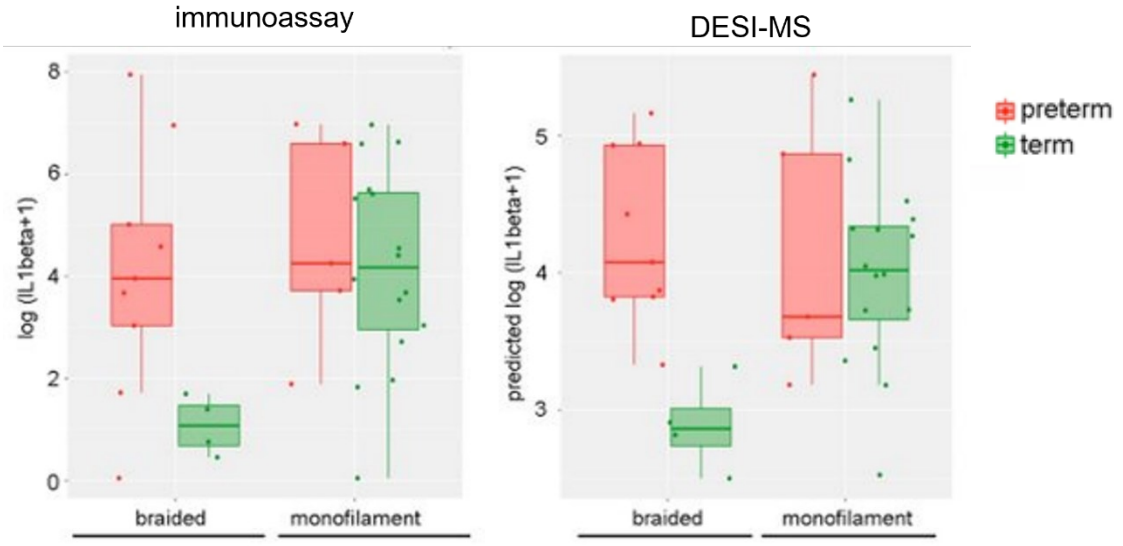


Figure 3.16 - Box and whiskers plot showing immunoassay measured or DESI-MS predicted levels of IL-1b in patient receiving cerclage with different materials. Preterm birth in women treated with cervical cerclage using braided cerclage material was associated with higher IL-1β levels compared to term birth outcomes (n = 13 pregnancies, 13 swabs), whereas no relationship between IL-1β levels measured or DESI-MS-predicted were observed with pregnancy outcome following cervical cerclage using monofilament material (n = 21 pregnancies, 21 swabs). All box and whisker plots are drawn with the lower, horizontal interior line, and upper bounds of the box representing the 25th percentile, median and 75th percentile values respectively. Whiskers extending from the lower or upper box to the position of the most extreme data point within ± 1.5 × IQR (interquartile range).

3.4 Discussion

The characterisation of cervicovaginal microbial communities for clinical purposes remains largely reliant on culture-based techniques and microscopy. Despite considerable advancements in metataxonomics and metagenomics-based approaches, the use of next generation sequencing is still largely limited to research contexts. This is partly due to the fact that these techniques are time-consuming and cost-prohibitive, but also because they only provide assessment of microbial composition and not that of host response, which is a critical determinant of disease pathology [303]. This is particularly important for the vaginal niche where bacterial composition has been shown to associate with pathology and disease in many populations, but not all. For example, the relationship between high microbial diversity and increased risk of preterm birth appears stronger in white populations than in black ones [233]. There is some evidence that this is determined by differences in tolerance and host response to specific bacteria that differs across ethnic groups [236]. These findings highlight the importance of analysing both vaginal microbiota composition and host response in considering the role of the vaginal microbiome in preterm birth risk phenotypes. The data presented in this chapter highlights direct swab analysis by DESI-MS as a novel and exciting method for rapid and simultaneous assessment of vaginal microbial composition and inflammation status in pregnancy populations.

Previous work on direct swab analysis by DESI-MS performed in our laboratory highlighted the impact that physical characteristics of swabs (e.g. deformation of swab shape) can have on the quality of resulting metabolic profiles [467]. Typically, loosening of rayon material from the swab tip or deformation of swab shape leads to the reduction of the signal to noise ratio of the resulting spectra. This occurs when loose material from the swab tip leads to increased ablation of the swab surface during DESI-MS, causing obstruction of the inlet capillary of the mass spectrometer. Deformed swab tip shapes can also negatively affect the geometry of the solvent flow and direction of secondary droplets into the inlet capillary. In addition to the physical characteristics of the swab, care was taken to assess blood contamination. Blood is a rich source of metabolites that could potentially bias the resulting metabolomic spectra of samples as has been reported in other metabolomics studies using NMR [468], [469] and MS [467], [470], [471]. Despite these concerns, the spectra obtained from swabs with either suboptimal physical characteristics or blood presence were included in the initial predictive models presented in this chapter. This decision also reflected the reality of these sample types-

vaginal bleeding is an important and relatively common symptom in pregnancy and swab tip damage may occur during storage. The capacity to predict microbial composition was found to be strong, despite including 'suboptimal' samples in the analysis. This highlights the robustness of this methodology. Future studies could still consider alternative sample and storage procedures that might reduce the chance of swab tip deformation. It is also possible that future models could be developed that would account for blood presence leading to improved capacity to predict microbiota composition in these samples. Other sources of potential confounding of the microbiota profiles included possible batch effects associated with the date of acquisition of the spectral data and the hospital where the samples was initially collected as well as BMI and ethnicity. However, these factors did not appear to be major drivers of variance in our dataset. A small proportion of the variation in the metabolic profile was associated with gestational age, which is consistent with previous reports [230], [472]. This provides confidence that changes detected by DESI-MS in the mucosal profiles across gestation were likely to reflect biologically relevant differences rather than technical sources of variation.

Characterisation of microbiota composition of samples analysed in this chapter was performed using gold-standard metataxonomics approaches. Consistent with many previous studies in the field, clustering of the relative abundance of bacteria allowed samples to be classified into well-defined vaginal CSTs. Integration of this data with DESI-MS profiles enabled the identification of a subset of m/z features that could be used to robustly predict the microbial composition of the cervicovaginal mucosa at different levels of taxonomic resolution. These findings are in line with earlier studies that have reported CVF metabolic signatures associated with microbial communities in the vaginal mucosa [327], [473]. However, in contrast to those methods, DESI-MS enables this detailed chemical information to be acquired without any sample preparation and within less than 3 minutes/sample. Prediction of vaginal microbial composition could be achieved with DESI-MS both at a genera level (*Lactobacillus* dominance versus depletion) and at a species level (CSTs contrasts). Furthermore, our findings were validated using a previously collected independent patient cohort. The same 113 features proved to be discriminant of *Lactobacillus* dominance and depletion also in samples from the VMET cohort, which was made up of patients with similar demographics and background PTB risk factors [199].

Many of the metabolites found to discriminate vaginal microbial compositions were consistent with those previously reported in the literature using GC or LC-MS and importantly, they have biologically plausible roles in mediating health and disease in the lower reproductive tract. For

example, higher levels of short- and long-chain fatty acids and biogenic amines were characteristic of high diversity composition, as has been reported in women with BV [327], [348]. Other metabolites increased in the LDEPL samples included hydroxycaproic acid, isoarachidic acid, phytenic acid, nevronic acid, oleic acid, phenyllactic acid and phospopantotenic acid. Features found to be higher in the LDOM group included isoleucyl-serine, Valyl-Alanine, N-acetylputrescine , 5-aminopentanamide, N-palmitoyl-cysteine and Thiomalic acid. Hydroxycaproic acid is an hydroxy fatty acid that has been previously linked to BV-related microorganisms like *Atopobium*, *Prevoltella* and *Gardnerella vaginalis* [474]. Arachidic acid, is part of a class of molecules called eicosanoids, whose production by mammalian cells in response to mechanical factors or chemical stimuli, such as cytokines or in response to pathogen invasion is well described [475]. Increasing levels of arachidic acid are also known to increase the level of prostaglandins. Prostaglandins are important mediators of inflammation, and at high levels, induce collagen remodelling in the fetal membranes providing a potential mechanistic link between high vaginal bacterial diversity and increased risk of PPROM and PTB [153], [476]. Similarly, fatty acids were found to be increased in LDEPL samples. These have previously been associated with BV and an increased risk of spontaneous PTB [324], [339]. Phenyllactic acid was also found to be at higher levels in LDEPL samples. This was surprising given that this molecule has previously been identified in supernatants from *Lactobacillus crispatus* cultures *in vitro* where it was proposed to have a bactericidal role [477]. However, it is possible that other vaginal taxa are also able to produce this compound yet this requires further investigation.

Lactobacillus dominated communities, specifically *L. crispatus* and *L. jensenii* have been reported to have strong positive correlations with several amino acids and dipeptides [327]. This was also observed in our findings with significantly higher levels of isoleucyl-serine and valyl-alanine detected in LDOM samples. N-acetylputrescine, which is often associated with BV and was also shown to be higher in the LDOM group. Since N-acetylputrescine is a downstream product of putrescine catabolism, an explanation for this observation could be the reported ability of *L. crispatus* to degrade putrescine [202]. Thiomalic acid was also significantly higher in LDOM samples. Recently the function of this molecule has been investigated in relation to glutathione peroxidase (GPX3) activity in mice [478]. GPX3 belongs to glutathione peroxidase family, and is a key antioxidant enzyme [479]. Markers of placental oxidative stress and senescence increase with gestational age. In their study, Xiu xu et al. reported that thiomalic acid inhibits the reduction of H_2O_2 by GPX3 [478]. It has been proposed

that one of the mechanisms that *Lactobacillus* species use to prevent pathogen colonisation is through the production of H_2O_2. By inhibiting H_2O_2 reduction, thiomalic acid may promote *Lactobacillus* species protection against pathogen invasion but this requires further investigation.

Given the importance of immune response to microbiota in shape health and disease states[139], [233], [236], [480], we next tested if inflammation status of the host in conjunction with their cervicovaginal bacterial composition could be rapidly assessed by DESI-MS. Currently, assessment of host response to vaginal pathogen colonisation relies upon subjective evaluation of clinical symptoms, microscopy (e.g. detection of clue cells) and sometimes measurement of circulating markers of inflammation (e.g. C-reactive protein). Data from a subset of the samples of VMET2 cohort that has been subjected to immunological profiling were integrated with the corresponding metabolomic profiles, which identified a total of 33 metabolites strongly correlated with a high or low status of inflammation in the host. In this way DESI-MS showed promising capacity to simultaneously capture information on the host immune response in addition to microbiota composition. Suboptimal microbial communities are known to activate immune responses in the host while *L. crispatus* dominated communities are for example associated with low levels of proinflammatory cytokines such as IL-1β [481]. The possibility to have readily available information in the two aspects has therefore clear translational advantages. The relationship between metabolites and immune response could be further demonstrated by comparing levels of known anti-inflammatory (e.g. lactate), or inflammatory (e.g. succinate and propionate) metabolites in the CVF with levels of key immune mediators. Interestingly, lactate levels as detected with DESI-MS in this patient cohort were not found to significantly differ among the main CST groups investigated. One possible explanation for this relates to the ecological function of lactate production and its importance in maintaining a low pH in the vaginal environment. Lactate is produced by both host cells and microbial cells and is also able to be produced by several non-*Lactobacillus* species such as *Atopobium*, *Megasphaera* and *Leptotrichia* [482]. This may explain why some vaginal microbiomes depleted of Lactobacillus species are apparently well tolerated by their host. The same does not appear to be true for SCFA and bioamines, which are not produced in high concentrations by *Lactobacillus* species and are strongly related with BV-associated taxa. Indeed, both propionate and succinate were found to be significantly higher in high-diversity vaginal CSTs in this study. Using a ratio of succinate/lactate and propionate/lactate therefore allowed for a more interpretable assessment of how metabolic output related to different

microbial communities might associated with immune marker levels in the vaginal niche. Higher levels of the succinate/lactate and propionate/lactate ratio were indeed shown to correspond with higher levels of inflammatory markers related to adaptive and innate immune response. These relationships were not always linear and differed for some of the immune markers, likely indicating different levels of activation for different inflammation pathways.

Consideration of inflammatory status is an important factor for determining specific treatments and for monitoring response [483]–[485]. For example, our group recently reported that cervical cerclage performed with braided material can induce vaginal bacterial dysbiosis, inflammatory activation and premature cervical ripening. These events are also associated with increased risk of preterm birth and do not occur in the same measure when monofilament material is used [139]. However, in some women, treatment does not appear to drive a strong change in the microbiota composition and/or lead to inflammatory activation. To determine if DESI-MS could eventually be used to rapidly assess this type of information, sub analyses were performed on samples collected from women receiving a cervical cerclage with different suture materials. Higher levels of inflammation markers were observed in patients receiving braided cerclage and delivering preterm compared to monofilament ones. Also, the levels predicted through DESI-MS would closely reflect the same relationship. This research holds significant importance as it emphasizes the need to measure these factors, as they directly impact the outcome, even when interventions are implemented. By assessing these variables, we gain valuable insights into the effectiveness of interventions and their implications.

In conclusion, we show that direct on-swab DESI-MS metabolic profiling of the cervicovaginal mucosa allows rapid and robust prediction of vaginal microbiota:host immune interactions. We have demonstrated that this approach has the potential to be used for identifying the risk of preterm birth and monitoring preventive treatments specifically for high-risk individuals. However, we also hypothesise that this technique can be applied to evaluate interactions between vaginal microbiota and the host in various clinical situations including both pregnant and non-pregnant women.

Chapter IV

4 THE CERVICOVAGINAL METABOLOME AND PRETERM BIRTH RISK STRATIFICATION

4.1 Introduction

In the previous chapter the use of metabolic profiling for the prediction of vaginal microbiota composition and host inflammatory status during pregnancy was investigated. Yet an infectious aetiology is only considered responsible for approximately 40% of all PTB cases [23]. Multiple other factors have been linked with an increased risk of PTB including maternal or fetal conditions, twin or multiple gestation, young or advanced maternal age, cervical anomalies, ethnicity, extremes of body-mass index, low socio-economic status, stress, excessive physical work, sexual activity, alcohol consumption and smoking habits [486]. However, the strongest risk factor for PTB in a current pregnancy is a history of PTB [100], [486]–[489]. Recurrent PTB (i.e. PTB occurring after a PTB in a previous pregnancy), represents 22.1% of total PTB cases [16]. Approximately 85% of women with one previous spontaneous PTB will deliver at term, but this is reduced to 70% for women with two previous preterm deliveries [99], [490]. This equates to an overall 2.5 fold increased risk of PTB in the current pregnancy [491]. This highlights genetic and epigenetic contributors to PTB risk phenotypes and justifies close clinical surveillance of these women in subsequent pregnancies.

Although PTB can be considered a syndrome with multiple underlying causal mechanisms, there exists a common set of phenotypic changes that precipitate both preterm and term birth. These include cervical ripening and dilatation, fetal membranes rupture and myometrial contractility. The detection of early changes in cervical length is widely used as a predictive tool to identify women at risk of PTB [492], [493]. This is commonly performed using ultrasound in the second trimester where women identified to have cervical length ≤ 25mm are considered at increased risk for PTB [494]–[496]. Cervical length measurements are also predictive of PTB in women who have undergone conisation of the cervix, which consists of the excision of a cone-shaped portion of the cervix to remove pre-cancerous lesions [497], [498]. This procedure is thought to weaken cervical structure that impairs its ability to support the developing fetus throughout gestation [102], [499], [500]. Current clinical interventions available to treat women with a short cervix include progesterone supplementation and cervical

cerclage [501]. Cerclage consists of the placement of a suture around the cervix providing mechanical support. When the suture is performed in women that present with bulging membranes and a dilated cervix, the procedure is referred to as “rescue cerclage”. This is because if left untreated, women are highly likely to deliver within 2-3 weeks of presentation [502]. Despite its wide use, evidence for the effectiveness of cervical cerclage to prevent PTB is variable. For example, Berghella et al. concluded that cerclage could be beneficial in women with singleton pregnancies, short cervix and experience of prior preterm birth [137]. Jorgensen et al. in a similar meta-analysis however, did not find any statistical significance for singleton pregnancies [503]. Both studies showed no benefit for multiple gestation pregnancies. Some studies have also showed a link between cerclage and increased risk of infection [134], [504]. In particular the use of braided suture material for the procedure has been linked to an increased risk of developing vaginal bacterial dysbiosis, inflammation and poorer pregnancy outcomes [139], [452], [505]. Moreover, cerclage is not recommended for women with active bleeding or sepsis, and it is unlikely to be successful for women who have advanced dilation and prominent membranes prolapse [502].

Currently there is no method available to predict early cervical shortening in pregnancy, nor identify those women who might benefit from treatment interventions. Additionally, tools enabling rapid assessment of vaginal microbial composition and inflammatory status after the insertion of a cerclage that could enable directed treatment strategies are lacking.

4.2 Hypothesis and aims

The work described in this chapter sets out to test the hypothesis that CVF metabolic profiles obtained by direct swab analysis by DESI-MS provide biochemical information useful for preterm risk stratification. To test this hypothesis, this chapter had several aims:

1. To determine if existing risk factors (e.g. a history of preterm birth, cervical treatment) are associated with specific DESI-MS swab metabolic signatures.

2. To investigate if DESI-MS swab metabolic profiles enable prediction of women who will experience cervical shortening and require an intervention during pregnancy.

3. To determine if DESI-MS swab metabolic profiles can be used to predict outcome following cervical cerclage

4. Determine if vaginal microbiota composition can be robustly predicted using direct swab analysis by DESI-MS following the insertion of a cervical cerclage.

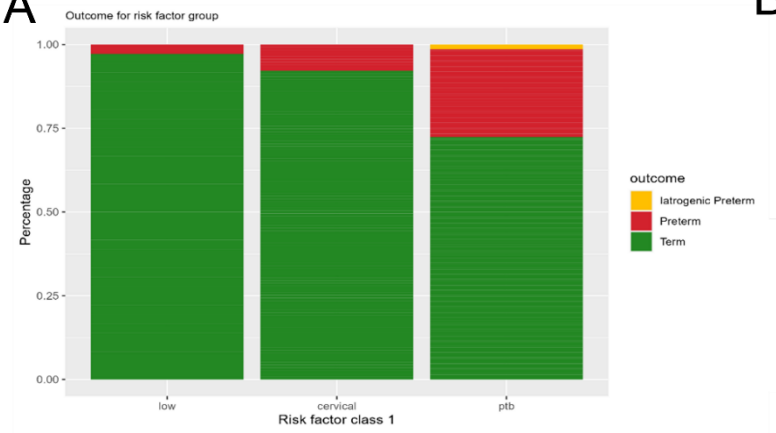
4.3 Results

4.3.1 Analysis of previous pregnancy history on PTB risk in the studied patient cohorts

To examine whether previous pregnancy history is an important risk factor for preterm birth in the index pregnancy for patients enrolled in this study, analysis of detailed clinical metadata of patients from both the VMET and VMET2 cohorts was performed. For VMET a total of 458 samples were collected during pregnancy from 160 subjects. For VMET2 a total of 573 samples were collected during pregnancy from 201 subjects.

Patients were grouped in three classes according to their preeminent PTB risk factor: "PTB" group represented patients with an history of previous PTB (n= 210); "cervical" includes patients that did not have an history of PTB but received cervical treatment for pre-cervical cancer (i.e. LLETZ or similar) (n= 115); "low" includes patients without risk factors related to their gynaecological history (n = 36). As seen in Figure 4.1, preterm birth rates in the index pregnancy were highest in women with previous PTB (38%), followed by women with cervical treatment (8%) and finally women without risk factors (3%) (chi square test p-value = 3.045e-5).

A



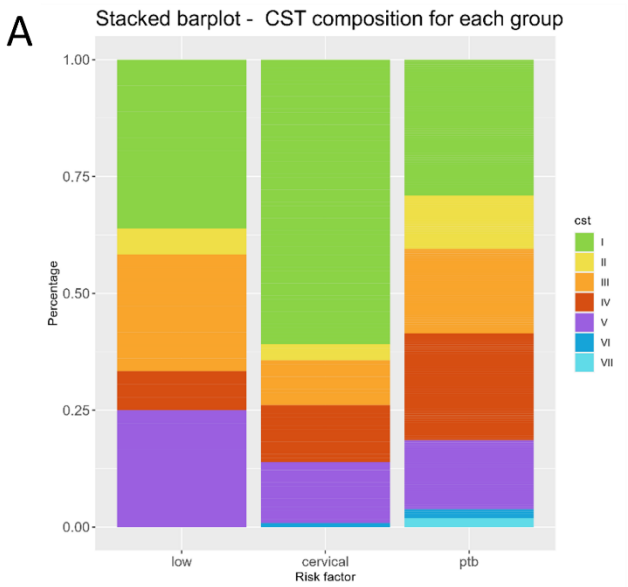
B

	IATROGENIC PRETERM	PRETERM	TERM
LOW	0	1	35
CERVICAL	0	9	106
PTB	3	55	152

Figure 4.1 - Bar plot showing percentages of different outcomes for each risk group. A) The bar plot illustrates the distribution of different outcomes across various risk groups. The rate of PTB was 38% for women with previous history of PTB, 8% for women with cervical treatment and 3% women without risk factors (chi square test p-value = 3.045e-5).B) Table with respective outcome values for each risk group.

Investigation of the vaginal microbiota composition of samples collected from these cohorts identified a significant difference in the prevalence of major CSTs on the basis of pregnancy history (chi square p-value = 1.64e-5) (Figure 4.2). Women that had a previous PTB were found to have lower rates of CSTI compared to women with a history of cervical treatment. Moreover, women with a previous history of PTB had higher prevalence of high diversity CST IV compared to women with cervical treatment or those at low risk.

A



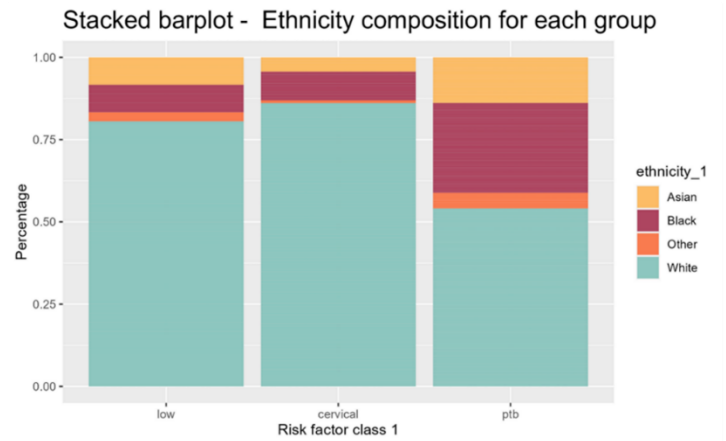
B

	I	II	III	IV	V	VI	VII
LOW	13	2	9	3	9	0	0
CERVICAL	70	4	11	14	15	1	0
PTB	61	24	38	48	31	4	4

Figure 4.2 – Prevalence of CST composition based on PTB birth risk category. A) The bar plot illustrates the distribution of different CST compositions across various risk groups. Lower incidence of CST I and higher prevalence of CST IV can be observed in the previous PTB risk group compared to the others (chi square test p-value = 1.64e-5). B) Table with respective CST values for each risk group.

As ethnicity is known to influence vaginal microbial composition, we next examined representation of major ethnic groups in women classified by pregnancy history risk factors. This analysis showed that Black women were over-represented in those women with a history of PTB (chi-square test p-value 9.17e-7) (Figure 4.3)

A



B

	ASIAN	BLACK	OTHER	WHITE
LOW	3	3	1	29
CERVICAL	5	10	1	99
PTB	29	57	10	113

Figure 4.3 – Bar plot showing percentages of different Ethnicity prevalence for each risk group. A) The bar plot illustrates the distribution of different ethnicities across various risk groups. In the previous PTB risk group there is a higher incidence of women of Black ethnicity compared to the other risk factors (chi square test p-value = 9.17e-7). B) Table with respective ethnicities for each risk group.

A sub analysis was next performed where prevalence of vaginal CSTs within each major PTB risk group was assessed on the basis of ethnicity. A higher representation of CST IV was observed among Black women compared to other ethnicities within women with a history of PTB. Due to limited sample sizes, it was not possible to conduct a similar comparison for other ethnicities or for the cervical risk group (Figure 4.4). To account for possible confounding of CST prevalence due to ethnicity, subsequent analyses presented in 4.3.2 are limited to white women.

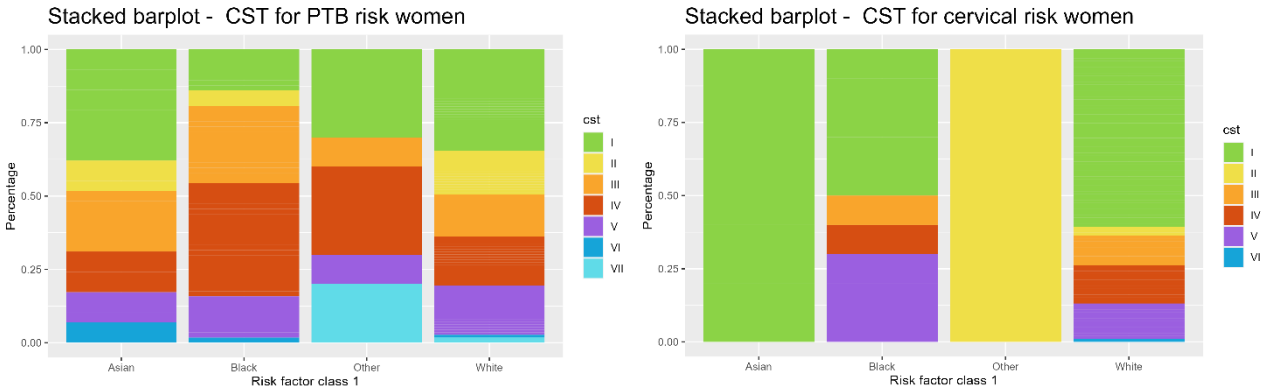


Figure 4.4 - Bar plot showing percentages of different CST composition for each risk group divided by ethnicity. The bar plots illustrates the distribution of CST across different ethnicities in women with previous PTB risk factor and women with cervical treatment risk factor.

4.3.2 The CVF metabolome and previous pregnancy history

As to allow comparison of metabolic profiles obtained from women in the VMET and VMET2 cohorts, common metabolic features were identified across the two datasets. These included metabolite features with the m/z values within a maximum error of 5ppm. Intensity values for matched features identified in negative ion mode were merged resulting in a total of 447 metabolites used for these analyses. For positive ion mode data, a total of 1050 common metabolite feature were merged and included in the analyses.

A preliminary untargeted analysis using PCA was performed to assess metabolic differences among the first sample collected in pregnancy from each patient (Figure 4.5). A total of n = 233 had at least 1 sample collected before 26 weeks gestation (chosen as it is the upper limit of the second trimester) included in the analysis. Eight (n =8) patients did not have a sample from before 26 weeks available and were therefore excluded. The first two components of the PCA capture cumulatively around 25% of the dataset variance both in positive and negative ion mode. When colouring the PCA according to the risk factor however, this aspect does not appear to be the main driver of the variance in the dataset.

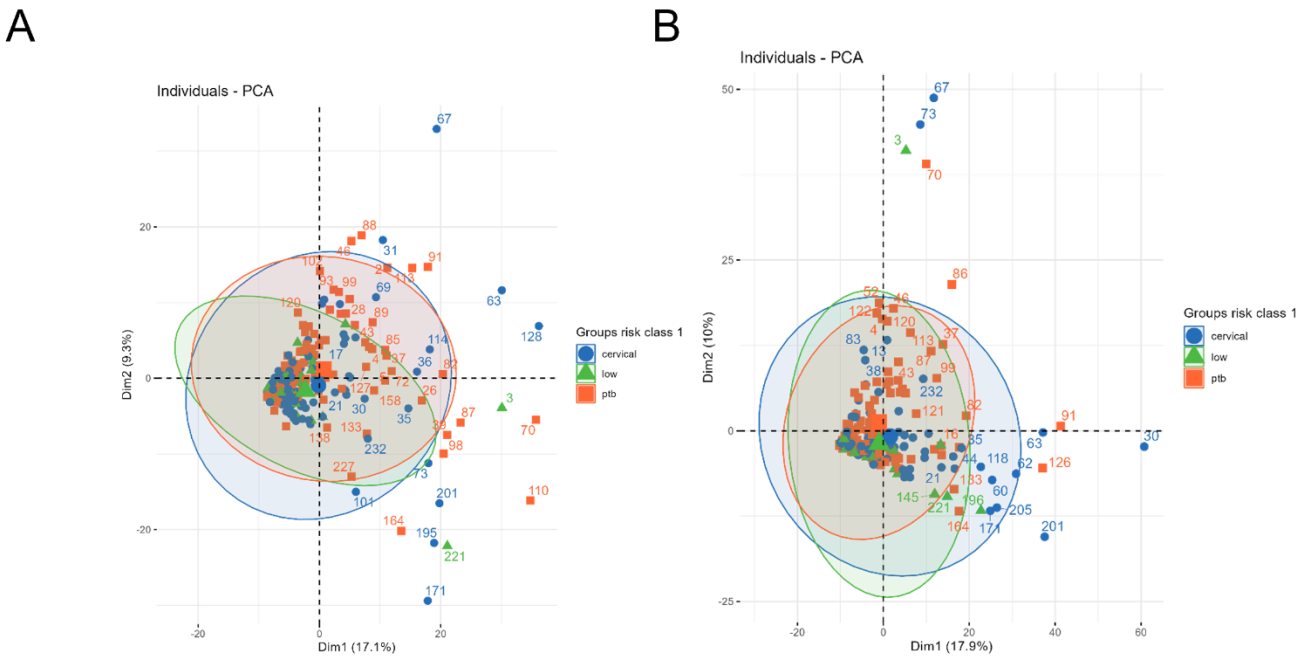


Figure 4.5 - PCA of DESI-MS metabolic profiles of CVF collected from white pregnant women in the VMET and VMET2 cohorts n = 233. Patients are classified as "low" risk (n=29), "cervical" risk (n = 93) or "ptb" indicating a history of ptb (n = 111). PCA of metabolic profiles generated from A) negative or B) positive ion mode data are presented..

The PCA was recoloured according to different potential confounders to enable investigation of how they may impact variance of the metabolic profiles i.e. CST, study – see Appendix figure 6). Observing the PCA plot, it was evident that some variance in the dataset could be attributed to the study cohort.

To determine if PTB risk factors could account for any variability within the dataset, PERMANOVA analysis was performed resulting. This indicated that risk factor accounted for a relatively small portion of the variance in the metabolic profiles (R^2 = 0.028; See Appendix Table 6). To investigate this further supervised modelling in the form of PLSDA analysis was performed using both positive and negative ion mode data (Figure 4.6) on samples collected from women with either cervical treatment or previous PTB. Examination of the resulting score plots showed separation between the metabolic profiles of the two patient for both the negative and positive ion mode data. The models parameters describe a reasonable capacity of correctly assign samples to the relative class. The positive model has a better sensitivity (pos = 0.58, neg= 0,48) but lower specificity (pos = 0.70, neg = 0.75) and in general both models are prone to identify false positives.

From the PLSDA model 56 features were identified through a linear model as differing significantly ($p < 0.05$) between women with different risk factors in negative ion mode and 81 in positive ion mode (Appendix Table 7). Six of these metabolites had previously been identified to significantly differ on the basis of microbial composition (see Chapter 3 and Appendix Table 1) and included Glycyl-Phenylalanine (m/z 221.0931), Phosphatidyl glycerol, (m/z 245.0433), Asparagine (m/z 131.0465), Leucyl-Glutammine (m/z 258.1459), Dimethoxyphenyl-hexadecylcarbamothioyl-propenamide (m/z 489.3147) , sulfanylheptenyl- Isotope Amino -sulfanylheptenyl-amino-oxooctadecanoic acid (m/z 456.3345). All the metabolites mentioned above have been identified using MS/MS with the exception of Asparagine and sulfanylheptenyl- Isotope Amino -sulfanylheptenyl-amino-oxooctadecanoic acid (please refer to Appendix Table 1 for further details).

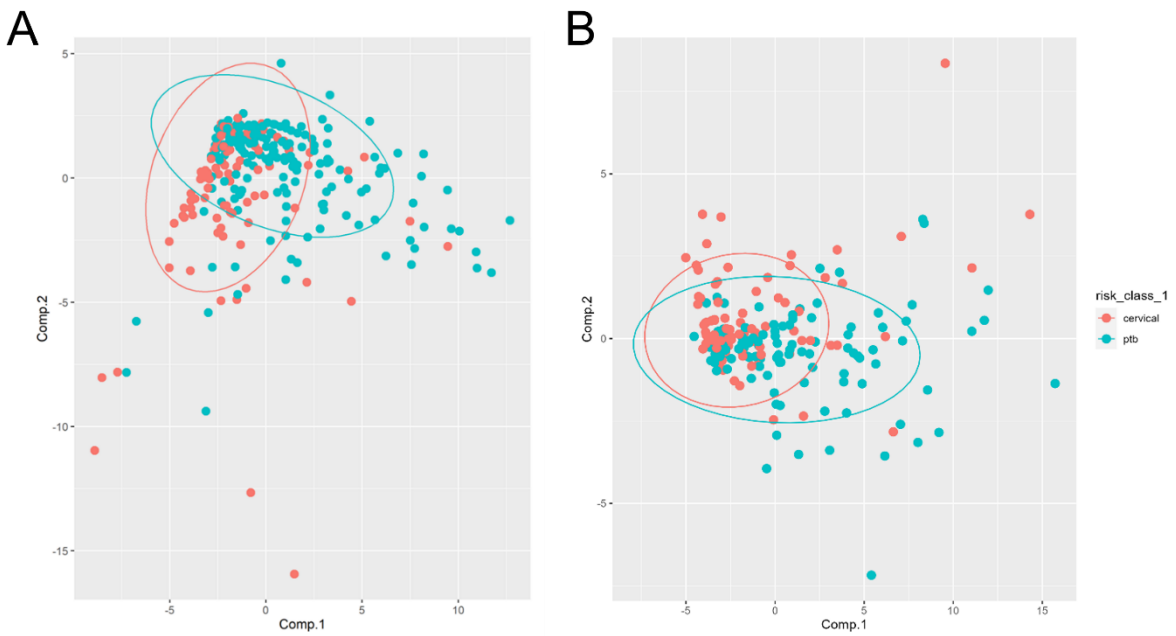


Figure 4.6 - PLS-DA model including all white patients from VMET and VMET2 cohorts n = 233 and two classes. A) PLS-DA plot generated from negative ion mode data: Sensitivity = 0.48, Specificity = 0.75, PPV = 0.61618223, NPV = 0.64, Precision= 0.62. B) PLS-DA plot generated from positive ion mode data: Sensitivity = 0.58, Specificity =0.70 , PPV= 0.61, NPV = 0.68, Precision =0.62

As women with a previous history of PTB have, by definition, a parity score > 1, a sub-analysis was performed to determine if parity could account for some of the observed differences in the metabolic profiles of women with a history of cervical treatment and those with a history or preterm birth. As seen in Figure 4.7, similar differences in the relative concentrations of the six reported metabolites were observed regardless of parity.

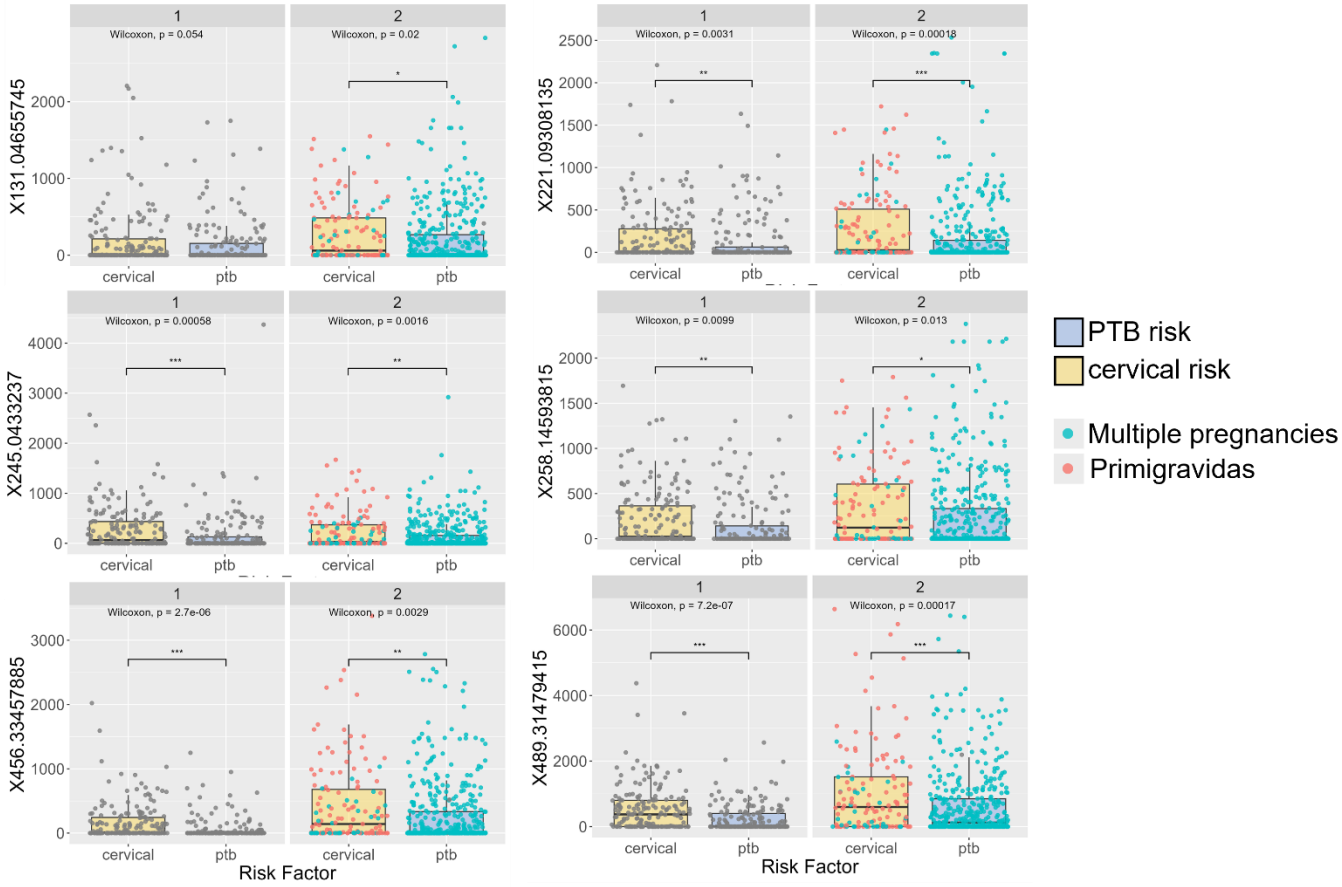


Figure 4.7 - Boxplots of the 6 shared significant features in VMET and VMET2 study. Samples from VMET 2 cohort are coloured according to parity score.

4.3.3 Assessment of the ability of DESI-MS derived CVF metabolic profiles to predict PTB

The ability of cervicovaginal fluid metabolic profiles to predict preterm birth as an outcome was also tested. For this analysis the first sample available per patient was selected from both VMET and VMET2 cohorts. A total of n = 346 women (n= 62 preterm and n= 282 term) had at least 1 sample collected before 26 weeks gestation (chosen as it is the upper limit of the second trimester). A PLS-DA model was performed to assess whether is possible to predict pregnancy outcome from early pregnancy samples. Despite resulting models indicating some separation in the multivariate metabolic profiles of preterm birth and term birth patients, the model parameters indicated weak fit for data generated from both positive and negative ion mode (Figure 4.8). Both positive and negative ion mode models present high specificity (neg = 0.96, pos=0.96) but low sensitivities (neg= 0.11, pos= 0.10). This results in models with a low balanced accuracy (neg = 0.53, pos = 0.54). Balanced accuracy is defined as (Sensitivity + Specificity) / 2 and a value close to 0.50 indicates a poor classification accuracy close to randomness.

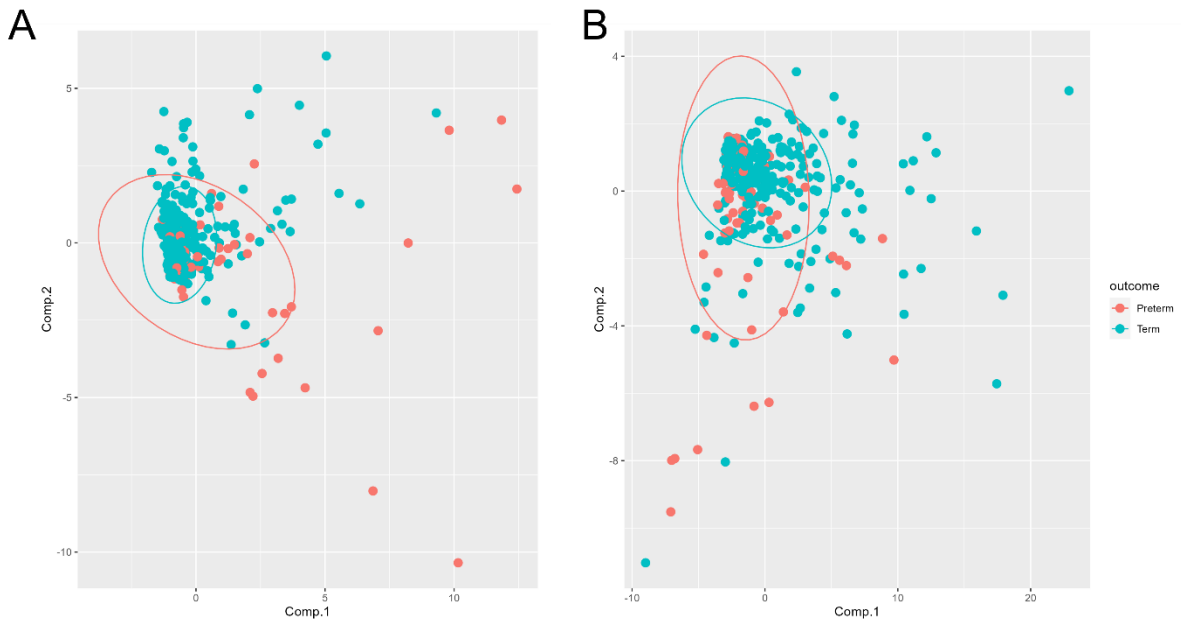


Figure 4.8 - PLS-DA model including all patients from VMET2 and VMET respectively. Trying to discriminate for outcome. A) PLS-DA plot generated from negative ion mode data: Sensitivity = 0.11, Specificity = 0.96, PPV = 0.44, NPV = 0.82, Precision= 0.44. B) PLS-DA plot generated from positive ion mode data: Sensitivity = 0.10, Specificity =0.96 , PPV= 0.41, NPV = 0.83, Precision =0.41

4.3.4 Prediction of cervical shortening and intervention response using DESI-MS swab profiling

The next objective of this study aimed to investigate potential CVF metabolic differences in relation to premature cervical shortening. For this, DESI-MS profiles acquired from longitudinal samples from patients in the VMET and VMET2 cohorts used (VMET n= 139, n=437 samples, VMET2 n=190, n= 581 samples) Within these patients a subgroup of "cases" and "controls" were defined. The "cases" group includes patients that during the gestation received either a cerclage or progesterone intervention due to cervical shortening ≤ 25mm as detected with vaginal ultrasound (VMET N= 78, n=251 samples, VMET2 N=69, n=231 samples). "Controls" did not experience pathological cervical shortening or any intervention during pregnancy (VMET n= 61, n=186 samples, VMET2 n=121, n= 350 samples). To test the ability of DESI-MS profiles to predict cervical shortening, the first sample available prior to intervention was used. The final number of "cases" samples that presented a sample collected before the intervention was 85 (VMET n = 34, VMET2 n=51). These were matched with 85 "controls". An example of the patients chosen with the relative sampling throughout gestational age and intervention received is shown in Figure 4.9.

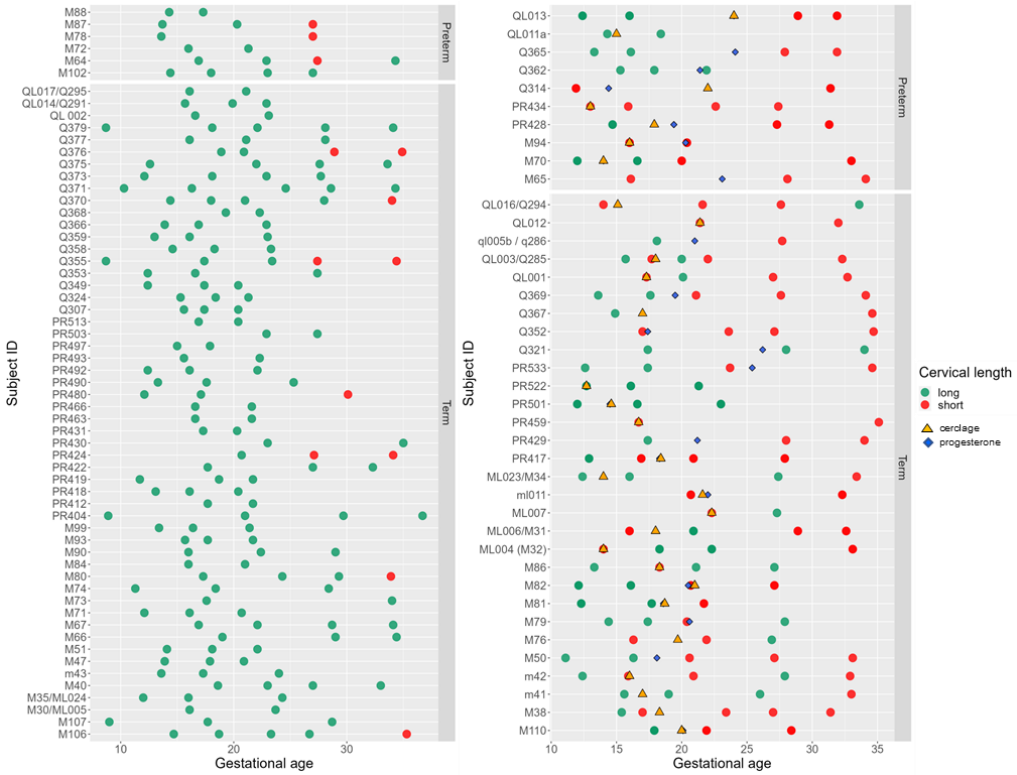


Figure 4.9 - VMET samples subset selected because of their intervention and availability of longitudinal data before and after intervention

Random forest modelling of DESI-MS data using negative ion mode features achieved an AUC = 0.65, (Sensitivity = 0.60, Specificity = 0.66) indicating moderate capacity of the metabolic profile to differentiate between samples collected from patients who would subsequently shorten, comparted to controls (Figure 4.10). Comparable results were obtained when a model was built with positive ion mode features (AUC = 0.68, Sensitivity = 0.64 and Specificity = 0.63).

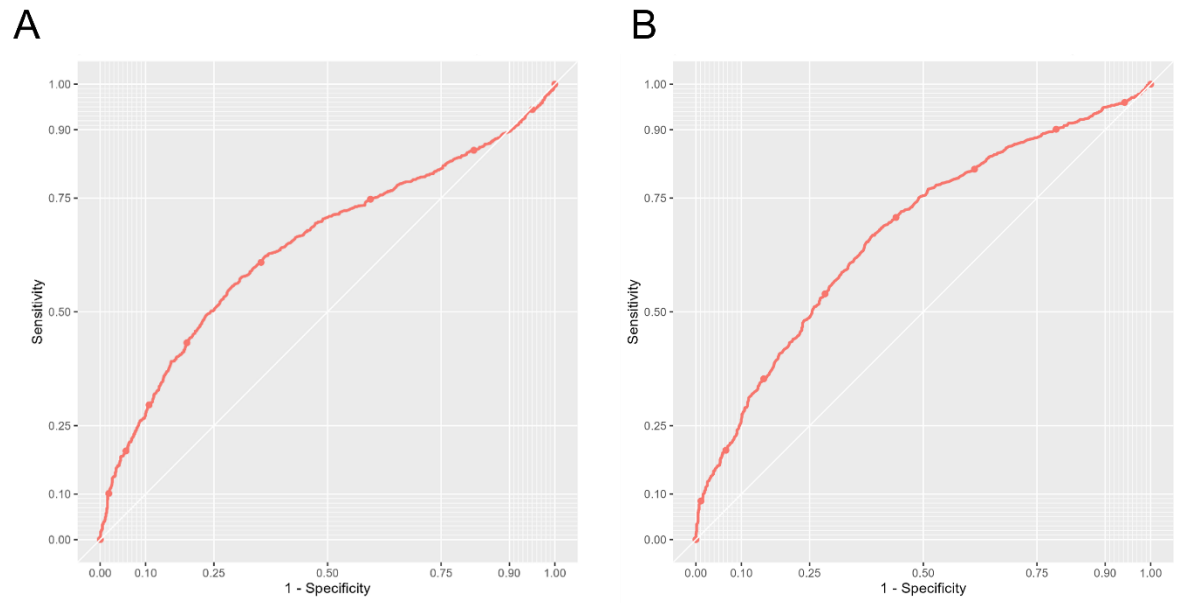


Figure 4.10 - ROC curve negative and positive ion mode obtained using only significant features for prediction. A) Model generated with negative ion mode data: AUC = 0.65, Sensitivity = 0.6, Specificity = 0.66, PPV = 0.64, NPV = 0.62, Precision = 0.64 B) Model generated with negative ion mode data: POS: AUC 0.68, , Sensitivity =0.64, Specificity 0.63, PPV = 0.64, NPV = 0.64 ,Precision n=0.64.

From the Random Forest model a total of 20 metabolic features in negative ion mode and 31 features in positive ion mode where found to be significantly different between patients that will subsequentially experience a cervical shortening and will require an intervention and controls. Further details about the specific m/z annotation and p-values can be found in Appendix table7.

4.3.5 Metabolic profiling to monitor cerclage efficacy

We next examined the capacity of DESI-MS metabolic profiles to identify patients with differing outcomes post-insertion of a cervical cerclage. From the "cases" in the VMET2 cohort n=14 subsequently shortened their cervix despite intervention (denominated "short") and n = 37 patients maintained a long cervix (> 25mm) after the intervention – denominated "long". This analysis was limited to patients experiencing shortening before the 28th week of gestation beyond which physiologically normal shortening of the cervix can begin to occur. Samples from the VMET cohort were not included in this analysis because the same granularity of information on cervical length after intervention was not available.

Looking firstly at the microbiota composition of these patients, no significant difference in genera or CST level compositional structures were observed (chi square test p-value =1 and p-value = 0.55 respectively) although there was a trend decrease in CST I and trend increase in patients with CSTIII in the 'short' group (Figure 4.11).

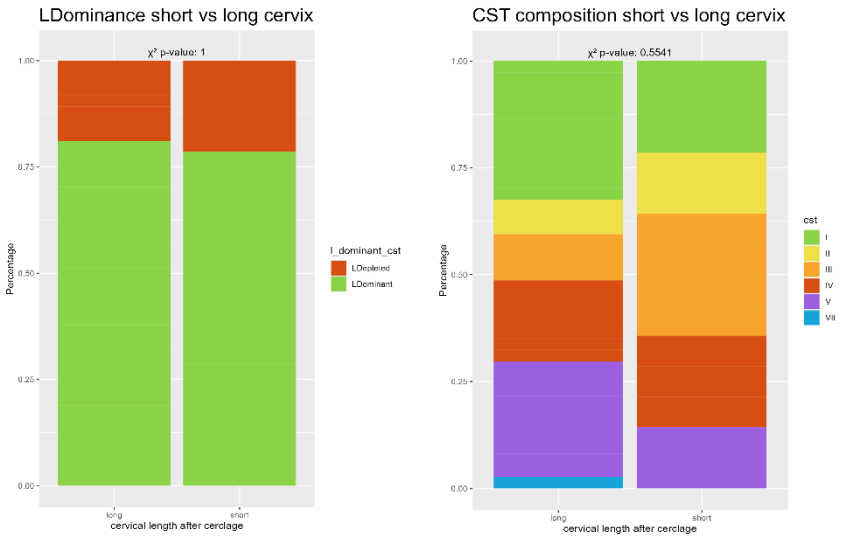


Figure 4.11 – Stacked bar plot representing microbial composition in patients with long or short cervix after cerclage intervention. Microbial composition does not significantly differ between patients with a long or short cervix after the cerclage intervention. No significant difference is observed at genera level (chi-square test p-value = 1) or at CST level (chi-square test p-value = 0.55).

PLSDA modelling of matching DESI-MS metabolic profiles collected in negative and positive ion mode indicated a moderate capacity to distinguish patients who would shorten despite intervention from those who would not (Figure 4.12), (Sensitivity (neg = 0.72, pos = 0.84) Specificity (neg = 0.3, pos = 0.27)).

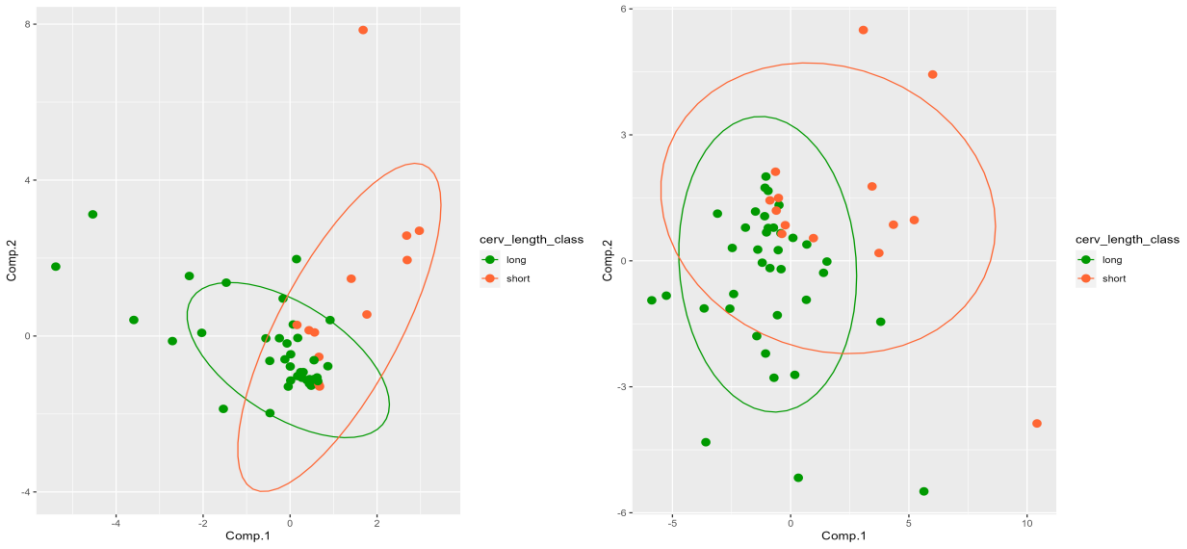


Figure 4.12 – PLSDA showing the difference between patients with long or short cervix after cerclage intervention. A) PLS-DA plot generated from negative ion mode data: Sensitivity = 0.72, Specificity = 0.3, PPV = 0.71, NPV = 0.29, Precision= 0.71. B) PLS-DA plot generated from positive ion mode data: Sensitivity = 0.85, Specificity =0.27 , PPV= 0.74, NPV = 0.67, Precision =0.74

4.3.6 Prediction of microbial composition after cerclage intervention

Microbiota composition post-cerclage has been previously linked to inflammatory host responses associated with poor pregnancy outcome. Therefore we lastly investigated whether direct swab profiling by DESI-MS could robustly predict microbiota composition post-cerclage insertion. The selection of samples for this analysis included patients that received both progesterone and cerclage or only cerclage where a swab sample was available after the cerclage intervention and before the 28th week of gestation (n = 55). One sample per patient was used for the analysis.

The distribution of CSTs in samples taken post-cerclage as determined using metataxonomics was: CST I n= 17, CST II n = 6, CST III n=9, CST IV n=10 and CST V n= 12. A random forest classifier model was built using the corresponding DESI-MS profiles and the performance of

the model to predict microbiota was assessed with ROC curve analyses. Models trained to predict microbiota composition at a genera level (Figure 4.13) had a strong predictive ability to classify *Lactobacillus* dominance from depletion and this could be achieved using data generated in negative or positive ion mode (negative: AUC = 0.92, sensitivity =0.51, specificity = 0.96; positive: AUC = 0.92, sensitivity =0.48, specificity = 0.96).

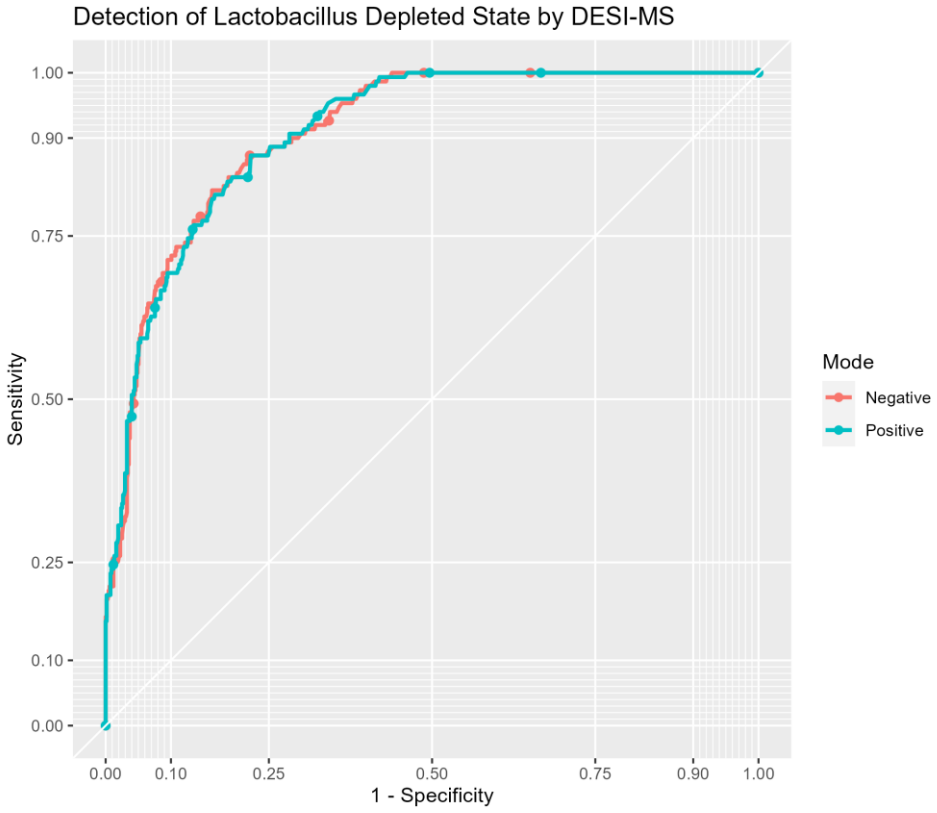


Figure 4.13 - ROC curve negative and positive ion mode predicting microbial composition at genera level for people who received a cerclage NEG: AUC = 0.92, Sensitivity = 0.51, Specificity = 0.96, POS: AUC = 0.92, sensitivity = 0.48 Specificity = 0.96

Similarly, the ability to predict microbiota composition at a CST level was very strong particularly in well represented CST-groups (Figure 4.14). A table describing the key performance characteristics of the model is presented (Table 4.1).

Table 4.1 - Random forest results for CSTs comparisons

Comparison	AUC (Negative)	Sensitivity (Negative)	Specificity (Negative)	AUC (Positive)	Sensitivity (Positive)	Specificity (Positive)
CST IV vs CSTI	0.98	0.92	0.89	0.98	0.89	0.88
CST IV vs CST III	0.93	0.86	0.91	0.94	0.85	0.93
CST III vs CST I	0.81	0.95	0.40	0.84	0.97	0.40
CST V vs CST I	0.84	0.68	0.83	0.85	0.69	0.82
CST V vs CST III	0.66	0.51	0.79	0.66	0.55	0.78

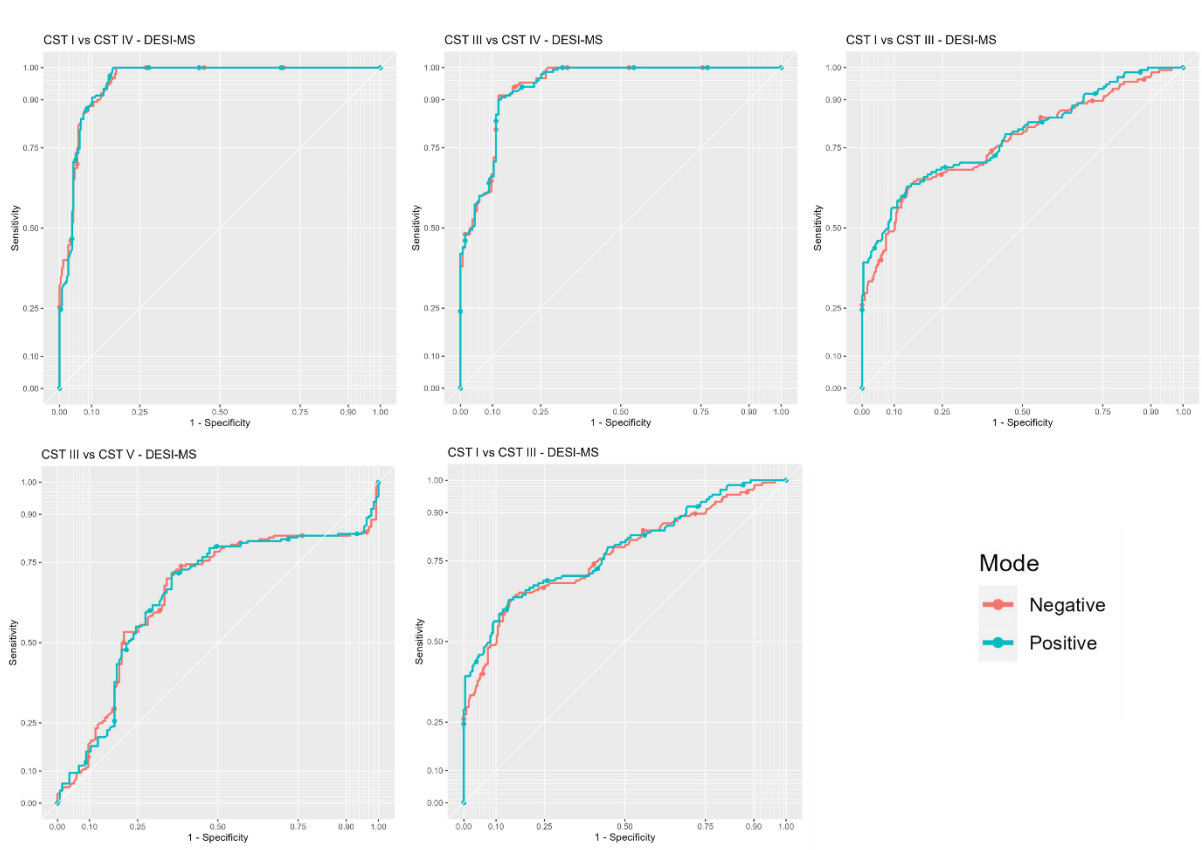


Figure 4.14 - ROC curve negative and positive ion mode showing prediction ability at different taxonomic levels CST I vs CST IV AUCneg = 0.98, sensitivity neg = 0.92, specificity neg =0.89, AUCpos = 0.98, sensitivity pos= 0.89 , specificity pos=0.88 . CST III vs CST IV AUCneg = 0.93, sensitivity neg = 0.86, specificity neg = 0.91, AUCpos = 0.94, sensitivity pos= 0.85, specificity pos= 0.93. CST I vs CST III AUCneg =0.81, sensitivity neg = 0.95 , specificity neg = 0.40, AUCpos = 0.84, sensitivity pos= 0.97, specificity pos= 0.40. CST I vs CST V AUCneg = 0.84, sensitivity neg =0.68, specificity neg =0.83, AUCpos = 0.85, sensitivity pos=0.69, specificity pos= 0.82. CST III vs CST V AUCneg = 0.66, sensitivity neg = 0.51, specificity neg = 0.79, AUCpos = 0.66, sensitivity pos=0.55, specificity pos= 0.78.

4.4 Discussion

A previous history of PTB is the leading risk factor for recurrence of PTB and is estimated to increase risk by approximately 2.5-fold [491]. However, the mechanisms underlying this association remain unclear. In a recent study, placental disorders in first and second births accounted for approximately 40% of recurrent extremely and very preterm births (<34 weeks gestation), and around 22% of recurrent late preterm births (34-37 weeks gestation) [506]. However, in an earlier study, among women with a previous spontaneous preterm birth, the rate of recurrence of PTB <35 weeks was the highest (around 65%) among women with a short cervix (<25 mm) and a positive fetal fibronectin test [507]. Moreover, there is substantial evidence to indicate that women who have previously experienced PPROM, are at an increased risk of PPROM in subsequent pregnancies. Data from the Collaborative Perinatal Project found that women who had experienced PPROM in a previous pregnancy had higher incidence of PPROM in subsequent pregnancies (21%) compared to women without a history of PPROM (4%) [508]. Similarly, Asrat et al. reported a 32% risk of recurrence in patients with a history of PPROM [509]. Ekwo et al. also reported a 5.5 fold increase in the risk of PPROM in women with PPROM in a previous pregnancy compared to controls [510]. Finally, Mercer et al. in their Preterm Prediction Study, reported an increased risk of PTB due to PPROM in the current pregnancy for women with a previous episode of PPROM but also that in general, among multiparous women, a previous preterm birth due to preterm PPROM was the primary risk factor for PPROM in a subsequent pregnancy [487]. The relationship between previous PPROM, short cervical length, positive fibronectin test, and recurrent PTB is still not fully understood, but there are hypothesis trying to justify this recurrence. It has been suggested that genetic or environmental factor could predispose women to repeat a PTB outcome with or without PPROM in subsequent pregnancies (e.g. carrying genetic variants which predispose to an excess of matrix degrading enzymes could predispose to premature cervical shortening and membrane ruptures) [511].

Women who have previously had a surgical treatment for cervical disease have approximately 2-fold increased risk of PTB in subsequent pregnancies[512]. The mechanism that links cervical surgery and risk of PTB is not clear. However, studies have demonstrated an association between the amount of cervical tissue removed during surgery and an increased risk of PTB [513]–[515]. The removal of excessive amount of tissue could cause accidental removal of healthy tissue altering cervical function with the consequent impairments of

antimicrobial defence mechanisms [516]. Cervical excisional procedures have also been hypothesised to allow pathogen ascension, leading to an inflammatory response with consequent cervical ripening and uterine contraction increase [516]. Another hypothesis is that the tissue removal causes a reduction in the mechanical support of the cervix [517]. It has also been hypothesised that LLETZ could cause a permanent shortening of the cervix although no study has reported evidence of this [518].

These observations collectively show that PTB can derive from a wide range of aetiologies and can therefore present different underlying mechanisms that can affect the response to different preventative treatments. Some studies have reported that cervical cerclage shows more benefits for patients with previous PTB history compared to progesterone [519]. On the other hand, in women with a previous cervical intervention the existing evidence does not support the use of cerclage for PTB prevention while data on progesterone is lacking [520]. It is therefore important to understand the underlying biology of PTB in order to develop new or more targeted intervention strategies.

Consistent with the literature, in the studied cohort described in this chapter, preterm birth prevalence was highest (38%) in women with a history of previous PTB. Women who had a history of cervical treatment has a prevalence of 8% whereas low risk women had 3%. While different risk factors involve substantially different underlying aetiologies and mechanistic pathways driving PTB, these often coalesce on a common phenotype, such as premature cervical ripening. Therefore, one of the hypotheses of the work described in this chapter was that the difference in underlying biology associated with risk factor would be reflected in the CVF metabolome. To test this we compared the first trimester CVF metabolic profiles of women with a history of PTB to those with cervical treatment. In agreement with the hypothesis, there were substantial differences identified in the CVF metabolome of these two groups. While preliminary unsupervised models indicated that early pregnancy CVF metabolic profiles obtained using DESI-MS in both negative and positive ion mode, are not strongly reflective of background preterm birth history and risk phenotypes, supervised models showed moderate discriminating abilities. This is expected considering that risk factor is not anticipated to be a primary determinant of the variance in this dataset. This result however, together with the result of the PERMANOVA analysis suggests that a portion of the metabolic variance is indeed related to previous pregnancy history.

Additional investigation is required to characterise the metabolites responsible for this variance and to understand their biological significance. However, a small number of metabolites (n = 6) were previously identified as associated with microbial composition as described in Chapter III. More specifically all of them were found to be higher in microbiotas characterized by *Lactobacillus* dominance. These findings reflect the difference in microbial composition that we observe among the different risk factors. Women with a previous PTB have less incidence of CST I and a higher prevalence of CST IV. In this group the metabolites identified were found in lower levels. In the cervical risk factor group on the other hand higher prevalence of CST I was reported.

The fact that we observed a higher prevalence of CST IV in patients with previous PTB history is in line with what observed in the literature. By definition, women who experienced a previous PTB, have undergone a prior pregnancy and their parity score is greater than 1. Parity has been linked to alterations in the vaginal microbial composition. Women that had previous pregnancies generally present a lower abundance of *L. crispatus*, and a higher presence of *L. iners* and *G. vaginalis* [521], [522].Moreover, in a recent study Romero et al. reported higher diversity in women with parity [523]. One suggested hypothesis to explain this association is the observed increase in diversity of the vaginal microbiota after live birth [441]. It is believed that this increase could be a cumulative phenomenon across multiple pregnancies [523].

Caution must be taken in interpreting our results on microbial composition differences. Upon examination of the ethnic distribution across different risk groups was noted a higher enrichment of black women in the previous PTB group. It is known from the literature that people with black or Hispanic ethnicity tend to have a higher proportion of high diversity bacterial communities [221], [524] . This however has only been observed in the US [525]. To date, the studies produced by our group and others in the UK have not shown the same difference [235], [441]. Furthermore, in black women this association has not been necessarily linked with higher risk of preterm birth or adverse outcome [526], [527].

For this reason, any comparison of the metabolic profiling should be adjusted for the patient ethnicity. Unfortunately, our study is constrained by a small representation of ethnicities other than white. It is not clear why the recruitment is weighted more on white women but one hypothesis could be that it reflects the demographic composition of the London boroughs where the samples were collected. Queen Charlotte's and Chelsea and St Mary's Hospitals are located in the Borough of Hammersmith and Fulham, Chelsea & Westminster Hospital is located in

the Borough of Kensington and Chelsea, and University College London Hospital is located in the Borough of Camden. These areas have low percentages of Black population (respectively 9.1%, 12.2% and 7.9%) according to the 2021 census [528]. Another hypothesis is that black communities and minorities in general, are object of racial discrimination that could impact their access to healthcare resources or undermine their trust and reliance on the healthcare system as a whole [529]. This could result in a lower rate of referrals to specialistic clinics. Another limitation of this study is the fact that comparisons were drawn mainly between risk groups, disregarding comparison with low risk population. This was due to power issues, since our study was mainly focused recruiting women referred to the clinic for high risk of PTB. Additionally we wanted to focus on examining differences in the mechanisms underlying main risk factors.

Having shown that there does appear to be a difference in the underlying aetiology, we next aimed to test whether or not that there was sufficient information present in the cervicovaginal metabolome that would enable early prediction of PTB as an outcome. However, our model indicates poor classification performance. Differences in metabolites were not examined due to the absence of robust variations in the overall metabolic profiles associated with the outcome. This observation aligns with the diverse possible aetiologies of PTB previously mentioned. Additionally, it should be noted that some of the women included in the study would have experienced PTB if they had not received an intervention. Accounting for these aspects in the model posed significant challenges.

Because of this high chance of generating spurious findings, we decided to focus on one of the main phenotypes of PTB. We attempted to see if CVF metabolome would be useful in predicting cervical shortening. We used the cervical shortening as a proxy marker for adverse pregnancy outcome. After carefully selecting samples to minimise any possible confounder in our analysis (one sample per patient, before intervention, matched for gestational age etc.) our models showed a moderate capacity to predict the likelihood of a subsequently shortening cervix. Our results indicate that DESI-MS metabolic profiling offers potential for stratifying women who will subsequently experience cervical shortening and require clinical intervention.

Cervical shortening and remodelling involves both structural and molecular changes [530]. On the molecular point of view, pro inflammatory mediators stimulate the production of MMPs and contribute to the breakdown of extracellular matrix components [530]. This breakdown potentially releases some key metabolites in the cervicovaginal fluid. Our results indicates that

early biochemical changes associated with pathophysiological shortening of the cervix are detectable in the cervicovaginal metabolome prior to clinically manifest shortening detectable by ultrasound.

Yet our model is not perfect and we are unable of accurately predict all the shortening cases. There are several reasons to explain why a better predictive model was not achieved. Firstly we attempted to account for the time from sampling to the time of the shortening event by selecting the last sample prior to shortening. However, there is potentially a range of latency times for cervical shortening and the patients included in the analysis might be experiencing a cervical shortening at different rates. It has been previously shown how cervical shortening rates differ in women with competent or incompetent cervices [531] and across different outcomes [532]. As the cervix progress its shortening, major changes in the metabolites released in the CVF can potentially alter the overall metabolic profile and obstacle an accurate prediction.

Additionally, another reason for the moderate predictive ability of our model, can relate to the underlying differences in the aetiology causing the shortening event. A subset of women are going to have an incompetent cervix due to mechanical issues while others may have a microbiota:host interaction that is causal of inflammation and subsequent cervical shortening [533]. These groups of women would have, yet again, a different set of metabolites and biochemicals that are predictive of the shortening event. And therefore, the presence of mixed underlying aetiologies, complicated the identification a subset of metabolites that are consistently predictive of cervical shortening. A different underlying aetiology is likely to explain also why cervical cerclage works only for a portion of women that present cervical shortening. Our models show that patients where cerclage keeps the cervix long, present a different CVF metabolic profile compared to patients that experience a shortening despite the treatment, although with inadequate predictive capacity.

Nevertheless, even if we cannot predict directly outcome or cerclage efficacy, there is now increasing evidence that the microbial response to cervical suture is a major determinant of the outcome of pregnancy [139], [199], [452].

The insertion of a stitch can in fact promote the growth of high diversity bacterial communities. For example Kindinger et al. showed how, depending on the filament material used, cerclage caused a shift towards vaginal microbiome dysbiosis and reduction of *Lactobacillus*

species [139]. Our objective was to understand whether or not having a suture would alter the cervical vaginal metabolome in a way that no longer allows us to reliably predict the microbiota composition in these patients. Our results suggest that cervical cerclage does not affect the ability of rapid on-swab metabolic profiling by DESI-MS to robustly predict vaginal microbiome composition. It is worth noting that the women included might have had different types of procedures, performed with different materials and some might have received antibiotic therapy. Despite all these variables DESI-MS confirmed to be a robust technique for microbiota composition prediction.

As a whole, these findings highlight that direct on swab metabolic profiling by DESI-MS is an innovative approach with the potential for the stratification of preterm birth risk phenotypes associated with previous pregnancy history and with cervical shortening. Also, this technique has the potential to be used as a tool for the monitoring and prediction of response to commonly used preventative interventions.

The ability to robustly predict microbiota composition and host response at the bedside using on-swab metabolic profiling through DESI-MS, would provide the clinicians with a rapid and objective information about women identified with pathogenic microbial compositions. If this information it is acquired post-suture, the clinician could benefit from more detail and real-time information about shifts in microbial composition and could prescribe more targeted antibiotic therapies. Similarly, women who do not appear to be responding to the treatment in a negative way, may be spared from unnecessary antibiotic risk. DESI-MS approach would also provide the capacity to take multiple samples and assess over a longer period of time the treatment response and determine whether or not additional clinical intervention is required.

One of the current limitations of introducing this type of technology into point of care diagnostics in a clinical setting is that the mass spectrometer that has been used in the studies up until now has a major drawback. Its large size renders it impractical to be placed on a mobile trolley, hindering the feasibility of genuine bedside care. Moreover, this highly sensitive machine is susceptible to environmental influences that can potentially affect the accuracy of results. Moving the instrument requires recalibrations, further adding to the logistical difficulties associated with its use. Nevertheless, ongoing efforts for the miniaturization of DESI-MS are currently underway and advancements in this area will be described in more detail in Chapter VI.

Chapter V

5 MASS SPECTROMETRY-BASED MEASUREMENT OF PROGESTERONE AND STEROIDS IN THE CERVICOVAGINAL NICHE DURING PREGNANCY

5.1 Introduction

In previous chapters the capacity of the CVF metabolome to reflect important risk factors of preterm birth such as host:microbiota interactions (Chapter 3) and previous history of PTB or cervical treatment (Chapter 4) were described. However, an infectious aetiology is associated with only approximately 40% of PTB cases [23] and many other mechanisms have been proposed to account for the remaining proportion of cases [534]. One of the most commonly purported mechanisms was proposed by Prof. Roberto Romero who formulated the hypothesis that a subset of preterm deliveries are caused by local progesterone deficiency, that is subclinical in nature and can only be observed through phenotypical changes in the cervix [95]. This hypothesis is based on the important role played by progesterone and other steroids in mediating pro-labour and pro-parturition pathways in gestational tissues [64]. For example, the binding of progesterone to nuclear receptors mediates cellular transcriptional activity targeting genes involved in the inflammatory response or contraction-associated proteins via inhibition of transcription factors such as NF-kB and AP-1 [73], [74]. This inhibition is achieved through the upregulation of the NF-kB inhibitor IkBa [73] and increased expression of ZEB-1, which suppresses the expression of contractile genes such as CX43, encoding a gap-junction protein, and oxytocin receptor. Through these mechanisms, progesterone maintains uterine quiescence and fosters an anti-inflammatory environment throughout the gestation, contributing to the successful maintenance of pregnancy until term.

In many other mammals progesterone levels drop at the end of pregnancy triggering the onset of labour, but in humans plasma levels of progesterone remain high throughout labour and decrease only after expulsion of the placenta [65]. Nevertheless, progesterone central role in mediating labour onset has been demonstrated by several studies. Blocking progesterone action at any stage during pregnancy (e.g. with a progesterone receptor antagonist such as RU486) results in the initiation of parturition and interruption of pregnancy in humans [76]–[78]. It has been proposed that progesterone signalling may be supressed towards the end of human gestation through changes in the expression of progesterone receptor isoforms with an increase

in the expression of PR-A, the functional antagonist of progesterone, leading to a 'functional withdrawal' of progesterone [535]. Another proposed mechanism involves the conversion of progesterone by 20αHSD in 20α- hydroxyprogesterone (20α- OHP), an inactive metabolite [80]. Abundance of 20αHSD protein and enzymatic activity in myometrium has been associated with the onset of labour and with the exposure of myometrial cells to pro-inflammatory and pro-labour stimuli [68], [536]. Accordingly, natural progesterone is widely used as a therapeutic to prevent PTB. Micronization of progesterone in oil, a process consisting of the attachment of small progesterone crystals to long chain fatty acids [537] allows for efficient absorption in the body[538]. Micronized progesterone was first described around 1980 but introduced in the clinical practice only in the late 1990s both in oral formulations and subsequently, in the form of vaginal gels and pessaries [128]. In 2015, the International Federation of Gynecology and Obstetrics (FIGO) released a series of guidelines for the use of progesterone in preventing preterm birth. In this document it is suggested the use of a daily dose of micronized progesterone for women that present a short cervical length (≤ 25mm) upon measurement with transvaginal ultrasound between 19 and 23 weeks of gestation [539].

Despite the widespread use of progesterone as a preventative therapy in the clinical practice, there remains conflicting evidence regarding its possible mechanism of action and its efficacy. In 2003 Fonseca et al. reported that the use of prophylactic vaginal progesterone suppositories decreases the rate of PTB and the frequency of uterine contractions [540]. In the same year Meis et al. reported that intramuscular injection of a synthetic progestogen (17-hydroxyprogesterone caproate) in women with previous preterm birth showed a reduction in recurrency compared with placebo treatment [132]. In 2013, a Cochrane review on the use progesterone for PTB prevention concluded that for women with a history of PTB, progesterone reduced the risk of PTB before 34 weeks, perinatal mortality, and incidence of low birthweight. For women with premature cervical shortening, progesterone was only effective in reducing PTB before 34 weeks [541]. However, in 2016 the double-masked placebo-controlled randomised OPPTIMUM trial conducted in 1228 women, failed to show any beneficial effects of vaginally administered progesterone on the rates of fetal death or PTB<34 weeks and moreover, reported that women with a short cervix did not respond any differently to any other women in the study [542]. Similar results were also found in the PROGRESS study the following year [543]. An overview of Cochrane systematic reviews in 2018 reported that intramuscular progesterone may actually increase risk of PTB in women with multiple pregnancy and more recently, the FDA advisory committee has recommended

17-hydroxyprogesterone caproate be withdrawn from the market for the prevention of preterm delivery [544].

In addition to the contradicting results around efficacy, the mechanistic rationale behind progesterone therapy for PTB prevention remains questionable given that progesterone supplementation does not change circulating levels [132], [545], [546] and levels of serum progesterone are actually slightly higher in the third trimester of women who subsequently deliver preterm [547], [548]. However, whether this is reflected in the levels of progesterone in the cervicovaginal niche remains unclear. In addition to progesterone, many other steroid hormones are involved in the regulation of labour. For example cortisol, which peaks in plasma concentrations at the beginning of labour, influences the production of prostaglandins in human amnion cells [549]–[551]. Prostaglandins, in turn promote myometrial contractility and cervical ripening. Cortisol is also involved in a positive feedback loop with CRH that increases placental production of estrogen [5]. Estrogens are also involved in the cervical ripening process through the activation of collagenase and elastase [49]. Androgens, during pregnancy, act as substrate for estrogen production in the placenta [552], [553]. They also have a direct role in cervical remodelling by enhancing collagenase activity and thus decreasing fibril collagen organization [554], [555]. Androgens have also been demonstrated to relax smooth muscle cells, in both animal models and humans [556]–[560]. Other studies have demonstrated similar relaxing effects of various androgens including testosterone, dihydrotestosterone, androstenedione and dehydroepiandrosterone sulphate (DHEA-S) also on the myometrium of non-pregnant rats and non-pregnant and pregnant humans [556], [561].

Currently, the investigation of steroid hormones in the context of the timing of parturition have been mainly limited to their systemic measurement in body fluids like urine, blood or plasma using immunoassays. To date, only one study has attempted to measure progesterone in cervicovaginal mucus through immunoassay [562]. However, this study was limited to a small number of poorly characterised patients at the end of pregnancy and reported no difference in levels of progesterone in women delivering term and preterm. Over the last two decades significant advancements in mass spectrometry (MS) have facilitated the routine analyses of steroid hormones in clinical and research settings [563]. MS presents several advantages compared to classical immunoassays despite currently being a slightly more costly technology. Immunoassays are hampered by issues of cross reactivity and relatively low specificity, particularly with multiplexed assays [564]. MS offers higher accuracy and direct *de novo*

identification and simultaneous measurement of multiple compounds in a single run. Comparisons of relative concentrations or ratios of compounds measured by MS have been shown to facilitate characterisation and diagnosis of several endocrine conditions [563], [565]. Endogenous steroids have been successfully measured both by GC-MS [566], [567], and high-performance liquid chromatography–tandem mass spectrometry (UPLC-MS/MS)[568]–[573] in serum, urine and even saliva. However, these assays have yet to be adapted for, and tested in CVF. Quantitation of steroid hormones in CVF could allow the hypothesis of progesterone deficiency in PTB to be tested.

5.2 Hypothesis and aims

We hypothesise that the levels of progesterone and other steroids in CVF can be measured by MS approaches and are reflective of systemic changes during pregnancy. Secondly, we hypothesise that CVF levels of progesterone and other steroid hormones differ in women who experience PTB compared to gestational age matched controls. To test these hypotheses, this chapter had several aims:

1. Assess the capacity of on-swab DESI-MS to detect and quantitatively measure progesterone in CVF samples.

2. Examine the capacity of UPLC-MS steroid assays to detect and measure sex steroids in cervicovaginal samples collected via soft cup devices.

3. To determine if steroid levels quantified in the CVF follow the same trend across gestational age as those reported in serum.

4. Measure and compare progesterone concentrations CVF in women experiencing PTB and gestation age matched controls.

5.3 Results

5.3.1 Progesterone measurement with DESI-MS

Detection and measurement of progesterone in CVF was first attempted using direct swab analysis by DESI-MS. For this experiment samples from the VMET2 cohort of 205 pregnant women at high risk of PTB were used. As mentioned in Chapter 3.1.1, swab samples (n= 573) were collected longitudinally during pregnancy and analysed by DESI-MS both in negative and positive ion modes. Progesterone was detected in positive ion mode DESI-MS data in the form of four different adducts: $[M+H]^+$ (m/z=315.231), $[M+Na]^+$ (m/z=337.337), $[M+K]^+$ (m/z=353.187), $[2M+Na]^+$ (m/z=651.443) (Figure 5.1). These adducts were identified putatively.

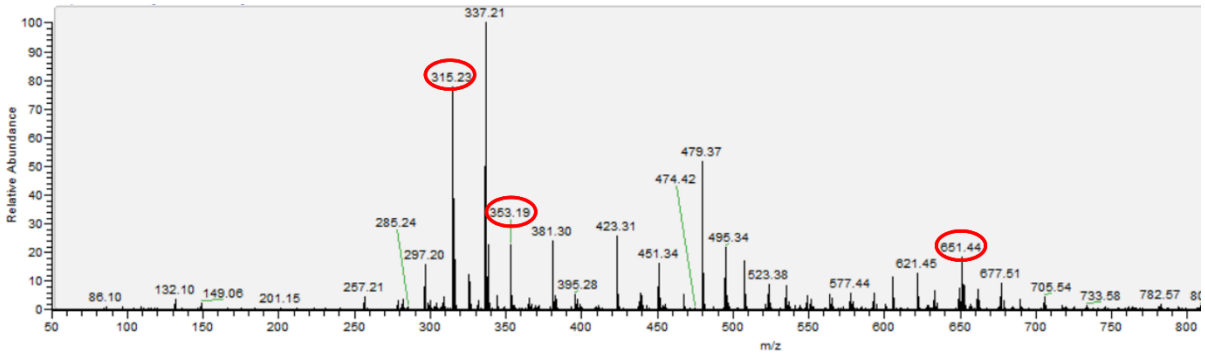


Figure 5.1- Example of a positive ion mode DESI-MS spectrum where some progesterone adducts are visible. In this example, 3 out of 4 identified adducts of progesterone are present at high intensity.

The protonated adduct could be observed near to the limit of detection (Figure 5.2) however examination of distribution of intensities of the peak across all samples suggested progesterone could only be identified with sufficient sensitivity in a relatively small subset of samples (n=175, 30%). Similar results were obtained for the other detected adducts (results not shown).

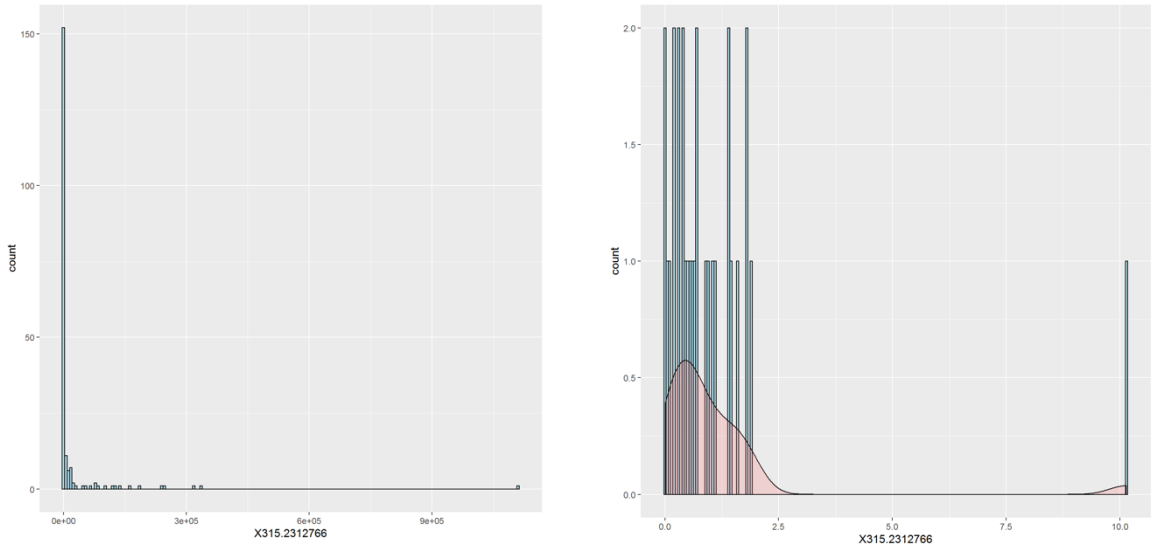


Figure 5.2- Histograms showing the distribution of the intensities of the $[M+H]^+$ adducts detected by DESI-MS. (A) Histogram showing that in the majority of samples where the adduct is detected its intensity is near zero. (B) Features with an intensity lower than 10 suggest a normal distribution curve cut by the limit of detection of DESI-MS.

5.3.2 Measurement of progesterone and related steroids in cervicovaginal fluid by UPLC-MS/MS steroids assay

We next investigated the possibility of detecting progesterone and other steroids using a UPLC-MS/MS assay designed for measurement of steroid hormones in urine and blood plasma. Sample preparation procedures were initially adapted to analyse cervicovaginal fluid collected from rayon swab samples. This involved extraction of water and lipid soluble metabolites from stored swabs. However, progesterone detection and recovery from the swab matrix was complicated by low cervicovaginal fluid yield, and high abundance of swab polymers (data not shown). Therefore, assaying was performed on CVF soft cup samples. The utilization of a soft cup collection device allows for the retrieval of larger volumes of cervicovaginal fluid while avoiding any potential interference from synthetic polymer matrices. The softcup samples were derived from a subset of patients enrolled in the VMET 2 study and included a small number of patients sampled longitudinally throughout their pregnancy. Exclusion criteria for these samples included women receiving progesterone therapy. This resulted in n = 30 samples suitable for the analysis. Of these, 20 patients were sampled cross sectionally and 5 longitudinally at 2 time points. All samples included in the study were collected between 12 and 24 weeks of gestational age and were divided in “low” (n = 14) or “high” (n = 11) risk of

PTB based on their previous pregnancy history (Figure 5.3). Table 5.1 provides key demographic and clinical information for the patients included in the study.

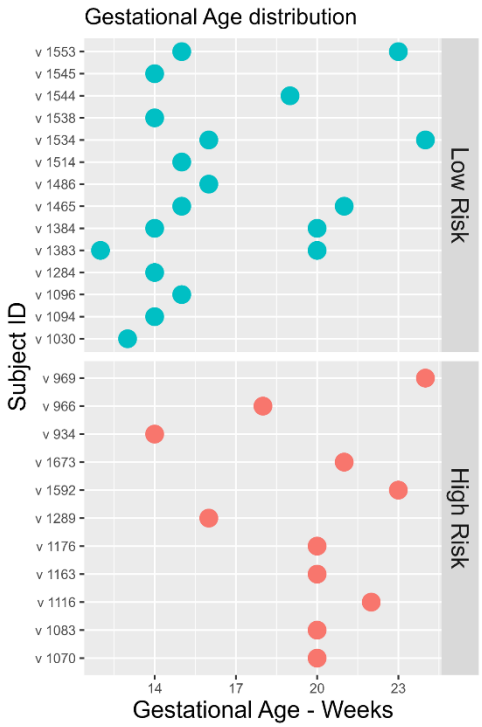


Figure 5.3 - Gestational age at time of soft cup sampling of a subset of VMET2 patients. Patients with a high (red) or low (green) background risk of PTB are sampled between 12th and 24th week of pregnancy.

Table 5.1 Demographic and pregnancy outcome information of patients providing soft cup samples

	Total	***Control***	***High Risk***
	N=25	**N=14**	**N=11**
Maternal Age (average)	33.4	33.3	33.8
BMI (average)	24.1	24.1	25.5
Ethnicity			
Black	5	0	5
Other	5	4	1
White	15	10	5
Parity (range)	0 - 3	0 -3	0 -3
Outcome			
Miscarriage	1	0	1
Preterm	9	0	9
Term	15	14	1
Gestational age at delivery (average)	34.35	36.07	27.5
Longitudinal Sampling			
No	20	9	11
Yes	5	5	0

As detailed in Chapter 2.4.5, the UPLC-MS/MS steroids assay was designed to quantify levels of 13 steroids. However, the concentration of the majority of targeted steroids in CVF were lower than the assay limit of detection (see Appendix table 8). Only progesterone, androstenedione, and cortisol were detected above limits of detection (0.005ng/ml, 0.01ng/ml, 0.05ng/ml respectively). One sample presented missing values (concentration estimate below < LLOQ). These were imputed by replacing with ½ of the LLOQ.

To obtain absolute concentrations (ng of steroid/gram of sample) serial dilution factors were applied to factor in the sample preparation procedure and concentrations were corrected using the wet weight of the sample at time of collection. A schematic representation of the workflow for the soft cup sample preparation procedure is shown in Figure 5.4.

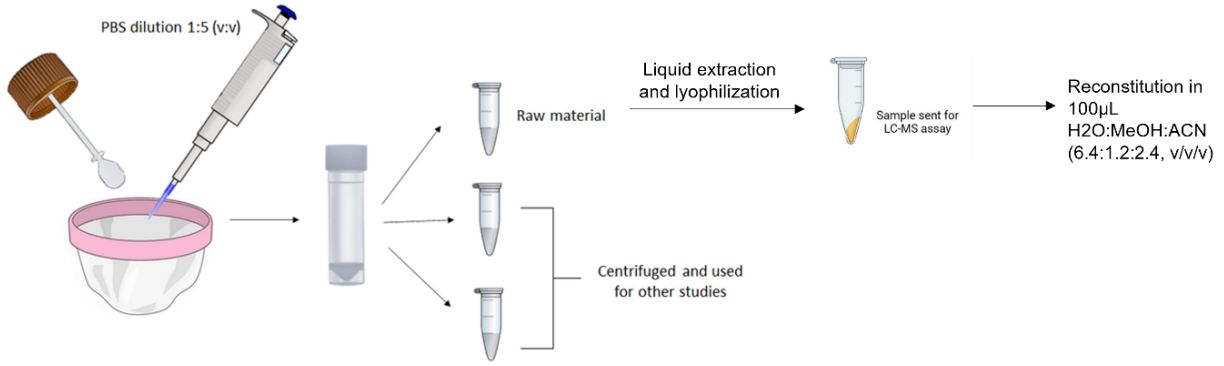


Figure 5.4 – Schematic representation of the soft cup sample preparation procedure from collection to analysis by UPLC-MS/MS.

CVF is diluted with a volume of PBS corresponding to 5 times the weight. This will achieve a dilution of 1:5 (v/v). PBS and vaginal fluids are mixed through pipetting and transferred in a clean 5ml tube. Raw uncentrifuged aliquots of the CVF were used for this study. Liquid extraction was performed and lyophilised extracts were reconstituted in 100µl of H2O: MeOH:CAN (6.4:1.2:2.4, v/v/v) prior to assaying.

The median (± SD) concentrations of the detected steroids in CVF collected from term pregnancy samples (n=19) were 6.61 ± 8.13 ng/g for progesterone, 4.16 ± 2.98 ng/g for cortisol, and 0.34 ± 0.21 ng/g for androstenedione (Figure 5.5).

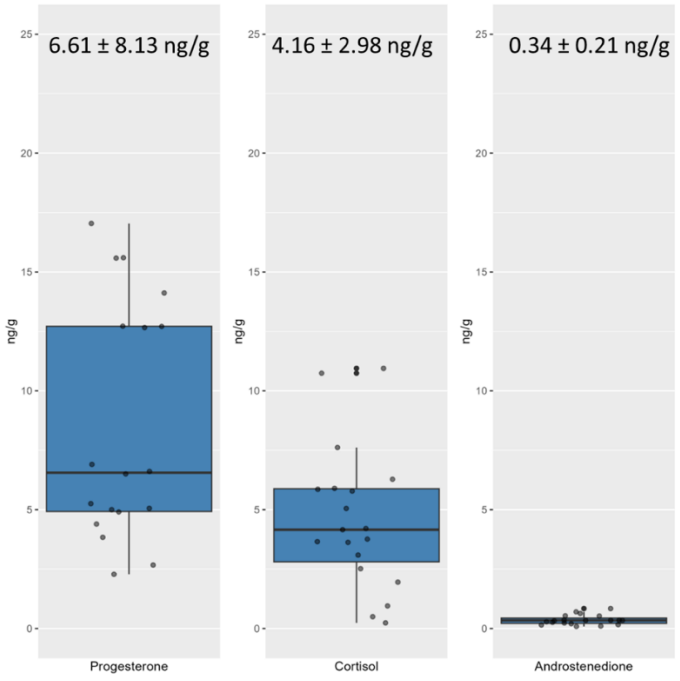


Figure 5.5 – Box plots representing values of progesterone, cortisol, androstenedione in the cervicovaginal mucosa. The values reported are median ± SD. Progesterone presented the highest values. Median levels of cortisol were approximately half that of progesterone and androstenedione presented the lowest concentration among the measured steroids.

We next examined the relationship between gestational age and the concentration of progesterone, androstenedione, and cortisol. This was initially tested with all available samples collected cross sectionally with the results of linear regressions between steroid concentrations and gestational age shown in Figure 5.6. A weak but significant positive correlation between progesterone levels and increasing gestational age was observed (R^2=0.15, p-value = 0.02). No significant association between gestational age and levels of cortisol or androstenedione were observed (cortisol R^2= -0.03, p-value = 0.83; androstenedione R^2= -0.02, p-value = 0.65).

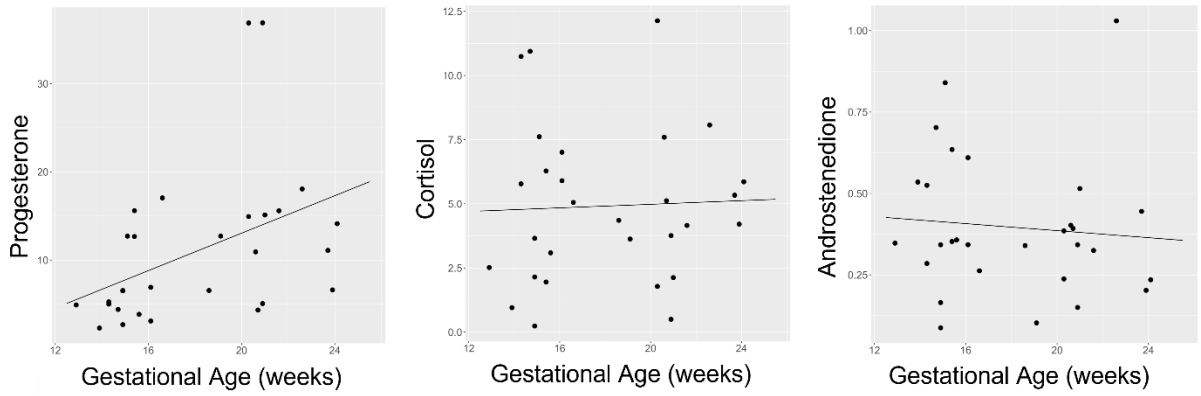


Figure 5.6 - Scatterplot with values of progesterone, cortisol, androstenedione in the cervicovaginal mucosa during gestation.

The figure illustrates the relationship between hormone levels and gestational age in the study population. A significant positive correlation was observed between progesterone levels and increasing gestational age (R2 = 0.15, p = 0.02). However, no significant association was found between gestational age and cortisol levels (R2 = -0.03, p = 0.83) or androstenedione levels (R2 = -0.02, p = 0.65).

Analysis of steroids levels for pregnancies where longitudinal samples were available (n=5 patients) showed a significant increase in progesterone levels (paired Wilcoxon test progesterone p-value = 0.03) but no significant relationship for cortisol and androstenedione, (paired Wilcoxon test cortisol p-value = 0.8, androstenedione p-value = 0.3) across gestation when looking within individual. Timepoint A is defined between 12 and 16 weeks, timepoint B is defined between 20 and 24 weeks. Results are shown in figure 5.7.

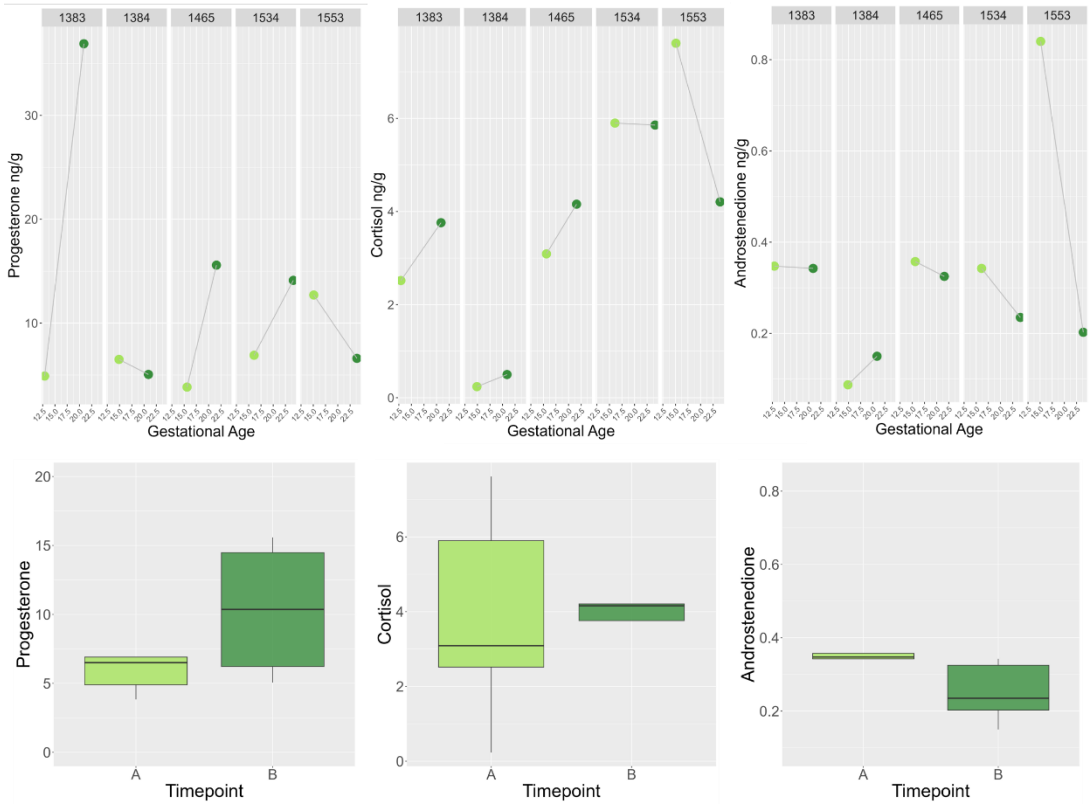


Figure 5.7 - Individuals with longitudinal sampling values of progesterone, cortisol, androstenedione, and difference between timepoint A and B. Mean of steroids levels at the two timepoints were compared with paired Wilcoxon-test: progesterone p-value = 0.03, cortisol p-value = 0.8 androstenedione p-value = 0.3.

Comparison of the levels of progesterone, cortisol, and androstenedione in patients with term pregnancies (n= 15) and preterm ones (n= 9) revealed no significant differences (Wilcoxon test progesterone p-value = 0.51, cortisol p-value = 0.17, androstenedione p-value = 0.57) (Figure 5.8).

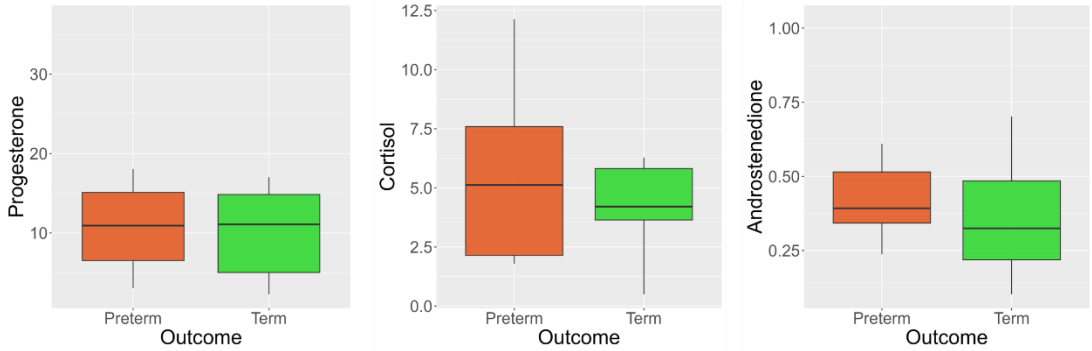


Figure 5.8 - Boxplot comparing values of progesterone, cortisol, androstenedione between term and preterm cases. Mean of steroids levels in the two groups were compared with Wilcoxon test: progesterone p-value = 0.51, cortisol p-value = 0.17, androstenedione p-value = 0.57.

5.4 Discussion

Steroid hormones have a central role in regulating human parturition. Given its important anti-inflammatory and pro-gestation role, it has been proposed that a deficiency of progesterone at a local level could drive PTB [95], [574]. This hypothesis is partly corroborated by some studies indicating that progesterone therapy prevents PTB in specific subgroups of pregnant women [132], [540]. However, neither its efficacy nor its mechanism of action have been definitively demonstrated. While systemic blood levels of progesterone remain high even in women who subsequently deliver preterm, it is feasible that local levels in the cervicovaginal environment are depleted. To begin to investigate this, we set out to explore the possibility of using different MS-based methods to detect and quantify levels of progesterone and other steroids in CVF. To begin we assessed the utility of direct swab analysis by DESI-MS to detect and measure progesterone on a large number of stored vaginal swab samples. However, the results indicated that progesterone could only be detected in a small subset of swabs. There are several possible explanations for this. Ambient ionisation-based methods are prone to ion suppression or matrix effects whereby the matrix components of a sampling device (in this case the rayon swab) can interact with the analyte and the instrument, leading to changes in the ionization efficiency, fragmentation pattern, and overall signal intensity [456]. Additionally, the structural characteristics of steroid hormones may lead to low ionization efficiency, with consequent reduced sensitivity [575]. A possible solution for this is chemical derivatization, which can be used to modify the functional groups in target molecules and improve the ionisation efficacy [576], however this requires further investigation.

To address issues regarding poor detection of progesterone by DESI-MS, I initially investigated used of targeted LC-MS based assays performed on metabolite swabs extracts. However, progesterone detection and recovery from the swab matrix was also poor and resulting spectra were contaminated high abundance of swab polymers. Contamination by polymers from the swab can likely be attributed to their stripping from the swab matrix during extraction in solvents used during the extraction procedure. Additionally, ion suppression and subsequently reduced ionisation efficiency and interference with MS analysis, especially when using liquid-liquid solvent extraction procedures may have further added to issues of detection. For these reasons, it was decided that the use CVF recovered from soft cups could provide a sample type more conducive to extraction and subsequent analysis by UPLC-MS.

When selecting CVF samples for analysis, care was taken to exclude patients receiving therapeutic progesterone. This was particularly important as patients recruited to this study were from high-risk preterm birth clinics where vaginal progesterone is a commonly used intervention. Unfortunately, this limited the number of samples that could be analysed in this pilot study. Regardless, the UPLC-MS/MS assay was found to successfully identify and quantify three out of the thirteen targeted steroids. There are several reasons as to why the remaining steroids could not be identified. One possibility is that these compounds undergo significant local metabolism, which might differ greatly from their systemic metabolism leading to concentration levels below the detection limit of the assay. Given that their concentrations are substantially higher in circulation, it is possible that these steroids do not play a significant functional role at the cervicovaginal mucosal interface but this was not explicitly examined here.

Consistent with what has been reported in blood samples, cross sectional analysis of CVF showed that progesterone levels increase in the cervicovaginal niche during the second trimester [97], [577], [578]. In contrast no significant change was observed for cortisol and androstenedione across this gestational age range. This is in contrast to what has been reported in serum where androstenedione levels have been shown to increase during pregnancy, peaking in the third trimester [552]. Similarly, maternal serum levels of cortisol increase during pregnancy from on average 141 ± 8 ng/ml at 5th week to 213 ± 5 ng/ml 20 weeks gestation [579]. One possible explanation for this difference might be that the relationship between these steroids and gestational age in the CVF could be non-linear and the model we have currently adopted does not reflect appropriately the nature of the biological trend. To investigate these trends more closely, longitudinal analyses were performed in a small subset of patients where two samples were collected at ~12 weeks and ~24 weeks of gestation. When comparing the paired analysis in patients sampled longitudinally with the cross-sectional analysis in relation with gestational age, our results showed some similarities. Progesterone showed increase with gestation consistently across the two analyses. Cortisol and androstenedione levels in CVF do not appear to have a significant relationship with gestational age. Caution must be taken in interpreting our data since longitudinal analysis was performed on only five patients and results might not be representative of a larger patient cohort population. Additionally, we can observe a wide interpatient variability. Within different individuals, the initial concentrations of each of these steroids are very variable. Despite this, at least for progesterone the longitudinal analysis indicates that there is a relatively consistent trajectory over time. This is encouraging

and indicates that this analysis should be performed on a larger population to validate our findings.

Finally, comparison of CVF levels of the detected steroids between term and preterm pregnancies failed to show any difference in their levels therefore refuting the hypothesis that progesterone deficiency drives PTB. These findings are broadly consistent with a recent cross sectional study by Lisnawati et al., who using immunoassay, reported that progesterone levels in cervicovaginal mucus of pregnant Indonesian women are slightly higher in women who subsequently deliver preterm compared to term pregnancies [562]. Moreover, their study identified a correlation between progesterone concentrations in CVF and plasma in PTB patients but not in healthy pregnancies. However, this study had several limitations that make interpretation and comparison of the findings challenging. For example, no exclusion criteria or any other clinical information was reported. The studied cohort included women at high risk of PTB, therefore it is plausible that some of the participants may have received therapeutic progesterone, which could have generated spurious estimates of progesterone concentrations. Furthermore, women were only sampled late in pregnancy (between 28-34 weeks of gestation), and it appears that reported concentrations were not corrected for weight of CVF collected, which makes comparison of their reported results with our own almost impossible. However, if we recalculate our findings and present them in the same units as Lisnawati et al. then a similar median value of cervicovaginal progesterone levels in healthy pregnancy is observed; 1.74 ng/ml versus 1.32 ng/ml. In other biofluids, including serum, comparison of steroid concentrations measured using either immunoassay or MS showed some differences in absolute values but congruent findings when slopes and trajectories of concentrations compared [580].

Our pilot study had several limitations that should be considered. Firstly, we only included a relatively small sample size of well phenotyped patients, which limits the power to detect more subtle differences in the concentrations of steroids in the CVF. Further, longitudinal sampling was limited to between ~12 weeks and ~24 weeks of gestation. It remains possible that functionally important changes in local steroid levels occurs outside of this timeframe.

In conclusion, progesterone, androstenedione, and cortisol are detectable in CVF. Progesterone concentration increases with gestational age, similarly to what is observed systemically. Finally, while validation in a larger cohort is required, our results indicate that preterm birth is not associated with CVF progesterone deficiency.

Chapter VI

6 ASSESSMENT OF CVF MICROBIOTA-HOST INTERACTIONS IN PREGNANT WOMEN RECEIVING LIVE BIOTHERAPEUTIC INTERVENTION

6.1 Introduction

As described in previous chapters, there is now substantial evidence that a highly-diverse vaginal microbiome depleted of *Lactobacillus* species increases the risk of infection and poor pregnancy outcomes, including PTB [23], [230], [235], [480]. Mechanistically, this is thought to involve enrichment of pathogenic bacteria that drive an inflammatory host response leading to premature activation of pro-labour and pro-parturition pathways [148]. In consideration of this, a potential strategy to prevent PTB involves modifying the vaginal microbiota composition. This could be achieved via the eradication of potential pathogens using antibiotics. While treating maternal infectious diseases (e.g. sexually transmitted diseases or upper respiratory tract infections) with antibiotic therapy can be an effective for pathogen eradication, off-target effects and increasing issues around antibiotic resistance limit this approach. Moreover, evidence supporting the use of antibiotic therapy to reduce PTB risk is highly inconsistent [581]–[584]. The majority of these studies have been conducted in pregnant women with BV and have been confounded by the selectively of antibiotics used, the timing of therapy commencement and differing routes of administration [585], [586] [242]. For example, the treatment of women with BV in mid-pregnancy is likely to be ineffective to prevent PTB as inflammatory pathway activation may have already started and antibiotic treatment could actually exacerbate inflammation through release of microbial components [587] . For example, lysis of bacteria by metronidazole may release endotoxins [147], which in turn activate inflammation in gestational tissues enhancing the inflammatory phenotype [149], [150]. Additionally, antibiotic therapy may be more effective against commensal species such as *Lactobacillus*, than against pathogens. The consequent depletion of *Lactobacillus* species could then leads to the inadvertent enrichment of pathogenic species in the vaginal niche [151], [153], [254]. The use of antibiotics in pregnancy has also been associated with adverse fetal outcomes. Vermuelen et al. reported that the use of clindamycin in pregnancy did not prevent PTB but resulted instead in an increase in neonatal infectious morbidity [588]. Antibiotics use in pregnancy has also been linked to elevated risk of necrotizing enterocolitis

[157] and cerebral palsy [584], [589] and later in life, the development of allergic conditions [590].

An alternative strategy for modifying the vaginal microbiota that is gaining attention is the use of live bio-therapeutic products. Live biotherapeutics are products containing live microorganisms and are designed to treat, cure, or prevent a disease or condition in humans [264]. Live biotherapeutics in the context of reproductive health generally aim to promote *Lactobacillus* species dominance in the vaginal niche. Several products have previously been tested for their ability to treat BV and UTIs and are most commonly administered orally. Often, these products include bacterial strains that have been derived from non-vaginal sources (e.g. food industry, breast milk, fecal samples) [591]–[594]. Although earlier studies using culture-based approaches suggesting that oral administration may increase vaginal *Lactobacillus* dominance [274], subsequent studies based on DNA sequencing methods, indicate that oral administration is not an effective strategy for vaginal colonisation [595], [596]. These limitations have led to the development of several new products that are vaginally administered, the most tested of which is LACTIN-V (Osel, California). This live biotherapeutic contains the strain *Lactobacillus crispatus* CTV-05, which was originally isolated from the vagina of a healthy woman [597], [598]. CTV-05 is capable of producing both L- and D- lactic acid, H_2O_2 and unlike most commercially available probiotic *Lactobacillus* strains, has been shown to adhere to vaginal epithelial cells and colonise the vagina [597], [599], [600]. This product has been previously used vaginally to treat BV recurrence in non-pregnant women and has an excellent safety and tolerance profile [601]. More specifically, LACTIN-V has been used in conjunction with vaginal metronidazole antibiotic therapy and resulted in a significant lower incidence of recurrent BV [255], [602] and in a sustained reduction in genital inflammation [603].

Our group recently performed a first in pregnancy trial (FLIP-1) exploring the safety, tolerability, and acceptability of LACTIN-V use in pregnant women at high risk of PTB. This study showed that LACTIN-V was safe and well tolerated by participants. Women found LACTIN-V easy and comfortable to use and the majority (30/31, 97%) would use LACTIN-V in future pregnancies. Although not powered to assess PTB as an outcome, the rate of PTB <34 weeks in women treated with LACTIN-V was 3.3% compared to 7% in historical cohort of 2,190 women at similar background PTB risk. As part of the trial, participants were carefully monitored and sampled across different timepoints including before, during and after treatment

with urine, blood, vaginal swabs, stools, cytobrush and soft cups CVF samples collected. The FLIP-1 trial therefore represented an ideal cohort of patients to further test the capacity of direct swab profiling by DESI-MS to rapidly assess and monitor vaginal microbiota-host interactions in pregnancy. However, a major limitation of transitioning the DESI-MS platform to a clinical setting is that is that it is currently coupled to a Thermofisher orbitrap mass spectrometer, which has a very large footprint and is not portable. The ACQUITY RDa Detector recently developed by Waters Corporation is a small-format, portable Time-of-Flight mass spectrometer that allows for acquisition of high-quality, accurate-mass data for routine small molecule applications. Importantly it is compliance-ready to satisfy FDA requirements for diagnostic applications [604], [605]. In the work presented in this chapter we apply direct on-swab metabolic profiling by ACQUITY RDa Detector to FLIP-1 cohort.

6.2 Hypothesis and aims

In this chapter, I set out to test the hypothesis that Direct Swab Profiling by DESI-MS using a small format portable mass-spectrometry platform enables effective prediction of vaginal microbiota composition and immune status in women receiving live-biotherapeutic intervention during pregnancy. Moreover, we hypothesised that pre-treatment CVF metabolic profile would enable stratification of women who can be successfully colonised with LACTIN-V from those who cannot. To test these hypothesis this chapter had the following aims:

1. Determine if vaginal microbiota composition can be robustly predicted using direct on-swab analysis by DESI-MS using RDa platform in women receiving LACTIN-V.

2. To investigate if DESI-MS swab metabolic profiles enable prediction of LACTIN-V colonisation of the vaginal mucosa.

3. To determine if DESI-MS swab metabolic profiles can be used to predict inflammation status in women receiving live biotherapeutic intervention

6.3 Results

6.3.1 Sample cohort

A total of 62 pregnant women at high risk of PTB were recruited and included in this study. Swab samples were collected longitudinally during pregnancy before and after a live biotherapeutic intervention as described in Chapter 2.1.2. A total of 362 vaginal samples were collected and analysed by DESI-MS in negative ion mode on a small format RDa mass spectrometer. Additionally, swab samples collected at timepoint 0, 1, 3 and 4 (n= 227) from the same cohort were used for metataxonomics analysis through 16s rRNA sequencing. Further information regarding the cohort demographic can be found in Table 6.1. The PTB rate in this high-risk population was of 19% (n = 12).

Table 6.1 – Demographic characteristics of FLIP cohort

	Total	***Term***	***Preterm***
	n= 62	**n= 50**	**n= 12**
Maternal Age (average)	33.2	33.2	34.05
BMI (average)	25.9	26.86	27.25
Ethnicity			
White	37	32	5
Other	6	6	0
Black	5	5	0
Asian	7	7	0
Parity (range)	0 - 3	0 -3	0 -3
Gestational age at delivery (average)	34.35	36.07	27.5
Blood Group			
A positive	20	20	0
B positive	12	11	1
O positive	17	12	5
AB positive	6	3	3
O negative	3	1	2
AB negative	1	1	0
A negative	3	2	1
Smoker			
No	50	40	10
Ex	10	8	2
Yes	2	2	0
Cerclage			
No	48	39	9
Yes	14	11	3
Progesterone			
No	45	36	9
Yes	17	14	3

6.3.2 Assessment of DESI-MS metabolic profiles acquired by RDa

Following data processing, an average of 3658 peaks in negative ion mode were detected per sample using the RDa ACQUITY DESI-MS platform. Unsupervised PCA was performed to determine if major sources of variation in the resulting metabolic profiles could be attributed to any non-biologically associated factors, such as batch-effects. This was performed initially on unscaled spectral data, and showed that the majority of variation in the data was not attributable to date of spectral acquisition (Figure 6.1). The cumulative variance explained by the first two component amounted to 14%.

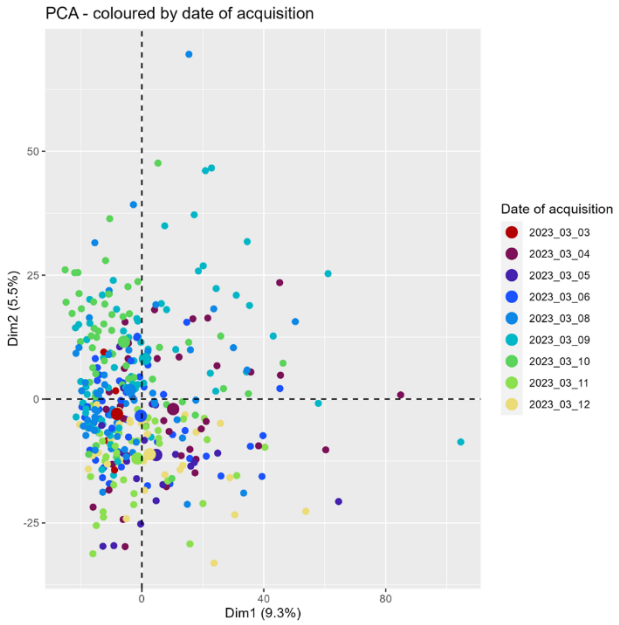


Figure 6.1 - PCA plot of all FLIP samples coloured by date of acquisition. The variance explained by the first two components of PCA reported was 14%.

Similarly, recolouring of the PCA score plots indicated that neither ethnicity or blood presence could account for the major source of variation in the resulting spectra profiles (Figure 6.2).

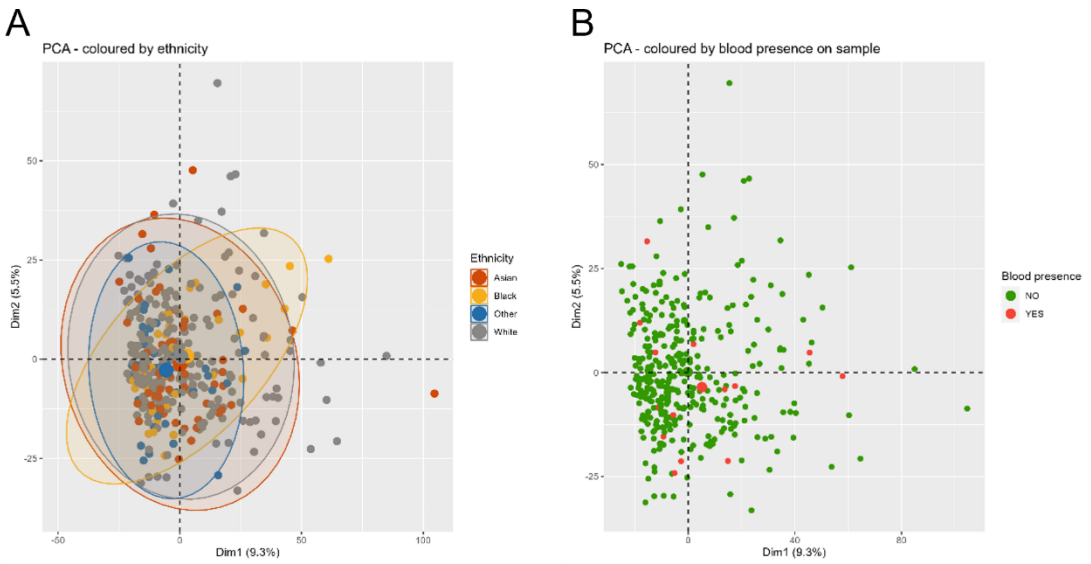


Figure 6.2 – PCA plots of all FLIP samples coloured by two main confounders. PCA plot recoloured according to ethnicity (A) and presence of blood contamination on the sample (B).

6.3.3 Microbial composition prediction using DESI-MS metabolic profiling

The DESI-MS profiling data was next integrated with metataxonomic data to examine the possibility that the microbiome was a source of metabolic variation in the CVF samples. For this, each sample was first assigned to vaginal CSTs using the VALENCIA clustering algorithm (see Chapter 2). A total of 292 samples were assigned to CST I (*L. crispatus*), 6 to CST II (*L. gasseri*) , 29 to CST III (*L. iners*), 24 to CST IV-B (high relative abundance of *G. vaginalis* and low relative abundance of BVAB1), 5 to CST IV-C (low relative abundance of *Lactobacillus* spp., *G. vaginalis*, *A. vaginae* and BVAB1 and increased diversity of other anaerobic bacteria) and 6 to CST V (*L. jensenii*). The samples were also classified at genus-level taxonomy into *Lactobacillus*-dominant or *Lactobacillus*-depleted. For the *Lactobacillus* dominant (LDOM) group samples from CST I, II, III and V were combined (n= 333 samples). *Lactobacillus* depleted (LDEPL) group includes samples from CST IV-B and CST IV-C (n= 29 samples. The metabolic features measured through DESI-MS in negative ion mode, were input in a linear mixed effect model to identify discriminant features between LDOM and LDEPL. After Benjamini-Hochberg correction, 410 metabolite features were found

significantly different between the two groups (p-value < 0.05). A subset of the significant features (n=58, p-value < 0.001) is shown in Figure 6.3.

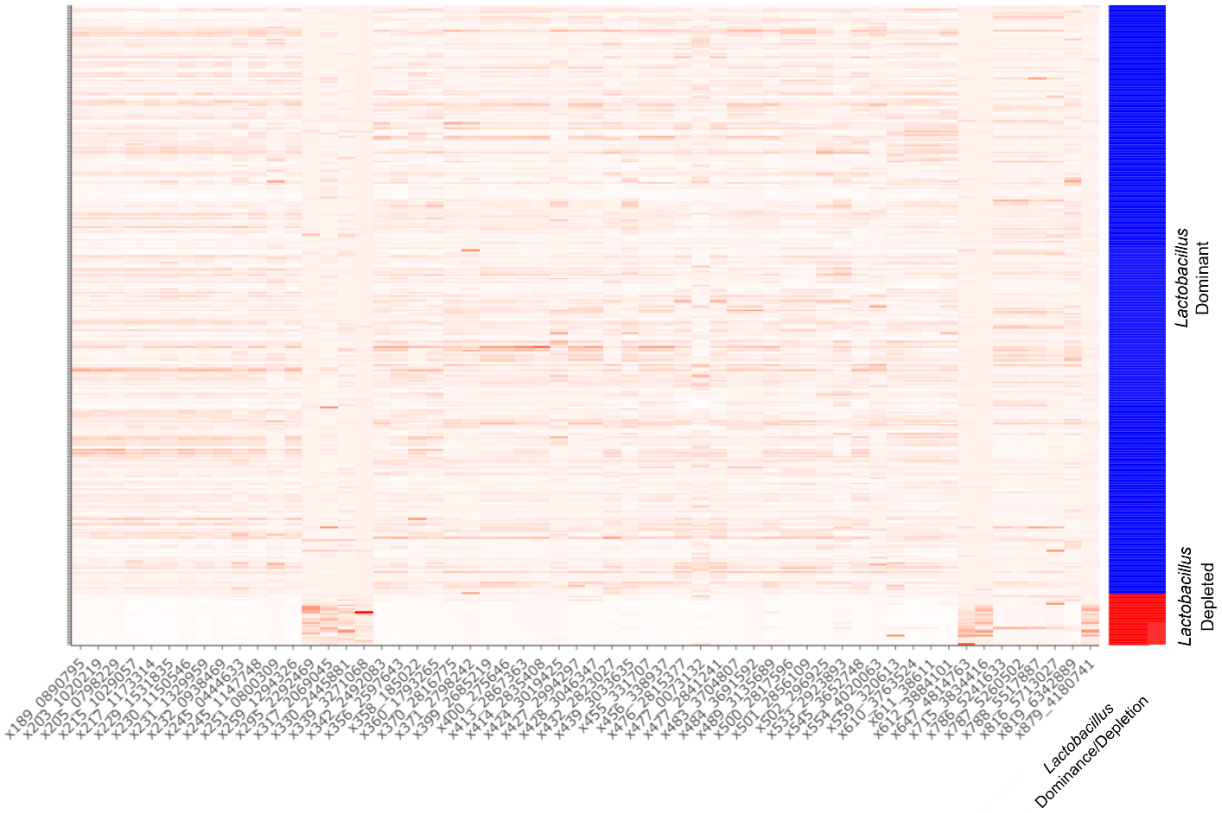


Figure 6.3 – Heatmap representing relative concentrations of DESI-MS derived metabolic features and their relationship with *Lactobacillus* dominance or depletion.

To enable comparison across the two DESI-MS platforms used during this PhD, the ppm error was calculated and used to identify metabolite features found to associate with LDOM and LDEPL in both the VMET and VMET2 cohorts previously described in Chapters 3 and 4. A total of 48 features were found to overlap between the platforms with an error range of ±20ppm. Further information about the 48 shared features, the relative ppm errors and the assignments is reported in Appendix table 10. The ID of 58% of the 48 shared features was confirmed with MS/MS experiments. An example of two features found to discriminate between LDOM and LDEPL are shown in Figure 6.4 together with a representation of their trajectory throughout gestation. One of these features N-acetylornithine (m/z 173.09) was found to be significantly enriched in LDOM samples independent of when in pregnancy the samples were collected. In contrast, Hydroxy caproic acid (m/z 131.07) was found to be significantly enriched in LDEPL samples.

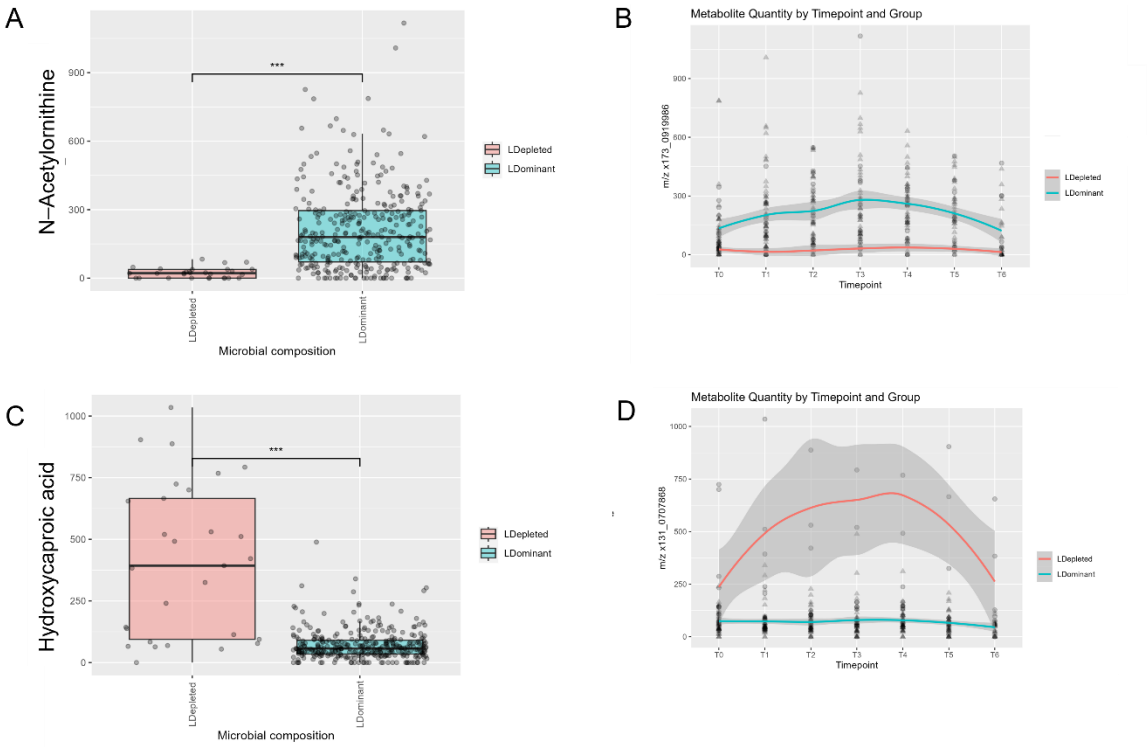


Figure 6.4 – Example of two metabolic features higher in LDOM or LDEPL and their change across gestation. N-Acetylornithine, is more abundant in LDOM group (p-value < 0.001) (A) independently of the timepoint of sampling (B). Conversely, Hydroxy caproic acid is higher in LDEPL group (p-value <0.001) (C), independently of the time of sampling (D). Both the features presented in this example were identified with MS/MS experiments.

The next objective of this work was to investigate whether the metabolic profile data generated by the RDa ACQUITY DESI-MS platform could be used to predict the microbial composition of CVF both at genera and CST levels. For this a random forest classifier model was built and the performance of the model assessed with ROC curve analysis. For the purpose of comparison the resulting ROC curve was plotted against the equivalent ROC curve generated from Orbitrap MS platform used in the VMET2 study (Figure 6.5). The model trained with FLIP data had an AUC of 0.98 indicating strong predictive ability to distinguish LDOM from LDEPL samples which was comparable to the AUC of the model trained with VMET2 data (0.97 for negative ion mode).

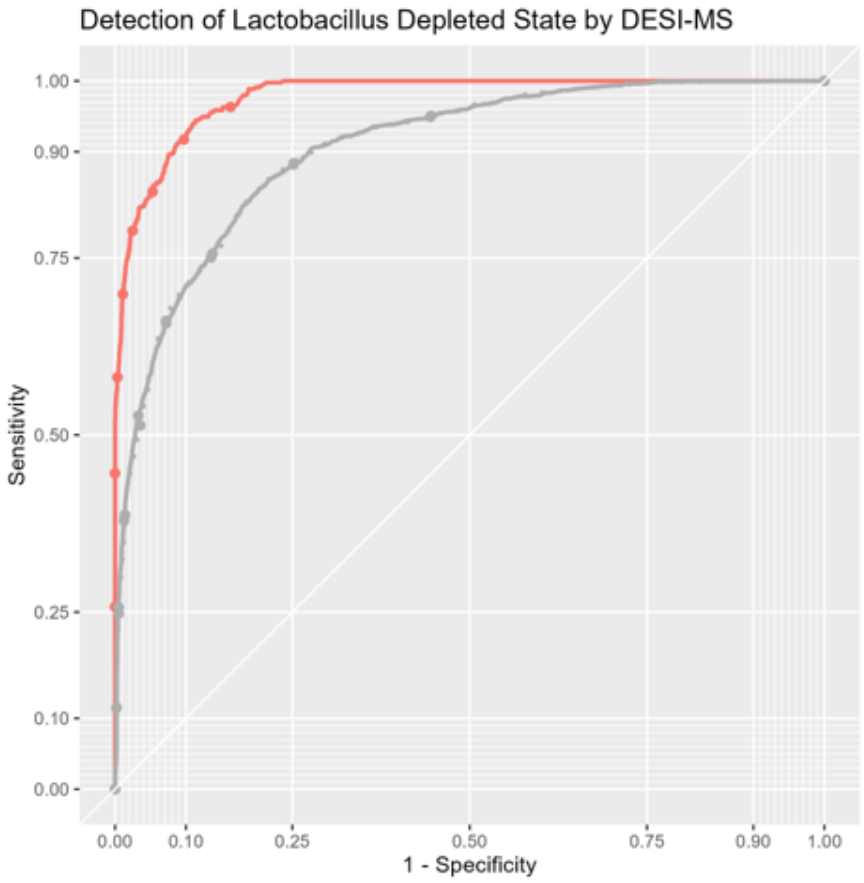


Figure 6.5– ROC curve showing prediction of Lactobacillus dominance or depletion status in FLIP samples and comparison with VMET2 data performance. FLIP AUC =0.98, Accuracy = 0.96, Sensitivity = 0.61, Specificity = 0.99, PPV = 0.95, NPV = 0.96, Precision = 0.95. VMET2 (neg/pos), AUC =0.97/0.90, Accuracy = 0.96/0.87, Sensitivity = 0.60/0.53, Specificity = 0.99/0.96, PPV = 0.94/0.79, NPV = 0.96/0.88, Precision = 0.94/0.79.

The capacity of the RDa ACQUITY DESI-MS platform to predict the major vaginal CSTs was also very strong in the FLIP cohort (Table 6.2). More specifically, the model trained for the contrasts between CSTI vs CSTIII had a AUC = 0.98 , the model for CSTI v CSTIV had a AUC = 0.98, and the contrast CSTII vs CSTIV had a AUC= 0.96. Poorer performance was observed for CST V due to a lack of power.

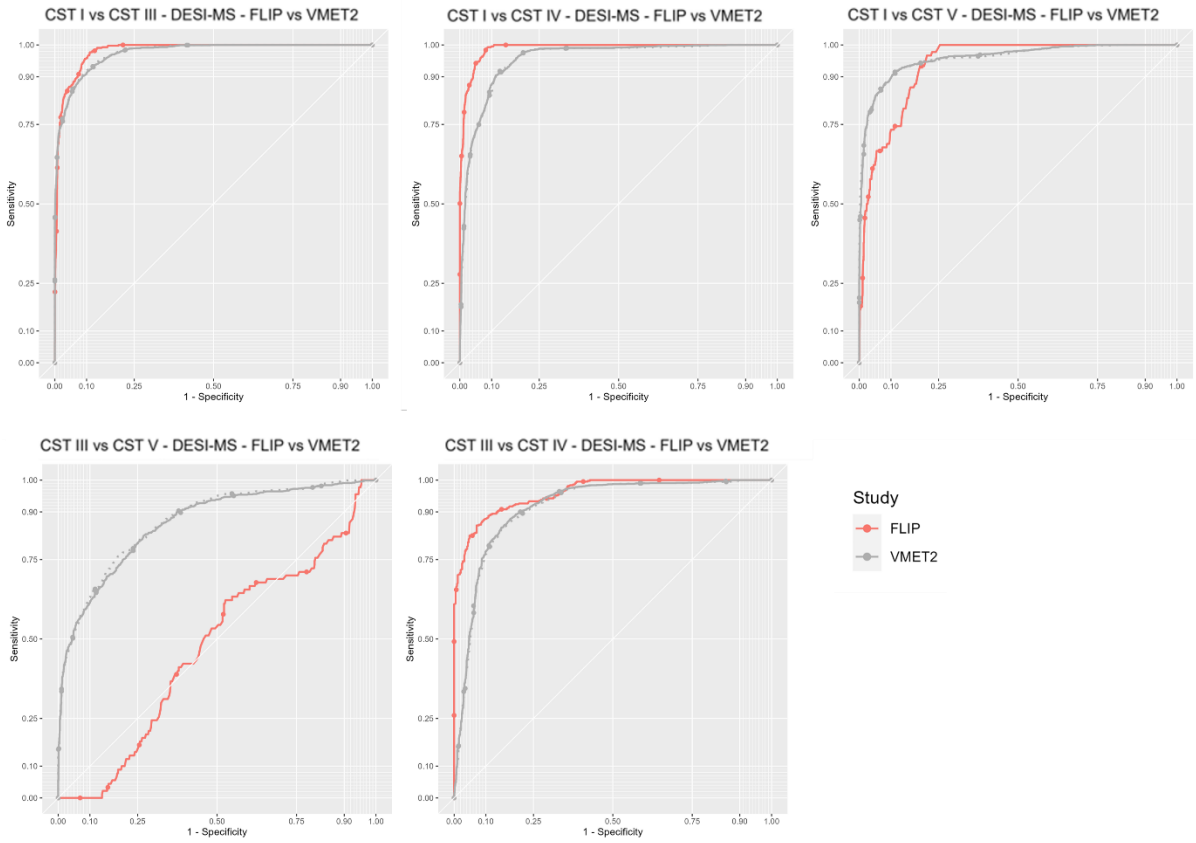


Figure 6.6 - ROC curve showing contrast of predictions of CSTs in FLIP samples and comparison with VMET2 data performance. FLIP results are shown in red, VMET2 results are shown in grey. Refer to table 6.2 for more information.

Table 6.2 AUC, Accuracy and Sensitivity for each CST contrast in FLIP study and VMET2 study.

		CST I vs IV	CST I vs III	CST I vs V	CST III vs IV	CST III vs V
FLIP	AUC	0.98	0.98	0.94	0.96	0.55
	Accuracy	0.96	0.95	0.97	0.87	0.81
	Sensitivity	0.99	0.99	0.01	0.87	0.97
VMET2	AUC	0.95	0.97	0.96	0.91	0.87
	Accuracy	0.88	0.91	0.93	0.84	0.79
	Sensitivity	0.92	0.97	0.75	0.83	0.9

6.3.4 Prediction of LACTIN-V colonisation by DESI-MS

Assessment of colonisation of *L. crispatus* following treatment with LACTIN V was performed using a combination of metataxonomic analyses and PCR. At the end of the one-week loading dose, a total of 49/61 patients were positive for the *L. crispatus* CTV-05. In the remaining patients (n=13) no *L. crispatus* was detected after the loading dose, either by PCR or by amplicon sequence variant mapping. The ability to predict colonisation response in patients by DESI-MS analysis of pre-treatment baseline vaginal swabs was then tested. A random forest model was trained using the DESI-MS metabolic profiles and a ROC curve analysis performed (Figure 6.7). For two of the patients, a baseline swab for DESI-MS analysis was not available therefore the final number of samples used in this analysis was 60 (colonisers, n = 47; non-colonisers, n=13). The resulting models indicated a good capacity to predict likely colonisations (AUC = 0.71), which was slightly improved if patients were pre-stratified on the basis of their microbial composition (AUC = 0.73 for patients with *Lactobacillus* dominant microbial composition at baseline; colonisers, n = 40 v non-colonisers, n=10). An equivalent analysis for *Lactobacillus* depleted samples at baseline could not be performed due to limited sample (n =10, colonisers = 7, non- colonisers = 3).

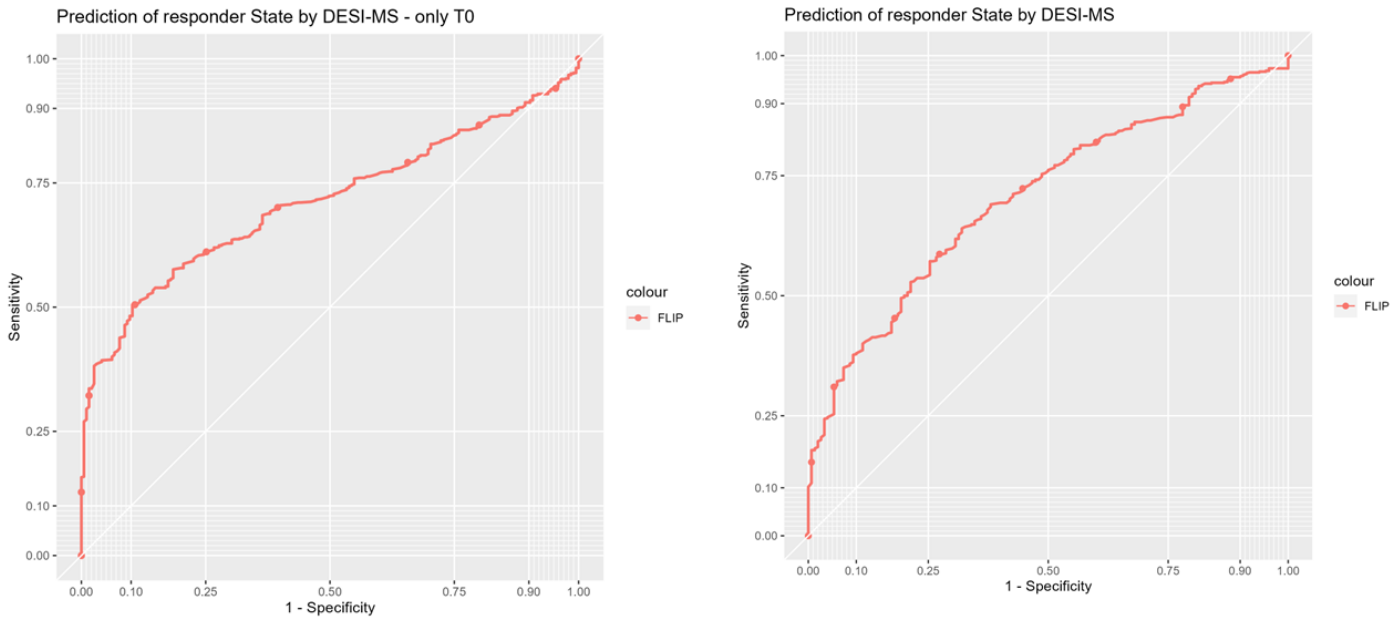


Figure 6.7 - ROC curve showing prediction of responder status in FLIP samples

A selection of metabolite features important in discriminating colonisers from non- colonisers is presented in Figure 6.8.

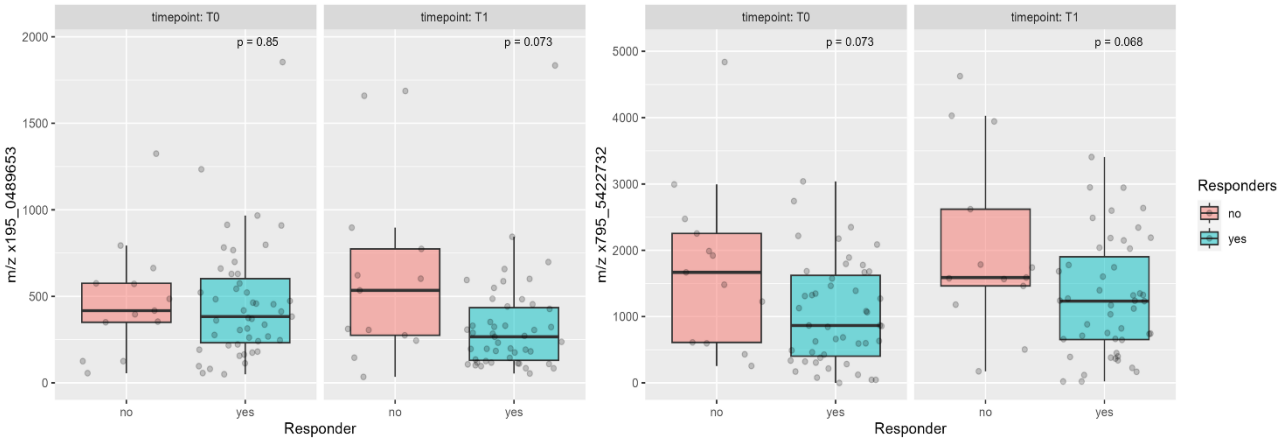


Figure 6.8 - Boxplots showing differences between timepoint 0 and timepoint 1 of metabolites in responders and non responders groups.

6.3.5 Inflammatory status prediction using DESI-MS metabolic profiling

DESI-MS profiles captured on the Orbitrap platform have been previously shown to provide the ability assess host inflammatory status (see Chapter 3 and Pruski et al., [199]) . Therefore we next explored the ability of the RDa ACQUITY DESI-MS platform to capture similar inflammatory response data at the cervicovaginal interface. For these analyses, data collected at T0 (pre-intervention), T1 (immediately after the loading dose), T3 (end of the maintenance dose) and T4 (8 weeks after final LACTIN-V dose) were used. At each of these timepoints, DESI-MS data was modelled against immune assay data for IL-8 and IL-1β. DESI-MS features predictive of log-transformed concentrations of each immune marker were identified using a random forest regression model. Cross-validation was performed to assess the robustness of the predictions, with a threshold of $R2 > 0.25$ indicating strong prediction capability. The results of the prediction demonstrated promising performance for IL-1β (CV R2 = 0.41) and IL-8 (CV R2 = 0.58). To further evaluate the ability of DESI-MS to predict the levels of these markers, log-transformed values obtained from the immune-assay measurements were plotted against the predicted values derived from DESI-MS features and a linear regression line fitted

(Figure 6.9). A robust correlation (IL-1β, CV R2 = 0.41 and IL-8, CV R2 = 0.58) between the measured immune marker concentrations and the predicted values was observed. Additionally, the samples were color-coded according to their classification as either LDOM or LDEPL as determined using metataxonomics and box plot of the predicted immune mediator values for each group was generated. The mean predicted concentrations of both IL-8 and IL1β were found to be significantly higher in women with LDEPL compared to LDOM.

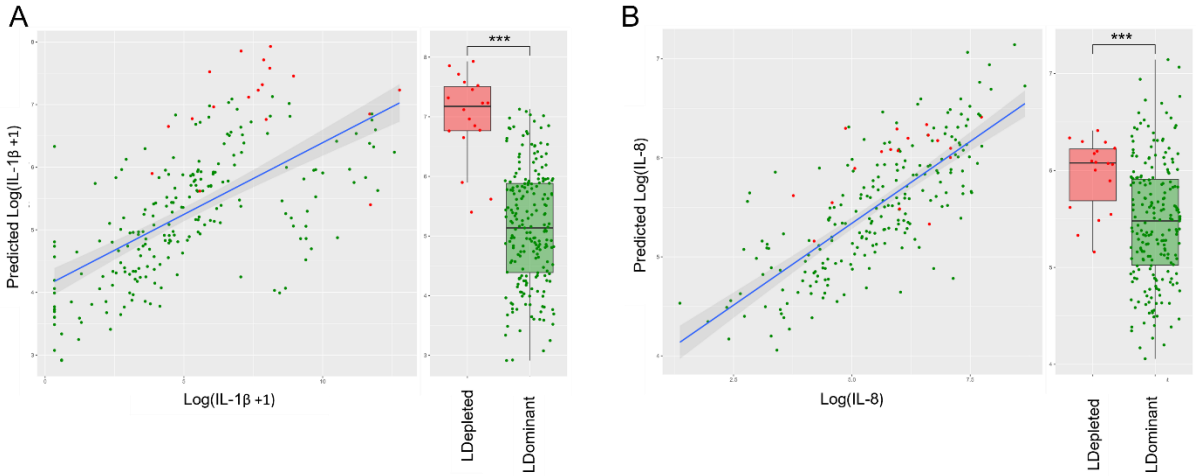


Figure 6.9 - Scatterplot showing relation between predicted and measured levels of main inflammation markers. (A) Predicted values of IL-1b plotted against measured levels (R2 = 0.41) and boxplot comparing levels in Lactobacillus dominant and Lactobacillus depleted (p-value = 1.1e-09), (B) Predicted values of IL-8 plotted against measured levels (R2 = 0.58) and boxplot comparing levels in Lactobacillus dominant and Lactobacillus depleted (p-value = 3.1e-05)

Metabolites strongly correlated with IL-8 and IL-1β were analysed to examine their levels in relation to microbial composition. The top negatively and positively correlated metabolites are shared across IL-8 and IL-1β and are m/z 823.3912 and m/z 896.5955. Their levels in LDOM and LDEPL patients are reported in Figure 6.10. Levels of m/z 896.5955, which is positively correlated with immune markers levels, are significantly higher in LDEPL group compared to LDOM groups (Wilcoxon test p-value = 0.014). On the other hand, m/z 823.3912 which is negatively correlated with IL-8 and IL-1β levels, appears higher in LDOM group despite the result not reaching statistical significance (Wilcoxon test p-value = 0.17).

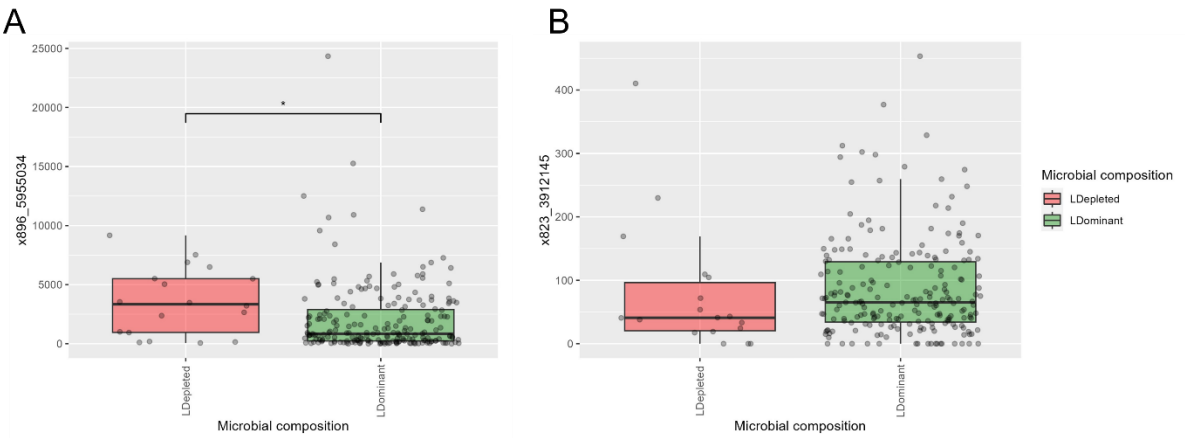


Figure 6.10 - Boxplots showing levels of main metabolites correlated with immune markers between LDOM and LDEPL samples. (A) Levels of m/z 896.5955 in LDEPL and LDOM samples. This metabolite is positively correlated with IL-1β and IL-8 levels and higher in LDEPL group (p-value 0.014). (B) Levels of m/z 823.3912 in LDEPL and LDOM samples. This metabolite is negatively correlated with IL-1β and IL-8 levels and higher in LDOM group (p-value 0.17).

6.4 Discussion

As previously described (Chapters 3 and Pruski et al. [199]), direct on swab metabolic profiling by DESI-MS enables robust prediction of CVF microbial composition and host inflammatory status. This highlights the potential of MS-based approaches to be used as point of care diagnostic devices that could substitute traditional chemical and biochemical measurements [606]. Despite these exciting results, the transition of DESI-MS to a clinical environment remains hampered by the size of the Orbitrap MS to which it is coupled. Like other comparable high resolution, high sensitivity MS instruments, it is non-portable. Recently the development of new ion sources with enhance resolution and sensitivity has led to the production of small-format mass spectrometers such as the RDa ACQUITY that is manufactured by Waters Corp [412]. This instrument has a small footprint (approximately the size of a microwave) and is portable but has yet to be tested for DESI-MS swab analysis. We began to test this platform by acquiring metabolic profiles of vaginal swabs collected from patients enrolled in the FLIP-1 trial, an exciting first in pregnancy study designed to assess the safety, tolerability and acceptability of LACTIN-V, a live biotherapeutic, in women at high-risk of PTB [607]. As reported recently, this trail showed that the LACTIN-V product is safe, well tolerated and highly accepted by women at high risk of PTB. Although not powered for clinical outcomes, the rate of early PTB in patients treated with LACTIN-V was 3.3% compared to 7% in a historical cohort of 2,190 women at similar background PTB risk [607]. Metataxonomic-based analysis indicated that LACTIN-V lead to the shift of non-*L. crispatus* (CSTI) dominated communities to *L. crispatus* (CSTI) dominance in 27/34 (80%) cases. *L. crispatus* (CSTI) dominance in pregnancy and been widely reported to provide protection against PTB [139], [233], [480], which may account for the observed low rates of PTB in the patient cohort.

Microbiota composition in these patients was initially assessed using ‘gold standard’ metataxonomics, which required substantial sample processing, DNA extraction, sequencing, and subsequent data analysis. This process currently takes days to weeks for completion, which renders the approach not compatible for rapid bed-side diagnostics of the vaginal microbiota. Currently, vaginal diagnostics in clinical settings are limited to culture and microscopy that provide a cheap but highly limited insight into microbiota composition and no insight into the inflammatory status of the host. In comparison, the acquisition of the vaginal metabolic profile on the ACQUITY RDa platform took only approximately 2 mins per sample and as demonstrated in the results, this provided sufficient information to accurately predict the

vaginal microbiota composition in the patient cohorts. Data were acquired only in negative ion mode. The acquisition of positive ion mode data was not possible at the time of our analysis due to ongoing optimisation procedures. As this was a new analytical platform, care was taken to assess the physical characteristics of the swabs (e.g., deformation of swab shape, blood presence) at the time of data acquisition [139]. Based on previous findings reported in Chapter 3, samples were analysed in small batches of 10 to minimise thawing, which reduced the likelihood of swab deformation and loss of material in this sample set. Additionally, swab samples with visible blood presence were noted, but despite previous reports of blood metabolites introducing bias in the resulting metabolomic spectra [467], [470], [471], unsupervised analyses of the resulting spectral profiles showed that major variation in the data was not attributable to blood contamination. This is an important consideration given that vaginal bleeding is a clinically relevant and a common symptom in pregnancy and may also occur during the sampling procedure. Other potential confounders such as acquisition date and ethnicity were also shown to not be the source of major variation in the metabolic profiles.

These findings provided confidence that underlying differences in the metabolic profiles were likely to be driven by underlying biology, of which we have previously shown involves microbe-host interactions. Consistent with this, 410 m/z features were found to significantly differ in concentration between *Lactobacillus* dominated and *Lactobacillus* depleted samples. This provided sufficient information to allow strong prediction of microbiota composition at genera-level, which is already clinically useful. However, strong prediction of microbiota composition was also achievable using the DESI-MS data at higher taxonomic resolution with major CSTs (CST I, III and IV) being readily distinguishable from each other. Indeed, the performance of the prediction models was comparable and in some instances better than the results obtained with the previous Orbitrap MS platform. The reason for improved performance is yet to be clarified and numerous different factors can contribute to this observation. However, one possibility could relate to the different analysers used for each platform. The RDa platform uses Time of Flight (TOF) to distinguish ion mass-to-charge ratio compared to the Orbtrap which uses the motion of ions trapped in an electrostatic field. Despite the higher resolution of the Orbitrap instrument, the TOF instrument detected significantly more metabolites compared to the Orbitrap. The reason for this result might be the dynamic range, particularly in the low abundance range. TOF mass spectrometers have in fact a superior dynamic range, allowing it to detect smaller peaks and low abundant metabolites more reliably

[608]. The dynamic range of Orbitrap instruments might be a limitation when analysing complex matrices such as the cervicovaginal fluid.

Despite these differences, a subset of MS features were found to consistently differ LDOM and LDEPL samples across both the Orbitrap and RDa platforms. This provides confidence that these metabolites genuinely reflect vaginal microbiota composition. While some of these features have been previously reported to be enriched in LDEPL and LDOM samples in non-pregnant women (e.g. short chain fatty acids and amino acids, respectively) [327], the identity of several other features remains unknown. Future studies that could focus on identification of these features and investigations into their function in the cervicovaginal niche are required. Regardless, these findings highlight the capacity of vaginal swab profiling by DESI-MS to provide rapid, robust and accurate prediction of the vaginal microbiota composition within a cohort of women undergoing therapeutic treatment designed to modulate their microbiome. The results also show that this is achievable on the portable and small format ACQUITY RDa platform, which opens up the possibility of using the technology as a point-of-care device in clinical settings. There are several clinical scenarios where such information could be utilised by clinicians and healthcare workers to direct treatment strategies. For example metabolic profiling could be used to detect microbiota composition changes early during pregnancy and personalise the therapeutic approach accordingly (e.g. use of targeted antibiotics or different intervention to modulate microbial composition). Also, these information could be used to monitor therapies in the gynaecological setting both in pregnancy and non-pregnancy contexts. Shifts in microbial composition could inform clinicians of the prognosis for treatment aimed at optimising microbial composition, such as live biotherapeutics or vaginal microbiome transplantation but also of the response to PTB preventative treatments such as cervical cerclage.

Although LACTIN-V treatment caused a positive shift toward *L. crispatus* dominance of the vaginal microbiota in a large portion of the patients enrolled in the study, strain-level analysis of the product using qPCR and analysis of ASVs unique to *L. crispatus* CTV-05, showed that some patients failed to colonise. We first hypothesised that failure to colonise may relate to patient compliance issues. Unfortunately, although patients provided a detailed study diary documenting times of application etc, the study was not designed to directly verify patient compliance. This could be assessed in future studies by investigating the presence or absence vaginal epithelial cells in returned, empty applicators. Nevertheless, the failure of LACTIN-V

to colonise some women has also been reported in earlier studies of non-pregnant women where the specific CTV-05 strain was not detectable after the loading dose phase [603]. For example, a study by Antonio et al. investigating the colonization of CTV-05 in sexually active women with no prior conditions should that in women were colonisation by CTV-05 failed, 85% were already colonized by an endogenous *L. crispatus* strain [609]. Therefore, we hypothesized that colonisation of the CTV-05 strain may be dependent on underling biological differences that are reflected in the CVF metabolome at baseline. Consistent with this, a random forest model built using DESI-MS features from subsequent 'colonisers' and 'non-colonisers' indicated a good capacity to distinguish these patients using their metabolic profiles. The ability to distinguish 'colonisers' and 'non-colonisers' was slightly improved when only patients with existing *Lactobacillus* dominance were sub-analysed. This may indicate that specific, existing vaginal strains have the capacity to out-compete the CTV-05 strain. Consistent with this notion, Ngugi et al. reported that in patients with BV, the presence of endogenous *L. crispatus* at enrolment was significantly associated with reduced colonisation by CTV-05 [610]. The use of direct swab analysis by DESI-MS to identify likely colonisers from non-colonisers could offer a rapid and efficient screening tool to improve selection of those patients most likely to response to a treatment intervention such as this.

Lastly, metabolic profiles were integrated with immune profiling data. Concentrations of inflammatory markers were measured at different timepoints and DESI-MS features were used to predict their levels. Our results showed promising performances, specifically predicting levels of pro-inflammatory markers IL-8 and IL-1β. IL-1β promotes collagenase production in the cervix facilitating the remodelling process [611]. Similarly, IL-8 is a potent chemokine for neutrophils, inducing activation and migration into the uterine cervical tissue. This in turn stimulates the release of collagenase and elastase in the tissue, resulting in cervical ripening [612]. Both IL-8 and IL-1β have consistently been associated with increased risk of spontaneous preterm delivery. A study by Puchner et al. reported that a unit increase in IL-1β level in women, corresponded to a 7.2 times increased risk of PTB [613]. Also, increased levels of IL-8 in the amniotic fluid were associated with intrauterine infection and increased risk of PTB in multiple studies [614]–[616]. Wennerholm et al, found that the levels of IL-8 in cervicovaginal fluid concentrations was associated with preterm delivery in twin pregnancies [617]. Our results confirm that metabolic features in the cervicovaginal fluid are strongly reflective of the inflammatory status of the patient and the sampling via swab guarantees a quick analysis through DESI-MS. Given the central role of these cytokines in the labour

mechanism and their association with PTB, it is evident the importance of a technique that could rapidly and reliably predict their levels. Rapid metabolic profiles can be translated into clinically useful information about the inflammation status of the patient and guide intervention strategies accordingly (e.g. monitoring of cerclage intervention). Additionally, our data confirms the association of these immune markers with microbial composition. Both IL-8 and IL-1β are higher in *Lactobacillus*-depleted communities.

The characterization of metabolites correlated with IL-1β and IL-8, unveiled an association between these immune markers and a group of metabolites. Their directionality confirms that highly diverse microbial communities have higher levels of metabolites positively correlated with inflammation markers and vice versa. Their mass range suggest they could be lipids and lipid metabolism alterations have been previously linked to the risk of PTB [618], [619]. However, caution must be taken in interpreting these findings since the metabolites have not been identified at the moment and further investigation is needed to confirm their identity and biological role.

Collectively these findings further support the validity of DESI-MS predictions and highlights the potential of this method in identifying and characterizing inflammation-associated phenotypes at the core of PTB risk. In conclusion, the results reported in this chapter confirm the potential of direct on-swab profiling by DESI-MS to predict microbial composition and inflammation status in the cervicovaginal niche also in the context of a live biotherapeutic intervention.

If proven suitable, the RDa platform could provide the means for Direct Swab Analyses by DESI-MS to be transitioned towards clinical care settings as a point-of-care (POC) diagnostic tool. The aim of POC testing is to significantly reduce the time between sample acquisition and analysis, and quickly lead to informed medical decisions. POC testing relies on smaller instruments, simplified procedures and user-friendly interfaces that can allow non-specialised personnel to perform the analyses directly at the patient's bedside, in the operating room or even at accident sites [620]. POC testing can increase the flexibility of receiving a diagnosis outside of traditional medical facilities where specific patients populations might be more easily reached [621]. This it is particularly important in the context of reproductive health, a topic that can be considered taboo for certain segments of the population. Having a tool that could potentially inform the clinician of vaginal microbial composition, inflammation status and likelihood of effectiveness of a therapy, could significantly improve the care provided for patients at risk of PTB.

Chapter VII

7 DISCUSSION

7.1 Background

Preterm birth remains the leading cause of death for children under the age of five [1], [2]. It affects approximately 11% of pregnancies worldwide and remains a pressing medical problem [4]. Even for children who survive, there are often long-term consequences such as impaired pulmonary functions, higher risks of cerebral palsy, and disabilities [5]. However, the understanding of the underlying causes of PTB and our ability to predict women at risk remains limited. This is due to the fact that PTB is a complex syndrome influenced by multiple pathological processes [3], [10]. There are several well-established risk factors of PTB including a history of PTB, maternal age, race, BMI, maternal drug and alcohol use, smoking status and socio-economic status [13], [14]. Unfortunately, risk factor assessment alone is not a reliable predictor of PTB as over 50% of preterm pregnancies have no risk factor present [103]. Cervical length measurement, although reliable and safe, may not fully predict all cases of PTB and may not provide timely information. Additionally, it is not routinely recommended for low-risk populations [105]–[107]. Thus, there is a necessity for an innovative approach that can rapidly and accurately diagnose and stratify PTB phenotypes.

Possible causal mechanisms driving PTB include infection, cervical incompetence, and local progesterone deficiency. It is estimated that around 40% of all PTB cases are linked with infectious aetiology, with ascending vaginal infection considered to be a common route of pathogen invasion [23], [83]. Recent studies have established an association between vaginal bacterial composition and risk of PTB. A depletion of *Lactobacillus* species and increased bacterial diversity in the vaginal niche are associated with an increased risk of PTB [234], [235]. Current techniques used to assess vaginal microbiota are costly, time-consuming, and do not provide information on the host:microbiota interactions, which are crucial for understanding PTB phenotypes. A potential solution lies in the study of the biochemical composition of the CVF that has been shown to reflect the microbial composition of the vaginal environment [304], [305], [328], [332]. Metabolomics, particularly through the newly developed direct on-swab DESI-MS approach, offers a rapid capacity to study and understand

microbial composition in mucosal surfaces [467]. This technique has the potential to be used in the POC context both for prediction and stratification of patients.

Given the close relationship between microbial composition and PTB risk, efforts have been made to modulate the bacterial communities in the vaginal niche. However, antibiotic treatment in pregnancy is sometimes detrimental both to commensal communities and to the health of the baby [153], [155]–[157], [254]. This has led to a focus on a positive modulation of microbiota through the use of live biotherapeutics [270], [271].

We hypothesize that the CVF biochemical composition holds information about microbiota and host interactions during pregnancy. The analysis of CVF using direct on-swab metabolic profiling by DESI-MS could rapidly capture this information that can be used to facilitate preterm birth risk prediction and stratification. Additionally, metabolic profiling could be used to predict microbial composition, inflammation status, to monitor preventative interventions and predict patient outcomes.

7.2 Summary of study findings

The work presented in Chapter III set out to test the capacity of the recently developed Direct on-swab metabolic profiling by DESI-MS to provide rapid assessment of the vaginal microbiota composition and inflammatory status. If successful, this would provide an alternative technology to the current culture-based and microscopy-based techniques used for vaginal diagnostics in clinical settings. These methods are slow, subjective and fail to provide any insight into inflammatory response in the host, which is a key driver of pathology. The results obtained clearly demonstrated that swab profiling by DESI-MS provides detailed metabolic information that could be used to accurately predict the microbial composition of CVF at different levels of taxonomic resolution. Importantly, these results were validated in an independent patient cohort, highlighting the robustness and reproducibility of the approach. Several discriminatory metabolites were identified by MS-MS approaches and their relationship with vaginal microbial composition was found to be consistent with previous literature or biologically plausible roles in lower reproductive tract health and disease [195], [348]. For example, higher levels of fatty acids and biogenic amines were associated with high diversity compositions, while *Lactobacillus*-dominated communities, have shown positive correlations with amino acids and dipeptides.

A major advantage of DESI-MS over other approaches is that it provides valuable insight into the microbiota:host interactions. Integration of immunological profiling data with metabolomic profiles identified 33 metabolites strongly correlated with the inflammation status of the host. This has important clinical implications. For example, recent data indicated that cervical cerclage with braided suture material can cause local inflammation and a shift in the composition of the vaginal microbiome [139]. Consistent with this, higher levels of inflammation markers as detected using DESI-MS were observed in patients receiving braided cerclage and those who subsequently delivered preterm. The ability to determine the inflammatory status of the cervicovaginal niche via a rapid swab analysis would be a major step forward in determining PTB risk phenotype but also for informing more treatments and monitoring response to interventions. For example, a potential application of DESI-MS is monitoring the response to antibiotic treatment during pregnancy. In situations where antibiotics are needed due to infections, their non-selective nature can potentially affect the vaginal microbiota negatively. By utilizing DESI-MS, it becomes feasible to easily assess any impact on the vaginal microbiota and evaluate the effectiveness of antibiotic interventions in a targeted manner. Also, a *Lactobacillus* depleted vaginal microbiome is a risk factor for PPROM in roughly 25% of cases [235]. CVF analysis could rapidly assess *Lactobacillus spp*. abundance and the presence of pathogens. Identification of this subset of patients could guide the use of antibiotic, prebiotic or live biotherapeutic interventions where appropriate.

In Chapter IV DESI-MS metabolic profiling was extended to the investigation of other possible PTB aetiologies that could be reflected in the CVF metabolome. Both previous preterm birth and history of surgical treatment for cervical disease represent significant risk factors for PTB, with an estimated 2.5-fold and 2-fold increase in risk respectively [3], [512]. The study aimed to investigate if the underlying biology associated with these risk factors is reflected in the CVF metabolome and in doing so, provide some insight into the underlying mechanisms. Supervised analyses of the metabolic features indicated moderate ability to discriminate patients on the basis of their pregnancy history, suggesting that a portion of the metabolic variance in the CVF is related to previous pregnancy history. It remains unclear how these metabolic differences underpin the biology related to previous pregnancy history. However recent data has highlighted that immune responses are altered from one pregnancy to the next [622]. For example, during a women's first pregnancy there is an accumulation of fetal-specific maternal CD8+ T cells that remain activated in the maternal system after birth. In subsequent pregnancies, memory T cell expression of PD-1 and LAG-3 drives tolerance of the developing

fetus. However, molecular disruptions that neutralize expression of these proteins can activate the fetal-specific CD8+ T cells, causing poor pregnancy outcomes (e.g. miscarriage or PTB) during subsequent, but not first pregnancies. It is feasible that a similar immunological 'memory' exists in the CVF that is reflected in the metabolic profile of CVF, but this remains to be proven.

Additionally, this aspect of the work indicated that differences in the CVF metabolome are moderately predictive of subsequent cervical shortening. While this finding requires validation in a larger patient cohort, it indicates the potential utility of swab profiling by DESI-MS to identify women at risk of cervical shortening early in their pregnancy. Given that the method can also provide insight into microbiota composition, this could enable information on underlying aetiology to be simultaneously provided thus guiding selection of the most appropriate interventions. Detection of biochemical changes in CVF associated with pathophysiological cervical shortening prior to clinically manifested shortening observed by ultrasound would also facilitate earlier intervention, which could potentially improve efficacy of existing treatments for cervical shortening [623], [624]. However, there are several limitations of the study. Apart from a relatively small sample size, there was also large variations in the timing and rate of cervical shortening observed in the patients, as well as likely different underlying aetiologies.

There are now several studies that have described a negative impact of cervical cerclage on the microbial composition and inflammation response in the cervicovaginal niche post-insertion in some women. This is particularly reported for patients receiving braided sutures [139], [452]. The ability to rapidly assess potential disruption of vaginal microbiota and host inflammatory response post-cerclage could provide valuable information for increased surveillance and subsequent clinical decisions. Analysis of CVF from patients with cervical cerclage *in situ* showed that direct on-swab metabolic profiling by DESI-MS remains a robust tool for predicting vaginal microbiome composition.

While these results collectively highlight its potential, an obstacle for the implementation of DESI-MS in clinical practice is the large size of the mass spectrometer, which limits its practicality as a point-of-care diagnostic tool. Efforts were made to address this as described in Chapter VI, where vaginal swabs collected from patients enrolled in the FLIP-1 trial were analysed using the small-format, portable mass spectrometry platform, ACQUITY RDa, manufactured by Waters Corp. In this trial a live biotherapeutic (LACTIN V) containing

CTV- 05 *Lactobacillus crispatus* strain was used for the first time in pregnancy. It showed excellent safety and tolerability in women at high risk of PTB and promisingly, its use was associated with low rates of PTB compared to historical cohorts [607]. Metataxonomic-based analysis of samples collected from women in the trial demonstrated that LACTIN-V shifted the microbiota towards *L. crispatus* (CSTI) dominated communities in 80% of cases. This was also readily demonstrated by DESI-MS analysis but with the major advantage of being achieved within 2 minutes per swab and without prior sample preparation. The resulting metabolic profiles allowed for strong prediction of microbial composition both at genera-level and CST-taxonomic resolution as well as strong prediction of inflammatory markers in the samples, with comparable performance to the previously used Orbitrap platform. A random forest model of the resulting metabolomic data also showed promise in distinguishing colonisers from non-colonisers, particularly in those patients with existing *Lactobacillus* dominance, suggesting that endogenous vaginal strains may out-compete the CTV-05 strain [609]. This provides another example of how on-swab metabolic profiling by DESI-MS could serve as a rapid and effective tool for patient stratification and treatment response in pregnancy.

It has been hypothesised that a local progesterone deficiency in the vaginal mucosa might be one of the underlying causes of a proportion of PTB cases [95]. As presented in Chapter V, we investigated the capacity to use mass spectrometry-based methods to detect and quantify levels of progesterone and other steroids in CVF. The initial approach of using on-swab DESI-MS showed that progesterone was only detectable in a minority of samples. This prompted redirection of our efforts towards other MS techniques. The adaptation of an existing UPLC-MS targeted assay to CVF samples collected in soft-cups allowed accurate quantification of progesterone, cortisol, and androstenedione. The results showed that progesterone levels increase during the second trimester of pregnancy, consistent with previous reports in blood samples [97], [577], [578]. However, no significant changes were observed for cortisol and androstenedione in CVF across the same gestational age range. When comparing CVF levels of detected steroids between term and preterm pregnancies we did not observe any significant difference, refuting the hypothesis of progesterone deficiency driving PTB. This pilot study had several limitations including a small sample size and limited longitudinal sampling. However, this study represents the first attempt of measuring CVF steroids using MS-based approaches and provides initial evidence against the local progesterone deficiency hypothesis.

7.3 Future Work

This thesis highlights direct on-swab metabolic profiling by DESI-MS as a novel and effective approach for rapidly assessing microbial composition and inflammation during pregnancy. However, the results generated during this project also identify areas for further investigation and methods development. As has been reported, small changes in swab orientation and distance from the MS inlet capillary can affect signal quality and intensity thus affecting reproducibly of the technique [467]. This was particularly observed when working on the Orbitrap platform, which was also limited by its size and footprint. The upgrade of the rotating swab holder used for small format ACQUITY RDa platform, addressed this issued to some extent, however deployment of the device in the clinical environment where it could be used by non-specialists requires further advancement towards a fully automated data-acquisition process. This could involve the insertion of the swab into an enclosed environment that isolates the sample from the external environment and analyses consistently the same region of the swab tip. Furthermore, the ACQUITY RDa platform required additional optimization for robust and reproducible acquisition of positive ion mode data. Simplifying the user interface of the DESI-MS platform would also facilitate its adoption in clinical settings.

While the work in this thesis was limited to the investigation of samples collected in pregnancy, direct on-swab metabolic profiling by DESI-MS has obvious application in non-pregnant women [467]. This includes testing for common gynaecological conditions such as BV and sexually transmitted infections. Moreover, the platform could be used for the study of other mucosal samples such as recently described by Ford et al. [625] who reported an application of nasal swab analysis for COVID-19 diagnosis.

As evidenced in some of the studies presented in this thesis, an important and necessary future step is to elucidate the biological role of potential biomarkers identified in the metabolite analysis. Future studies should incorporate experiments involving tandem mass spectrometry (MS/MS) for accurate metabolite assignment. The results of MS/MS experiments would provide valuable insights into the specific biological functions of these potential biomarkers.

7.4 Conclusion

The work presented in this thesis shows that direct on-swab metabolic profiling by DESI-MS holds great potential as a point-of-care diagnostic tool for rapid, simultaneous assessment of microbial composition and host inflammatory status. Moreover, the data generated by the method provides important mechanistic insight into the interplay between the microbial community of the cervicovaginal niche and host response. As demonstrated, this type of information may be useful for preterm birth risk phenotyping and patient stratification, such as the early identification of women at risk of experiencing cervical shortening or for monitoring microbiota-host responses to treatment interventions like cervical cerclage or live biotherapeutics. The capacity to achieve robust prediction of bacterial communities and inflammatory status from vaginal swabs using a small-format, portable DESI-MS platform, represents a significant advancement towards the application of this rapidly advancing technology in the clinical environment. It is envisaged that in pregnancy, bedside assessment of vaginal microbiota-host interactions could provide objective insight into underlying aetiologies allowing earlier and more directed and personalised therapeutic strategies for the prevention of PTB. Given the recognised importance of the vaginal microbiome in other areas of women's health, it is likely that diagnostic information provided by DESI-MS will have even wider applications beyond pregnancy.

References

[1] J. P. Vogel, S. Chawanpaiboon, A.-B. Moller, K. Watananirun, M. Bonet, and P. Lumbiganon, 'The global epidemiology of preterm birth', *Best Pract. Res. Clin. Obstet. Gynaecol.*, vol. 52, pp. 3–12, Oct. 2018, doi: 10.1016/j.bpobgyn.2018.04.003.

[2] S. Chawanpaiboon *et al.*, 'Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis', *Lancet Glob. Health*, vol. 7, no. 1, pp. e37–e46, Jan. 2019, doi: 10.1016/S2214-109X(18)30451-0.

[3] R. L. Goldenberg, J. F. Culhane, J. D. Iams, and R. Romero, 'Epidemiology and causes of preterm birth', *The Lancet*, vol. 371, no. 9606, pp. 75–84, Jan. 2008, doi: 10.1016/S0140-6736(08)60074-4.

[4] H. Blencowe *et al.*, 'National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications', *Lancet Lond. Engl.*, vol. 379, no. 9832, pp. 2162–2172, Jun. 2012, doi: 10.1016/S0140-6736(12)60820-4.

[5] S. Beck *et al.*, 'The worldwide incidence of preterm birth: a systematic review of maternal mortality and morbidity', *Bull. World Health Organ.*, vol. 88, no. 1, pp. 31–38, Jan. 2010, doi: 10.2471/BLT.08.062554.

[6] B. Hamilton, J. Martin, and M. Osterman, 'Births: Provisional Data for 2020', National Center for Health Statistics, May 2021. doi: 10.15620/cdc:104993.

[7] Office for National Statistics (ONS), 'Birth characteristics in England and Wales - Office for National Statistics', *statistical bulletin*, Jan. 19, 2023. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/birthcharacteristicsinenglandandwales/2021#cite-this-statistical-bulletin (accessed May 29, 2023).

[8] N. Morisaki *et al.*, 'Impact of stillbirths on international comparisons of preterm birth rates: a secondary analysis of the WHO multi-country survey of Maternal and Newborn Health', *BJOG Int. J. Obstet. Gynaecol.*, vol. 124, no. 9, pp. 1346–1354, Aug. 2017, doi: 10.1111/1471-0528.14548.

[9] World Health Organization, 'Born too soon: the global action report on preterm birth', p. 112, 2012.

[10] R. Romero, S. K. Dey, and S. J. Fisher, 'Preterm labor: One syndrome, many causes', *Science*, vol. 345, no. 6198, pp. 760–765, Aug. 2014, doi: 10.1126/science.1251816.

[11] S. Saigal and L. W. Doyle, 'An overview of mortality and sequelae of preterm birth from infancy to adulthood', *The Lancet*, vol. 371, no. 9608, pp. 261–269, Jan. 2008, doi: 10.1016/S0140-6736(08)60136-1.

[12] C. I. Pravia and M. Benny, 'Long-term consequences of prematurity', *Cleve. Clin. J. Med.*, vol. 87, no. 12, pp. 759–767, Dec. 2020, doi: 10.3949/ccjm.87a.19108.

[13] H. A. Frey and M. A. Klebanoff, 'The epidemiology, etiology, and costs of preterm birth', *Semin. Fetal. Neonatal Med.*, vol. 21, no. 2, pp. 68–73, Apr. 2016, doi: 10.1016/j.siny.2015.12.011.

[14] S. W. Wen, G. Smith, Q. Yang, and M. Walker, 'Epidemiology of preterm birth and neonatal outcome', *Semin. Fetal. Neonatal Med.*, vol. 9, no. 6, pp. 429–435, Dec. 2004, doi: 10.1016/j.siny.2004.04.002.

[15] J. Klumper, A. C. J. Ravelli, C. Roos, A. Abu-Hanna, and M. A. Oudijk, 'Deprived neighborhoods and spontaneous preterm birth: A national cohort study', *Eur. J. Obstet. Gynecol. Reprod. Biol.*, vol. 274, pp. 88–95, Jul. 2022, doi: 10.1016/j.ejogrb.2022.05.012.

[16] B. Kazemier *et al.*, 'Impact of obstetric history on the risk of spontaneous preterm birth in singleton and multiple pregnancies: a systematic review', *BJOG Int. J. Obstet. Gynaecol.*, vol. 121, no. 10, pp. 1197–1208, Sep. 2014, doi: 10.1111/1471-0528.12896.
[17] G. C. S. Smith and J. P. Pell, 'Teenage pregnancy and risk of adverse perinatal outcomes associated with first and second births: population based retrospective cohort study', *BMJ*, vol. 323, no. 7311, pp. 476–476, Sep. 2001, doi: 10.1136/bmj.323.7311.476.
[18] M. Carolan, 'Maternal age ≥45 years and maternal and perinatal outcomes: A review of the evidence', *Midwifery*, vol. 29, no. 5, pp. 479–489, May 2013, doi: 10.1016/j.midw.2012.04.001.
[19] U. Waldenström, V. Aasheim, A. B. V. Nilsen, S. Rasmussen, H. J. Pettersson, and E. Shytt, 'Adverse Pregnancy Outcomes Related to Advanced Maternal Age Compared With Smoking and Being Overweight', *Obstet. Gynecol.*, vol. 123, no. 1, pp. 104–112, Jan. 2014, doi: 10.1097/AOG.0000000000000062.
[20] J. Schaaf, S. Liem, B. Mol, A. Abu-Hanna, and A. Ravelli, 'Ethnic and Racial Disparities in the Risk of Preterm Birth: A Systematic Review and Meta-Analysis', *Am. J. Perinatol.*, vol. 30, no. 06, pp. 433–450, Oct. 2012, doi: 10.1055/s-0032-1326988.
[21] Z. A.-F. Kistka *et al.*, 'Racial disparity in the frequency of recurrence of preterm birth', *Am. J. Obstet. Gynecol.*, vol. 196, no. 2, p. 131.e1-131.e6, Feb. 2007, doi: 10.1016/j.ajog.2006.06.093.
[22] T. Kiserud *et al.*, 'The World Health Organization Fetal Growth Charts: A Multinational Longitudinal Study of Ultrasound Biometric Measurements and Estimated Fetal Weight', *PLOS Med.*, vol. 14, no. 1, p. e1002220, Jan. 2017, doi: 10.1371/journal.pmed.1002220.
[23] H. C. G. Nadeau, A. Subramaniam, and W. W. Andrews, 'Infection and preterm birth', *Semin. Fetal. Neonatal Med.*, vol. 21, no. 2, pp. 100–105, Apr. 2016, doi: 10.1016/j.siny.2015.12.008.
[24] J. B. Liao, C. S. Buhimschi, and E. R. Norwitz, 'Normal Labor: Mechanism and Duration', *Obstet. Gynecol. Clin. North Am.*, vol. 32, no. 2, pp. 145–164, Jun. 2005, doi: 10.1016/j.ogc.2005.01.001.
[25] O. Shynlova, L. Nadeem, J. Zhang, C. Dunk, and S. Lye, 'Myometrial activation: Novel concepts underlying labor', *Placenta*, vol. 92, pp. 28–36, Mar. 2020, doi: 10.1016/j.placenta.2020.02.005.
[26] Liggins G, 'Cervical ripening as an inflammatory reaction', *Ellwood Anderson Eds Cervix Pregnancy Labour Clin. Biochem. Investig.*, 1981.
[27] R. Migale *et al.*, 'Modeling hormonal and inflammatory contributions to preterm and term labor using uterine temporal transcriptomics', *BMC Med.*, vol. 14, no. 1, p. 86, Dec. 2016, doi: 10.1186/s12916-016-0632-4.
[28] S. Lui *et al.*, 'Delineating differential regulatory signatures of the human transcriptome in the choriodecidua and myometrium at term labor†,‡', *Biol. Reprod.*, vol. 98, no. 3, pp. 422–436, Mar. 2018, doi: 10.1093/biolre/iox186.
[29] G. C. Sharp, J. L. Hutchinson, N. Hibbert, T. C. Freeman, P. T. K. Saunders, and J. E. Norman, 'Transcription Analysis of the Myometrium of Labouring and Non-Labouring Women', *PLOS ONE*, vol. 11, no. 5, p. e0155413, May 2016, doi: 10.1371/journal.pone.0155413.
[30] S. Lim *et al.*, 'Nuclear factor kappa B activation occurs in the amnion prior to labour onset and modulates the expression of numerous labour associated genes', *PloS One*, vol. 7, no. 4, p. e34707, 2012, doi: 10.1371/journal.pone.0034707.

[31] J. R. G. Challis, S. G. Matthews, W. Gibb, and S. J. Lye, 'Endocrine and Paracrine Regulation of Birth at Term and Preterm*', *Endocr. Rev.*, vol. 21, no. 5, pp. 514–550, Oct. 2000, doi: 10.1210/edrv.21.5.0407.
[32] C. Kanellopoulos-Langevin, S. M. Caucheteux, P. Verbeke, and D. M. Ojcius, 'Tolerance of the fetus by the maternal immune system: role of inflammatory mediators at the feto-maternal interface.', *Reprod. Biol. Endocrinol.*, vol. 1, no. 1, p. 121, 2003, doi: 10.1186/1477-7827-1-121.
[33] O. Shynlova, Y.-H. Lee, K. Srikhajon, and S. J. Lye, 'Physiologic Uterine Inflammation and Labor Onset: Integration of Endocrine and Mechanical Signals', *Reprod. Sci.*, vol. 20, no. 2, pp. 154–167, Feb. 2013, doi: 10.1177/1933719112446084.
[34] R. Menon, 'Initiation of human parturition: signaling from senescent fetal tissues via extracellular vesicle mediated paracrine mechanism', *Obstet. Gynecol. Sci.*, vol. 62, no. 4, p. 199, 2019, doi: 10.5468/ogs.2019.62.4.199.
[35] T. Geach, 'Fetal signalling initiates parturition', *Nat. Rev. Endocrinol.*, vol. 11, no. 9, pp. 505–505, Sep. 2015, doi: 10.1038/nrendo.2015.115.
[36] J. Geng, C. Huang, and S. Jiang, 'Roles and regulation of the matrix metalloproteinase system in parturition: R OLES AND R EGULATION OF MMP S YSTEM IN P ARTURITION', *Mol. Reprod. Dev.*, vol. 83, no. 4, pp. 276–286, Apr. 2016, doi: 10.1002/mrd.22626.
[37] J. Zhang, O. Shynlova, S. Sabra, A. Bang, L. Briollais, and S. J. Lye, 'Immunophenotyping and activation status of maternal peripheral blood leukocytes during pregnancy and labour, both term and preterm', *J. Cell. Mol. Med.*, vol. 21, no. 10, pp. 2386–2402, Oct. 2017, doi: 10.1111/jcmm.13160.
[38] R. Romero, J. Espinoza, L. F. Gonçalves, J. P. Kusanovic, L. A. Friel, and J. K. Nien, 'Inflammation in preterm and term labour and delivery', *Semin. Fetal. Neonatal Med.*, vol. 11, no. 5, pp. 317–326, Oct. 2006, doi: 10.1016/j.siny.2006.05.001.
[39] S. P. Sivarajasingam, N. Imami, and M. R. Johnson, 'Myometrial cytokines and their role in the onset of labour', *J. Endocrinol.*, vol. 231, no. 3, pp. R101–R119, Dec. 2016, doi: 10.1530/JOE-16-0157.
[40] J. McLaren, D. J. Taylor, and S. C. Bell, 'Increased concentration of pro–matrix metalloproteinase 9 in term fetal membranes overlying the cervix before labor: Implications for membrane remodeling and rupture', *Am. J. Obstet. Gynecol.*, vol. 182, no. 2, pp. 409–416, Feb. 2000, doi: 10.1016/S0002-9378(00)70232-8.
[41] R. Romero *et al.*, 'Prevalence and Clinical Significance of Sterile Intra-amniotic Inflammation in Patients with Preterm Labor and Intact Membranes', *Am. J. Reprod. Immunol.*, vol. 72, no. 5, pp. 458–474, Nov. 2014, doi: 10.1111/aji.12296.
[42] A. R. Mohan, J. A. Loudon, and P. R. Bennett, 'Molecular and biochemical mechanisms of preterm labour', *Semin. Fetal. Neonatal Med.*, vol. 9, no. 6, pp. 437–444, Dec. 2004, doi: 10.1016/j.siny.2004.08.001.
[43] J. A. Majzoub, 'Corticotropin-releasing hormone physiology', *Eur. J. Endocrinol.*, vol. 155, no. suppl_1, pp. S71–S76, Nov. 2006, doi: 10.1530/eje.1.02247.
[44] M. McLean, A. Bisits, J. Davies, R. Woods, P. Lowry, and R. Smith, 'A placental clock controlling the length of human pregnancy', *Nat. Med.*, vol. 1, no. 5, pp. 460–463, May 1995, doi: 10.1038/nm0595-460.
[45] M. Ilicic, T. Zakar, and J. W. Paul, 'The Regulation of Uterine Function During Parturition: an Update and Recent Advances', *Reprod. Sci.*, vol. 27, no. 1, pp. 3–28, Jan. 2020, doi: 10.1007/s43032-019-00001-y.
[46] J. R. G. Challis and S. Hooper, 'Birth: outcome of a positive cascade', *Baillières Clin. Endocrinol. Metab.*, vol. 3, no. 3, pp. 781–793, Nov. 1989, doi: 10.1016/S0950-351X(89)80053-9.

[47] D. K. Grammatopoulos and E. W. Hillhouse, ‘Role of corticotropin-releasing hormone in onset of labour’, *The Lancet*, vol. 354, no. 9189, pp. 1546–1549, Oct. 1999, doi: 10.1016/S0140-6736(99)03418-2.
[48] R. G. Carroll, ‘Female Reproductive System’, in *Elsevier’s Integrated Physiology*, Elsevier, 2007, pp. 177–187. doi: 10.1016/B978-0-323-04318-2.50020-0.
[49] S. Vannuccini, C. Bocchi, F. M. Severi, J. R. Challis, and F. Petraglia, ‘Endocrinology of human parturition’, *Ann. Endocrinol.*, vol. 77, no. 2, pp. 105–113, Jun. 2016, doi: 10.1016/j.ando.2016.04.025.
[50] X. Fang, S. Wong, and B. F. Mitchell, ‘Relationships among sex steroids, oxytocin, and their receptors in the rat uterus during late gestation and at parturition.’, *Endocrinology*, vol. 137, no. 8, pp. 3213–3219, Aug. 1996, doi: 10.1210/endo.137.8.8754742.
[51] P. W. Nathanielsz, ‘Comparative studies on the initiation of labor’, *Eur. J. Obstet. Gynecol. Reprod. Biol.*, vol. 78, no. 2, pp. 127–132, Jun. 1998, doi: 10.1016/S0301-2115(98)00058-X.
[52] T. Mori and J. Kitawaki, ‘Chapter 7 - Role of estrogen and estrogen-related factors in endometriosis’, in *Immunology of Endometriosis*, K. Koga, Ed., in Reproductive Immunology. Academic Press, 2022, pp. 105–119. doi: 10.1016/B978-0-12-820661-4.00022-4.
[53] G. W. Theobald, A. Graham, J. Campbell, P. D. Gange, and W. J. Driscoll, ‘Use of Post-pituitary Extract in Obstetrics’, *BMJ*, vol. 2, no. 4567, pp. 123–127, Jul. 1948, doi: 10.1136/bmj.2.4567.123.
[54] G. Gimpl and F. Fahrenholz, ‘The Oxytocin Receptor System: Structure, Function, and Regulation’, *Physiol. Rev.*, vol. 81, no. 2, pp. 629–683, Apr. 2001, doi: 10.1152/physrev.2001.81.2.629.
[55] T. Higuchi, Y. Tadokoro, K. Honda, and H. Negoro, ‘Detailed analysis of blood oxytocin levels during suckling and parturition in the rat’, *J. Endocrinol.*, vol. 110, no. 2, pp. 251–256, Aug. 1986, doi: 10.1677/joe.0.1100251.
[56] A.-R. Fuchs, R. Romero, D. Keefe, M. Parra, E. Oyarzun, and E. Behnke, ‘Oxytocin secretion and human parturition: Pulse frequency and duration increase during spontaneous labor in women’, *Am. J. Obstet. Gynecol.*, vol. 165, no. 5, pp. 1515–1523, Nov. 1991, doi: 10.1016/0002-9378(91)90399-C.
[57] S. Thornton, J. M. Davison, and P. H. Baylis, ‘Plasma oxytocin during the first and second stages of spontaneous human labour’, *Acta Endocrinol. (Copenh.)*, vol. 126, no. 5, pp. 425–429, May 1992, doi: 10.1530/acta.0.1260425.
[58] J. A. Keelan, M. Blumenstein, R. J. A. Helliwell, T. A. Sato, K. W. Marvin, and M. D. Mitchell, ‘Cytokines, Prostaglandins and Parturition—A Review’, *Placenta*, vol. 24, pp. S33–S46, Apr. 2003, doi: 10.1053/plac.2002.0948.
[59] F. Vilella *et al.*, ‘PGE $_2$ and PGF $_2$ α Concentrations in Human Endometrial Fluid as Biomarkers for Embryonic Implantation’, *J. Clin. Endocrinol. Metab.*, vol. 98, no. 10, pp. 4123–4132, Oct. 2013, doi: 10.1210/jc.2013-2205.
[60] M. E. Carsten, ‘Prostaglandins and cellular calcium transport in the pregnant human uterus’, *Am. J. Obstet. Gynecol.*, vol. 117, no. 6, pp. 824–832, Nov. 1973, doi: 10.1016/0002-9378(73)90499-7.
[61] W. Y. Chan, I. Berezin, and E. E. Daniel, ‘Effects of inhibition of prostaglandin synthesis on uterine oxytocin receptor concentration and myometrial gap junction density in parturient rats’, *Biol. Reprod.*, vol. 39, no. 5, pp. 1117–1128, Dec. 1988, doi: 10.1095/biolreprod39.5.1117.

[62] C. D. Cuppett and S. N. Caritis, ‘Uterine Contraction Agents and Tocolytics’, in *Clinical Pharmacology During Pregnancy*, Elsevier, 2013, pp. 307–330. doi: 10.1016/B978-0-12-386007-1.00019-2.
[63] R. Bakker, S. Pierce, and D. Myers, ‘The role of prostaglandins E1 and E2, dinoprostone, and misoprostol in cervical ripening and the induction of labor: a mechanistic approach’, *Arch. Gynecol. Obstet.*, vol. 296, no. 2, pp. 167–179, Aug. 2017, doi: 10.1007/s00404-017-4418-5.
[64] A. Rokas, S. Mesiano, O. Tamam, A. LaBella, G. Zhang, and L. Muglia, ‘Developing a theoretical evolutionary framework to solve the mystery of parturition initiation’, *eLife*, vol. 9, p. e58343, doi: 10.7554/eLife.58343.
[65] A. Csapo, ‘Progesterone “block”’, *Am. J. Anat.*, vol. 98, no. 2, pp. 273–291, 1956, doi: 10.1002/aja.1000980206.
[66] S. Mesiano, Y. Wang, and E. R. Norwitz, ‘Progesterone Receptors in the Human Pregnancy Uterus: Do they Hold the Key to Birth Timing?’, *Reprod. Sci.*, vol. 18, no. 1, pp. 6–19, Jan. 2011, doi: 10.1177/1933719110382922.
[67] I. R. Young, ‘The Comparative Physiology of Parturition in Mammals’, *Endocrinol. Parturition*, vol. 27, pp. 10–30, 2001, doi: 10.1159/000061036.
[68] L. Nadeem, O. Shynlova, E. Matysiak-Zablocki, S. Mesiano, X. Dong, and S. Lye, ‘Molecular evidence of functional progesterone withdrawal in human myometrium’, *Nat. Commun.*, vol. 7, p. 11565, May 2016, doi: 10.1038/ncomms11565.
[69] M. A. Schneider, M. C. Davies, and J. W. Honour, ‘The timing of placental competence in pregnancy after oocyte donation’, *Fertil. Steril.*, vol. 59, no. 5, pp. 1059–1064, May 1993, doi: 10.1016/S0015-0282(16)55928-7.
[70] S. Taraborrelli, ‘Physiology, production and action of progesterone’, *Acta Obstet. Gynecol. Scand.*, vol. 94, pp. 8–16, Nov. 2015, doi: 10.1111/aogs.12771.
[71] M. S. Brown and J. L. Goldstein, ‘Receptor-Mediated Control of Cholesterol Metabolism: Study of human mutants has disclosed how cells regulate a substance that is both vital and lethal.’, *Science*, vol. 191, no. 4223, pp. 150–154, Jan. 1976, doi: 10.1126/science.174194.
[72] A. A. Bremer and W. L. Miller, ‘Chapter 13 - Regulation of Steroidogenesis’, in *Cellular Endocrinology in Health and Disease*, A. Ulloa-Aguirre and P. M. Conn, Eds., Boston: Academic Press, 2014, pp. 207–227. doi: 10.1016/B978-0-12-408134-5.00013-5.
[73] D. B. Hardy, B. A. Janowski, D. R. Corey, and C. R. Mendelson, ‘Progesterone Receptor Plays a Major Antiinflammatory Role in Human Myometrial Cells by Antagonism of Nuclear Factor-κB Activation of Cyclooxygenase 2 Expression’, *Mol. Endocrinol.*, vol. 20, no. 11, pp. 2724–2733, Nov. 2006, doi: 10.1210/me.2006-0112.
[74] X. Dong, C. Yu, O. Shynlova, J. R. G. Challis, P. S. Rennie, and S. J. Lye, ‘p54nrb Is a Transcriptional Corepressor of the Progesterone Receptor that Modulates Transcription of the Labor-Associated Gene, Connexin 43 (Gja1)’, *Mol. Endocrinol.*, vol. 23, no. 8, pp. 1147–1160, Aug. 2009, doi: 10.1210/me.2008-0357.
[75] Guoyang Luo *et al.*, ‘Progesterone Inhibits Basal and TNF-α-Induced Apoptosis in Fetal Membranes: A Novel Mechanism to Explain Progesterone-Mediated Prevention of Preterm Birth’, *Reprod. Sci.*, vol. 17, no. 6, pp. 532–539, Jun. 2010, doi: 10.1177/1933719110363618.
[76] O. M. Avrech, A. Golan, Z. Weinraub, I. Bukovsky, and E. Caspi, ‘Mifepristone (RU486) alone or in combination with a prostaglandin analogue for termination of early pregnancy: a review’, *Fertil. Steril.*, vol. 56, no. 3, pp. 385–393, Sep. 1991, doi: 10.1016/S0015-0282(16)54527-0.

[77] K. Chwalisz, K. Stöckemann, U. Fuhrmann, K. H. Fritzemeier, A. Einspanier, and R. E. Garfield, 'Mechanism of Action of Antiprogestins in the Pregnant Uterus', *Ann. N. Y. Acad. Sci.*, vol. 761, no. 1, pp. 202–223, 1995, doi: 10.1111/j.1749-6632.1995.tb31380.x.
[78] K. Chwalisz, 'The use of progesterone antagonists for cervical ripening and as an adjunct to labour and delivery', *Hum. Reprod. Oxf. Engl.*, vol. 9 Suppl 1, pp. 131–161, Jun. 1994, doi: 10.1093/humrep/9.suppl_1.131.
[79] J. Loudon, S. Sooranna, P. Bennett, and M. Johnson, 'Changes in progesterone receptor isoforms in cultured human uterine smooth muscle cells with pregnancy, mechanical stretch and labour at term', *J. Obstet. Gynaecol.*, vol. 23, pp. 59–59, 2003, doi: 10.1080/718591773.
[80] B. Runnebaum and J. Zander, 'Progesterone and 20 alpha-dihydroprogesterone in human myometrium during pregnancy', *Acta Endocrinol. Suppl. (Copenh.)*, vol. 150, pp. 3–45, 1971.
[81] R. P. Piekorz, S. Gingras, A. Hoffmeyer, J. N. Ihle, and Y. Weinstein, 'Regulation of progesterone levels during pregnancy and parturition by signal transducer and activator of transcription 5 and 20alpha-hydroxysteroid dehydrogenase', *Mol. Endocrinol. Baltim. Md*, vol. 19, no. 2, pp. 431–440, Feb. 2005, doi: 10.1210/me.2004-0302.
[82] G. Mendz, N. Kaakoush, and J. Quinlivan, 'Bacterial aetiological agents of intra-amniotic infections and preterm birth in pregnant women', *Front. Cell. Infect. Microbiol.*, vol. 3, 2013, Accessed: May 22, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/fcimb.2013.00058
[83] R. Romero and M. Mazor, 'Infection and preterm labor', *Clin. Obstet. Gynecol.*, vol. 31, no. 3, pp. 553–584, Sep. 1988, doi: 10.1097/00003081-198809000-00006.
[84] R. Romero, R. Gómez, T. Chaiworapongsa, G. Conoscenti, J. C. Kim, and Y. M. Kim, 'The role of infection in preterm labour and delivery', *Paediatr. Perinat. Epidemiol.*, vol. 15 Suppl 2, pp. 41–56, Jul. 2001, doi: 10.1046/j.1365-3016.2001.00007.x.
[85] K. Benirschke, 'Routes and types of infection in the fetus and the newborn', *AMA J. Dis. Child.*, vol. 99, pp. 714–721, Jun. 1960, doi: 10.1001/archpedi.1960.02070030716003.
[86] H. E. Jones *et al.*, 'Differing prevalence and diversity of bacterial species in fetal membranes from very preterm and term labor', *PloS One*, vol. 4, no. 12, p. e8205, Dec. 2009, doi: 10.1371/journal.pone.0008205.
[87] R. Romero *et al.*, 'Infection and labor. V. Prevalence, microbiology, and clinical significance of intraamniotic infection in women with preterm labor and intact membranes', *Am. J. Obstet. Gynecol.*, vol. 161, no. 3, pp. 817–824, Sep. 1989, doi: 10.1016/0002-9378(89)90409-2.
[88] R. Romero *et al.*, 'Infection and labor. VI. Prevalence, microbiology, and clinical significance of intraamniotic infection in twin gestations with preterm labor', *Am. J. Obstet. Gynecol.*, vol. 163, no. 3, pp. 757–761, Sep. 1990, doi: 10.1016/0002-9378(90)91063-i.
[89] L. Sykes, D. A. MacIntyre, X. J. Yap, T. G. Teoh, and P. R. Bennett, 'The Th1:Th2 Dichotomy of Pregnancy and Preterm Labour', *Mediators Inflamm.*, vol. 2012, pp. 1–12, 2012, doi: 10.1155/2012/967629.
[90] T. Bertran *et al.*, 'Slight Pro-Inflammatory Immunomodulation Properties of Dendritic Cells by *Gardnerella vaginalis* : The "Invisible Man" of Bacterial Vaginosis?', *J. Immunol. Res.*, vol. 2016, pp. 1–13, 2016, doi: 10.1155/2016/9747480.
[91] F. Amjadi, E. Salehi, M. Mehdizadeh, and R. Aflatoonian, 'Role of the innate immunity in female reproductive tract', *Adv. Biomed. Res.*, vol. 3, p. 1, Jan. 2014, doi: 10.4103/2277-9175.124626.

[92] T. M. Lindström and P. R. Bennett, ‘The role of nuclear factor kappa B in human labour’, *Reprod. Camb. Engl.*, vol. 130, no. 5, pp. 569–581, Nov. 2005, doi: 10.1530/rep.1.00197.
[93] G. C. Sharp, H. Ma, P. T. K. Saunders, and J. E. Norman, ‘A computational model of lipopolysaccharide-induced nuclear factor kappa B activation: a key signalling pathway in infection-induced preterm labour’, *PloS One*, vol. 8, no. 7, p. e70180, 2013, doi: 10.1371/journal.pone.0070180.
[94] E. Carmina, F. Z. Stanczyk, and R. A. Lobo, ‘Chapter 34 - Evaluation of Hormonal Status’, in *Yen and Jaffe’s Reproductive Endocrinology (Eighth Edition)*, J. F. Strauss and R. L. Barbieri, Eds., Philadelphia: Elsevier, 2019, pp. 887-915.e4. doi: 10.1016/B978-0-323-47912-7.00034-2.
[95] R. Romero, ‘Vaginal Progesterone to Reduce the Rate of Preterm Birth and Neonatal Morbidity: A Solution at Last’, *Womens Health*, vol. 7, no. 5, pp. 501–504, Sep. 2011, doi: 10.2217/WHE.11.60.
[96] M. Mazor *et al.*, ‘Human preterm birth is associated with systemic and local changes in progesterone/17β-estradiol ratios’, *Am. J. Obstet. Gynecol.*, vol. 171, no. 1, pp. 231–236, Jul. 1994, doi: 10.1016/0002-9378(94)90474-X.
[97] C. W. Ku, J. C. Allen Jr, S. M. Lek, M. L. Chia, N. S. Tan, and T. C. Tan, ‘Serum progesterone distribution in normal pregnancies compared to pregnancies complicated by threatened miscarriage from 5 to 13 weeks gestation: a prospective cohort study’, *BMC Pregnancy Childbirth*, vol. 18, no. 1, p. 360, Sep. 2018, doi: 10.1186/s12884-018-2002-z.
[98] G. S. Berkowitz and E. Papiernik, ‘Epidemiology of preterm birth’, *Epidemiol. Rev.*, vol. 15, no. 2, pp. 414–443, 1993, doi: 10.1093/oxfordjournals.epirev.a036128.
[99] L. S. Bakketeig, H. J. Hoffman, and E. E. Harley, ‘The tendency to repeat gestational age and birth weight in successive births’, *Am. J. Obstet. Gynecol.*, vol. 135, no. 8, pp. 1086–1103, Dec. 1979, doi: 10.1016/0002-9378(79)90742-7.
[100] B. M. Mercer *et al.*, ‘The preterm prediction study: effect of gestational age and cause of preterm birth on subsequent obstetric outcome. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network’, *Am. J. Obstet. Gynecol.*, vol. 181, no. 5 Pt 1, pp. 1216–1221, Nov. 1999, doi: 10.1016/s0002-9378(99)70111-0.
[101] F. Bruinsma and M. Quinn, ‘The risk of preterm birth following treatment for precancerous changes in the cervix: a systematic review and meta-analysis’, *BJOG Int. J. Obstet. Gynaecol.*, vol. 118, no. 9, pp. 1031–1041, 2011, doi: 10.1111/j.1471-0528.2011.02944.x.
[102] L. Poon, M. Savvas, D. Zamblera, E. Skyfta, and K. Nicolaides, ‘Large loop excision of transformation zone and cervical length in the prediction of spontaneous preterm delivery’, *BJOG Int. J. Obstet. Gynaecol.*, vol. 119, no. 6, pp. 692–698, 2012, doi: 10.1111/j.1471-0528.2011.03203.x.
[103] G. A. Dekker, S. Y. Lee, R. A. North, L. M. McCowan, N. A. B. Simpson, and C. T. Roberts, ‘Risk factors for preterm birth in an international prospective cohort of nulliparous women’, *PloS One*, vol. 7, no. 7, p. e39154, 2012, doi: 10.1371/journal.pone.0039154.
[104] K. M. Orzechowski, R. C. Boelig, J. K. Baxter, and V. Berghella, ‘A universal transvaginal cervical length screening program for preterm birth prevention’, *Obstet. Gynecol.*, vol. 124, no. 3, pp. 520–525, Sep. 2014, doi: 10.1097/AOG.0000000000000428.
[105] M. S. Esplin *et al.*, ‘Predictive Accuracy of Serial Transvaginal Cervical Lengths and Quantitative Vaginal Fetal Fibronectin Levels for Spontaneous Preterm Birth Among

Nulliparous Women', *JAMA*, vol. 317, no. 10, pp. 1047–1056, Mar. 2017, doi: 10.1001/jama.2017.1373.
[106] S. L. Bloom and K. J. Leveno, 'Unproven Technologies in Maternal-Fetal Medicine and the High Cost of US Health Care', *JAMA*, vol. 317, no. 10, pp. 1025–1026, Mar. 2017, doi: 10.1001/jama.2017.1218.
[107] A. M. Friedman, N. Schwartz, J. Ludmir, S. Parry, J. A. Bastek, and H. M. Sehdev, 'Can transabdominal ultrasound identify women at high risk for short cervical length?', *Acta Obstet. Gynecol. Scand.*, vol. 92, no. 6, pp. 637–641, 2013, doi: 10.1111/aogs.12111.
[108] N. L. Hezelgrave and A. H. Shennan, 'Quantitative Fetal Fibronectin to Predict Spontaneous Preterm Birth: A Review', *Womens Health*, vol. 12, no. 1, pp. 121–128, Jan. 2016, doi: 10.2217/whe.15.74.
[109] D. S. Abbott *et al.*, 'Quantitative fetal fibronectin to predict preterm birth in asymptomatic women at high risk', *Obstet. Gynecol.*, vol. 125, no. 5, pp. 1168–1176, May 2015, doi: 10.1097/AOG.0000000000000754.
[110] A. M. Peaceman *et al.*, 'Fetal fibronectin as a predictor of preterm birth in patients with symptoms: a multicenter trial', *Am. J. Obstet. Gynecol.*, vol. 177, no. 1, pp. 13–18, Jul. 1997, doi: 10.1016/s0002-9378(97)70431-9.
[111] S. N. Deshpande *et al.*, 'Rapid fetal fibronectin testing to predict preterm birth in women with symptoms of premature labour: a systematic review and cost analysis', *Health Technol. Assess. Winch. Engl.*, vol. 17, no. 40, pp. 1–138, Sep. 2013, doi: 10.3310/hta17400.
[112] H. Honest *et al.*, 'Screening to prevent spontaneous preterm birth: systematic reviews of accuracy and effectiveness literature with economic modelling', *Health Technol. Assess. Winch. Engl.*, vol. 13, no. 43, pp. 1–627, Sep. 2009, doi: 10.3310/hta13430.
[113] Y. Sadovsky and S. A. Friedman, 'Fetal fibronectin and preterm labor', *N. Engl. J. Med.*, vol. 326, no. 10, p. 709, Mar. 1992, doi: 10.1056/NEJM199203053261014.
[114] Y. J. Heng, S. Liong, M. Permezel, G. E. Rice, M. K. W. Di Quinzio, and H. M. Georgiou, 'Human cervicovaginal fluid biomarkers to predict term and preterm labor', *Front. Physiol.*, vol. 6, p. 151, May 2015, doi: 10.3389/fphys.2015.00151.
[115] A. Conde-Agudelo, 'Birth Spacing and Risk of Adverse Perinatal Outcomes', p. 15.
[116] S. M. Lee *et al.*, 'The clinical significance of a positive Amnisure test in women with term labor with intact membranes', *J. Matern.-Fetal Neonatal Med. Off. J. Eur. Assoc. Perinat. Med. Fed. Asia Ocean. Perinat. Soc. Int. Soc. Perinat. Obstet.*, vol. 22, no. 4, pp. 305–310, Apr. 2009, doi: 10.1080/14767050902801694.
[117] K. Flood and F. D. Malone, 'Prevention of preterm birth', *Semin. Fetal. Neonatal Med.*, vol. 17, no. 1, pp. 58–63, Feb. 2012, doi: 10.1016/j.siny.2011.08.001.
[118] Y. S. Han, E. H. Ha, H. S. Park, Y. J. Kim, and S. S. Lee, 'Relationships between pregnancy outcomes, biochemical markers and pre-pregnancy body mass index', *Int. J. Obes. 2005*, vol. 35, no. 4, pp. 570–577, Apr. 2011, doi: 10.1038/ijo.2010.162.
[119] H. M. Salihu, M. H. Aliyu, B. J. Pierre-Louis, and G. R. Alexander, 'Levels of excess infant deaths attributable to maternal smoking during pregnancy in the United States', *Matern. Child Health J.*, vol. 7, no. 4, pp. 219–227, Dec. 2003, doi: 10.1023/a:1027319517405.
[120] J. Lumley, C. Chamberlain, T. Dowswell, S. Oliver, L. Oakley, and L. Watson, 'Interventions for promoting smoking cessation during pregnancy', *Cochrane Database Syst. Rev.*, no. 3, p. CD001055, Jul. 2009, doi: 10.1002/14651858.CD001055.pub3.

[121] T. O. Scholl, ‘Iron status during pregnancy: setting the stage for mother and infant’, *Am. J. Clin. Nutr.*, vol. 81, no. 5, pp. 1218S-1222S, May 2005, doi: 10.1093/ajcn/81.5.1218.
[122] N. Milman, T. Paszkowski, I. Cetin, and C. Castelo-Branco, ‘Supplementation during pregnancy: beliefs and science’, *Gynecol. Endocrinol.*, vol. 32, no. 7, pp. 509–516, Jul. 2016, doi: 10.3109/09513590.2016.1149161.
[123] C. Y. Spong, ‘Prediction and Prevention of Recurrent Spontaneous Preterm Birth’, *Obstet. Gynecol.*, vol. 110, no. 2 Part 1, pp. 405–415, Aug. 2007, doi: 10.1097/01.AOG.0000275287.08520.4a.
[124] P. Piette, ‘The history of natural progesterone, the never-ending story’, *Climacteric*, vol. 21, no. 4, pp. 308–314, Jul. 2018, doi: 10.1080/13697137.2018.1462792.
[125] L. Sykes and P. R. Bennett, ‘Efficacy of progesterone for prevention of preterm birth’, *Best Pract. Res. Clin. Obstet. Gynaecol.*, vol. 52, pp. 126–136, Oct. 2018, doi: 10.1016/j.bpobgyn.2018.08.006.
[126] G. C. Di Renzo, ‘Concerns about the review of vaginal progesterone and the vaginal first-pass effect’, *Climacteric*, vol. 22, no. 1, pp. 105–105, Jan. 2019, doi: 10.1080/13697137.2018.1491544.
[127] A. Coomarasamy *et al.*, ‘Micronized vaginal progesterone to prevent miscarriage: a critical evaluation of randomized evidence’, *Am. J. Obstet. Gynecol.*, vol. 223, no. 2, pp. 167–176, Aug. 2020, doi: 10.1016/j.ajog.2019.12.006.
[128] G. C. Di Renzo, V. Tosto, and V. Tsibizova, ‘Progesterone: History, facts, and artifacts’, *Best Pract. Res. Clin. Obstet. Gynaecol.*, vol. 69, pp. 2–12, Nov. 2020, doi: 10.1016/j.bpobgyn.2020.07.012.
[129] E. B. da Fonseca, R. E. Bittar, M. H. B. Carvalho, and M. Zugaib, ‘Prophylactic administration of progesterone by vaginal suppository to reduce the incidence of spontaneous preterm birth in women at increased risk: a randomized placebo-controlled double-blind study’, *Am. J. Obstet. Gynecol.*, vol. 188, no. 2, pp. 419–424, Feb. 2003, doi: 10.1067/mob.2003.41.
[130] E. B. Fonseca, E. Celik, M. Parra, M. Singh, K. H. Nicolaides, and Fetal Medicine Foundation Second Trimester Screening Group, ‘Progesterone and the risk of preterm birth among women with a short cervix’, *N. Engl. J. Med.*, vol. 357, no. 5, pp. 462–469, Aug. 2007, doi: 10.1056/NEJMoa067815.
[131] S. S. Hassan *et al.*, ‘Vaginal progesterone reduces the rate of preterm birth in women with a sonographic short cervix: a multicenter, randomized, double-blind, placebo-controlled trial’, *Ultrasound Obstet. Gynecol. Off. J. Int. Soc. Ultrasound Obstet. Gynecol.*, vol. 38, no. 1, pp. 18–31, Jul. 2011, doi: 10.1002/uog.9017.
[132] P. J. Meis *et al.*, ‘Prevention of recurrent preterm delivery by 17 alpha-hydroxyprogesterone caproate’, *N. Engl. J. Med.*, vol. 348, no. 24, pp. 2379–2385, Jun. 2003, doi: 10.1056/NEJMoa035140.
[133] M. D. van Zijl, B. Koullali, B. W. Mol, E. Pajkrt, and M. A. Oudijk, ‘Prevention of preterm delivery: current challenges and future prospects’, *Int. J. Womens Health*, vol. 8, pp. 633–645, Dec. 2016, doi: 10.2147/IJWH.S89317.
[134] Z. Alfirevic, T. Stampalija, and N. Medley, ‘Cervical stitch (cerclage) for preventing preterm birth in singleton pregnancy’, *Cochrane Database Syst. Rev.*, vol. 2017, no. 6, p. CD008991, Jun. 2017, doi: 10.1002/14651858.CD008991.pub3.
[135] V. Berghella, T. J. Rafael, J. M. Szychowski, O. A. Rust, and J. Owen, ‘Cerclage for short cervix on ultrasonography in women with singleton gestations and previous preterm birth: a meta-analysis’, *Obstet. Gynecol.*, vol. 117, no. 3, pp. 663–671, Mar. 2011, doi: 10.1097/AOG.0b013e31820ca847.

[136] A. Shennan, L. Story, B. Jacobsson, W. A. Grobman, and the F. W. G. for P. Birth, 'FIGO good practice recommendations on cervical cerclage for prevention of preterm birth', *Int. J. Gynecol. Obstet.*, vol. 155, no. 1, pp. 19–22, 2021, doi: 10.1002/ijgo.13835.
[137] V. Berghella, A. O. Odibo, M. S. To, O. A. Rust, and S. M. Althuisius, 'Cerclage for short cervix on ultrasonography: meta-analysis of trials using individual patient-level data', *Obstet. Gynecol.*, vol. 106, no. 1, pp. 181–189, Jul. 2005, doi: 10.1097/01.AOG.0000168435.17200.53.
[138] R. C. Boelig and V. Berghella, 'Current options for mechanical prevention of preterm birth', *Semin. Perinatol.*, vol. 41, no. 8, pp. 452–460, Dec. 2017, doi: 10.1053/j.semperi.2017.08.003.
[139] L. M. Kindinger *et al.*, 'Relationship between vaginal microbial dysbiosis, inflammation, and pregnancy outcomes in cervical cerclage', *Sci. Transl. Med.*, vol. 8, no. 350, p. 350ra102, Aug. 2016, doi: 10.1126/scitranslmed.aag1026.
[140] R. Romero, F. Gotsch, B. Pineles, and J. P. Kusanovic, 'Inflammation in pregnancy: its roles in reproductive physiology, obstetrical complications, and fetal injury', *Nutr. Rev.*, vol. 65, no. 12 Pt 2, pp. S194-202, Dec. 2007, doi: 10.1111/j.1753-4887.2007.tb00362.x.
[141] M. Bellad *et al.*, 'Clindamycin to reduce preterm birth in a low resource setting: a randomised placebo-controlled clinical trial', *BJOG Int. J. Obstet. Gynaecol.*, vol. 125, no. 12, pp. 1601–1609, 2018, doi: 10.1111/1471-0528.15290.
[142] R. L. Goldenberg *et al.*, 'The HPTN 024 Study: The efficacy of antibiotics to prevent chorioamnionitis and preterm birth', *Am. J. Obstet. Gynecol.*, vol. 194, no. 3, pp. 650–661, Mar. 2006, doi: 10.1016/j.ajog.2006.01.004.
[143] R. F. Lamont, 'Spontaneous preterm labour that leads to preterm birth: An update and personal reflection', *Placenta*, vol. 79, pp. 21–29, Apr. 2019, doi: 10.1016/j.placenta.2019.03.010.
[144] D. Subtil *et al.*, 'Early clindamycin for bacterial vaginosis in pregnancy (PREMEVA): a multicentre, double-blind, randomised controlled trial', *Lancet Lond. Engl.*, vol. 392, no. 10160, pp. 2171–2179, Nov. 2018, doi: 10.1016/S0140-6736(18)31617-9.
[145] M. A. Klebanoff and R. M. Brotman, 'Treatment of bacterial vaginosis to prevent preterm birth', *The Lancet*, vol. 392, no. 10160, pp. 2141–2142, Nov. 2018, doi: 10.1016/S0140-6736(18)32115-9.
[146] L. Sykes, D. MacIntyre, T. Teoh, and P. Bennett, 'Targeting Immune Activation in the Prevention of Pre-term Labour', *Eur. Obstet. Gynaecol.*, vol. 6, pp. 100–106, Dec. 2011.
[147] A.-M. Morency and E. Bujold, 'The effect of second-trimester antibiotic therapy on the rate of preterm birth', *J. Obstet. Gynaecol. Can. JOGC J. Obstet. Gynecol. Can. JOGC*, vol. 29, no. 1, pp. 35–44, Jan. 2007, doi: 10.1016/s1701-2163(16)32350-7.
[148] D. A. MacIntyre, L. Sykes, T. G. Teoh, and P. R. Bennett, 'Prevention of preterm labour via the modulation of inflammatory pathways', *J. Matern.-Fetal Neonatal Med. Off. J. Eur. Assoc. Perinat. Med. Fed. Asia Ocean. Perinat. Soc. Int. Soc. Perinat. Obstet.*, vol. 25 Suppl 1, pp. 17–20, Apr. 2012, doi: 10.3109/14767058.2012.666114.
[149] D. A. MacIntyre *et al.*, 'Activator protein 1 is a key terminal mediator of inflammation-induced preterm labor in mice', *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.*, vol. 28, no. 5, pp. 2358–2368, May 2014, doi: 10.1096/fj.13-247783.
[150] R. Migale *et al.*, 'Specific Lipopolysaccharide Serotypes Induce Differential Maternal and Neonatal Inflammatory Responses in a Murine Model of Preterm Labor', *Am. J. Pathol.*, vol. 185, no. 9, pp. 2390–2401, Sep. 2015, doi: 10.1016/j.ajpath.2015.05.015.

[151] R. F. Lamont, ‘Advances in the Prevention of Infection-Related Preterm Birth’, *Front. Immunol.*, vol. 6, p. 566, 2015, doi: 10.3389/fimmu.2015.00566.
[152] M. J. Ferris, A. Masztal, K. E. Aldridge, J. D. Fortenberry, P. L. Fidel, and D. H. Martin, ‘Association of Atopobium vaginae, a recently described metronidazole resistant anaerobe, with bacterial vaginosis’, *BMC Infect. Dis.*, vol. 4, p. 5, Feb. 2004, doi: 10.1186/1471-2334-4-5.
[153] R. G. Brown *et al.*, ‘Vaginal dysbiosis increases risk of preterm fetal membrane rupture, neonatal sepsis and is exacerbated by erythromycin’, *BMC Med.*, vol. 16, no. 1, p. 9, Jan. 2018, doi: 10.1186/s12916-017-0999-x.
[154] D. G. Bostwick, J. Woody, C. Hunt, and W. Budd, ‘Antimicrobial resistance genes and modelling of treatment failure in bacterial vaginosis: clinical study of 289 symptomatic women’, *J. Med. Microbiol.*, vol. 65, no. 5, pp. 377–386, 2016, doi: 10.1099/jmm.0.000236.
[155] R. H. Beigi, M. N. Austin, L. A. Meyn, M. A. Krohn, and S. L. Hillier, ‘Antimicrobial resistance associated with the treatment of bacterial vaginosis’, *Am. J. Obstet. Gynecol.*, vol. 191, no. 4, pp. 1124–1129, Oct. 2004, doi: 10.1016/j.ajog.2004.05.033.
[156] S. Kenyon *et al.*, ‘Has publication of the results of the ORACLE Children Study changed practice in the UK?’, *BJOG Int. J. Obstet. Gynaecol.*, vol. 117, no. 11, pp. 1344–1349, Oct. 2010, doi: 10.1111/j.1471-0528.2010.02661.x.
[157] S. L. Kenyon, D. J. Taylor, W. Tarnow-Mordi, and ORACLE Collaborative Group, ‘Broad-spectrum antibiotics for preterm, prelabour rupture of fetal membranes: the ORACLE I randomised trial. ORACLE Collaborative Group’, *Lancet Lond. Engl.*, vol. 357, no. 9261, pp. 979–988, Mar. 2001, doi: 10.1016/s0140-6736(00)04233-1.
[158] A. Iliopoulou, R. N. Thin, and P. Turner, ‘Fluorimetric and microbiological assays of erythromycin concentrations in plasma and vaginal washings.’, *Br. J. Vener. Dis.*, vol. 57, no. 4, pp. 263–267, Aug. 1981.
[159] T. Kuriyama *et al.*, ‘Antimicrobial susceptibility of 800 anaerobic isolates from patients with dentoalveolar infection to 13 oral antibiotics’, *Oral Microbiol. Immunol.*, vol. 22, no. 4, pp. 285–288, Aug. 2007, doi: 10.1111/j.1399-302X.2007.00365.x.
[160] M. Desaki, H. Okazaki, T. Sunazuka, S. Omura, K. Yamamoto, and H. Takizawa, ‘Molecular Mechanisms of Anti-Inflammatory Action of Erythromycin in Human Bronchial Epithelial Cells: Possible Role in the Signaling Pathway That Regulates Nuclear Factor-κB Activation’, *Antimicrob. Agents Chemother.*, vol. 48, no. 5, pp. 1581–1585, May 2004, doi: 10.1128/AAC.48.5.1581-1585.2004.
[161] G. Berg *et al.*, ‘Microbiome definition re-visited: old concepts and new challenges’, *Microbiome*, vol. 8, no. 1, p. 103, Dec. 2020, doi: 10.1186/s40168-020-00875-0.
[162] Whipps JM, Lewis K, Cooke RC., *Mycoparasitism and plant disease control*, vol. Fungi in Biological Control Systems. Manchester University Press, 1988.
[163] R. Sender, S. Fuchs, and R. Milo, ‘Revised Estimates for the Number of Human and Bacteria Cells in the Body’, *PLOS Biol.*, vol. 14, no. 8, p. e1002533, Aug. 2016, doi: 10.1371/journal.pbio.1002533.
[164] T. Uzbay, ‘Germ-free animal experiments in the gut microbiota studies’, *Curr. Opin. Pharmacol.*, vol. 49, pp. 6–10, Dec. 2019, doi: 10.1016/j.coph.2019.03.016.
[165] C. Alam *et al.*, ‘Effects of a germ-free environment on gut immune regulation and diabetes progression in non-obese diabetic (NOD) mice’, *Diabetologia*, vol. 54, no. 6, pp. 1398–1406, Jun. 2011, doi: 10.1007/s00125-011-2097-5.
[166] K. Shimizu, Y. Muranaka, R. Fujimura, H. Ishida, S. Tazume, and T. Shimamura, ‘Normalization of Reproductive Function in Germfree Mice Following Bacterial Contamination.’, *Exp. Anim.*, vol. 47, no. 3, pp. 151–158, 1998, doi: 10.1538/expanim.47.151.

[167] J. Chow, S. M. Lee, Y. Shen, A. Khosravi, and S. K. Mazmanian, ‘Chapter 8 - Host–Bacterial Symbiosis in Health and Disease’, in *Advances in Immunology*, S. Fagarasan and A. Cerutti, Eds., in Mucosal Immunity, vol. 107. Academic Press, 2010, pp. 243–274. doi: 10.1016/B978-0-12-381300-8.00008-3.
[168] A. M. Cole, ‘Innate host defense of human vaginal and cervical mucosae’, *Curr. Top. Microbiol. Immunol.*, vol. 306, pp. 199–230, 2006.
[169] L. Howe, R. Wiggins, P. W. Soothill, M. R. Millar, P. J. Horner, and A. P. Corfield, ‘Mucinase and sialidase activity of the vaginal microflora: implications for the pathogenesis of preterm labour’, *Int. J. STD AIDS*, vol. 10, no. 7, pp. 442–447, Jul. 1999, doi: 10.1258/0956462991914438.
[170] Y. Goto *et al.*, ‘Innate lymphoid cells regulate intestinal epithelial cell glycosylation’, *Science*, vol. 345, no. 6202, p. 1254009, Sep. 2014, doi: 10.1126/science.1254009.
[171] P. C. Kashyap *et al.*, ‘Genetically dictated change in host mucus carbohydrate landscape exerts a diet-dependent effect on the gut microbiota’, *Proc. Natl. Acad. Sci.*, vol. 110, no. 42, pp. 17059–17064, Oct. 2013, doi: 10.1073/pnas.1306070110.
[172] S. L. Taylor, M. A. McGuckin, S. Wesselingh, and G. B. Rogers, ‘Infection’s Sweet Tooth: How Glycans Mediate Infection and Disease Susceptibility’, *Trends Microbiol.*, vol. 26, no. 2, pp. 92–101, Feb. 2018, doi: 10.1016/j.tim.2017.09.011.
[173] S. Ali *et al.*, ‘Secretor Polymorphism and Human Immunodeficiency Virus Infection in Senegalese Women’, *J. Infect. Dis.*, vol. 181, no. 2, pp. 737–739, Feb. 2000, doi: 10.1086/315234.
[174] D. Hu *et al.*, ‘Associations of *FUT2* and *FUT3* gene polymorphisms with Crohn’s disease in Chinese patients: *FUT2* and *FUT3* in Crohn’s disease’, *J. Gastroenterol. Hepatol.*, vol. 29, no. 10, pp. 1778–1785, Oct. 2014, doi: 10.1111/jgh.12599.
[175] S. Kundu *et al.*, ‘The effect of secretor status and the vaginal microbiome on birth outcome’. medRxiv, p. 2021.11.24.21266804, Nov. 24, 2021. doi: 10.1101/2021.11.24.21266804.
[176] L. Weinstein, M. Bogin, J. H. Howard, and B. B. Finkelstone, ‘A survey of the vaginal flora at various ages, with special reference to the Döderlein bacillus’, *Am. J. Obstet. Gynecol.*, vol. 32, no. 2, pp. 211–218, Aug. 1936, doi: 10.1016/S0002-9378(36)90120-5.
[177] S. A. Ibrahim, ‘Lactic Acid Bacteria: Lactobacillus spp.: Other Species’, in *Reference Module in Food Science*, Elsevier, 2016. doi: 10.1016/B978-0-08-100596-5.00857-X.
[178] T. Zotta, E. Parente, and A. Ricciardi, ‘Aerobic metabolism in the genus Lactobacillus: impact on stress response and potential applications in the food industry’, *J. Appl. Microbiol.*, vol. 122, no. 4, pp. 857–869, Apr. 2017, doi: 10.1111/jam.13399.
[179] S. Yildirim *et al.*, ‘Primate vaginal microbiomes exhibit species specificity without universal Lactobacillus dominance’, *ISME J.*, vol. 8, no. 12, pp. 2431–2444, Dec. 2014, doi: 10.1038/ismej.2014.90.
[180] E. Amabebe and D. O. C. Anumba, ‘The Vaginal Microenvironment: The Physiologic Role of Lactobacilli’, *Front. Med.*, vol. 5, 2018, Accessed: Mar. 18, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/fmed.2018.00181
[181] S. S. Witkin, H. Mendes-Soares, I. M. Linhares, A. Jayaram, W. J. Ledger, and L. J. Forney, ‘Influence of Vaginal Bacteria and D - and L -Lactic Acid Isomers on Vaginal Extracellular Matrix Metalloproteinase Inducer: Implications for Protection against Upper Genital Tract Infections’, *mBio*, vol. 4, no. 4, pp. e00460-13, Aug. 2013, doi: 10.1128/mBio.00460-13.
[182] D. E. O’Hanlon, T. R. Moench, and R. A. Cone, ‘Vaginal pH and microbicidal lactic acid when lactobacilli dominate the microbiota’, *PloS One*, vol. 8, no. 11, p. e80074, 2013, doi: 10.1371/journal.pone.0080074.

[183] D. E. O'Hanlon, T. R. Moench, and R. A. Cone, 'In vaginal fluid, bacteria associated with bacterial vaginosis can be suppressed with lactic acid but not hydrogen peroxide', *BMC Infect. Dis.*, vol. 11, p. 200, Jul. 2011, doi: 10.1186/1471-2334-11-200.
[184] E. R. Kashket, 'Bioenergetics of lactic acid bacteria: cytoplasmic pH and osmotolerance*', *FEMS Microbiol. Rev.*, vol. 3, no. 3, pp. 233–244, Sep. 1987, doi: 10.1111/j.1574-6968.1987.tb02463.x.
[185] H.-L. Alakomi, E. Skyttä, M. Saarela, T. Mattila-Sandholm, K. Latva-Kala, and I. M. Helander, 'Lactic Acid Permeabilizes Gram-Negative Bacteria by Disrupting the Outer Membrane', *Appl. Environ. Microbiol.*, vol. 66, no. 5, pp. 2001–2005, May 2000.
[186] G. Tachedjian, M. Aldunate, C. S. Bradshaw, and R. A. Cone, 'The role of lactic acid production by probiotic Lactobacillus species in vaginal health', *Res. Microbiol.*, vol. 168, no. 9, pp. 782–792, Nov. 2017, doi: 10.1016/j.resmic.2017.04.001.
[187] A. C. Hearps *et al.*, 'Vaginal lactic acid elicits an anti-inflammatory response from human cervicovaginal epithelial cells and inhibits production of pro-inflammatory mediators associated with HIV acquisition', *Mucosal Immunol.*, vol. 10, no. 6, Art. no. 6, Nov. 2017, doi: 10.1038/mi.2017.27.
[188] R. F. Lamont, 'Infection in the prediction and antibiotics in the prevention of spontaneous preterm labour and preterm birth', *BJOG Int. J. Obstet. Gynaecol.*, vol. 110 Suppl 20, pp. 71–75, Apr. 2003, doi: 10.1016/s1470-0328(03)00034-x.
[189] P. A. Cadieux, J. Burton, E. Devillard, and G. Reid, 'Lactobacillus by-products inhibit the growth and virulence of uropathogenic Escherichia coli', *J. Physiol. Pharmacol. Off. J. Pol. Physiol. Soc.*, vol. 60 Suppl 6, pp. 13–18, Dec. 2009.
[190] E. Motevaseli *et al.*, 'The Effect of Lactobacillus crispatus and Lactobacillus rhamnosusCulture Supernatants on Expression of Autophagy Genes and HPV E6 and E7 Oncogenes in The HeLa Cell Line', *Cell J. Yakhteh*, vol. 17, no. 4, pp. 601–607, 2016.
[191] B. Huang, J. M. Fettweis, J. P. Brooks, K. K. Jefferson, and G. A. Buck, 'The Changing Landscape of the Vaginal Microbiome', *Clin. Lab. Med.*, vol. 34, no. 4, pp. 747–761, Dec. 2014, doi: 10.1016/j.cll.2014.08.006.
[192] J. Ravel *et al.*, 'Vaginal microbiome of reproductive-age women', *Proc. Natl. Acad. Sci.*, vol. 108, no. supplement_1, pp. 4680–4687, Mar. 2011, doi: 10.1073/pnas.1002611107.
[193] R. Romero *et al.*, 'The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women', *Microbiome*, vol. 2, no. 1, p. 4, Feb. 2014, doi: 10.1186/2049-2618-2-4.
[194] M. France, M. Alizadeh, S. Brown, B. Ma, and J. Ravel, 'Towards a deeper understanding of the vaginal microbiota', *Nat. Microbiol.*, vol. 7, no. 3, Art. no. 3, Mar. 2022, doi: 10.1038/s41564-022-01083-2.
[195] S. Srinivasan *et al.*, 'Bacterial Communities in Women with Bacterial Vaginosis: High Resolution Phylogenetic Analyses Reveal Relationships of Microbiota to Clinical Criteria', *PLOS ONE*, vol. 7, no. 6, p. e37818, Jun. 2012, doi: 10.1371/journal.pone.0037818.
[196] J. H. H. M. van de Wijgert *et al.*, 'The Vaginal Microbiota: What Have We Learned after a Decade of Molecular Characterization?', *PLOS ONE*, vol. 9, no. 8, p. e105998, Aug. 2014, doi: 10.1371/journal.pone.0105998.
[197] T. Drell *et al.*, 'Characterization of the Vaginal Micro- and Mycobiome in Asymptomatic Reproductive-Age Estonian Women', *PLOS ONE*, vol. 8, no. 1, p. e54379, Jan. 2013, doi: 10.1371/journal.pone.0054379.

[198] M. T. France *et al.*, 'VALENCIA: a nearest centroid classification method for vaginal microbial communities based on composition', *Microbiome*, vol. 8, no. 1, p. 166, Nov. 2020, doi: 10.1186/s40168-020-00934-6.
[199] P. Pruski *et al.*, 'Direct on-swab metabolic profiling of vaginal microbiome host interactions during pregnancy and preterm birth', *Nat. Commun.*, vol. 12, no. 1, Art. no. 1, Oct. 2021, doi: 10.1038/s41467-021-26215-w.
[200] S. Lee *et al.*, 'Community State Types of Vaginal Microbiota and Four Types of Abnormal Vaginal Microbiota in Pregnant Korean Women', *Front. Public Health*, vol. 8, p. 507024, Oct. 2020, doi: 10.3389/fpubh.2020.507024.
[201] W. J. Y. Chee, S. Y. Chew, and L. T. L. Than, 'Vaginal microbiota and the potential of Lactobacillus derivatives in maintaining vaginal health', *Microb. Cell Factories*, vol. 19, no. 1, p. 203, Dec. 2020, doi: 10.1186/s12934-020-01464-4.
[202] M. T. France, H. Mendes-Soares, and L. J. Forney, 'Genomic Comparisons of Lactobacillus crispatus and Lactobacillus iners Reveal Potential Ecological Drivers of Community Composition in the Vagina', *Appl. Environ. Microbiol.*, vol. 82, no. 24, pp. 7063–7073, Dec. 2016, doi: 10.1128/AEM.02385-16.
[203] M. J. Ferris, J. Norori, M. Zozaya-Hinchliffe, and D. H. Martin, 'Cultivation-Independent Analysis of Changes in Bacterial Vaginosis Flora Following Metronidazole Treatment', *J. Clin. Microbiol.*, vol. 45, no. 3, pp. 1016–1018, Mar. 2007, doi: 10.1128/JCM.02085-06.
[204] L. Goodfellow *et al.*, 'Vaginal bacterial load in the second trimester is associated with early preterm birth recurrence: a nested case–control study', *BJOG Int. J. Obstet. Gynaecol.*, vol. 128, no. 13, pp. 2061–2072, Dec. 2021, doi: 10.1111/1471-0528.16816.
[205] P. Hay, 'Bacterial vaginosis', *F1000Research*, vol. 6, p. 1761, Sep. 2017, doi: 10.12688/f1000research.11417.1.
[206] N. Zheng, R. Guo, J. Wang, W. Zhou, and Z. Ling, 'Contribution of Lactobacillus iners to Vaginal Health and Diseases: A Systematic Review', *Front. Cell. Infect. Microbiol.*, vol. 11, 2021, Accessed: Mar. 18, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/fcimb.2021.792787
[207] R. Thapa, S. Ray, and P. A. Keyel, 'Interaction of Macrophages and Cholesterol-Dependent Cytolysins: The Impact on Immune Response and Cellular Survival', *Toxins*, vol. 12, no. 9, p. 531, Aug. 2020, doi: 10.3390/toxins12090531.
[208] R. M. Brotman *et al.*, 'Bacterial Vaginosis Assessed by Gram Stain and Diminished Colonization Resistance to Incident Gonococcal, Chlamydial, and Trichomonal Genital Infection', *J. Infect. Dis.*, vol. 202, no. 12, pp. 1907–1915, Dec. 2010, doi: 10.1086/657320.
[209] H. L. Martin Jr. *et al.*, 'Vaginal Lactobacilli, Microbial Flora, and Risk of Human Immunodeficiency Virus Type 1 and Sexually Transmitted Disease Acquisition', *J. Infect. Dis.*, vol. 180, no. 6, pp. 1863–1868, Dec. 1999, doi: 10.1086/315127.
[210] C. Gosmann *et al.*, 'Lactobacillus-Deficient Cervicovaginal Bacterial Communities Are Associated with Increased HIV Acquisition in Young South African Women', *Immunity*, vol. 46, no. 1, pp. 29–37, Jan. 2017, doi: 10.1016/j.immuni.2016.12.013.
[211] R. S. McClelland *et al.*, 'Evaluation of the association between the concentrations of key vaginal bacteria and the increased risk of HIV acquisition in African women from five cohorts: a nested case-control study', *Lancet Infect. Dis.*, vol. 18, no. 5, pp. 554–564, May 2018, doi: 10.1016/S1473-3099(18)30058-6.
[212] C. R. Cohen *et al.*, 'Bacterial vaginosis associated with increased risk of female-to-male HIV-1 transmission: a prospective cohort analysis among African couples', *PLoS Med.*, vol. 9, no. 6, p. e1001251, 2012, doi: 10.1371/journal.pmed.1001251.

[213] H. Borgdorff *et al.*, 'Lactobacillus-dominated cervicovaginal microbiota associated with reduced HIV/STI prevalence and genital HIV viral load in African women', *ISME J.*, vol. 8, no. 9, pp. 1781–1793, Sep. 2014, doi: 10.1038/ismej.2014.26.
[214] R. van Houdt, B. Ma, S. M. Bruisten, A. G. C. L. Speksnijder, J. Ravel, and H. J. C. de Vries, 'Lactobacillus iners-dominated vaginal microbiota is associated with increased susceptibility to Chlamydia trachomatis infection in Dutch women: a case–control study', *Sex. Transm. Infect.*, vol. 94, no. 2, pp. 117–123, Mar. 2018, doi: 10.1136/sextrans-2017-053133.
[215] J. E. Lee *et al.*, 'Association of the vaginal microbiota with human papillomavirus infection in a Korean twin cohort', *PloS One*, vol. 8, no. 5, p. e63514, 2013, doi: 10.1371/journal.pone.0063514.
[216] J. Norenhag, J. Du, M. Olovsson, H. Verstraelen, L. Engstrand, and N. Brusselaers, 'The vaginal microbiota, human papillomavirus and cervical dysplasia: a systematic review and network meta-analysis', *BJOG Int. J. Obstet. Gynaecol.*, vol. 127, no. 2, pp. 171–180, 2020, doi: 10.1111/1471-0528.15854.
[217] N. Brusselaers, S. Shrestha, J. van de Wijgert, and H. Verstraelen, 'Vaginal dysbiosis and the risk of human papillomavirus and cervical cancer: systematic review and meta-analysis', *Am. J. Obstet. Gynecol.*, vol. 221, no. 1, pp. 9-18.e8, Jul. 2019, doi: 10.1016/j.ajog.2018.12.011.
[218] X. Zhou *et al.*, 'The vaginal bacterial communities of Japanese women resemble those of women in other racial groups', *FEMS Immunol. Med. Microbiol.*, vol. 58, no. 2, pp. 169–181, Mar. 2010, doi: 10.1111/j.1574-695X.2009.00618.x.
[219] R. Martín, N. Soberón, F. Vázquez, and J. E. Suárez, 'La microbiota vaginal: composición, papel protector, patología asociada y perspectivas terapéuticas', *Enfermedades Infecc. Microbiol. Clínica*, vol. 26, no. 3, pp. 160–167, Mar. 2008, doi: 10.1157/13116753.
[220] C. Kenyon, R. Colebunders, and T. Crucitti, 'The global epidemiology of bacterial vaginosis: a systematic review', *Am. J. Obstet. Gynecol.*, vol. 209, no. 6, pp. 505–523, Dec. 2013, doi: 10.1016/j.ajog.2013.05.006.
[221] J. M. Fettweis *et al.*, 'Differences in vaginal microbiome in African American women versus women of European ancestry', *Microbiology*, vol. 160, no. Pt 10, pp. 2272–2282, Oct. 2014, doi: 10.1099/mic.0.081034-0.
[222] L. E. Janssen, R. J. T. Verduin, C. J. M. de Groot, M. A. Oudijk, and M. A. de Boer, 'The association between vaginal hygiene practices and spontaneous preterm birth: A case-control study', *PloS One*, vol. 17, no. 6, p. e0268248, 2022, doi: 10.1371/journal.pone.0268248.
[223] D. P. Misra and B. Trabert, 'Vaginal douching and risk of preterm birth among African American women', *Am. J. Obstet. Gynecol.*, vol. 196, no. 2, p. 140.e1-140.e8, Feb. 2007, doi: 10.1016/j.ajog.2006.10.880.
[224] A. Barrientos-Durán, A. Fuentes-López, A. de Salazar, J. Plaza-Díaz, and F. García, 'Reviewing the Composition of Vaginal Microbiota: Inclusion of Nutrition and Probiotic Factors in the Maintenance of Eubiosis', *Nutrients*, vol. 12, no. 2, Art. no. 2, Feb. 2020, doi: 10.3390/nu12020419.
[225] K. Gliniewicz *et al.*, 'Comparison of the Vaginal Microbiomes of Premenopausal and Postmenopausal Women', *Front. Microbiol.*, vol. 10, 2019, Accessed: May 22, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/fmicb.2019.00193
[226] T. Shiraishi *et al.*, 'Influence of menstruation on the microbiota of healthy women's labia minora as analyzed using a 16S rRNA gene-based clone library method', *Jpn. J. Infect. Dis.*, vol. 64, no. 1, pp. 76–80, 2011.

[227] C. M. Mitchell, D. N. Fredricks, R. L. Winer, and L. Koutsky, ‘Effect of Sexual Debut on Vaginal Microbiota in a Cohort of Young Women’, *Obstet. Gynecol.*, vol. 120, no. 6, pp. 1306–1313, Dec. 2012.
[228] R. M. Brotman *et al.*, ‘Association between cigarette smoking and the vaginal microbiota: a pilot study’, *BMC Infect. Dis.*, vol. 14, p. 471, Aug. 2014, doi: 10.1186/1471-2334-14-471.
[229] R. Romero *et al.*, ‘The vaginal microbiota of pregnant women who subsequently have spontaneous preterm labor and delivery and those with a normal delivery at term’, *Microbiome*, vol. 2, no. 1, p. 18, Dec. 2014, doi: 10.1186/2049-2618-2-18.
[230] D. B. DiGiulio *et al.*, ‘Temporal and spatial variation of the human microbiota during pregnancy’, *Proc. Natl. Acad. Sci.*, vol. 112, no. 35, pp. 11060–11065, Sep. 2015, doi: 10.1073/pnas.1502875112.
[231] M. J. Peelen *et al.*, ‘The influence of the vaginal microbiota on preterm birth: A systematic review and recommendations for a minimum dataset for future research’, *Placenta*, vol. 79, pp. 30–39, Apr. 2019, doi: 10.1016/j.placenta.2019.03.011.
[232] R. W. Hyman *et al.*, ‘Diversity of the vaginal microbiome correlates with preterm birth’, *Reprod. Sci. Thousand Oaks Calif*, vol. 21, no. 1, pp. 32–40, Jan. 2014, doi: 10.1177/1933719113488838.
[233] B. J. Callahan *et al.*, ‘Replication and refinement of a vaginal microbial signature of preterm birth in two racially distinct cohorts of US women’, *Proc. Natl. Acad. Sci. U. S. A.*, vol. 114, no. 37, pp. 9966–9971, Sep. 2017, doi: 10.1073/pnas.1705899114.
[234] M. J. Stout, Y. Zhou, K. M. Wylie, P. I. Tarr, G. A. Macones, and M. G. Tuuli, ‘Early pregnancy vaginal microbiome trends and preterm birth’, *Am. J. Obstet. Gynecol.*, vol. 217, no. 3, p. 356.e1-356.e18, Sep. 2017, doi: 10.1016/j.ajog.2017.05.030.
[235] R. G. Brown *et al.*, ‘Establishment of vaginal microbiota composition in early pregnancy and its association with subsequent preterm prelabor rupture of the fetal membranes’, *Transl. Res.*, vol. 207, pp. 30–43, May 2019, doi: 10.1016/j.trsl.2018.12.005.
[236] M. A. Elovitz *et al.*, ‘Cervicovaginal microbiota and local immune response modulate the risk of spontaneous preterm delivery’, *Nat. Commun.*, vol. 10, no. 1, p. 1305, Mar. 2019, doi: 10.1038/s41467-019-09285-9.
[237] A. C. Freitas *et al.*, ‘Increased richness and diversity of the vaginal microbiota and spontaneous preterm birth’, *Microbiome*, vol. 6, no. 1, p. 117, Jun. 2018, doi: 10.1186/s40168-018-0502-8.
[238] D. B. Nelson, H. Shin, J. Wu, and M. G. Dominguez-Bello, ‘The Gestational Vaginal Microbiome and Spontaneous Preterm Birth among Nulliparous African American Women’, *Am. J. Perinatol.*, vol. 33, no. 9, pp. 887–893, Jul. 2016, doi: 10.1055/s-0036-1581057.
[239] L. Petricevic *et al.*, ‘Characterisation of the vaginal Lactobacillus microbiota associated with preterm delivery’, *Sci. Rep.*, vol. 4, p. 5136, May 2014, doi: 10.1038/srep05136.
[240] L. M. Kindinger *et al.*, ‘The interaction between vaginal microbiota, cervical length, and vaginal progesterone treatment for preterm birth risk’, *Microbiome*, vol. 5, no. 1, p. 6, Jan. 2017, doi: 10.1186/s40168-016-0223-9.
[241] H. Verstraelen, R. Verhelst, G. Claeys, E. De Backer, M. Temmerman, and M. Vaneechoutte, ‘Longitudinal analysis of the vaginal microflora in pregnancy suggests that L. crispatus promotes the stability of the normal vaginal microflora and that L. gasseri and/or L. iners are more conducive to the occurrence of abnormal vaginal microflora’, *BMC Microbiol.*, vol. 9, p. 116, Jun. 2009, doi: 10.1186/1471-2180-9-116.

[242] E. Bayar, P. R. Bennett, D. Chan, L. Sykes, and D. A. MacIntyre, ‘The pregnancy microbiome and preterm birth’, *Semin. Immunopathol.*, vol. 42, no. 4, pp. 487–499, Aug. 2020, doi: 10.1007/s00281-020-00817-w.
[243] V. C. Allport, D. Pieber, D. M. Slater, R. Newton, J. O. White, and P. R. Bennett, ‘Human labour is associated with nuclear factor-kappaB activity which mediates cyclo-oxygenase-2 expression and is involved with the “functional progesterone withdrawal”’, *Mol. Hum. Reprod.*, vol. 7, no. 6, pp. 581–586, Jun. 2001, doi: 10.1093/molehr/7.6.581.
[244] M. Piersanti and S. J. Lye, ‘Increase in messenger ribonucleic acid encoding the myometrial gap junction protein, connexin-43, requires protein synthesis and is associated with increased expression of the activator protein-1, c-fos’, *Endocrinology*, vol. 136, no. 8, pp. 3571–3578, Aug. 1995, doi: 10.1210/endo.136.8.7628395.
[245] R. Osmers *et al.*, ‘Origin of cervical collagenase during parturition’, *Am. J. Obstet. Gynecol.*, vol. 166, no. 5, pp. 1455–1460, May 1992, doi: 10.1016/0002-9378(92)91619-l.
[246] S. Khanjani, V. Terzidou, M. R. Johnson, and P. R. Bennett, ‘NFκB and AP-1 Drive Human Myometrial IL8 Expression’, *Mediators Inflamm.*, vol. 2012, p. 504952, 2012, doi: 10.1155/2012/504952.
[247] A. R. Belt, J. J. Baldassare, M. Molnár, R. Romero, and F. Hertelendy, ‘The nuclear transcription factor NF-kappaB mediates interleukin-1beta-induced expression of cyclooxygenase-2 in human myometrial cells’, *Am. J. Obstet. Gynecol.*, vol. 181, no. 2, pp. 359–366, Aug. 1999, doi: 10.1016/s0002-9378(99)70562-4.
[248] D. A. Kniss, B. Rovin, R. H. Fertel, and P. D. Zimmerman, ‘Blockade NF-kappaB activation prohibits TNF-alpha-induced cyclooxygenase-2 gene expression in ED27 trophoblast-like cells’, *Placenta*, vol. 22, no. 1, pp. 80–89, Jan. 2001, doi: 10.1053/plac.2000.0591.
[249] J. M. Gonzalez, C.-W. Franzke, F. Yang, R. Romero, and G. Girardi, ‘Complement activation triggers metalloproteinases release inducing cervical remodeling and preterm birth in mice’, *Am. J. Pathol.*, vol. 179, no. 2, pp. 838–849, Aug. 2011, doi: 10.1016/j.ajpath.2011.04.024.
[250] A. C. Hearps *et al.*, ‘Vaginal lactic acid elicits an anti-inflammatory response from human cervicovaginal epithelial cells and inhibits production of pro-inflammatory mediators associated with HIV acquisition’, *Mucosal Immunol.*, vol. 10, no. 6, pp. 1480–1490, Nov. 2017, doi: 10.1038/mi.2017.27.
[251] A. Hearps *et al.*, ‘Lactic Acid, a Vaginal Microbiota Metabolite, Elicits an Anti-inflammatory Response from Vaginal and Cervical Epithelial Cells’, *AIDS Res. Hum. Retroviruses*, vol. 30, no. S1, pp. A238–A239, Oct. 2014, doi: 10.1089/aid.2014.5527.abstract.
[252] S. S. Witkin, H. Mendes-Soares, I. M. Linhares, A. Jayaram, W. J. Ledger, and L. J. Forney, ‘Influence of vaginal bacteria and D- and L-lactic acid isomers on vaginal extracellular matrix metalloproteinase inducer: implications for protection against upper genital tract infections’, *mBio*, vol. 4, no. 4, pp. e00460-13, Aug. 2013, doi: 10.1128/mBio.00460-13.
[253] L. Rahkonen *et al.*, ‘Factors affecting matrix metalloproteinase-8 levels in the vaginal and cervical fluids in the first and second trimester of pregnancy’, *Hum. Reprod. Oxf. Engl.*, vol. 24, no. 11, pp. 2693–2702, Nov. 2009, doi: 10.1093/humrep/dep284.
[254] M. J. Ferris, A. Masztal, K. E. Aldridge, J. D. Fortenberry, P. L. Fidel, and D. H. Martin, ‘Association of Atopobium vaginae, a recently described metronidazole resistant anaerobe, with bacterial vaginosis’, *BMC Infect. Dis.*, vol. 4, p. 5, Feb. 2004, doi: 10.1186/1471-2334-4-5.

[255] C. R. Cohen *et al.*, ‘Randomized Trial of Lactin-V to Prevent Recurrence of Bacterial Vaginosis’, *N. Engl. J. Med.*, vol. 382, no. 20, pp. 1906–1915, May 2020, doi: 10.1056/NEJMoa1915254.
[256] R. Russo, E. Karadja, and F. De Seta, ‘Evidence-based mixture containing Lactobacillus strains and lactoferrin to prevent recurrent bacterial vaginosis: a double blind, placebo controlled, randomised clinical trial’, *Benef. Microbes*, vol. 10, no. 1, pp. 19–26, Feb. 2019, doi: 10.3920/BM2018.0075.
[257] G. R. Gibson and M. B. Roberfroid, ‘Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics’, *J. Nutr.*, vol. 125, no. 6, pp. 1401–1412, Jun. 1995, doi: 10.1093/jn/125.6.1401.
[258] S. Yang, G. Reid, J. R. G. Challis, S. O. Kim, G. B. Gloor, and A. D. Bocking, ‘Is There a Role for Probiotics in the Prevention of Preterm Birth?’, *Front. Immunol.*, vol. 6, 2015, Accessed: May 20, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/fimmu.2015.00062
[259] M. France *et al.*, ‘Insight into the ecology of vaginal bacteria through integrative analyses of metagenomic and metatranscriptomic data’, *Genome Biol.*, vol. 23, Mar. 2022, doi: 10.1186/s13059-022-02635-9.
[260] V. Rousseau, J. P. Lepargneur, C. Roques, M. Remaud-Simeon, and F. Paul, ‘Prebiotic effects of oligosaccharides on selected vaginal lactobacilli and pathogenic microorganisms’, *Anaerobe*, vol. 11, no. 3, pp. 145–153, Jun. 2005, doi: 10.1016/j.anaerobe.2004.12.002.
[261] J. Grimoud *et al.*, ‘In vitro screening of probiotic lactic acid bacteria and prebiotic glucooligosaccharides to select effective synbiotics’, *Anaerobe*, vol. 16, no. 5, pp. 493–500, Oct. 2010, doi: 10.1016/j.anaerobe.2010.07.005.
[262] Food and Agriculture Organization of the United Nations and World Health Organization, Eds., *Probiotics in food: health and nutritional properties and guidelines for evaluation*. in FAO food and nutrition paper, no. 85. Rome: Food and Agriculture Organization of the United Nations : World Health Organization, 2006.
[263] M. Cordaillat-Simmons, A. Rouanet, and B. Pot, ‘Live biotherapeutic products: the importance of a defined regulatory framework’, *Exp. Mol. Med.*, vol. 52, no. 9, Art. no. 9, Sep. 2020, doi: 10.1038/s12276-020-0437-6.
[264] ‘Early Clinical Trials with Live Biotherapeutic Products: Chemistry, Manufacturing, and Control Information; Guidance for Industry’.
[265] A.-C. Chen, M.-Y. Chung, J. H. Chang, and H.-C. Lin, ‘Pathogenesis Implication for Necrotizing Enterocolitis Prevention in Preterm Very-Low-Birth-Weight Infants’, *J. Pediatr. Gastroenterol. Nutr.*, vol. 58, no. 1, pp. 7–11, Jan. 2014, doi: 10.1097/MPG.0b013e3182a7dc74.
[266] J. Z. Goldenberg *et al.*, ‘Probiotics for the prevention of Clostridium difficile-associated diarrhea in adults and children’, *Cochrane Database Syst. Rev.*, no. 5, 2013, doi: 10.1002/14651858.CD006095.pub3.
[267] Y. Yang, Y. Guo, Q. Kan, X. G. Zhou, X. Y. Zhou, and Y. Li, ‘A meta-analysis of probiotics for preventing necrotizing enterocolitis in preterm neonates’, *Braz. J. Med. Biol. Res.*, vol. 47, pp. 804–810, Sep. 2014, doi: 10.1590/1414-431X20143857.
[268] P. M. Grin, P. M. Kowalewska, W. Alhazzan, and A. E. Fox-Robichaud, ‘Lactobacillus for preventing recurrent urinary tract infections in women: meta-analysis’, *Can. J. Urol.*, vol. 20, no. 1, pp. 6607–6614, Feb. 2013.
[269] N. Kirihara, M. Kamitomo, T. Tabira, T. Hashimoto, H. Taniguchi, and T. Maeda, ‘Effect of probiotics on perinatal outcome in patients at high risk of preterm birth’, *J. Obstet. Gynaecol. Res.*, vol. 44, no. 2, pp. 241–247, 2018, doi: 10.1111/jog.13497.

[270] A. Jarde *et al.*, 'Pregnancy outcomes in women taking probiotics or prebiotics: a systematic review and meta-analysis', *BMC Pregnancy Childbirth*, vol. 18, no. 1, p. 14, Jan. 2018, doi: 10.1186/s12884-017-1629-5.
[271] A. Homayouni *et al.*, 'Effects of Probiotics on the Recurrence of Bacterial Vaginosis: A Review', *J. Low. Genit. Tract Dis.*, vol. 18, no. 1, pp. 79–86, Jan. 2014, doi: 10.1097/LGT.0b013e31829156ec.
[272] G. Reid, D. Beuerman, C. Heinemann, and A. W. Bruce, 'Probiotic Lactobacillus dose required to restore and maintain a normal vaginal flora', *FEMS Immunol. Med. Microbiol.*, vol. 32, no. 1, pp. 37–41, Dec. 2001, doi: 10.1111/j.1574-695X.2001.tb00531.x.
[273] P. Mirmonsef *et al.*, 'Free Glycogen in Vaginal Fluids Is Associated with Lactobacillus Colonization and Low Vaginal pH', *PLOS ONE*, vol. 9, no. 7, p. e102467, Jul. 2014, doi: 10.1371/journal.pone.0102467.
[274] G. Reid *et al.*, 'Oral use of Lactobacillus rhamnosus GR-1 and L. fermentum RC-14 significantly alters vaginal flora: randomized, placebo-controlled trial in 64 healthy women', *FEMS Immunol. Med. Microbiol.*, vol. 35, no. 2, pp. 131–134, Mar. 2003, doi: 10.1016/S0928-8244(02)00465-0.
[275] K. Anukam *et al.*, 'Augmentation of antimicrobial metronidazole therapy of bacterial vaginosis with oral probiotic Lactobacillus rhamnosus GR-1 and Lactobacillus reuteri RC-14: randomized, double-blind, placebo controlled trial', *Microbes Infect.*, vol. 8, no. 6, pp. 1450–1454, May 2006, doi: 10.1016/j.micinf.2006.01.003.
[276] A. López-Moreno and M. Aguilera, 'Vaginal Probiotics for Reproductive Health and Related Dysbiosis: Systematic Review and Meta-Analysis', *J. Clin. Med.*, vol. 10, no. 7, p. 1461, Apr. 2021, doi: 10.3390/jcm10071461.
[277] K. C. Anukam, E. Osazuwa, G. I. Osemene, F. Ehigiagbe, A. W. Bruce, and G. Reid, 'Clinical study comparing probiotic Lactobacillus GR-1 and RC-14 with metronidazole vaginal gel to treat symptomatic bacterial vaginosis', *Microbes Infect.*, vol. 8, no. 12–13, pp. 2772–2776, Oct. 2006, doi: 10.1016/j.micinf.2006.08.008.
[278] P. Mastromarino *et al.*, 'Effectiveness of Lactobacillus-containing vaginal tablets in the treatment of symptomatic bacterial vaginosis', *Clin. Microbiol. Infect. Off. Publ. Eur. Soc. Clin. Microbiol. Infect. Dis.*, vol. 15, no. 1, pp. 67–74, Jan. 2009, doi: 10.1111/j.1469-0691.2008.02112.x.
[279] D. Parent *et al.*, 'Therapy of bacterial vaginosis using exogenously-applied Lactobacilli acidophili and a low dose of estriol: a placebo-controlled multicentric clinical trial', *Arzneimittelforschung.*, vol. 46, no. 1, pp. 68–73, Jan. 1996.
[280] A. Hallén, C. Jarstrand, and C. Påhlson, 'Treatment of bacterial vaginosis with lactobacilli', *Sex. Transm. Dis.*, vol. 19, no. 3, pp. 146–148, Jun. 1992, doi: 10.1097/00007435-199205000-00007.
[281] F. Vicariotto, L. Mogna, and M. Del Piano, 'Effectiveness of the two microorganisms Lactobacillus fermentum LF15 and Lactobacillus plantarum LP01, formulated in slow-release vaginal tablets, in women affected by bacterial vaginosis: a pilot study', *J. Clin. Gastroenterol.*, vol. 48 Suppl 1, pp. S106-112, Dec. 2014, doi: 10.1097/MCG.0000000000000226.
[282] Imperial College London, 'Feasibility, Safety, Tolerance and Assessment of LACTIN-V Use in a Cohort of Pregnant Women at Risk of Preterm Birth', clinicaltrials.gov, Clinical trial registration NCT03992534, May 2020. Accessed: May 25, 2022. [Online]. Available: https://clinicaltrials.gov/ct2/show/NCT03992534
[283] M. Sharma *et al.*, 'An Insight into Vaginal Microbiome Techniques', *Life*, vol. 11, no. 11, p. 1229, Nov. 2021, doi: 10.3390/life11111229.

[284] C. Colonna and M. Steelman, ‘Amsel Criteria’, in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2022. Accessed: Apr. 07, 2022. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK542319/
[285] R. Amsel, P. A. Totten, C. A. Spiegel, K. C. Chen, D. Eschenbach, and K. K. Holmes, ‘Nonspecific vaginitis. Diagnostic criteria and microbial and epidemiologic associations’, *Am. J. Med.*, vol. 74, no. 1, pp. 14–22, Jan. 1983, doi: 10.1016/0002-9343(83)91112-9.
[286] N. Tripathi and A. Sapra, ‘Gram Staining’, in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2022. Accessed: Apr. 07, 2022. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK562156/
[287] R. P. Nugent, M. A. Krohn, and S. L. Hillier, ‘Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation’, *J. Clin. Microbiol.*, vol. 29, no. 2, pp. 297–301, Feb. 1991, doi: 10.1128/jcm.29.2.297-301.1991.
[288] R. Bhujel, S. K. Mishra, S. K. Yadav, K. D. Bista, and K. Parajuli, ‘Comparative study of Amsel’s criteria and Nugent scoring for diagnosis of bacterial vaginosis in a tertiary care hospital, Nepal’, *BMC Infect. Dis.*, vol. 21, no. 1, p. 825, Aug. 2021, doi: 10.1186/s12879-021-06562-1.
[289] J. A. Simoes, M. G. Discacciati, E. M. Brolazo, P. M. Portugal, D. V. Dini, and M. C. M. Dantas, ‘Clinical diagnosis of bacterial vaginosis’, *Int. J. Gynaecol. Obstet. Off. Organ Int. Fed. Gynaecol. Obstet.*, vol. 94, no. 1, pp. 28–32, Jul. 2006, doi: 10.1016/j.ijgo.2006.04.013.
[290] E. A. Wolf *et al.*, ‘Culturomics Approaches Expand the Diagnostic Accuracy for Sexually Transmitted Infections’, *Int. J. Mol. Sci.*, vol. 22, no. 19, p. 10815, Oct. 2021, doi: 10.3390/ijms221910815.
[291] P. C. Y. Woo, S. K. P. Lau, J. L. L. Teng, H. Tse, and K.-Y. Yuen, ‘Then and now: use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories’, *Clin. Microbiol. Infect.*, vol. 14, no. 10, pp. 908–934, Oct. 2008, doi: 10.1111/j.1469-0691.2008.02070.x.
[292] J. M. Janda and S. L. Abbott, ‘16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls’, *J. Clin. Microbiol.*, vol. 45, no. 9, pp. 2761–2764, Sep. 2007, doi: 10.1128/JCM.01228-07.
[293] J. Patel, ‘16S rRNA gene sequencing for bacterial pathogen identification in the clinical laboratory’, *Mol. Diagn.*, vol. 6, no. 4, pp. 313–321, Dec. 2001, doi: 10.1054/modi.2001.29158.
[294] Y. Van de Peer, S. Chapelle, and R. De Wachter, ‘A quantitative map of nucleotide substitution rates in bacterial rRNA’, *Nucleic Acids Res.*, vol. 24, no. 17, pp. 3381–3391, Sep. 1996, doi: 10.1093/nar/24.17.3381.
[295] J. A. Frank, C. I. Reich, S. Sharma, J. S. Weisbaum, B. A. Wilson, and G. J. Olsen, ‘Critical Evaluation of Two Primers Commonly Used for Amplification of Bacterial 16S rRNA Genes’, *Appl. Environ. Microbiol.*, vol. 74, no. 8, pp. 2461–2470, Apr. 2008, doi: 10.1128/AEM.02272-07.
[296] A. Schmalenberger, F. Schwieger, and C. C. Tebbe, ‘Effect of primers hybridizing to different evolutionarily conserved regions of the small-subunit rRNA gene in PCR-based microbial community analyses and genetic profiling’, *Appl. Environ. Microbiol.*, vol. 67, no. 8, pp. 3557–3563, Aug. 2001, doi: 10.1128/AEM.67.8.3557-3563.2001.
[297] M. Tärnberg, T. Jakobsson, J. Jonasson, and U. Forsum, ‘Identification of randomly selected colonies of lactobacilli from normal vaginal fluid by pyrosequencing of the 16S rDNA variable V1 and V3 regions’, *APMIS*, vol. 110, no. 11, pp. 802–810, Nov. 2002, doi: 10.1034/j.1600-0463.2002.1101106.x.

[298] R. Hummelen *et al.*, 'Deep Sequencing of the Vaginal Microbiota of Women with HIV', *PLOS ONE*, vol. 5, no. 8, p. e12078, Aug. 2010, doi: 10.1371/journal.pone.0012078.
[299] J. D. Silverman *et al.*, 'Measuring and mitigating PCR bias in microbiota datasets', *PLoS Comput. Biol.*, vol. 17, no. 7, p. e1009113, Jul. 2021, doi: 10.1371/journal.pcbi.1009113.
[300] G. D. Correia, J. R. Marchesi, and D. A. MacIntyre, 'Moving beyond DNA: towards functional analysis of the vaginal microbiome by non-sequencing-based methods', *Curr. Opin. Microbiol.*, vol. 73, p. 102292, Jun. 2023, doi: 10.1016/j.mib.2023.102292.
[301] S. K. Hilton *et al.*, 'Metataxonomic and Metagenomic Approaches vs. Culture-Based Techniques for Clinical Pathology', *Front. Microbiol.*, vol. 7, 2016, Accessed: Jun. 01, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fmicb.2016.00484
[302] F. A. Baltoumas *et al.*, 'Exploring microbial functional biodiversity at the protein family level—From metagenomic sequence reads to annotated protein clusters', *Front. Bioinforma.*, vol. 3, 2023, Accessed: Jun. 01, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fbinf.2023.1157956
[303] P. Pruski *et al.*, 'Assessment of microbiota:host interactions at the vaginal mucosa interface', *Methods San Diego Calif*, vol. 149, pp. 74–84, Oct. 2018, doi: 10.1016/j.ymeth.2018.04.022.
[304] C. Ceccarani *et al.*, 'Diversity of vaginal microbiome and metabolome during genital infections', *Sci. Rep.*, vol. 9, no. 1, p. 14095, Oct. 2019, doi: 10.1038/s41598-019-50410-x.
[305] M. Fu, X. Zhang, Y. Liang, S. Lin, W. Qian, and S. Fan, 'Alterations in Vaginal Microbiota and Associated Metabolome in Women with Recurrent Implantation Failure', *mBio*, vol. 11, no. 3, pp. e03242-19, Jun. 2020, doi: 10.1128/mBio.03242-19.
[306] G. R. Huggins and G. Preti, 'Volatile constituents of human vaginal secretions', *Am. J. Obstet. Gynecol.*, vol. 126, no. 1, pp. 129–136, Sep. 1976, doi: 10.1016/0002-9378(76)90477-4.
[307] L. L. Klein, K. R. Jonscher, M. J. Heerwagen, R. S. Gibbs, and J. L. McManaman, 'Shotgun proteomic analysis of vaginal fluid from women in late pregnancy', *Reprod. Sci. Thousand Oaks Calif*, vol. 15, no. 3, pp. 263–273, Apr. 2008, doi: 10.1177/1933719107311189.
[308] J. C. M. Tsibris, 'Cervical Mucus', in *Gynecologic Endocrinology*, J. J. Gold and J. B. Josimovich, Eds., Boston, MA: Springer US, 1987, pp. 175–183. doi: 10.1007/978-1-4613-2157-6_10.
[309] P. D. Thomas *et al.*, 'PANTHER: a library of protein families and subfamilies indexed by function', *Genome Res.*, vol. 13, no. 9, pp. 2129–2141, Sep. 2003, doi: 10.1101/gr.772403.
[310] D. M. Good *et al.*, 'Body fluid proteomics for biomarker discovery: lessons from the past hold the key to success in the future', *J. Proteome Res.*, vol. 6, no. 12, pp. 4549–4555, Dec. 2007, doi: 10.1021/pr070529w.
[311] S. Hu, J. A. Loo, and D. T. Wong, 'Human body fluid proteome analysis', *Proteomics*, vol. 6, no. 23, pp. 6326–6353, Dec. 2006, doi: 10.1002/pmic.200600284.
[312] A. Schmidt and R. Aebersold, 'High-accuracy proteome maps of human body fluids', *Genome Biol.*, vol. 7, no. 11, p. 242, Nov. 2006, doi: 10.1186/gb-2006-7-11-242.
[313] G. Zegels, G. A. A. Van Raemdonck, E. P. Coen, W. A. A. Tjalma, and X. W. M. Van Ostade, 'Comprehensive proteomic analysis of human cervical-vaginal fluid using

colposcopy samples', *Proteome Sci.*, vol. 7, p. 17, Apr. 2009, doi: 10.1186/1477-5956-7-17.
[314] T. M. Nelson *et al.*, 'Cigarette smoking is associated with an altered vaginal tract metabolomic profile', *Sci. Rep.*, vol. 8, no. 1, Art. no. 1, Jan. 2018, doi: 10.1038/s41598-017-14943-3.
[315] R. Cruickshank and A. Sharman, 'The Biology of the Vagina in the Human Subject', *BJOG Int. J. Obstet. Gynaecol.*, vol. 41, no. 2, pp. 208–226, Apr. 1934, doi: 10.1111/j.1471-0528.1934.tb08759.x.
[316] D. P. Wolf, J. E. Sokoloski, and M. Litt, 'Composition and functioin of human cervical mucus', *Biochim. Biophys. Acta*, vol. 630, no. 4, pp. 545–558, Jul. 1980, doi: 10.1016/0304-4165(80)90009-4.
[317] A. T. Gregoire, O. Kandil, and W. J. Ledger, 'The glycogen content of human vaginal epithelial tissue', *Fertil. Steril.*, vol. 22, no. 1, pp. 64–68, Jan. 1971, doi: 10.1016/s0015-0282(16)37989-4.
[318] W. B. Ayre, 'THE GLYCOGEN-ESTROGEN RELATIONSHIP IN THE VAGINAL TRACT', *J. Clin. Endocrinol. Metab.*, vol. 11, no. 1, pp. 103–110, Jan. 1951, doi: 10.1210/jcem-11-1-103.
[319] R. Tester and F. H. Al-Ghazzewi, 'Intrinsic and extrinsic carbohydrates in the vagina: A short review on vaginal glycogen', *Int. J. Biol. Macromol.*, vol. 112, pp. 203–206, Jun. 2018, doi: 10.1016/j.ijbiomac.2018.01.166.
[320] H. K. Haefner and C. P. Crum, 'Benign Conditions of the Vagina', in *Diagnostic Gynecologic and Obstetric Pathology*, Elsevier, 2018, pp. 258–274. doi: 10.1016/B978-0-323-44732-4.00011-X.
[321] D. J. Anderson, J. Marathe, and J. Pudney, 'The Structure of the Human Vaginal Stratum Corneum and its Role in Immune Defense', *Am. J. Reprod. Immunol.*, vol. 71, no. 6, pp. 618–623, 2014, doi: 10.1111/aji.12230.
[322] P. Mirmonsef *et al.*, 'Glycogen Levels in Undiluted Genital Fluid and Their Relationship to Vaginal pH, Estrogen, and Progesterone', *PLOS ONE*, vol. 11, no. 4, p. e0153553, Apr. 2016, doi: 10.1371/journal.pone.0153553.
[323] G. T. Spear *et al.*, 'Human α-amylase Present in Lower-Genital- Tract Mucosal Fluid Processes Glycogen to Support Vaginal Colonization by Lactobacillus', p. 10.
[324] S. Baldewijns, M. Sillen, I. Palmans, P. Vandecruys, P. Van Dijck, and L. Demuyser, 'The Role of Fatty Acid Metabolites in Vaginal Health and Disease: Application to Candidiasis', *Front. Microbiol.*, vol. 12, 2021, Accessed: Apr. 08, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/fmicb.2021.705779
[325] J. M. Macklaim, A. D. Fernandes, J. M. Di Bella, J.-A. Hammond, G. Reid, and G. B. Gloor, 'Comparative meta-RNA-seq of the vaginal microbiota and differential expression by Lactobacillus iners in health and dysbiosis', *Microbiome*, vol. 1, no. 1, p. 12, Apr. 2013, doi: 10.1186/2049-2618-1-12.
[326] G. Petrou and T. Crouzier, 'Mucins as multifunctional building blocks of biomaterials', *Biomater. Sci.*, vol. 6, no. 9, pp. 2282–2297, 2018, doi: 10.1039/C8BM00471D.
[327] S. Srinivasan *et al.*, 'Metabolic Signatures of Bacterial Vaginosis', *mBio*, vol. 6, no. 2, pp. e00204-15, doi: 10.1128/mBio.00204-15.
[328] C. Noecker *et al.*, 'Metabolic Model-Based Integration of Microbiome Taxonomic and Metabolomic Profiles Elucidates Mechanistic Links between Ecological and Metabolic Variation', *mSystems*, vol. 1, no. 1, pp. e00013-15, doi: 10.1128/mSystems.00013-15.
[329] X. Zhou, S. J. Bent, M. G. Schneider, C. C. Davis, M. R. Islam, and L. J. Y. 2004 Forney, 'Characterization of vaginal microbial communities in adult healthy women

using cultivation-independent methods', *Microbiology*, vol. 150, no. 8, pp. 2565–2573, doi: 10.1099/mic.0.26905-0.
[330] M. A. R. Vinolo, H. G. Rodrigues, R. T. Nachbar, and R. Curi, '17 - Modulation of inflammatory and immune responses by short-chain fatty acids', in *Diet, Immunity and Inflammation*, P. C. Calder and P. Yaqoob, Eds., in Woodhead Publishing Series in Food Science, Technology and Nutrition. Woodhead Publishing, 2013, pp. 435–458. doi: 10.1533/9780857095749.3.435.
[331] C. J. Yeoman *et al.*, 'A Multi-Omic Systems-Based Approach Reveals Metabolic Markers of Bacterial Vaginosis and Insight into the Disease', *PLOS ONE*, vol. 8, no. 2, p. e56111, Feb. 2013, doi: 10.1371/journal.pone.0056111.
[332] M. A. Krohn, S. L. Hillier, and D. A. Eschenbach, 'Comparison of methods for diagnosing bacterial vaginosis among pregnant women.', *J. Clin. Microbiol.*, vol. 27, no. 6, pp. 1266–1271, Jun. 1989.
[333] P. Piot, E. Van Dyck, P. Godts, and J. Vanderheyden, 'The vaginal microbial flora in non-specific vaginitis', *Eur. J. Clin. Microbiol.*, vol. 1, no. 5, pp. 301–306, Oct. 1982, doi: 10.1007/BF02019976.
[334] C. A. Spiegel, R. Amsel, D. Eschenbach, F. Schoenknecht, and K. K. Holmes, 'Anaerobic bacteria in nonspecific vaginitis', *N. Engl. J. Med.*, vol. 303, no. 11, pp. 601–607, Sep. 1980, doi: 10.1056/NEJM198009113031102.
[335] R. Stanek, R. E. Gain, D. D. Glover, and B. Larsen, 'High performance ion exclusion chromatographic characterization of the vaginal organic acids in women with bacterial vaginosis', *Biomed. Chromatogr. BMC*, vol. 6, no. 5, pp. 231–235, Oct. 1992, doi: 10.1002/bmc.1130060506.
[336] H. Wolrath, U. Forsum, P. G. Larsson, and H. Borén, 'Analysis of bacterial vaginosis-related amines in vaginal fluid by gas chromatography and mass spectrometry', *J. Clin. Microbiol.*, vol. 39, no. 11, pp. 4026–4031, Nov. 2001, doi: 10.1128/JCM.39.11.4026-4031.2001.
[337] J. Downes, M. Munson, and W. G. Wade, 'Dialister invisus sp. nov., isolated from the human oral cavity', *Int. J. Syst. Evol. Microbiol.*, vol. 53, no. Pt 6, pp. 1937–1940, Nov. 2003, doi: 10.1099/ijs.0.02640-0.
[338] A. N. Chaudry *et al.*, 'Analysis of Vaginal Acetic Acid in Patients Undergoing Treatment for Bacterial Vaginosis', *J. Clin. Microbiol.*, vol. 42, no. 11, pp. 5170–5175, Nov. 2004, doi: 10.1128/JCM.42.11.5170-5175.2004.
[339] M. Aldunate *et al.*, 'Antimicrobial and immune modulatory effects of lactic acid and short chain fatty acids produced by vaginal microbiota associated with eubiosis and bacterial vaginosis', *Front. Physiol.*, vol. 6, 2015, Accessed: Apr. 20, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/fphys.2015.00164
[340] D. J. Delgado-Diaz, D. Tyssen, J. A. Hayward, R. Gugasyan, A. C. Hearps, and G. Tachedjian, 'Distinct Immune Responses Elicited From Cervicovaginal Epithelial Cells by Lactic Acid and Short Chain Fatty Acids Associated With Optimal and Non-optimal Vaginal Microbiota', *Front. Cell. Infect. Microbiol.*, vol. 9, 2020, Accessed: May 22, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/fcimb.2019.00446
[341] J. M. Sanches *et al.*, 'Vaginal lipidomics of women with vulvovaginal candidiasis and cytolytic vaginosis: A non-targeted LC-MS pilot study', *PloS One*, vol. 13, no. 8, p. e0202401, 2018, doi: 10.1371/journal.pone.0202401.
[342] T. M. Nelson, J.-L. C. Borgogna, R. M. Brotman, J. Ravel, S. T. Walk, and C. J. Yeoman, 'Vaginal biogenic amines: biomarkers of bacterial vaginosis or precursors to vaginal dysbiosis?', *Front. Physiol.*, vol. 6, 2015, Accessed: Apr. 28, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/fphys.2015.00253

[343] P. Shah and E. Swiatlo, ‘A multifaceted role for polyamines in bacterial pathogens’, *Mol. Microbiol.*, vol. 68, no. 1, pp. 4–16, 2008, doi: 10.1111/j.1365-2958.2008.06126.x.
[344] A. E. Pegg, ‘Mammalian polyamine metabolism and function’, *IUBMB Life*, vol. 61, no. 9, pp. 880–894, Sep. 2009, doi: 10.1002/iub.230.
[345] D. L. Cruden and R. P. Galask, ‘Reduction of Trimethylamine Oxide to Trimethylamine by Mobiluncus Strains Isolated from Patients with Bacterial Vaginosis’, *Microb. Ecol. Health Dis.*, vol. 1, no. 2, pp. 95–100, Jan. 1988, doi: 10.3109/08910608809140187.
[346] H. Wolrath, H. Borén, A. Hallén, and U. Forsum, ‘Trimethylamine content in vaginal secretion and its relation to bacterial vaginosis’, *APMIS Acta Pathol. Microbiol. Immunol. Scand.*, vol. 110, no. 11, pp. 819–824, Nov. 2002, doi: 10.1034/j.1600-0463.2002.1101108.x.
[347] A. McMillan *et al.*, ‘A multi-platform metabolomics approach identifies highly specific biomarkers of bacterial diversity in the vagina of pregnant and non-pregnant women’, *Sci. Rep.*, vol. 5, no. 1, Art. no. 1, Sep. 2015, doi: 10.1038/srep14174.
[348] B. Vitali, F. Cruciani, G. Picone, C. Parolin, G. Donders, and L. Laghi, ‘Vaginal microbiome and metabolome highlight specific signatures of bacterial vaginosis’, *Eur. J. Clin. Microbiol. Infect. Dis.*, vol. 34, no. 12, pp. 2367–2376, Dec. 2015, doi: 10.1007/s10096-015-2490-y.
[349] K. Dettmer, P. A. Aronov, and B. D. Hammock, ‘MASS SPECTROMETRY-BASED METABOLOMICS’, *Mass Spectrom. Rev.*, vol. 26, no. 1, pp. 51–78, 2007, doi: 10.1002/mas.20108.
[350] D. S. Wishart *et al.*, ‘HMDB 4.0: the human metabolome database for 2018’, *Nucleic Acids Res.*, vol. 46, no. Database issue, pp. D608–D617, Jan. 2018, doi: 10.1093/nar/gkx1089.
[351] J. K. Nicholson and I. D. Wilson, ‘Opinion: understanding “global” systems biology: metabonomics and the continuum of metabolism’, *Nat. Rev. Drug Discov.*, vol. 2, no. 8, pp. 668–676, Aug. 2003, doi: 10.1038/nrd1157.
[352] G. G. Harrigan *et al.*, ‘Application of high-throughput Fourier-transform infrared spectroscopy in toxicology studies: contribution to a study on the development of an animal model for idiosyncratic toxicity’, *Toxicol. Lett.*, vol. 146, no. 3, pp. 197–205, Feb. 2004, doi: 10.1016/j.toxlet.2003.09.011.
[353] H. E. Johnson, D. Broadhurst, D. B. Kell, M. K. Theodorou, R. J. Merry, and G. W. Griffith, ‘High-throughput metabolic fingerprinting of legume silage fermentations via Fourier transform infrared spectroscopy and chemometrics’, *Appl. Environ. Microbiol.*, vol. 70, no. 3, pp. 1583–1592, Mar. 2004, doi: 10.1128/AEM.70.3.1583-1592.2004.
[354] J. H. F. Bothwell and J. L. Griffin, ‘An introduction to biological nuclear magnetic resonance spectroscopy’, *Biol. Rev.*, vol. 86, no. 2, pp. 493–510, May 2011, doi: 10.1111/j.1469-185X.2010.00157.x.
[355] N. V. Reo, ‘Nmr-Based Metabolomics’, *Drug Chem. Toxicol.*, vol. 25, no. 4, pp. 375–382, Jan. 2002, doi: 10.1081/DCT-120014789.
[356] A. Bauermeister, ‘Mass spectrometry-based metabolomics in microbiome investigations’, p. 18.
[357] E. de Hoffmann and V. Stroobant, *Mass spectrometry: principles and applications*, 3rd ed. Chichester, West Sussex, England ; Hoboken, NJ: J. Wiley, 2007.
[358] Z.-G. Gong, J. Hu, X. Wu, and Y.-J. Xu, ‘The Recent Developments in Sample Preparation for Mass Spectrometry-Based Metabolomics’, *Crit. Rev. Anal. Chem.*, vol. 47, no. 4, pp. 325–331, Jul. 2017, doi: 10.1080/10408347.2017.1289836.

[359] D. Volmer and L. L. Jessome, ‘Ion Suppression: A Major Concern in Mass Spectrometry’, *LCGC N. Am.*, vol. 24, no. 5, pp. 498–510, May 2006.
[360] W. Zhou, S. Yang, and P. G. Wang, ‘Matrix effects and application of matrix effect factor’, *Bioanalysis*, vol. 9, no. 23, pp. 1839–1844, Dec. 2017, doi: 10.4155/bio-2017-0214.
[361] S. K. Grebe and R. J. Singh, ‘LC-MS/MS in the Clinical Laboratory - Where to From Here?’, *Clin. Biochem. Rev.*, vol. 32, no. 1, pp. 5–31, Feb. 2011.
[362] E. Watson and G. Reid, ‘Metabolomics as a clinical testing method for the diagnosis of vaginal dysbiosis’, *Am. J. Reprod. Immunol.*, vol. 80, no. 2, p. e12979, 2018, doi: 10.1111/aji.12979.
[363] A. Agin *et al.*, ‘Metabolomics – an overview. From basic principles to potential biomarkers (part 1)’, *Médecine Nucl.*, vol. 40, no. 1, pp. 4–10, Feb. 2016, doi: 10.1016/j.mednuc.2015.12.006.
[364] D. J. Beale *et al.*, ‘Review of recent developments in GC–MS approaches to metabolomics-based research’, *Metabolomics*, vol. 14, no. 11, p. 152, Nov. 2018, doi: 10.1007/s11306-018-1449-2.
[365] C. Bhardwaj and L. Hanley, ‘Ion sources for mass spectrometric identification and imaging of molecular species’, *Nat Prod Rep*, vol. 31, no. 6, pp. 756–767, 2014, doi: 10.1039/C3NP70094A.
[366] J. J. Pitt, ‘Principles and Applications of Liquid Chromatography-Mass Spectrometry in Clinical Biochemistry’, *Clin. Biochem. Rev.*, vol. 30, no. 1, pp. 19–34, Feb. 2009.
[367] L. Konermann, ‘A Simple Model for the Disintegration of Highly Charged Solvent Droplets during Electrospray Ionization’, *J. Am. Soc. Mass Spectrom.*, vol. 20, no. 3, pp. 496–506, Mar. 2009, doi: 10.1016/j.jasms.2008.11.007.
[368] J. Bastrup, S. Birkelund, A. A. Asuni, C. Volbracht, and A. Stensballe, ‘Dual strategy for reduced signal-suppression effects in matrix-assisted laser desorption/ionization mass spectrometry imaging’, *Rapid Commun. Mass Spectrom.*, vol. 33, no. 22, pp. 1711–1721, Nov. 2019, doi: 10.1002/rcm.8521.
[369] E. Tobolkina, ‘New analytical tools combining gel electrophoresis and mass spectrometry’, 2014. doi: 10.13140/RG.2.1.4866.5125.
[370] A. T. Kicman, M. C. Parkin, and R. K. Iles, ‘An introduction to mass spectrometry based proteomics—Detection and characterization of gonadotropins and related molecules’, *Mol. Cell. Endocrinol.*, vol. 260–262, pp. 212–227, Jan. 2007, doi: 10.1016/j.mce.2006.02.022.
[371] A. G. Brenton and A. R. Godfrey, ‘Accurate mass measurement: terminology and treatment of data’, *J. Am. Soc. Mass Spectrom.*, vol. 21, no. 11, pp. 1821–1835, Nov. 2010, doi: 10.1016/j.jasms.2010.06.006.
[372] V. Gold, Ed., *The IUPAC Compendium of Chemical Terminology: The Gold Book*, 4th ed. Research Triangle Park, NC: International Union of Pure and Applied Chemistry (IUPAC), 2019. doi: 10.1351/goldbook.
[373] G. Hart-Smith and S. J. Blanksby, ‘Mass Analysis’, in *Mass Spectrometry in Polymer Chemistry*, John Wiley & Sons, Ltd, 2011, pp. 5–32. doi: 10.1002/9783527641826.ch1.
[374] R. E. March, ‘An Introduction to Quadrupole Ion Trap Mass Spectrometry’, *J. Mass Spectrom.*, vol. 32, no. 4, pp. 351–369, 1997, doi: 10.1002/(SICI)1096-9888(199704)32:4<351::AID-JMS512>3.0.CO;2-Y.
[375] A. M. Haag, ‘Mass Analyzers and Mass Spectrometers’, in *Modern Proteomics – Sample Preparation, Analysis and Practical Applications*, H. Mirzaei and M. Carrasco, Eds., in Advances in Experimental Medicine and Biology. Cham: Springer International Publishing, 2016, pp. 157–169. doi: 10.1007/978-3-319-41448-5_7.

[376] B. A. Mamyrin, ‘Time-of-flight mass spectrometry (concepts, achievements, and prospects)’, *Int. J. Mass Spectrom.*, vol. 206, no. 3, pp. 251–266, Mar. 2001, doi: 10.1016/S1387-3806(00)00392-4.
[377] F. Hillenkamp and M. Karas, ‘[12] Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization’, in *Methods in Enzymology*, in Mass Spectrometry, vol. 193. Academic Press, 1990, pp. 280–295. doi: 10.1016/0076-6879(90)93420-P.
[378] null Makarov, ‘Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis’, *Anal. Chem.*, vol. 72, no. 6, pp. 1156–1162, Mar. 2000, doi: 10.1021/ac991131p.
[379] A. Makarov, E. Denisov, O. Lange, and S. Horning, ‘Dynamic range of mass accuracy in LTQ orbitrap hybrid mass spectrometer’, *J. Am. Soc. Mass Spectrom.*, vol. 17, no. 7, pp. 977–982, Jul. 2006, doi: 10.1016/j.jasms.2006.03.006.
[380] R. G. Cooks, Z. Ouyang, Z. Takats, and J. M. Wiseman, ‘Ambient Mass Spectrometry’, *Science*, vol. 311, no. 5767, pp. 1566–1570, Mar. 2006, doi: 10.1126/science.1119426.
[381] C. R. Ferreira, K. E. Yannell, A. K. Jarmusch, V. Pirro, Z. Ouyang, and R. G. Cooks, ‘Ambient Ionization Mass Spectrometry for Point-of-Care Diagnostics and Other Clinical Measurements’, *Clin. Chem.*, vol. 62, no. 1, pp. 99–110, Jan. 2016, doi: 10.1373/clinchem.2014.237164.
[382] V. Pirro, A. K. Jarmusch, M. Vincenti, and R. G. Cooks, ‘Direct drug analysis from oral fluid using medical swab touch spray mass spectrometry’, *Anal. Chim. Acta*, vol. 861, pp. 47–54, Feb. 2015, doi: 10.1016/j.aca.2015.01.008.
[383] K. S. Kerian, A. K. Jarmusch, and R. Graham Cooks, ‘Touch spray mass spectrometry for in situ analysis of complex samples’, *Analyst*, vol. 139, no. 11, pp. 2714–2720, 2014, doi: 10.1039/C4AN00548A.
[384] A. K. Jarmusch, V. Pirro, K. S. Kerian, and R. G. Cooks, ‘Detection of strep throat causing bacterium directly from medical swabs by touch spray-mass spectrometry’, *The Analyst*, vol. 139, no. 19, pp. 4785–4789, Oct. 2014, doi: 10.1039/c4an00959b.
[385] Z. Ling *et al.*, ‘Molecular analysis of the diversity of vaginal microbiota associated with bacterial vaginosis’, *BMC Genomics*, vol. 11, p. 488, Sep. 2010, doi: 10.1186/1471-2164-11-488.
[386] K.-C. Schäfer *et al.*, ‘In vivo, in situ tissue analysis using rapid evaporative ionization mass spectrometry’, *Angew. Chem. Int. Ed Engl.*, vol. 48, no. 44, pp. 8240–8242, 2009, doi: 10.1002/anie.200902546.
[387] K. Hiraoka, K. Nishidate, K. Mori, D. Asakawa, and S. Suzuki, ‘Development of probe electrospray using a solid needle’, *Rapid Commun. Mass Spectrom. RCM*, vol. 21, no. 18, pp. 3139–3144, 2007, doi: 10.1002/rcm.3201.
[388] G. A. Gómez-Ríos and J. Pawliszyn, ‘Development of Coated Blade Spray Ionization Mass Spectrometry for the Quantitation of Target Analytes Present in Complex Matrices’, *Angew. Chem. Int. Ed.*, vol. 53, no. 52, pp. 14503–14507, 2014, doi: 10.1002/anie.201407057.
[389] M. Doué, G. Dervilly-Pinel, A. Gicquiau, K. Pouponneau, F. Monteau, and B. Le Bizec, ‘High throughput identification and quantification of anabolic steroid esters by atmospheric solids analysis probe mass spectrometry for efficient screening of drug preparations’, *Anal. Chem.*, vol. 86, no. 12, pp. 5649–5655, Jun. 2014, doi: 10.1021/ac501072g.
[390] P.-K. So, T.-T. Ng, H. Wang, B. Hu, and Z.-P. Yao, ‘Rapid detection and quantitation of ketamine and norketamine in urine and oral fluid by wooden-tip electrospray

ionization mass spectrometry', *The Analyst*, vol. 138, no. 8, pp. 2239–2243, Apr. 2013, doi: 10.1039/c3an36641c.
[391] C.-H. Lin, W.-C. Liao, H.-K. Chen, and T.-Y. Kuo, 'Paper spray-MS for bioanalysis', *Bioanalysis*, vol. 6, no. 2, pp. 199–208, Jan. 2014, doi: 10.4155/bio.13.310.
[392] Y. Su, H. Wang, J. Liu, P. Wei, R. G. Cooks, and Z. Ouyang, 'Quantitative paper spray mass spectrometry analysis of drugs of abuse', *The Analyst*, vol. 138, no. 16, p. 4443, 2013, doi: 10.1039/c3an00934c.
[393] N. E. Manicke, P. Abu-Rabie, N. Spooner, Z. Ouyang, and R. G. Cooks, 'Quantitative Analysis of Therapeutic Drugs in Dried Blood Spot Samples by Paper Spray Mass Spectrometry: An Avenue to Therapeutic Drug Monitoring', *J. Am. Soc. Mass Spectrom.*, vol. 22, no. 9, pp. 1501–1507, Sep. 2011, doi: 10.1007/s13361-011-0177-x.
[394] R. D. Espy *et al.*, 'Paper Spray and Extraction Spray Mass Spectrometry for the Direct and Simultaneous Quantification of Eight Drugs of Abuse in Whole Blood', *Anal. Chem.*, vol. 86, no. 15, pp. 7712–7718, Aug. 2014, doi: 10.1021/ac5016408.
[395] R. D. Espy, N. E. Manicke, Z. Ouyang, and R. G. Cooks, 'Rapid analysis of whole blood by paper spray mass spectrometry for point-of-care therapeutic drug monitoring', *The Analyst*, vol. 137, no. 10, p. 2344, 2012, doi: 10.1039/c2an35082c.
[396] S. A. Borden, A. Saatchi, G. W. Vandergrift, J. Palaty, M. Lysyshyn, and C. G. Gill, 'A new quantitative drug checking technology for harm reduction: Pilot study in Vancouver, Canada using paper spray mass spectrometry', *Drug Alcohol Rev.*, vol. 41, no. 2, pp. 410–418, Feb. 2022, doi: 10.1111/dar.13370.
[397] R. Garrett, C. M. Rezende, and D. R. Ifa, 'Coffee origin discrimination by paper spray mass spectrometry and direct coffee spray analysis', *Anal. Methods*, vol. 5, no. 21, p. 5944, 2013, doi: 10.1039/c3ay41247d.
[398] P. da Silva Ferreira, D. Fernandes de Abreu e Silva, R. Augusti, and E. Piccin, 'Forensic analysis of ballpoint pen inks using paper spray mass spectrometry', *The Analyst*, vol. 140, no. 3, pp. 811–819, 2015, doi: 10.1039/C4AN01617C.
[399] A. M. Hamid *et al.*, 'Rapid Discrimination of Bacteria by Paper Spray Mass Spectrometry', *Anal. Chem.*, vol. 86, no. 15, pp. 7500–7507, Aug. 2014, doi: 10.1021/ac501254b.
[400] B. Fatou *et al.*, 'In vivo Real-Time Mass Spectrometry for Guided Surgery Application', *Sci. Rep.*, vol. 6, no. 1, Art. no. 1, May 2016, doi: 10.1038/srep25919.
[401] J. Balog *et al.*, 'Intraoperative tissue identification using rapid evaporative ionization mass spectrometry', *Sci. Transl. Med.*, vol. 5, no. 194, p. 194ra93, Jul. 2013, doi: 10.1126/scitranslmed.3005623.
[402] Z. Takáts, J. M. Wiseman, B. Gologan, and R. G. Cooks, 'Mass spectrometry sampling under ambient conditions with desorption electrospray ionization', *Science*, vol. 306, no. 5695, pp. 471–473, Oct. 2004, doi: 10.1126/science.1104404.
[403] Z. Takats, N. Strittmatter, and J. S. McKenzie, 'Ambient Mass Spectrometry in Cancer Research', in *Advances in Cancer Research*, Elsevier, 2017, pp. 231–256. doi: 10.1016/bs.acr.2016.11.011.
[404] K. M. Alsante, S. W. Baertschi, M. Coutant, B. L. Marquez, T. R. Sharp, and T. C. Zelesky, '3 - Degradation and Impurity Analysis for Pharmaceutical Drug Candidates', in *Separation Science and Technology*, S. Ahuja and S. Scypinski, Eds., in Handbook of Modern Pharmaceutical Analysis, vol. 10. Academic Press, 2011, pp. 59–169. doi: 10.1016/B978-0-12-375680-0.00003-6.
[405] S. Guenther *et al.*, 'Spatially resolved metabolic phenotyping of breast cancer by desorption electrospray ionization mass spectrometry', *Cancer Res.*, vol. 75, no. 9, pp. 1828–1837, May 2015, doi: 10.1158/0008-5472.CAN-14-2258.

[406] A. Wójtowicz and R. Wietecha-Posłuszny, ‘DESI-MS analysis of human fluids and tissues for forensic applications’, *Appl. Phys. A*, vol. 125, no. 5, p. 312, Apr. 2019, doi: 10.1007/s00339-019-2564-2.
[407] N. V. Schwab, M. O. Ore, M. N. Eberlin, S. Morin, and D. R. Ifa, ‘Functionalized porous silicon surfaces as DESI-MS substrates for small molecules analysis’, *Anal. Chem.*, vol. 86, no. 23, pp. 11722–11726, Dec. 2014, doi: 10.1021/ac503252g.
[408] J. M. Wiseman, C. A. Evans, C. L. Bowen, and J. H. Kennedy, ‘Direct analysis of dried blood spots utilizing desorption electrospray ionization (DESI) mass spectrometry’, *The Analyst*, vol. 135, no. 4, pp. 720–725, Apr. 2010, doi: 10.1039/b922329k.
[409] A. Tata *et al.*, ‘Rapid Detection of Necrosis in Breast Cancer with Desorption Electrospray Ionization Mass Spectrometry’, *Sci. Rep.*, vol. 6, no. 1, Art. no. 1, Oct. 2016, doi: 10.1038/srep35374.
[410] D. Calligaris *et al.*, ‘Application of desorption electrospray ionization mass spectrometry imaging in breast cancer margin analysis’, *Proc. Natl. Acad. Sci.*, vol. 111, no. 42, pp. 15184–15189, Oct. 2014, doi: 10.1073/pnas.1408129111.
[411] L. S. Eberlin *et al.*, ‘Cholesterol Sulfate Imaging in Human Prostate Cancer Tissue by Desorption Electrospray Ionization Mass Spectrometry’, *Anal. Chem.*, vol. 82, no. 9, pp. 3430–3434, May 2010, doi: 10.1021/ac9029482.
[412] S. Banerjee *et al.*, ‘Diagnosis of prostate cancer by desorption electrospray ionization mass spectrometric imaging of small metabolites and lipids’, *Proc. Natl. Acad. Sci.*, vol. 114, no. 13, pp. 3334–3339, Mar. 2017, doi: 10.1073/pnas.1700677114.
[413] K. S. Kerian *et al.*, ‘Differentiation of prostate cancer from normal tissue in radical prostatectomy specimens by desorption electrospray ionization and touch spray ionization mass spectrometry’, *Analyst*, vol. 140, no. 4, pp. 1090–1098, 2015, doi: 10.1039/C4AN02039A.
[414] V. Pirro, A. K. Jarmusch, C. M. Alfaro, E. M. Hattab, A. A. Cohen-Gadol, and R. Graham Cooks, ‘Utility of neurological smears for intrasurgical brain cancer diagnostics and tumour cell percentage by DESI-MS’, *Analyst*, vol. 142, no. 3, pp. 449–454, 2017, doi: 10.1039/C6AN02645A.
[415] A. K. Jarmusch, V. Pirro, Z. Baird, E. M. Hattab, A. A. Cohen-Gadol, and R. G. Cooks, ‘Lipid and metabolite profiles of human brain tumors by desorption electrospray ionization-MS’, *Proc. Natl. Acad. Sci.*, vol. 113, no. 6, pp. 1486–1491, Feb. 2016, doi: 10.1073/pnas.1523306113.
[416] L. S. Eberlin *et al.*, ‘Classifying Human Brain Tumors by Lipid Imaging with Mass Spectrometry’, *Cancer Res.*, vol. 72, no. 3, pp. 645–654, Jan. 2012, doi: 10.1158/0008-5472.CAN-11-2465.
[417] S. Santagata *et al.*, ‘Intraoperative mass spectrometry mapping of an onco-metabolite to guide brain tumor surgery’, *Proc. Natl. Acad. Sci.*, vol. 111, no. 30, pp. 11121–11126, Jul. 2014, doi: 10.1073/pnas.1404724111.
[418] L.-H. Li, H.-Y. Hsieh, and C.-C. Hsu, ‘Clinical Application of Ambient Ionization Mass Spectrometry’, *Mass Spectrom.*, vol. 6, no. 2, pp. S0060–S0060, 2017, doi: 10.5702/massspectrometry.S0060.
[419] N. Abbassi-Ghadi *et al.*, ‘Repeatability and reproducibility of desorption electrospray ionization-mass spectrometry (DESI-MS) for the imaging analysis of human cancer tissue: a gateway for clinical applications’, *Anal. Methods*, vol. 7, no. 1, pp. 71–80, Dec. 2014, doi: 10.1039/C4AY01770F.
[420] Y. Song, N. Talaty, W. A. Tao, Z. Pan, and R. G. Cooks, ‘Rapid ambient mass spectrometric profiling of intact, untreated bacteria using desorption electrospray ionization’, *Chem. Commun.*, no. 1, pp. 61–63, Dec. 2006, doi: 10.1039/B615724F.

[421] J. Bartel, J. Krumsiek, and F. J. Theis, ‘STATISTICAL METHODS FOR THE ANALYSIS OF HIGH-THROUGHPUT METABOLOMICS DATA’, *Comput. Struct. Biotechnol. J.*, vol. 4, no. 5, p. e201301009, Jan. 2013, doi: 10.5936/csbj.201301009.
[422] M. Bhattacharya, L. Fuhrman, A. Ingram, K. W. Nickerson, and T. Conway, ‘Single-run separation and detection of multiple metabolic intermediates by anion-exchange high-performance liquid chromatography and application to cell pool extracts prepared from Escherichia coli’, *Anal. Biochem.*, vol. 232, no. 1, pp. 98–106, Nov. 1995, doi: 10.1006/abio.1995.9954.
[423] H. Tweeddale, L. Notley-McRobb, and T. Ferenci, ‘Effect of slow growth on metabolism of Escherichia coli, as revealed by global metabolite pool (“metabolome”) analysis’, *J. Bacteriol.*, vol. 180, no. 19, pp. 5109–5116, Oct. 1998, doi: 10.1128/JB.180.19.5109-5116.1998.
[424] J. Trygg and T. Lundstedt, ‘Chemometrics Techniques for Metabonomics’, in *The Handbook of Metabonomics and Metabolomics*, Elsevier, 2007, pp. 171–199. doi: 10.1016/B978-044452841-4/50007-2.
[425] E. Saccenti, H. Hoefsloot, A. Smilde, J. Westerhuis, and M. Hendriks, ‘Reflections on univariate and multivariate analysis of metabolomics data’, *Metabolomics*, vol. 10, Dec. 2013, doi: 10.1007/s11306-013-0598-6.
[426] A. Sreekumar *et al.*, ‘Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression’, *Nature*, vol. 457, no. 7231, pp. 910–914, Feb. 2009, doi: 10.1038/nature07762.
[427] E. Altmaier, S. L. Ramsay, A. Graber, H.-W. Mewes, K. M. Weinberger, and K. Suhre, ‘Bioinformatics analysis of targeted metabolomics--uncovering old and new tales of diabetic mice under medication’, *Endocrinology*, vol. 149, no. 7, pp. 3478–3489, Jul. 2008, doi: 10.1210/en.2007-1747.
[428] Y. Takahashi *et al.*, ‘Improved metabolomic data-based prediction of depressive symptoms using nonlinear machine learning with feature selection’, *Transl. Psychiatry*, vol. 10, no. 1, Art. no. 1, May 2020, doi: 10.1038/s41398-020-0831-9.
[429] H. Yamamoto *et al.*, ‘Dimensionality reduction for metabolome data using PCA, PLS, OPLS, and RFDA with differential penalties to latent variables’, *Chemom. Intell. Lab. Syst.*, vol. 98, no. 2, pp. 136–142, Oct. 2009, doi: 10.1016/j.chemolab.2009.05.006.
[430] I. Jolliffe, ‘Principal Component Analysis’, in *Wiley StatsRef: Statistics Reference Online*, N. Balakrishnan, T. Colton, B. Everitt, W. Piegorsch, F. Ruggeri, and J. L. Teugels, Eds., 1st ed.Wiley, 2014. doi: 10.1002/9781118445112.stat06472.
[431] S. Lamichhane, P. Sen, A. M. Dickens, T. Hyötyläinen, and M. Orešič, ‘An Overview of Metabolomics Data Analysis: Current Tools and Future Perspectives’, in *Comprehensive Analytical Chemistry*, Elsevier, 2018, pp. 387–413. doi: 10.1016/bs.coac.2018.07.001.
[432] J. J. Jansen, S. Smit, H. C. J. Hoefsloot, and A. K. Smilde, ‘The photographer and the greenhouse: how to analyse plant metabolomics data: How to Analyse Plant Metabolomics Data’, *Phytochem. Anal.*, vol. 21, no. 1, pp. 48–60, Jan. 2010, doi: 10.1002/pca.1181.
[433] E. Want and P. Masson, ‘Processing and analysis of GC/LC-MS-based metabolomics data’, *Methods Mol. Biol. Clifton NJ*, vol. 708, pp. 277–298, 2011, doi: 10.1007/978-1-61737-985-7_17.
[434] T. Mehmood, H. Martens, S. Sæbø, J. Warringer, and L. Snipen, ‘A Partial Least Squares based algorithm for parsimonious variable selection’, *Algorithms Mol. Biol.*, vol. 6, no. 1, p. 27, Dec. 2011, doi: 10.1186/1748-7188-6-27.

[435] P. S. Gromski *et al.*, ‘A tutorial review: Metabolomics and partial least squares-discriminant analysis – a marriage of convenience or a shotgun wedding’, *Anal. Chim. Acta*, vol. 879, pp. 10–23, Jun. 2015, doi: 10.1016/j.aca.2015.02.012.
[436] J. A. Westerhuis *et al.*, ‘Assessment of PLSDA cross validation’, *Metabolomics*, vol. 4, no. 1, pp. 81–89, Mar. 2008, doi: 10.1007/s11306-007-0099-6.
[437] M. W. Browne, ‘Cross-Validation Methods’, *J. Math. Psychol.*, vol. 44, no. 1, pp. 108–132, Mar. 2000, doi: 10.1006/jmps.1999.1279.
[438] T. Chen *et al.*, ‘Random Forest in Clinical Metabolomics for Phenotypic Discrimination and Biomarker Selection’, *Evid. Based Complement. Alternat. Med.*, vol. 2013, pp. 1–11, 2013, doi: 10.1155/2013/298183.
[439] D. Amaratunga, J. Cabrera, and Y.-S. Lee, ‘Enriched random forests’, *Bioinformatics*, vol. 24, no. 18, pp. 2010–2014, Sep. 2008, doi: 10.1093/bioinformatics/btn356.
[440] L. Breiman, ‘Random Forests’, *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
[441] D. A. MacIntyre *et al.*, ‘The vaginal microbiome during pregnancy and the postpartum period in a European population’, *Sci. Rep.*, vol. 5, p. 8988, Mar. 2015, doi: 10.1038/srep08988.
[442] A. Sundquist *et al.*, ‘Bacterial flora-typing with targeted, chip-based Pyrosequencing’, *BMC Microbiol.*, vol. 7, no. 1, p. 108, Nov. 2007, doi: 10.1186/1471-2180-7-108.
[443] J. J. Kozich, S. L. Westcott, N. T. Baxter, S. K. Highlander, and P. D. Schloss, ‘Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform’, *Appl. Environ. Microbiol.*, vol. 79, no. 17, pp. 5112–5120, Sep. 2013, doi: 10.1128/AEM.01043-13.
[444] Y. Wang and P.-Y. Qian, ‘Conservative Fragments in Bacterial 16S rRNA Genes and Primer Design for 16S Ribosomal DNA Amplicons in Metagenomic Studies’, *PLOS ONE*, vol. 4, no. 10, p. e7401, Oct. 2009, doi: 10.1371/journal.pone.0007401.
[445] C. Quast *et al.*, ‘The SILVA ribosomal RNA gene database project: improved data processing and web-based tools’, *Nucleic Acids Res.*, vol. 41, no. D1, pp. D590–D596, Jan. 2013, doi: 10.1093/nar/gks1219.
[446] J. M. Fettweis *et al.*, ‘Species-level classification of the vaginal microbiome’, *BMC Genomics*, vol. 13, no. 8, p. S17, Dec. 2012, doi: 10.1186/1471-2164-13-S8-S17.
[447] J. R. Cole *et al.*, ‘Ribosomal Database Project: data and tools for high throughput rRNA analysis’, *Nucleic Acids Res.*, vol. 42, no. Database issue, pp. D633-642, Jan. 2014, doi: 10.1093/nar/gkt1244.
[448] ‘Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data’. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed Sep. 13, 2022).
[449] E. Bolyen *et al.*, ‘Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2’, *Nat. Biotechnol.*, vol. 37, no. 8, Art. no. 8, Aug. 2019, doi: 10.1038/s41587-019-0209-9.
[450] M. I. Love, W. Huber, and S. Anders, ‘Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2’, *Genome Biol.*, vol. 15, no. 12, p. 550, Dec. 2014, doi: 10.1186/s13059-014-0550-8.
[451] P. J. McMurdie and S. Holmes, ‘phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data’, *PLOS ONE*, vol. 8, no. 4, p. e61217, Apr. 2013, doi: 10.1371/journal.pone.0061217.
[452] D. Chan *et al.*, ‘Microbial-driven preterm labour involves crosstalk between the innate and adaptive immune response’, *Nat. Commun.*, vol. 13, no. 1, Art. no. 1, Feb. 2022, doi: 10.1038/s41467-022-28620-1.

[453] M. N. Anahtar *et al.*, ‘Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract’, *Immunity*, vol. 42, no. 5, pp. 965–976, May 2015, doi: 10.1016/j.immuni.2015.04.019.
[454] V. Jespers *et al.*, ‘A longitudinal analysis of the vaginal microbiota and vaginal immune mediators in women from sub-Saharan Africa’, *Sci. Rep.*, vol. 7, no. 1, Art. no. 1, Sep. 2017, doi: 10.1038/s41598-017-12198-6.
[455] E. A. Kremleva and A. V. Sgibnev, ‘Proinflammatory Cytokines as Regulators of Vaginal Microbiota’, *Bull. Exp. Biol. Med.*, vol. 162, no. 1, pp. 75–78, Nov. 2016, doi: 10.1007/s10517-016-3549-1.
[456] P. Pruski, ‘Development of novel mass spectrometric methods for point-of-care mucosal diagnostics’, Sep. 2019, doi: 10.25560/96407.
[457] C. Mc *et al.*, ‘A cross-platform toolkit for mass spectrometry and proteomics’, *Nat. Biotechnol.*, vol. 30, no. 10, Oct. 2012, doi: 10.1038/nbt.2377.
[458] S. Gibb and K. Strimmer, ‘MALDIquant: a versatile R package for the analysis of mass spectrometry data’, *Bioinformatics*, vol. 28, no. 17, pp. 2270–2271, Sep. 2012, doi: 10.1093/bioinformatics/bts447.
[459] T. Bald, J. Barth, A. Niehues, M. Specht, M. Hippler, and C. Fufezan, ‘pymzML—Python module for high-throughput bioinformatics on mass spectrometry data’, *Bioinformatics*, vol. 28, no. 7, pp. 1052–1053, Apr. 2012, doi: 10.1093/bioinformatics/bts066.
[460] J. H. H. M. van de Wijgert, ‘The vaginal microbiome and sexually transmitted infections are interlinked: Consequences for treatment and prevention’, *PLOS Med.*, vol. 14, no. 12, p. e1002478, Dec. 2017, doi: 10.1371/journal.pmed.1002478.
[461] C. S. da Silva, S. J. Adad, M. A. Hazarabedian de Souza, A. C. Macêdo Barcelos, A. P. Sarreta Terra, and E. F. C. Murta, ‘Increased frequency of bacterial vaginosis and Chlamydia trachomatis in pregnant women with human papillomavirus infection’, *Gynecol. Obstet. Invest.*, vol. 58, no. 4, pp. 189–193, 2004, doi: 10.1159/000079822.
[462] N. H. Liversedge, A. Turner, P. J. Horner, S. D. Keay, J. M. Jenkins, and M. G. Hull, ‘The influence of bacterial vaginosis on in-vitro fertilization and embryo implantation during assisted reproduction treatment’, *Hum. Reprod. Oxf. Engl.*, vol. 14, no. 9, pp. 2411–2415, Sep. 1999, doi: 10.1093/humrep/14.9.2411.
[463] G. Işik, Ş. Demirezen, H. G. Dönmez, and M. S. Beksaç, ‘Bacterial vaginosis in association with spontaneous abortion and recurrent pregnancy losses’, *J. Cytol. Indian Acad. Cytol.*, vol. 33, no. 3, pp. 135–140, 2016, doi: 10.4103/0970-9371.188050.
[464] P. Mirmonsef, M. R. Zariffard, D. Gilbert, H. Makinde, A. L. Landay, and G. T. Spear, ‘Short-chain fatty acids induce pro-inflammatory cytokine production alone and in combination with toll-like receptor ligands’, *Am. J. Reprod. Immunol. N. Y. N 1989*, vol. 67, no. 5, pp. 391–400, May 2012, doi: 10.1111/j.1600-0897.2011.01089.x.
[465] C. R. Cavaglieri, A. Nishiyama, L. C. Fernandes, R. Curi, E. A. Miles, and P. C. Calder, ‘Differential effects of short-chain fatty acids on proliferation and production of pro- and anti-inflammatory cytokines by cultured lymphocytes’, *Life Sci.*, vol. 73, no. 13, pp. 1683–1690, Aug. 2003, doi: 10.1016/s0024-3205(03)00490-9.
[466] S. Al-Mushrif, A. Eley, and B. M. Jones, ‘Inhibition of chemotaxis by organic acids from anaerobes may prevent a purulent response in bacterial vaginosis’, *J. Med. Microbiol.*, vol. 49, no. 11, pp. 1023–1030, Nov. 2000, doi: 10.1099/0022-1317-49-11-1023.
[467] P. Pruski *et al.*, ‘Medical Swab Analysis Using Desorption Electrospray Ionization Mass Spectrometry: A Noninvasive Approach for Mucosal Diagnostics’, *Anal. Chem.*, vol. 89, no. 3, pp. 1540–1550, Feb. 2017, doi: 10.1021/acs.analchem.6b03405.

[468] C. Lucas-Torres *et al.*, 'General Guidelines for Sample Preparation Strategies in HR-μMAS NMR-based Metabolomics of Microscopic Specimens', *Metabolites*, vol. 10, no. 2, Art. no. 2, Feb. 2020, doi: 10.3390/metabo10020054.
[469] K. Njoku, C. J. J. Sutton, A. D. Whetton, and E. J. Crosbie, 'Metabolomic Biomarkers for Detection, Prognosis and Identifying Recurrence in Endometrial Cancer', *Metabolites*, vol. 10, no. 8, p. 314, Jul. 2020, doi: 10.3390/metabo10080314.
[470] K. Laks, T. Kirsipuu, T. Dmitrijeva, A. Salumets, and P. Palumaa, 'Assessment of Blood Contamination in Biological Fluids Using MALDI-TOF MS', *Protein J.*, vol. 35, no. 3, pp. 171–176, Jun. 2016, doi: 10.1007/s10930-016-9657-y.
[471] K. Barkovits *et al.*, 'Blood Contamination in CSF and Its Impact on Quantitative Analysis of Alpha-Synuclein', *Cells*, vol. 9, no. 2, p. E370, Feb. 2020, doi: 10.3390/cells9020370.
[472] L. Laghi *et al.*, 'Vaginal metabolic profiles during pregnancy: Changes between first and second trimester', *PLoS ONE*, vol. 16, no. 4, p. e0249925, Apr. 2021, doi: 10.1371/journal.pone.0249925.
[473] J. C. Borgogna *et al.*, 'The vaginal metabolome and microbiota of cervical HPV-positive and HPV-negative women: a cross-sectional analysis', *BJOG Int. J. Obstet. Gynaecol.*, vol. 127, no. 2, pp. 182–192, Jan. 2020, doi: 10.1111/1471-0528.15981.
[474] B. Xiao, D. A, H. Qin, L. Mi, and D. Zhang, 'Correlation Analysis of Vaginal Microbiome Changes and Bacterial Vaginosis Plus Vulvovaginal Candidiasis Mixed Vaginitis Prognosis', *Front. Cell. Infect. Microbiol.*, vol. 12, p. 860589, Mar. 2022, doi: 10.3389/fcimb.2022.860589.
[475] J. S. Goodwin and J. Ceuppens, 'Regulation of the immune response by prostaglandins', *J. Clin. Immunol.*, vol. 3, no. 4, pp. 295–315, Oct. 1983, doi: 10.1007/BF00915791.
[476] K. R. Dhanasekar, B. Shilpa, N. Gomathy, and S. Kundavi, 'Prenatal Probiotics: The Way Forward in Prevention of Preterm Birth', *J. Clin. Gynecol. Obstet.*, vol. 8, no. 3, Art. no. 3, Sep. 2019, doi: 10.14740/jcgo.v8i3.571.
[477] O. Abdul-Rahim *et al.*, 'Phenyl-Lactic Acid Is an Active Ingredient in Bactericidal Supernatants of Lactobacillus crispatus', *J. Bacteriol.*, vol. 203, no. 19, pp. e00360-21, doi: 10.1128/JB.00360-21.
[478] X. Xu *et al.*, 'Differential expression and anti-oxidant function of glutathione peroxidase 3 in mouse uterus during decidualization', *FEBS Lett.*, vol. 588, no. 9, pp. 1580–1589, May 2014, doi: 10.1016/j.febslet.2014.02.043.
[479] X. Jin *et al.*, 'Proteomics analysis of human placenta reveals glutathione metabolism dysfunction as the underlying pathogenesis for preeclampsia', *Biochim. Biophys. Acta BBA - Proteins Proteomics*, vol. 1865, no. 9, pp. 1207–1214, Sep. 2017, doi: 10.1016/j.bbapap.2017.07.003.
[480] J. M. Fettweis *et al.*, 'The vaginal microbiome and preterm birth', *Nat. Med.*, vol. 25, no. 6, Art. no. 6, Jun. 2019, doi: 10.1038/s41591-019-0450-2.
[481] S. R. Hedges, F. Barrientes, R. A. Desmond, and J. R. Schwebke, 'Local and systemic cytokine levels in relation to changes in vaginal flora', *J. Infect. Dis.*, vol. 193, no. 4, pp. 556–562, Feb. 2006, doi: 10.1086/499824.
[482] I. M. Linhares, P. C. Giraldo, and E. C. Baracat, 'Vaginal bacterial flora: up to date', *Rev. Assoc. Médica Bras.*, vol. 56, pp. 370–374, 2010, doi: 10.1590/S0104-42302010000300026.
[483] E. Y. Jung, K. H. Park, S. Y. Lee, A. Ryu, J. K. Joo, and J. W. Park, 'Predicting outcomes of emergency cerclage in women with cervical insufficiency using inflammatory markers in maternal blood and amniotic fluid', *Int. J. Gynecol. Obstet.*, vol. 132, no. 2, pp. 165–169, Feb. 2016, doi: 10.1016/j.ijgo.2015.07.011.

[484] S. Y. Kim, K. H. Park, H. J. Kim, Y. M. Kim, K. Ahn, and K.-N. Lee, 'Inflammation-related immune proteins in maternal plasma as potential predictive biomarkers for rescue cerclage outcome in women with cervical insufficiency', *Am. J. Reprod. Immunol.*, vol. 88, no. 1, p. e13557, 2022, doi: 10.1111/aji.13557.
[485] K.-N. Lee, K. H. Park, Y. M. Kim, I. Cho, and T. E. Kim, 'Prediction of emergency cerclage outcomes in women with cervical insufficiency: The role of inflammatory, angiogenic, and extracellular matrix-related proteins in amniotic fluid', *PLOS ONE*, vol. 17, no. 5, p. e0268291, May 2022, doi: 10.1371/journal.pone.0268291.
[486] D. M. Ferrero *et al.*, 'Cross-Country Individual Participant Analysis of 4.1 Million Singleton Births in 5 Countries with Very High Human Development Index Confirms Known Associations but Provides No Biologic Explanation for 2/3 of All Preterm Births', *PLOS ONE*, vol. 11, no. 9, p. e0162506, Sep. 2016, doi: 10.1371/journal.pone.0162506.
[487] B. M. Mercer *et al.*, 'The Preterm Prediction Study: prediction of preterm premature rupture of membranes through clinical findings and ancillary testing. The National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network', *Am. J. Obstet. Gynecol.*, vol. 183, no. 3, pp. 738–745, Sep. 2000, doi: 10.1067/mob.2000.106766.
[488] D. Drassinower, S. G. Običan, Z. Siddiq, D. Heller, C. Gyamfi-Bannerman, and A. M. Friedman, 'Does the clinical presentation of a prior preterm birth predict risk in a subsequent pregnancy?', *Am. J. Obstet. Gynecol.*, vol. 213, no. 5, p. 686.e1–7, Nov. 2015, doi: 10.1016/j.ajog.2015.07.029.
[489] V. H. Gonzalez-Quintero *et al.*, 'Influence of gestational age and reason for prior preterm birth on rates of recurrent preterm delivery', *Am. J. Obstet. Gynecol.*, vol. 205, no. 3, p. 275.e1–5, Sep. 2011, doi: 10.1016/j.ajog.2011.06.043.
[490] R. A. Carr-Hill and M. H. Hall, 'The repetition of spontaneous preterm labour', *Br. J. Obstet. Gynaecol.*, vol. 92, no. 9, pp. 921–928, Sep. 1985, doi: 10.1111/j.1471-0528.1985.tb03071.x.
[491] R. L. Goldenberg *et al.*, 'What we have learned about the predictors of preterm birth', *Semin. Perinatol.*, vol. 27, no. 3, pp. 185–193, Jun. 2003, doi: 10.1016/S0146-0005(03)00017-X.
[492] E. Papiernik, J. Bouyer, D. Collin, G. Winisdoerffer, and J. Dreyfus, 'Precocious cervical ripening and preterm labor', *Obstet. Gynecol.*, vol. 67, no. 2, pp. 238–242, Feb. 1986, doi: 10.1097/00006250-198602000-00014.
[493] T. M. Stubbs, J. P. Van Dorsten, and M. C. Miller, 'The preterm cervix and preterm labor: relative risks, predictive values, and change over time', *Am. J. Obstet. Gynecol.*, vol. 155, no. 4, pp. 829–834, Oct. 1986, doi: 10.1016/s0002-9378(86)80031-x.
[494] V. Berghella, A. Roman, C. Daskalakis, A. Ness, and J. K. Baxter, 'Gestational Age at Cervical Length Measurement and Incidence of Preterm Birth', *Obstet. Gynecol.*, vol. 110, no. 2 Part 1, pp. 311–317, Aug. 2007, doi: 10.1097/01.AOG.0000270112.05025.1d.
[495] J. Owen *et al.*, 'Multicenter randomized trial of cerclage for preterm birth prevention in high-risk women with shortened midtrimester cervical length', *Am. J. Obstet. Gynecol.*, vol. 201, no. 4, p. 375.e1–8, Oct. 2009, doi: 10.1016/j.ajog.2009.08.015.
[496] J. Grimes-Dennis and V. Berghella, 'Cervical length and prediction of preterm delivery', *Curr. Opin. Obstet. Gynecol.*, vol. 19, no. 2, pp. 191–195, Apr. 2007, doi: 10.1097/GCO.0b013e3280895dd3.
[497] L. M. Kindinger *et al.*, 'Preterm Birth Prevention Post-Conization: A Model of Cervical Length Screening with Targeted Cerclage', *PLoS ONE*, vol. 11, no. 11, p. e0163793, Nov. 2016, doi: 10.1371/journal.pone.0163793.

[498] V. Berghella, L. Pereira, A. Gariepy, and G. Simonazzi, ‘Prior cone biopsy: prediction of preterm birth by cervical ultrasound’, *Am. J. Obstet. Gynecol.*, vol. 191, no. 4, pp. 1393–1397, Oct. 2004, doi: 10.1016/j.ajog.2004.06.087.
[499] K. Haram, J. H. S. Mortensen, and A.-L. Wollen, ‘Preterm delivery: an overview’, *Acta Obstet. Gynecol. Scand.*, vol. 82, no. 8, pp. 687–704, 2003, doi: 10.1034/j.1600-0412.2003.00218.x.
[500] M. Kyrgiou, G. Koliopoulos, P. Martin-Hirsch, M. Arbyn, W. Prendiville, and E. Paraskevaidis, ‘Obstetric outcomes after conservative treatment for intraepithelial or early invasive cervical lesions: systematic review and meta-analysis’, *Lancet Lond. Engl.*, vol. 367, no. 9509, pp. 489–498, Feb. 2006, doi: 10.1016/S0140-6736(06)68181-6.
[501] M. S. To *et al.*, ‘Cervical cerclage for prevention of preterm delivery in woman with short cervix:randomised controlled trial’, *The Lancet*, vol. 363, no. 9424, pp. 1849–1853, Jun. 2004, doi: 10.1016/S0140-6736(04)16351-4.
[502] R. Brown *et al.*, ‘Prospective observational study of vaginal microbiota pre- and post-rescue cervical cerclage’, *BJOG Int. J. Obstet. Gynaecol.*, vol. 126, no. 7, pp. 916–925, 2019, doi: 10.1111/1471-0528.15600.
[503] A. L. Jorgensen, Z. Alfirevic, C. Tudur Smith, P. R. Williamson, and cerclage IPD Meta-analysis Group, ‘Cervical stitch (cerclage) for preventing pregnancy loss: individual patient data meta-analysis’, *BJOG Int. J. Obstet. Gynaecol.*, vol. 114, no. 12, pp. 1460–1476, Dec. 2007, doi: 10.1111/j.1471-0528.2007.01515.x.
[504] M. Quinn, ‘Final report of the MRC/RCOG randomised controlled trial of cervical cerclage’, *Br. J. Obstet. Gynaecol.*, vol. 100, no. 12, pp. 1154–1155, Dec. 1993, doi: 10.1111/j.1471-0528.1993.tb15198.x.
[505] V. H. Morton *et al.*, ‘Monofilament suture versus braided suture thread to improve pregnancy outcomes after vaginal cervical cerclage (C-STICH): a pragmatic randomised, controlled, phase 3, superiority trial’, *The Lancet*, vol. 400, no. 10361, pp. 1426–1436, Oct. 2022, doi: 10.1016/S0140-6736(22)01808-6.
[506] T. Tingleff, Å. Vikanes, S. Räisänen, L. Sandvik, G. Murzakanova, and K. Laine, ‘Risk of preterm birth in relation to history of preterm birth: a population-based registry study of 213 335 women in Norway’, *BJOG Int. J. Obstet. Gynaecol.*, vol. 129, no. 6, pp. 900–907, 2022, doi: 10.1111/1471-0528.17013.
[507] J. D. Iams *et al.*, ‘The Preterm Prediction Study: recurrence risk of spontaneous preterm birth. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network’, *Am. J. Obstet. Gynecol.*, vol. 178, no. 5, pp. 1035–1040, May 1998, doi: 10.1016/s0002-9378(98)70544-7.
[508] R. L. Naeye, ‘Factors that predispose to premature rupture of the fetal membranes’, *Obstet. Gynecol.*, vol. 60, no. 1, pp. 93–98, Jul. 1982.
[509] T. Asrat *et al.*, ‘Rate of recurrence of preterm premature rupture of membranes in consecutive pregnancies’, *Am. J. Obstet. Gynecol.*, vol. 165, no. 4 Pt 1, pp. 1111–1115, Oct. 1991, doi: 10.1016/0002-9378(91)90481-6.
[510] E. E. Ekwo, C. A. Gosselink, and A. Moawad, ‘Unfavorable outcome in penultimate pregnancy and premature rupture of membranes in successive pregnancy’, *Obstet. Gynecol.*, vol. 80, no. 2, pp. 166–172, Aug. 1992.
[511] R. Romero and G. Tromp, ‘High-dimensional biology in obstetrics and gynecology: functional genomics in microarray studies’, *Am. J. Obstet. Gynecol.*, vol. 195, no. 2, pp. 360–363, Aug. 2006, doi: 10.1016/j.ajog.2006.06.077.
[512] P. Sasieni *et al.*, ‘Risk of preterm birth following surgical treatment for cervical disease: executive summary of a recent symposium’, *BJOG Int. J. Obstet. Gynaecol.*, vol. 123, no. 9, pp. 1426–1429, 2016, doi: 10.1111/1471-0528.13839.

[513] S. Khalid *et al.*, 'The thickness and volume of LLETZ specimens can predict the relative risk of pregnancy-related morbidity', *BJOG Int. J. Obstet. Gynaecol.*, vol. 119, no. 6, pp. 685–691, May 2012, doi: 10.1111/j.1471-0528.2011.03252.x.
[514] L. Sadler, A. Saftlas, W. Wang, M. Exeter, J. Whittaker, and L. McCowan, 'Treatment for cervical intraepithelial neoplasia and risk of preterm delivery', *JAMA*, vol. 291, no. 17, pp. 2100–2106, May 2004, doi: 10.1001/jama.291.17.2100.
[515] B. Noehr, A. Jensen, K. Frederiksen, A. Tabor, and S. K. Kjaer, 'Depth of cervical cone removed by loop electrosurgical excision procedure and subsequent risk of spontaneous preterm delivery', *Obstet. Gynecol.*, vol. 114, no. 6, pp. 1232–1238, Dec. 2009, doi: 10.1097/AOG.0b013e3181bf1ef2.
[516] J. A. Svare *et al.*, 'The relationship between prior cervical conization, cervical microbial colonization and preterm premature rupture of the membranes', *Eur. J. Obstet. Gynecol. Reprod. Biol.*, vol. 47, no. 1, pp. 41–45, Oct. 1992, doi: 10.1016/0028-2243(92)90212-h.
[517] M. Arbyn *et al.*, 'Perinatal mortality and other severe adverse pregnancy outcomes associated with treatment of cervical intraepithelial neoplasia: meta-analysis', *BMJ*, vol. 337, p. a1284, Sep. 2008, doi: 10.1136/bmj.a1284.
[518] R. L. Fischer, G. Sveinbjornsson, and C. Hansen, 'Cervical sonography in pregnant women with a prior cone biopsy or loop electrosurgical excision procedure', *Ultrasound Obstet. Gynecol.*, vol. 36, no. 5, pp. 613–617, 2010, doi: 10.1002/uog.7682.
[519] S.-W. Wang, L.-L. Ma, S. Huang, L. Liang, and J.-R. Zhang, 'Role of Cervical Cerclage and Vaginal Progesterone in the Treatment of Cervical Incompetence with/without Preterm Birth History', *Chin. Med. J. (Engl.)*, vol. 129, no. 22, pp. 2670–2675, Nov. 2016, doi: 10.4103/0366-6999.193451.
[520] M. Grabovac, A. M. Lewis-Mikhael, and S. D. McDonald, 'Interventions to Try to Prevent Preterm Birth in Women With a History of Conization: A Systematic Review and Meta-analyses', *J. Obstet. Gynaecol. Can.*, vol. 41, no. 1, pp. 76-88.e7, Jan. 2019, doi: 10.1016/j.jogc.2018.04.026.
[521] D. Nasioudis *et al.*, 'Influence of Pregnancy History on the Vaginal Microbiome of Pregnant Women in their First Trimester', *Sci. Rep.*, vol. 7, p. 10201, Aug. 2017, doi: 10.1038/s41598-017-09857-z.
[522] K. Kervinen *et al.*, 'Parity and gestational age are associated with vaginal microbiota composition in term and late term pregnancies', *eBioMedicine*, vol. 81, Jul. 2022, doi: 10.1016/j.ebiom.2022.104107.
[523] R. Romero *et al.*, 'The vaginal microbiota of pregnant women varies with gestational age, maternal age, and parity'. bioRxiv, p. 2023.02.26.530121, Feb. 27, 2023. doi: 10.1101/2023.02.26.530121.
[524] X. Zhou *et al.*, 'Differences in the composition of vaginal microbial communities found in healthy Caucasian and black women', *ISME J.*, vol. 1, no. 2, Art. no. 2, Jun. 2007, doi: 10.1038/ismej.2007.12.
[525] S. Sun *et al.*, 'Race, the Vaginal Microbiome, and Spontaneous Preterm Birth', *mSystems*, vol. 7, no. 3, pp. e00017-22, May 2022, doi: 10.1128/msystems.00017-22.
[526] E. Gulavi *et al.*, 'Vaginal microbiota in women with spontaneous preterm labor versus those with term labor in Kenya: a case control study', *BMC Microbiol.*, vol. 22, no. 1, p. 270, Nov. 2022, doi: 10.1186/s12866-022-02681-0.
[527] N. C. A. Juliana, R. P. H. Peters, S. Al-Nasiry, A. E. Budding, S. A. Morré, and E. Ambrosino, 'Composition of the vaginal microbiota during pregnancy in women living in sub-Saharan Africa: a PRISMA-compliant review', *BMC Pregnancy Childbirth*, vol. 21, p. 596, Sep. 2021, doi: 10.1186/s12884-021-04072-1.

[528] ‘Ethnic group, England and Wales - Office for National Statistics’. https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/bulletins/ethnicgroupenglandandwales/census2021 (accessed May 20, 2023).
[529] K. R. Van Daalen *et al.*, ‘Racial discrimination and adverse pregnancy outcomes: a systematic review and meta-analysis’, *BMJ Glob. Health*, vol. 7, no. 8, p. e009227, Aug. 2022, doi: 10.1136/bmjgh-2022-009227.
[530] S. M. Yellon, ‘Contributions to the dynamics of cervix remodeling prior to term and preterm birth’, *Biol. Reprod.*, vol. 96, no. 1, pp. 13–23, Jan. 2017, doi: 10.1095/biolreprod.116.142844.
[531] E. R. Guzman, C. Mellon, A. M. Vintzileos, C. V. Ananth, C. Walters, and K. Gipson, ‘Longitudinal assessment of endocervical canal length between 15 and 24 weeks’ gestation in women at risk for pregnancy loss or preterm birth’, *Obstet. Gynecol.*, vol. 92, no. 1, pp. 31–37, Jul. 1998, doi: 10.1016/s0029-7844(98)00120-3.
[532] J. D. Iams, D. Cebrik, C. Lynch, N. Behrendt, and A. Das, ‘The rate of cervical change and the phenotype of spontaneous preterm birth’, *Am. J. Obstet. Gynecol.*, vol. 205, no. 2, p. 130.e1–6, Aug. 2011, doi: 10.1016/j.ajog.2011.05.021.
[533] M. Thakur and K. Mahajan, ‘Cervical Insufficiency’, in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2023. Accessed: Jun. 20, 2023. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK525954/
[534] J.-M. Moutquin, ‘Classification and heterogeneity of preterm birth’, *BJOG Int. J. Obstet. Gynaecol.*, vol. 110, no. s20, pp. 30–33, 2003, doi: 10.1046/j.1471-0528.2003.00021.x.
[535] S. Mesiano and T. Welsh, ‘PROGESTERONE AND THE CONTROL OF HUMAN PREGNANCY AND PARTURITION’, *Fetal Matern. Med. Rev.*, vol. 20, no. 2, pp. 97–118, May 2009, doi: 10.1017/S096553950900240X.
[536] K. C. Williams, N. E. Renthal, J. C. Condon, R. D. Gerard, and C. R. Mendelson, ‘MicroRNA-200a serves a key role in the decline of progesterone receptor function leading to term and preterm labor’, *Proc. Natl. Acad. Sci.*, vol. 109, no. 19, pp. 7529–7534, May 2012, doi: 10.1073/pnas.1200650109.
[537] W. S. Maxson and J. T. Hargrove, ‘Bioavailability of oral micronized progesterone’, *Fertil. Steril.*, vol. 44, no. 5, pp. 622–626, Nov. 1985, doi: 10.1016/S0015-0282(16)48977-6.
[538] J. T. Hargrove, W. S. Maxson, and A. C. Wentz, ‘Absorption of oral progesterone is influenced by vehicle and particle size’, *Am. J. Obstet. Gynecol.*, vol. 161, no. 4, pp. 948–951, 1989, doi: 10.1016/0002-9378(89)90759-X.
[539] FIGO Working Group on Best Practice in Maternal-Fetal Medicine, ‘Best practice in maternal-fetal medicine’, *Int. J. Gynecol. Obstet.*, vol. 128, no. 1, pp. 80–82, Jan. 2015, doi: 10.1016/j.ijgo.2014.10.011.
[540] E. B. da Fonseca, R. E. Bittar, M. H. B. Carvalho, and M. Zugaib, ‘Prophylactic administration of progesterone by vaginal suppository to reduce the incidence of spontaneous preterm birth in women at increased risk: A randomized placebo-controlled double-blind study’, *Am. J. Obstet. Gynecol.*, vol. 188, no. 2, pp. 419–424, Feb. 2003, doi: 10.1067/mob.2003.41.
[541] J. M. Dodd, L. Jones, V. Flenady, R. Cincotta, and C. A. Crowther, ‘Prenatal administration of progesterone for preventing preterm birth in women considered to be at risk of preterm birth’, *Cochrane Database Syst. Rev.*, no. 7, 2013, doi: 10.1002/14651858.CD004947.pub3.
[542] J. E. Norman *et al.*, ‘Vaginal progesterone prophylaxis for preterm birth (the OPPTIMUM study): a multicentre, randomised, double-blind trial’, *The Lancet*, vol. 387, no. 10033, pp. 2106–2116, May 2016, doi: 10.1016/S0140-6736(16)00350-0.

[543] C. A. Crowther *et al.*, ‘Vaginal progesterone pessaries for pregnant women with a previous preterm birth to prevent neonatal respiratory distress syndrome (the PROGRESS Study): A multicentre, randomised, placebo-controlled trial’, *PLOS Med.*, vol. 14, no. 9, p. e1002390, Sep. 2017, doi: 10.1371/journal.pmed.1002390.
[544] ‘ACOG Statement on FDA Committee Recommendation to Withdraw 17p Hydroxyprogesterone Caproate’. https://www.acog.org/en/news/news-releases/2022/10/statement-on-fda-committee-recommendation-to-withdraw-17p-hydroxyprogesterone-caproate (accessed Jun. 26, 2023).
[545] J. E. Norman *et al.*, ‘Effect of prolonged in vivo administration of progesterone in pregnancy on myometrial gene expression, peripheral blood leukocyte activation, and circulating steroid hormone levels’, *Reprod. Sci. Thousand Oaks Calif*, vol. 18, no. 5, pp. 435–446, May 2011, doi: 10.1177/1933719110395404.
[546] E. R. Norwitz and A. B. Caughey, ‘Progesterone Supplementation and the Prevention of Preterm Birth’, *Rev. Obstet. Gynecol.*, vol. 4, no. 2, pp. 60–72, 2011.
[547] S. Mesiano, ‘Myometrial Progesterone Responsiveness’, *Semin. Reprod. Med.*, vol. 25, no. 1, pp. 5–13, Jan. 2007, doi: 10.1055/s-2006-956771.
[548] T. Feng *et al.*, ‘The association between serum progesterone level and preterm delivery’, *Int. J. Gynaecol. Obstet. Off. Organ Int. Fed. Gynaecol. Obstet.*, vol. 142, no. 3, pp. 308–314, Sep. 2018, doi: 10.1002/ijgo.12564.
[549] B. R. Carr, C. R. Parker, J. D. Madden, P. C. MacDonald, and J. C. Porter, ‘Maternal plasma adrenocorticotropin and cortisol relationships throughout human pregnancy’, *Am. J. Obstet. Gynecol.*, vol. 139, no. 4, pp. 416–422, Feb. 1981, doi: 10.1016/0002-9378(81)90318-5.
[550] W. Chaim and M. Mazor, ‘The relationship between hormones and human parturition’, *Arch. Gynecol. Obstet.*, vol. 262, no. 1, pp. 43–51, Nov. 1998, doi: 10.1007/s004040050226.
[551] F. A. Potestio, T. Zakar, and D. M. Olson, ‘Glucocorticoids stimulate prostaglandin synthesis in human amnion cells by a receptor-mediated mechanism’, *J. Clin. Endocrinol. Metab.*, vol. 67, no. 6, pp. 1205–1210, Dec. 1988, doi: 10.1210/jcem-67-6-1205.
[552] S. Makieva, P. T. K. Saunders, and J. E. Norman, ‘Androgens in pregnancy: roles in parturition’, *Hum. Reprod. Update*, vol. 20, no. 4, pp. 542–559, Jul. 2014, doi: 10.1093/humupd/dmu008.
[553] R. Smith, ‘Parturition’, *N. Engl. J. Med.*, vol. 356, no. 3, pp. 271–283, Jan. 2007, doi: 10.1056/NEJMra061360.
[554] M. Mochizuki and T. Maruo, ‘Effect of dehydroepiandrosterone sulfate on uterine cervical ripening in late pregnancy’, *Acta Physiol. Hung.*, vol. 65, no. 3, pp. 267–274, 1985.
[555] M. Mochizuki, T. Honda, M. Deguchi, H. Morikawa, and S. Tojo, ‘A study on the effect of dehydroepiandrosterone sulfate on so-called cervical ripening’, *Acta Obstet. Gynecol. Scand.*, vol. 57, no. 5, pp. 397–401, 1978, doi: 10.3109/00016347809156518.
[556] M. Perusquía, E. Navarrete, J. Jasso-Kamel, and L. M. Montaño, ‘Androgens induce relaxation of contractile activity in pregnant human myometrium at term: a nongenomic action on L-type calcium channels’, *Biol. Reprod.*, vol. 73, no. 2, pp. 214–221, Aug. 2005, doi: 10.1095/biolreprod.104.036954.
[557] L. M. Montaño, E. Calixto, A. Figueroa, E. Flores-Soto, V. Carbajal, and M. Perusquía, ‘Relaxation of androgens on rat thoracic aorta: testosterone concentration dependent agonist/antagonist L-type Ca2+ channel activity, and 5beta-dihydrotestosterone restricted to L-type Ca2+ channel blockade’, *Endocrinology*, vol. 149, no. 5, pp. 2517–2526, May 2008, doi: 10.1210/en.2007-1288.

[558] V. Kouloumenta, A. Hatziefthimiou, E. Paraskeva, K. Gourgoulianis, and P. A. Molyvdas, 'Non-genomic effect of testosterone on airway smooth muscle', *Br. J. Pharmacol.*, vol. 149, no. 8, pp. 1083–1091, Dec. 2006, doi: 10.1038/sj.bjp.0706936.
[559] J. Bordallo *et al.*, 'Modulatory role of endogenous androgens on airway smooth muscle tone in isolated guinea-pig and bovine trachea; involvement of beta2-adrenoceptors, the polyamine system and external calcium', *Eur. J. Pharmacol.*, vol. 601, no. 1–3, pp. 154–162, Dec. 2008, doi: 10.1016/j.ejphar.2008.10.039.
[560] M. Chevalier, G. Gilbert, P. Lory, R. Marthan, J. F. Quignard, and J. P. Savineau, 'Dehydroepiandrosterone (DHEA) inhibits voltage-gated T-type calcium channels', *Biochem. Pharmacol.*, vol. 83, no. 11, pp. 1530–1539, Jun. 2012, doi: 10.1016/j.bcp.2012.02.025.
[561] M. Perusquía, G. Campos, J. L. Corona, and C. Kubli-Garfias, 'Antagonism by 5-reduced steroids of the tonic and phasic contractions induced by serotonin in the isolated rat uterus', *Proc. West. Pharmacol. Soc.*, vol. 34, pp. 395–398, 1991.
[562] Y. Lisnawati and N. Wibowo, 'Correlation of cervical progesterone levels to plasma progesterone levels in normal pregnancy and preterm labor: A cross-sectional study', *F1000Research*, vol. 8, p. 189, Oct. 2022, doi: 10.12688/f1000research.17525.2.
[563] F. Z. Stanczyk and N. J. Clarke, 'Advantages and challenges of mass spectrometry assays for steroid hormones', *J. Steroid Biochem. Mol. Biol.*, vol. 121, no. 3, pp. 491–495, Aug. 2010, doi: 10.1016/j.jsbmb.2010.05.001.
[564] J. M. Boyd and S. M. H. Sadrzadeh, 'Chapter 14 - Limitations of immunoassays for screening of drugs of abuse in urine: issues of false positive and false negative results', in *Accurate Results in the Clinical Laboratory (Second Edition)*, A. Dasgupta and J. L. Sepulveda, Eds., Elsevier, 2019, pp. 233–242. doi: 10.1016/B978-0-12-813776-5.00014-5.
[565] S. J. Soldin and O. P. Soldin, 'Steroid Hormone Analysis by Tandem Mass Spectrometry', *Clin. Chem.*, vol. 55, no. 6, pp. 1061–1066, Jun. 2009, doi: 10.1373/clinchem.2007.100008.
[566] B. G. Wolthers and G. P. B. Kraan, 'Clinical applications of gas chromatography and gas chromatography–mass spectrometry of steroids', *J. Chromatogr. A*, vol. 843, no. 1, pp. 247–274, May 1999, doi: 10.1016/S0021-9673(99)00153-3.
[567] R. J. Santen *et al.*, 'Superiority of gas chromatography/tandem mass spectrometry assay (GC/MS/MS) for estradiol for monitoring of aromatase inhibitor therapy', *Steroids*, vol. 72, no. 8, pp. 666–671, Jul. 2007, doi: 10.1016/j.steroids.2007.05.003.
[568] T. Guo, M. Chan, and S. J. Soldin, 'Steroid profiles using liquid chromatography-tandem mass spectrometry with atmospheric pressure photoionization source', *Arch. Pathol. Lab. Med.*, vol. 128, no. 4, pp. 469–475, Apr. 2004, doi: 10.5858/2004-128-469-SPULCM.
[569] T. Guo, R. L. Taylor, R. J. Singh, and S. J. Soldin, 'Simultaneous determination of 12 steroids by isotope dilution liquid chromatography-photospray ionization tandem mass spectrometry', *Clin. Chim. Acta Int. J. Clin. Chem.*, vol. 372, no. 1–2, pp. 76–82, Oct. 2006, doi: 10.1016/j.cca.2006.03.034.
[570] X. Xu, J. M. Roman, H. J. Issaq, L. K. Keefer, T. D. Veenstra, and R. G. Ziegler, 'Quantitative measurement of endogenous estrogens and estrogen metabolites in human serum by liquid chromatography-tandem mass spectrometry', *Anal. Chem.*, vol. 79, no. 20, pp. 7813–7821, Oct. 2007, doi: 10.1021/ac070494j.
[571] R. E. Nelson, S. K. Grebe, D. J. OKane, and R. J. Singh, 'Liquid chromatography-tandem mass spectrometry assay for simultaneous measurement of estradiol and estrone in human plasma', *Clin. Chem.*, vol. 50, no. 2, pp. 373–384, Feb. 2004, doi: 10.1373/clinchem.2003.025478.

[572] R. Draisci, L. Palleschi, E. Ferretti, L. Lucentini, and P. Cammarata, ‘Quantitation of anabolic hormones and their metabolites in bovine serum and urine by liquid chromatography–tandem mass spectrometry’, *J. Chromatogr. A*, vol. 870, no. 1, pp. 511–522, Feb. 2000, doi: 10.1016/S0021-9673(99)01293-5.
[573] J. Liu *et al.*, ‘Quantification of 10 steroid hormones in human saliva from Chinese adult volunteers’, *J. Int. Med. Res.*, vol. 46, no. 4, pp. 1414–1427, Apr. 2018, doi: 10.1177/0300060517752733.
[574] S. K. Kota *et al.*, ‘Endocrinology of parturition’, *Indian J. Endocrinol. Metab.*, vol. 17, no. 1, pp. 50–59, 2013, doi: 10.4103/2230-8210.107841.
[575] Q. Liu, Q. Chi, R.-T. Fan, H.-D. Tian, and X. Wang, ‘Quantitative-Profiling Method of Serum Steroid Hormones by Hydroxylamine-Derivatization HPLC–MS’, *Nat. Prod. Bioprospecting*, vol. 9, no. 3, pp. 201–208, Apr. 2019, doi: 10.1007/s13659-019-0204-3.
[576] G. Huang, H. Chen, X. Zhang, R. G. Cooks, and Z. Ouyang, ‘Rapid Screening of Anabolic Steroids in Urine by Reactive Desorption Electrospray Ionization’, *Anal. Chem.*, vol. 79, no. 21, pp. 8327–8332, Nov. 2007, doi: 10.1021/ac0711079.
[577] H. Schock *et al.*, ‘Hormone concentrations throughout uncomplicated pregnancies: a longitudinal study’, *BMC Pregnancy Childbirth*, vol. 16, p. 146, Jul. 2016, doi: 10.1186/s12884-016-0937-5.
[578] null Van Molenh, ‘DETERMINATION OF PLASMA PROGESTERONE DURING PREGNANCY’, *Clin. Chim. Acta Int. J. Clin. Chem.*, vol. 8, pp. 943–953, Nov. 1963, doi: 10.1016/0009-8981(63)90019-6.
[579] T. Vrijkotte, S. de Rooij, T. Roseboom, and T. Twickler, ‘Maternal serum cortisol levels during pregnancy differ by fetal sex’, *Psychoneuroendocrinology*, vol. 149, p. 105999, Mar. 2023, doi: 10.1016/j.psyneuen.2022.105999.
[580] J. F. Dorgan, T. R. Fears, R. P. McMahon, L. Aronson Friedman, B. H. Patterson, and S. F. Greenhut, ‘Measurement of steroid sex hormones in serum: a comparison of radioimmunoassay and mass spectrometry’, *Steroids*, vol. 67, no. 3, pp. 151–158, Mar. 2002, doi: 10.1016/S0039-128X(01)00147-7.
[581] J. C. Hauth, R. L. Goldenberg, W. W. Andrews, M. B. DuBard, and R. L. Copper, ‘Reduced incidence of preterm delivery with metronidazole and erythromycin in women with bacterial vaginosis’, *N. Engl. J. Med.*, vol. 333, no. 26, pp. 1732–1736, Dec. 1995, doi: 10.1056/NEJM199512283332603.
[582] H. M. McDonald *et al.*, ‘Impact of metronidazole therapy on preterm birth in women with bacterial vaginosis flora (Gardnerella vaginalis): a randomised, placebo controlled trial’, *Br. J. Obstet. Gynaecol.*, vol. 104, no. 12, pp. 1391–1397, Dec. 1997, doi: 10.1111/j.1471-0528.1997.tb11009.x.
[583] R. F. Lamont, S. L. B. Duncan, D. Mandal, and P. Bassett, ‘Intravaginal clindamycin to reduce preterm birth in women with abnormal genital tract flora’, *Obstet. Gynecol.*, vol. 101, no. 3, pp. 516–522, Mar. 2003, doi: 10.1016/s0029-7844(02)03054-5.
[584] S. Hantoushzadeh, R. Anvari Aliabad, and A. H. Norooznezhad, ‘Antibiotics, Inflammation, and Preterm Labor: A Missed Conclusion’, *J. Inflamm. Res.*, vol. 13, pp. 245–254, May 2020, doi: 10.2147/JIR.S248382.
[585] R. F. Lamont, C.-L. Nhan-Chang, J. D. Sobel, K. Workowski, A. Conde-Agudelo, and R. Romero, ‘Treatment of abnormal vaginal flora in early pregnancy with clindamycin for the prevention of spontaneous preterm birth: a systematic review and metaanalysis’, *Am. J. Obstet. Gynecol.*, vol. 205, no. 3, pp. 177–190, Sep. 2011, doi: 10.1016/j.ajog.2011.03.047.

[586] P. Brocklehurst, A. Gordon, E. Heatley, and S. J. Milan, ‘Antibiotics for treating bacterial vaginosis in pregnancy’, *Cochrane Database Syst. Rev.*, no. 1, p. CD000262, Jan. 2013, doi: 10.1002/14651858.CD000262.pub4.
[587] A. M. VanHook, ‘Antibiotic-induced inflammation’, *Sci. Signal.*, vol. 15, no. 745, p. eade1683, Aug. 2022, doi: 10.1126/scisignal.ade1683.
[588] G. M. Vermeulen and H. W. Bruinse, ‘Prophylactic administration of clindamycin 2% vaginal cream to reduce the incidence of spontaneous preterm birth in women with an increased recurrence risk: a randomised placebo-controlled double-blind trial’, *Br. J. Obstet. Gynaecol.*, vol. 106, no. 7, pp. 652–657, Jul. 1999, doi: 10.1111/j.1471-0528.1999.tb08363.x.
[589] V. Flenady, G. Hawley, O. M. Stock, S. Kenyon, and N. Badawi, ‘Prophylactic antibiotics for inhibiting preterm labour with intact membranes’, *Cochrane Database Syst. Rev.*, no. 12, p. CD000246, Dec. 2013, doi: 10.1002/14651858.CD000246.pub2.
[590] H. E. Jakobsson, C. Jernberg, A. F. Andersson, M. Sjölund-Karlsson, J. K. Jansson, and L. Engstrand, ‘Short-term antibiotic treatment has differing long-term impacts on the human throat and gut microbiome’, *PloS One*, vol. 5, no. 3, p. e9836, Mar. 2010, doi: 10.1371/journal.pone.0009836.
[591] Y. H. Shim, S. J. Lee, and J. W. Lee, ‘Antimicrobial activity of lactobacillus strains against uropathogens’, *Pediatr. Int.*, vol. 58, no. 10, pp. 1009–1013, 2016, doi: 10.1111/ped.12949.
[592] F. A. Ayeni *et al.*, ‘Inhibition of uropathogens by lactic acid bacteria isolated from dairy foods and cow’s intestine in western Nigeria’, *Arch. Microbiol.*, vol. 191, no. 8, pp. 639–648, Aug. 2009, doi: 10.1007/s00203-009-0492-9.
[593] M. Gotteland *et al.*, ‘Probiotic screening and safety evaluation of Lactobacillus strains from plants, artisanal goat cheese, human stools, and breast milk’, *J. Med. Food*, vol. 17, no. 4, pp. 487–495, Apr. 2014, doi: 10.1089/jmf.2013.0030.
[594] Z. Mei and D. Li, ‘The role of probiotics in vaginal health’, *Front. Cell. Infect. Microbiol.*, vol. 12, p. 963868, Jul. 2022, doi: 10.3389/fcimb.2022.963868.
[595] E. Yefet, R. Colodner, M. Strauss, Y. Gam Ze Letova, and Z. Nachum, ‘A Randomized Controlled Open Label Crossover Trial to Study Vaginal Colonization of Orally Administered Lactobacillus Reuteri RC-14 and Rhamnosus GR-1 in Pregnant Women at High Risk for Preterm Labor’, *Nutrients*, vol. 12, no. 4, p. 1141, Apr. 2020, doi: 10.3390/nu12041141.
[596] S. Yang *et al.*, ‘Effect of Oral Probiotic Lactobacillus rhamnosus GR-1 and Lactobacillus reuteri RC-14 on the Vaginal Microbiota, Cytokines and Chemokines in Pregnant Women’, *Nutrients*, vol. 12, no. 2, p. 368, Jan. 2020, doi: 10.3390/nu12020368.
[597] M. A. D. Antonio and S. L. Hillier, ‘DNA Fingerprinting of Lactobacillus crispatus Strain CTV-05 by Repetitive Element Sequence-Based PCR Analysis in a Pilot Study of Vaginal Colonization’, *J. Clin. Microbiol.*, vol. 41, no. 5, pp. 1881–1887, May 2003, doi: 10.1128/JCM.41.5.1881-1887.2003.
[598] M. A. Antonio, S. E. Hawes, and S. L. Hillier, ‘The identification of vaginal Lactobacillus species and the demographic and microbiologic characteristics of women colonized by these species’, *J. Infect. Dis.*, vol. 180, no. 6, pp. 1950–1956, Dec. 1999, doi: 10.1086/315109.
[599] L. Kwok, A. E. Stapleton, W. E. Stamm, S. L. Hillier, C. L. Wobbe, and K. Gupta, ‘Adherence of Lactobacillus crispatus to vaginal epithelial cells from women with or without a history of recurrent urinary tract infection’, *J. Urol.*, vol. 176, no. 5, pp. 2050–2054; discussion 2054, Nov. 2006, doi: 10.1016/j.juro.2006.07.014.

[600] A. Hemmerling *et al.*, ‘Phase 2a study assessing colonization efficiency, safety, and acceptability of Lactobacillus crispatus CTV-05 in women with bacterial vaginosis’, *Sex. Transm. Dis.*, vol. 37, no. 12, pp. 745–750, Dec. 2010, doi: 10.1097/OLQ.0b013e3181e50026.
[601] L. A. Lagenaur *et al.*, ‘Connecting the Dots: Translating the Vaginal Microbiome Into a Drug’, *J. Infect. Dis.*, vol. 223, no. Suppl 3, pp. S296–S306, Dec. 2020, doi: 10.1093/infdis/jiaa676.
[602] H. A, W. Mr, C. G, and C. Cr, ‘Response to antibiotic treatment of bacterial vaginosis predicts the effectiveness of LACTIN-v (lactobacillus crispatus CTV-05) in the prevention of recurrent disease’, *Am. J. Obstet. Gynecol.*, vol. 228, no. 2, p. S777, Feb. 2023, doi: 10.1016/j.ajog.2022.11.123.
[603] E. Armstrong *et al.*, ‘Sustained effect of LACTIN-V (Lactobacillus crispatus CTV-05) on genital immunology following standard bacterial vaginosis treatment: results from a randomised, placebo-controlled trial’, *Lancet Microbe*, vol. 3, no. 6, pp. e435–e442, Jun. 2022, doi: 10.1016/S2666-5247(22)00043-X.
[604] C. Henry and P. D. Rainville, ‘Demonstrating the Waters ACQUITY RDa Detector as an Ease of Use Solution for Routine Accurate Mass Measurements in Forced Degradation Studies’.
[605] waters, ‘ACQUITY RDa Detector and RemoteAnalyzer for Walk-Up Accurate Mass Confirmation Studies’, *www.waters.com*. http://www.waters.com/waters/library.htm?cid=135077027&lid=135087878 (accessed Jun. 22, 2023).
[606] Z. Yan, C. Cheng, and S. Liu, ‘Applications of Mass Spectrometry in Analyses of Steroid Hormones’, in *LC-MS in Drug Bioanalysis*, Q. A. Xu and T. L. Madden, Eds., Boston, MA: Springer US, 2012, pp. 251–286. doi: 10.1007/978-1-4614-3828-1_10.
[607] E. Bayar *et al.*, *Safety, tolerability and acceptability of Lactobacillus crispatus CTV-05 (LACTIN-V) in pregnant women at high-risk of preterm birth*. 2022. doi: 10.21203/rs.3.rs-1418870/v1.
[608] A. Kaufmann and S. Walker, ‘Comparison of linear intrascan and interscan dynamic ranges of Orbitrap and ion-mobility time-of-flight mass spectrometers’, *Rapid Commun. Mass Spectrom.*, vol. 31, no. 22, pp. 1915–1926, 2017, doi: 10.1002/rcm.7981.
[609] M. A. D. Antonio, L. A. Meyn, P. J. Murray, B. Busse, and S. L. Hillier, ‘Vaginal colonization by probiotic Lactobacillus crispatus CTV-05 is decreased by sexual activity and endogenous Lactobacilli’, *J. Infect. Dis.*, vol. 199, no. 10, pp. 1506–1513, May 2009, doi: 10.1086/598686.
[610] B. M. Ngugi *et al.*, ‘Effects of bacterial vaginosis-associated bacteria and sexual intercourse on vaginal colonization with the probiotic Lactobacillus crispatus CTV-05’, *Sex. Transm. Dis.*, vol. 38, no. 11, pp. 1020–1027, Nov. 2011, doi: 10.1097/OLQ.0b013e3182267ac4.
[611] A. Ito *et al.*, ‘Human recombinant interleukin-1α-mediated stimulation of procollagenase production and suppression of biosynthesis of tissue inhibitor of metalloproteinases in rabbit uterine cervical fibroblasts’, *FEBS Lett.*, vol. 234, no. 2, pp. 326–330, 1988, doi: 10.1016/0014-5793(88)80109-1.
[612] N. Kanayama and T. Terao, ‘The relationship between granulocyte elastase-like activity of cervical mucus and cervical maturation’, *Acta Obstet. Gynecol. Scand.*, vol. 70, no. 1, pp. 29–34, 1991, doi: 10.3109/00016349109006174.
[613] K. Puchner *et al.*, ‘Mid-trimester Amniotic Fluid Interleukins (IL-1β, IL-10 and IL-18) as Possible Predictors of Preterm Delivery’, *In Vivo*, vol. 25, no. 1, pp. 141–148, Jan. 2011.

[614] R. Figueroa, D. Garry, A. Elimian, K. Patel, P. B. Sehgal, and N. Tejani, ‘Evaluation of amniotic fluid cytokines in preterm labor and intact membranes’, *J. Matern.-Fetal Neonatal Med. Off. J. Eur. Assoc. Perinat. Med. Fed. Asia Ocean. Perinat. Soc. Int. Soc. Perinat. Obstet.*, vol. 18, no. 4, pp. 241–247, Oct. 2005, doi: 10.1080/13506120500223241.
[615] M. Kacerovsky *et al.*, ‘Value of amniotic fluid interleukin-8 for the prediction of histological chorioamnionitis in preterm premature rupture of membranes’, *Neuro Endocrinol. Lett.*, vol. 30, no. 6, pp. 733–738, 2009.
[616] R. Romero, M. Ceska, C. Avila, M. Mazor, E. Behnke, and I. Lindley, ‘Neutrophil attractant/activating peptide-1 / interleukin-8 in term and preterm parturition’, *Am. J. Obstet. Gynecol.*, vol. 165, no. 4, pp. 813–820, Oct. 1991, doi: 10.1016/0002-9378(91)90422-N.
[617] U.-B. Wennerholm *et al.*, ‘Interleukin-1α, interleukin-6 and interleukin-8 in cervico/vaginal secretion for screening of preterm birth in twin gestation’, *Acta Obstet. Gynecol. Scand.*, vol. 77, no. 5, pp. 508–514, Jan. 1998, doi: 10.1080/j.1600-0412.1998.770507.x.
[618] A.-C. Morillon *et al.*, ‘Association between phospholipid metabolism in plasma and spontaneous preterm birth: a discovery lipidomic analysis in the cork pregnancy cohort’, *Metabolomics*, vol. 16, no. 2, p. 19, Feb. 2020, doi: 10.1007/s11306-020-1639-6.
[619] M. T. Aung *et al.*, ‘Maternal lipidomic signatures in relation to spontaneous preterm birth and large-for-gestational age neonates’, *Sci. Rep.*, vol. 11, no. 1, Art. no. 1, Apr. 2021, doi: 10.1038/s41598-021-87472-9.
[620] P. B. Luppa, C. Müller, A. Schlichtiger, and H. Schlebusch, ‘Point-of-care testing (POCT): Current techniques and future perspectives’, *TrAC Trends Anal. Chem.*, vol. 30, no. 6, pp. 887–898, Jun. 2011, doi: 10.1016/j.trac.2011.01.019.
[621] N. Yared, ‘The Role of Point-of-Care Testing in Specific Populations’, *Clin. Lab. Med.*, vol. 43, no. 2, pp. 181–187, Jun. 2023, doi: 10.1016/j.cll.2023.02.004.
[622] J. M. Kinder *et al.*, ‘CD8+ T Cell Functional Exhaustion Overrides Pregnancy-Induced Fetal Antigen Alloimmunization’, *Cell Rep.*, vol. 31, no. 12, p. 107784, Jun. 2020, doi: 10.1016/j.celrep.2020.107784.
[623] S. M. Leow *et al.*, ‘Preterm birth prediction in asymptomatic women at mid-gestation using a panel of novel protein biomarkers: the Prediction of PreTerm Labor (PPeTaL) study’, *Am. J. Obstet. Gynecol. MFM*, vol. 2, no. 2, p. 100084, May 2020, doi: 10.1016/j.ajogmf.2019.100084.
[624] L. Di Salvo, T. Cade, P. Sheehan, H. M. Georgiou, M. Di Quinzio, and S. P. Brennecke, ‘Identification of biochemical biomarkers associated with premature cervical shortening in high-risk, asymptomatic pregnant women: a retrospective data analysis’, *J. Obstet. Gynaecol.*, vol. 43, no. 1, p. 2212299, Dec. 2023, doi: 10.1080/01443615.2023.2212299.
[625] L. Ford *et al.*, ‘Rapid detection of SARS-CoV2 by Ambient Mass Spectrometry Techniques’. medRxiv, p. 2020.10.07.20207647, Oct. 12, 2020. doi: 10.1101/2020.10.07.20207647.

Appendix

Appendix table 1 - Statistically significant metabolite features found associated with microbial composition in the VMET2 patient cohort with reported ppm mass error, putative annotation, MS/MS, and sub class of metabolic compounds.

Appendix table 1 – Complete list of the 113 m/z features found significantly different in *Lactobacillus* dominance and depletion groups.

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	117.0562	117.0561	117.0557	-4.2	-3.78	Hydroxyvaleric acid	C5H10O3	M-H-	NO	3	Fatty acids and conjugates	LDEPL	8.78E-11	7.6	8.60E-21	10.8
-	118.0231	118.0231	118.0227	-3.82	-3.38	Isotope Succinic acid	(13C)C3H6O4	M-H-	NO	3	Dicarboxylic acids and derivatives	LDEPL	7.31E-03	3.4	5.70E-04	4.2
-	128.0357	128.0351	128.0353	-2.78	1.28	Oxoproline	C5H7NO3	M-H-	YES	2	Amino acids, peptides, analogues	LDOM	1.36E-02	-3.2	6.91E-07	-5.9
-	129.1035	129.1036	129.1033	-1.74	-2.52	N-Acetylputrescine	C6H14N2O	M-H-	YES	2	Carboximidic acids	LDOM	1.24E-06	-5.8	6.84E-03	-3.4
-	130.0877	130.0877	130.0874	-2.22	-2.54	Leucine	C6H13NO2	M-H-	YES	2	Amino acids, peptides, analogues	LDOM	3.03E-04	-4.4	2.13E-03	-3.8

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	131.0465	131.0466	131.0462	-2.62	-2.84	Asparagine	C4H8N2O3	M-H-	NO	3	Amino acids, peptides, analogues	LDOM	9.26E-05	-4.7	4.22E-03	-3.6
-	131.0717	131.0717	131.0714	-2.47	-2.05	Hydroxycaproic acid	C6H12O3	M-H-	YES	2	Fatty acids and conjugates	LDEPL	2.75E-14	9.1	4.63E-23	11.8
-	132.0306	132.0306	132.0302	-2.73	-2.96	Aspartic acid	C4H7NO4	M-H-	YES	2	Amino acids, peptides, analogues	LDOM	2.49E-03	-3.8	1.90E-05	-5.1
-	145.0622	145.0621	145.0619	-1.96	-1.72	Glutamine	C5H10N2O3	M-H-	NO	3	Amino acids, peptides, analogues	LDOM	4.34E-04	-4.3	1.05E-02	-3.2
-	146.0462	146.0462	146.0459	-2.11	-1.84	Glutamate	C5H9NO4	M-H-	YES	2	Amino acids, peptides, analogues	LDOM	9.43E-03	-3.3	6.99E-07	-5.9

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	147.0302	147.0302	147.0299	-2.32	-2.27	Hydroxyglutaric acid	C5H8O5	M-H-	YES	2	Beta hydroxy acids and derivatives	LDEPL	4.64E-04	4.3	1.21E-02	3.2
-	148.9915	148.9915	148.9914	-0.51	-0.87	Thiomalic acid	C4H6O4S	M-H-	YES	2		LDOM	1.31E-08	-6.8	1.70E-06	-5.6
-	151.0264	151.0264	151.0261	-2.22	-2.26	Oxypurinol	C5H4N4O2	M-H-	YES	2	Purines and purine derivatives	LDEPL	2.11E-05	5.1	3.34E-05	4.9
-	164.0721	164.072	164.0717	-2.27	-1.74	Phenylalanine	C9H11NO2	M-H-	NO	3	Amino acids, peptides, and analogues	LDOM	6.99E-05	-4.8	1.99E-03	-3.8
-	165.056	165.0561	165.0557	-1.86	-2.22	Phenyllactic acid	C9H10O3	M-H-	YES	2		LDEPL	1.98E-03	3.8	4.22E-03	3.6
-	173.0934	173.0934	173.0932	-1.29	-1.13	N2-Acetyl-ornithine	C7H14N2O3	M-H-	YES	2	Amino acids, peptides, and analogues	LDOM	4.35E-05	-4.9	9.52E-05	-4.7

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	187.109	187.1089	187.1088	-1	-0.77	Valyl-Alanine	C8H16N2O3	M-H-	YES	2	Amino acids, peptides, and analogues	LDOM	1.68E-06	-5.7	1.11E-06	-5.7
-	201.0382	201.0382	201.038	-0.85	-0.87	Phenyl sulfoxide	C12H10OS	M-H-	NO	3		LDOM	2.94E-02	-2.9	8.60E-04	-4.1
-	**201.1246**	201.1243	201.1245	0.65	0.94	Leucyl-Alanine	C9H18N2O3	M-H-	YES	2	Amino acids, peptides, and analogues	LDOM	4.56E-03	-3.6	7.57E-04	-4.1
-	203.1039	203.1038	203.1037	-0.91	-0.28	Valyl-Serine	C8H16N2O4	M-H-	YES	2	Amino acids, peptides, and analogues	LDOM	1.05E-05	-5.2	1.08E-03	-4
-	217.1195	217.1196	217.1194	-0.59	-0.7	Leucyl-Serine	C9H18N2O4	M-H-	YES	2	Amino acids, peptides, and analogues	LDOM	4.38E-08	-6.5	3.92E-07	-6

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	221.0124	221.0123	221.0125	0.39	0.7	Carboxysulfanyl-pentanedioic acid	C7H10O6S	M-H-	YES	2		LDOM	1.13E-02	-3.2	4.65E-04	-4.3
-	221.0928	221.0934	221.0932	1.79	-0.72	Glycyl-Phenylalanine	C11H14N2O3	M-H-	YES	2	Amino acids, peptides, and analogues	LDOM	7.78E-03	-3.4	1.93E-02	-3
-	231.1351	231.1352	231.135	-0.41	-0.85	Leucyl-Threonine	C10H20N2O4	M-H-	YES	2	Amino acids, peptides, and analogues	LDOM	5.70E-04	-4.2	3.65E-03	-3.6
-	244.1303	244.1302	244.1303	-0.09	0.24	Leucyl-glycyl-glycine	C10H19N3O4	M-H-	YES	2	Amino acids, peptides, and analogues	LDOM	7.92E-04	-4.1	1.91E-03	-3.8
-	245.0432	245.0434	245.0432	-0.17	-0.83	Phosphatidyl glycerol	C6H15O8P	M-H-	YES	2		LDOM	9.63E-03	-3.3	2.61E-03	-3.7

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	245.1143	245.1145	245.1143	-0.12	-0.63	Valyl-Glutamine	C10H18N2O5	M-H-	YES	2	Amino acids, peptides, and analogues	LDOM	1.67E-02	-3.1	5.20E-03	-3.5
-	251.0773	251.0773	251.0772	-0.55	-0.22	Methoxymethyl methyl-glucopyranosiduronate	C9H16O8	M-H-	NO	3		LDOM	1.95E-03	-3.8	3.61E-05	-4.9
-	258.146	258.1459	258.1459	-0.39	0.09	Leucyl-Glutamine	C11H21N3O4	M-H-	YES	2	Amino acids, peptides, and analogues	LDOM	2.86E-04	-4.4	1.74E-04	-4.5
-	259.13	259.1299	259.1299	-0.39	0.05	Glutamyl-Leucine	C11H20N2O5	M-H-	YES	2	Amino acids, peptides, and analogues	LDOM	7.99E-03	-3.4	1.68E-03	-3.9

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	263.14	263.1401	263.1401	0.31	0.11	Valyl-Phenylalanine	C14H20N2O3	M-H-	YES	2	Amino acids, peptides, and analogues	LDOM	1.70E-03	-3.9	2.58E-02	-2.9
-	268.2366	268.2364	268.2364	-0.92	-0.15	Isotope Heptadecylenic acid (C17:1)	(13C)C16H32O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	4.74E-06	5.5	1.72E-04	4.5
-	277.2174	277.217	277.2173	-0.38	1.02	Linolenic Acid (C18:3)	C18H30O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	1.15E-02	3.2	1.01E-05	5.2
-	279.2329	279.2331	279.233	0.25	-0.29	Linoleic acid (C18:2)	C18H32O2	M-H-	YES	2	Fatty acids and conjugates	LDEPL	1.48E-02	3.1	1.08E-06	5.8
-	280.2364	280.2362	280.2364	-0.11	0.54	Isotope Linoleic acid (C18:2)	(13C)C17H32O2	M-H-	YES	2	Fatty acids and conjugates	LDEPL	1.23E-02	3.2	4.10E-06	5.4
-	281.2486	281.2487	281.2486	0.1	-0.46	Oleic Acid (C18:1)	C18H34O2	M-H-	YES	2	Fatty acids and conjugates	LDEPL	2.70E-03	3.7	3.92E-07	6

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	282.2519	282.2518	282.252	0.17	0.49	Isotope Oleic acid (18:1)	(13C)C17H34O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	1.50E-02	3.1	1.46E-03	3.9
-	283.2642	283.2643	283.2643	0.26	-0.08	Stearic acid (18:0)	C18H36O2	M-H-	YES	2	Fatty acids and conjugates	LDEPL	7.48E-03	3.4	1.52E-06	5.7
-	284.2675	284.2676	284.2677	0.45	0.05	Isotope Stearic acid (C18:0)	(13C)C17H36O3	M-H-	YES	2	Fatty acids and conjugates	LDEPL	6.78E-03	3.4	8.60E-04	4.1
-	293.2121	293.2112	293.2122	0.29	3.37	Oxo-octadecadienoic acid	C18H30O3	M-H-	NO	3	Fatty acids and conjugates	LDEPL	4.76E-06	5.5	2.13E-03	3.8
-	295.2279	295.2277	295.2279	0.06	0.7	HODE	C18H32O3	M-H-	NO	3	Lineolic acids and derivatives	LDEPL	1.89E-05	5.1	1.08E-06	5.7
-	295.2643	295.264	295.2643	0	0.9	Nonadecenoic acid (C19:1)	C19H36O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	2.86E-06	5.6	1.19E-07	6.2
-	297.2799	297.2797	297.2799	-0.03	0.51	Nonadecanoic acid (C19:0)	C19H38O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	5.86E-07	6	1.69E-14	8.9

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	307.2643	307.2645	307.2643	0.12	-0.67	Eicosadienoic acid (C20:2)	C20H36O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	7.33E-06	5.3	1.54E-14	9
-	309.2799	309.2798	309.2799	-0.14	0.25	Phytenic acid (C20:1)	C20H38O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	7.24E-04	4.1	1.58E-06	5.6
-	310.2833	310.2831	310.2833	0	0.46	Isotope Phytenic acid (C20:1)	(13C)C19H38O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	3.48E-11	7.8	3.29E-13	8.5
-	**311.222**	311.2225	311.2228	2.47	0.91	Dihydroxy-octadecadienoic acid	C18H32O4	M-H-	NO	3	Fatty acids and conjugates	LDEPL	5.89E-03	3.5	4.58E-02	2.7
-	311.2955	311.2955	311.2956	0.23	0.27	Eicosanoic acid, Arachidic Acid (C20:0)	C20H40O2	M-H-	YES	2	Fatty acids and conjugates	LDEPL	3.52E-04	4.3	1.70E-06	5.6
-	**312.2985**	**312.2985**	312.299	1.46	1.62	Isotope Eicosanoic acid (C20:0)	(13C)C19H40O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	1.43E-06	5.8	3.84E-04	4.3

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	337.3113	337.3108	337.3112	-0.34	1.12	Docosenoic acid (C22:1)	C22H42O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	2.32E-11	8	1.52E-16	9.7
-	339.3268	339.327	339.3269	0.17	-0.22	Docosanoic acid (C22:0)	C22H44O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	8.31E-06	5.3	2.65E-12	8.1
-	**342.2471**	342.2472	342.2472	0.12	-0.12	N-palmitoyl-cysteine	C19H35NO2S	M-H-	YES	2	Amino acids, peptides, and analogues	LDOM	1.32E-09	-7.1	1.56E-12	-8.2
-	343.2509	343.2508	343.2506	-1.08	-0.77	Isotope N-palmitoyl-cysteine	(13C)C18H35NO2S	M-H-	NO	3	Amino acids, peptides, and analogues	LDOM	4.87E-06	-5.5	2.45E-06	-5.5
-	353.3424	353.3429	353.3425	0.38	-1.16	Tricosanoic acid (C23:0)	C23H46O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	1.35E-07	6.3	2.04E-06	5.6

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	365.2808	365.2805	365.281	0.52	1.41	N-palmitoyl glutamine	C21H40N2O4	M-H2O-H-	NO	3	Amino acids, peptides, and analogues	LDOM	1.59E-08	-6.7	4.93E-05	-4.8
-	365.3425	365.3423	365.3425	-0.01	0.55	Nervonic acid (C24:1)	C24H46O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	4.34E-04	4.3	1.74E-06	5.6
-	366.3458	366.347	366.3459	0.17	-3.03	Isotope Nervonic acid (C24:1)	(13C)C23H46O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	1.88E-05	5.1	7.01E-07	5.8
-	367.3581	367.358	367.3582	0.39	0.45	Lignoceric acid (C24:0)	C24H48O2	M-H-	YES	2	Fatty acids and conjugates	LDEPL	2.38E-02	3	5.87E-03	3.4
-	**368.2629**	368.2625	368.2629	0.03	1.16	N-palmitoyl methionine	C21H41NO3S	M-H2O-H-	NO	3	Amino acids, peptides, and analogues	LDOM	4.34E-04	-4.3	3.81E-02	-2.8

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	368.3615	368.3615	368.3616	0.06	0.05	Isotope Lignoceric acid (C24:0)	(13C)C23H48O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	2.87E-06	5.6	2.77E-09	7
-	370.2784	370.2784	370.2785	0.14	0.34	Octadecenylamino-sulfanylpropanoic acid	C21H41NO2S	M-H-	NO	3		LDOM	1.89E-05	-5.1	2.35E-07	-6.1
-	381.3736	381.3739	381.3738	0.43	-0.32	Pentacosanoic acid (C25:0)	C25H50O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	2.82E-03	3.7	1.22E-02	3.2
-	393.3738	393.3741	393.3738	-0.08	-0.73	Ximenic acid (C26:1)	C26H50O2	M-H-	YES	2	Fatty acids and conjugates	LDEPL	7.99E-03	3.4	1.00E-07	6.3
-	395.3893	395.3895	395.3895	0.45	-0.04	Cerotic acid (C26:0)	C26H52O2	M-H-	NO	3	Fatty acids and conjugates	LDEPL	2.94E-03	3.7	4.30E-05	4.9

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	399.1838	399.185	399.1845	1.82	-1.33	Amino-aminohexyl-diaminomethylideneamino-pentanamide	C12H28N6O9	M-H-	NO	3		LDEPL	1.08E-02	3.3	1.38E-04	4.6
-	399.2684	399.2688	399.2687	0.76	-0.21	Oxosulfanyl-oxazolidinyl-octadecanamide	C21H40N2O3S	M-H-	YES	2		LDOM	4.76E-06	-5.4	9.84E-13	-8.2
-	400.2721	400.2726	400.2721	-0.04	-1.41	Isotope Oxosulfanyl-oxazolidinyl-octadecanamide	(13C)C20H40N2O3S	M-H-	NO	3		LDOM	2.54E-04	-4.4	8.27E-05	-4.7
-	413.2838	413.2847	413.2843	-1.21	0.97	Methyl-hexadecylcarbamothioylamino-oxobutanoate	C22H42N2O3S	M-H-	NO	3		LDOM	1.11E-06	-5.8	3.84E-04	-4.3

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	427.2999	427.2992	427.3	0.24	1.84	Ethyl-hexadecylcarbamothioylamino-oxobutanoate	C23H44N2O3S	M-H-	NO	3		LDOM	3.68E-08	-6.5	4.69E-06	-5.4
-	438.299	438.2985	438.299	-0.06	1.23	LysoPE(O-16:0)	C21H46NO6P	M-H-	YES	2	Glycerophosphoethanolamines	LDOM	8.26E-05	-4.7	3.92E-07	-6
-	439.3019	439.3015	439.3024	1.05	2.02	Isotope LysoPE(O-16:0)	(13C)C20H46NO6P	M-H-	YES	2	Glycerophosphoethanolamines	LDOM	5.99E-06	-5.4	3.34E-04	-4.3
-	441.3153	441.3157	441.3156	0.61	-0.25	Propyl-hexadecylcarbamothioylamino-oxobutanoate	C24H46N2O3S	M-H-	NO	3		LDOM	3.47E-07	-6.1	3.30E-02	-2.8
-	**450.2628**	**450.2627**	450.2626	0.47	0.18	LysoPE(16:1)	C21H42NO7P	M-H-	NO	3	Glycerophosphoethanolamines	LDOM	3.63E-03	-3.6	2.95E-04	-4.4

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	455.3309	455.3313	455.3313	0.77	0.1	Amino-sulfanylheptenyl-amino-oxooctadecanoic acid	C25H48N2O3S	M-H-	NO	3		LDOM	2.28E-12	-8.3	1.64E-11	-7.8
-	456.3342	456.335	456.3347	1.01	-0.68	Isotope Amino-sulfanylheptenyl-amino-oxooctadecanoic acid	(13C)C24H48N2O3S	M-H-	YES	2		LDOM	1.42E-06	-5.8	1.70E-03	-3.9
-	476.2784	476.2776	476.278	-0.75	0.93	Lyso-PE(18:2)	C23H44NO7P	M-H-	YES	2	Glycerophosphoethanolamines	LDOM	3.93E-02	-2.7	4.31E-04	-4.3
-	478.2937	478.2951	478.2939	0.51	-2.49	Lyso-PE(18:1)	C23H46NO7P	M-H-	YES	2	Glycerophosphoethanolamines	LDOM	2.06E-02	-3	5.95E-03	-3.4
-	483.2727	483.2733	483.2728	0.23	-0.98	Lyso-PG(16:0)	C22H45O9P	M-H-	YES	2	Glycerophosphoglycerols	LDOM	3.72E-03	-3.6	3.44E-02	-2.8

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	**483.3623**	483.3622	483.3626	0.66	0.76	Octadecanyloxy-propoxy-thiazolyl-propanamide	C27H52N2O3S	M-H-	NO	3		LDOM	5.50E-06	-5.4	8.85E-03	-3.3
-	489.3148	489.3148	489.3156	1.64	1.65	Dimethoxyphenyl-hexadecylcarbamothioyl-propenamide	C28H46N2O3S	M-H-	NO	3		LDOM	8.78E-11	-7.6	1.09E-08	-6.7
-	500.2785	500.2781	500.2783	-0.38	0.4	Lyso-PE(20:4)	C25H44NO7P	M-H-	YES	2	Glycerophosphoethanolamines	LDOM	9.41E-04	-4	1.05E-04	-4.6
-	524.2982	**524.2995**	524.2994	2.21	0.13	Lyso-PS(18:0)	C24H48NO9P	M-H-	YES	2	Glycerophosphoserines	LDOM	1.57E-04	-4.6	3.31E-03	-3.6
-	565.5211	**565.5203**	565.5201	-1.81	-0.35	Octadecyl-oxy-octadecanoic acid (SAHSA)	C36H70O4	M-H-	NO	3	Fatty acids and conjugates	LDEPL	1.95E-03	3.8	8.17E-07	5.8

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
-	653.4929	653.4927	653.4917	-1.83	-1.5	DG(36:3)	C39H70O5	M+Cl-	NO	3	Lineolic acids and derivatives	LDEPL	2.11E-05	5.1	2.54E-02	2.9
-	712.5302	712.5292	712.5287	-2.17	-0.69	PC(O-14:1/18:2)	C40H76NO7P	M-H-	YES	2	Glycerophosphocholines	LDEPL	2.42E-02	2.9	1.02E-02	3.3
-	**760.5141**	760.5139	760.5134	-0.92	-0.68	PS(16:0/18:1)	C40H76NO10P	M-H-	YES	2	Glycerophosphoserines	LDOM	6.64E-06	-5.4	2.72E-02	-2.9
-	771.5204	771.5165	771.5182	-2.85	2.15	PG(18:1/18:2)	C42H77O10P	M-H-	YES	2	Glycerophosphoglycerols	LDOM	1.09E-04	-4.7	2.11E-03	-3.8
-	**788.5451**	**788.5455**	788.5447	-0.53	-0.95	PS(18:0/18:1)	C42H80NO10P	M-H-	YES	2	Glycerophosphoserines	LDOM	2.05E-05	-5.1	1.12E-03	-4
+	85.0282	85.0285	85.0284	1.96	-1.2	Butynoate	C4H4O2	M+H+	NO	3	Carbonyl compounds	LDOM	3.24E-04	-4.7	1.70E-04	-4.7
+	89.1072	89.1074	89.1073	0.8	-0.95	Putrescine	C4H12N2	M+H+	NO	3	Amines	LDEPL	1.12E-02	3.8	1.19E-12	8.1
+	95.0602	95.0604	95.0604	2.32	-0.3	Pyridinamin	C5H6N2	M+H+	YES	2	Phenylpyridines	LDEPL	4.25E-05	5.2	9.51E-12	7.8
+	**100.0756**	100.0757	100.0757	1.1	-0.13	Pyrrolidinecarboxaldehyde	C5H9NO	M+H+	NO	3		LDEPL	2.71E-02	3.4	1.27E-03	4.1

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
+	103.123	103.123	103.123	-0.08	-0.46	Cadaverine	C5H14N2	M+H+	NO	3	Amines	LDEPL	4.25E-05	5.2	2.54E-25	12.1
+	105.07	105.0699	105.0699	-0.64	0.03	Styrene	C8H8	M+H+	YES	2	Styrenes	LDEPL	6.65E-07	6.1	8.92E-20	10.4
+	106.0731	106.0733	106.0733	1.7	-0.49	Isotope Styrene	(13C)C7H8	M+H+	YES	2	Styrenes	LDEPL	4.05E-06	5.7	1.19E-12	8.1
+	112.0867	112.087	112.0869	1.36	-0.91	Histamine	C5H9N3	M+H+	YES	2	Amines	LDEPL	4.23E-02	3.2	1.24E-06	5.7
+	120.0807	120.0809	120.0808	1.22	-0.76	Vinylaniline	C8H9N	M+H+	YES	2	Aniline and substituted anilines	LDOM	3.85E-05	-5.3	1.47E-12	-8.2
+	121.0646	121.0649	121.0648	1.82	-0.89	Tolualdehyde	C8H8O	M+H+	YES	2	Phenylacetaldehydes	LDEPL	1.32E-08	6.9	2.77E-17	9.7
+	122.0679	122.0683	122.0682	1.97	-1.28	Isotope Tolualdehyde	(13C)C7H8O	M+H+	NO	3	Phenylacetaldehydes	LDEPL	2.05E-04	4.8	3.11E-15	9.1
+	122.0962	122.0965	122.0964	1.71	-0.56	Phenylethylamine	C8H11N	M+H+	YES	2	Phenethylamines	LDEPL	5.41E-11	7.9	3.70E-20	10.6
+	132.1017	132.102	132.1019	1.25	-1.01	Leucine	C6H13NO2	M+H+	YES	2	Amino acids, peptides, and analogues	LDOM	4.70E-08	-6.7	3.89E-22	-11.5

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
+	138.091	138.0914	138.0913	1.98	-1.04	Tyramine	C8H11NO	M+H+	YES	2	Phenethyl amines	LDEPL	8.31E-07	6	2.22E -15	9
+	**144.0806**	144.0807	144.0808	1.18	1.04	Aminonaphthal ene	C10H9N	M+H+	YES	2		LDEPL	3.54E-06	5.7	3.05E -07	6
+	145.1333	145.1337	145.1335	1.4	-1.59	N-Acetylcadaverin e	C7H16N2O	M+H+	YES	2	Carboxyli c acid derivative s	LDEPL	3.21E-02	3.3	1.26E -17	9.7
+	147.1128	147.1129	147.1128	-0.19	-0.47	Lysine	C6H14N2O2	M+H+	NO	3	Amino acids, peptides, and analogues	LDOM	3.37E-02	-3.3	6.46E -06	-5.4
+	148.0601	148.0604	148.0604	2.3	-0.18	Glutamate	C5H9NO4	M+H+	YES	2	Amino acids, peptides, and analogues	LDOM	1.55E-02	-3.7	2.86E -04	-4.5
+	166.086	166.0863	166.0863	1.66	0.21	Phenylalanine	C9H11NO2	M+H+	YES	2	Amino acids, peptides, and analogues	LDOM	3.82E-07	-6.2	8.15E -15	-9

+/-	*m/z* measured VMET2	*m/z* measured VMET	Theoretical *m/z*	ppm error VMET2	ppm error VMET	Metabolite Annotation	Sum formula	Adduct	MS/MS		Sub Class	Elevated in	VMET2 q-value (BH)	VMET2 t-ratio	q-value (BH)	t-ratio
+	175.1192	175.119	175.119	-0.88	0.05	Arginine	C6H14N4O2	M+H+	YES	2	Amino acids, peptides, and analogues	LDOM	3.94E-02	-3.3	1.07E-04	-4.8
+	344.2645	344.2637				Unknown		M+H+	NO	3		LDOM	2.32E-02	-3.5	3.22E-03	-3.8
+	401.3445	401.346				Unknown			NO	3		LDOM	4.34E-02	-3.2	1.32E-02	-3.4
+	426.3569	426.3581	426.3578	2.12	-0.65	Vaccenyl carnitine	C25H47NO4	M+H+	YES	2	Quaternary ammonium salts	LDOM	2.03E-02	-3.5	9.90E-04	-4.2
+	429.376	429.3773	429.3761	0.34	-2.79	Docosanylsulfanylbutanoic acid	C26H52O2S	M+H+	NO	3	NA	LDOM	1.29E-02	-3.7	3.06E-03	-3.9
+	527.1574	527.1586	527.1583	1.72	-0.65	Maltotriose	C18H32O16	M+Na+	YES	2	Carbohydrates and carbohydrate conjugates	LDEPL	1.74E-07	6.4	4.23E-03	3.8

Appendix table 2 – Summary of PERMANOVA results. The proportion of metabolome variance explained (R2) per covariate term (rows) is reported for all metabolic profiling datasets, alongside the degrees of freedom (Df), sum-of-squares (SumOfSqs), PERMANOVA Pseudo F-statistic (F),and P-value (Pr(>F)). The "Residual" term represents the remaining unexplained variance. Significance levels for the F-test P-value are highlighted as follows: "***" P-value < 0.001; "**" P-value < 0.01; "*" P-value < 0.05. 0.001 is the smallest numerical estimate

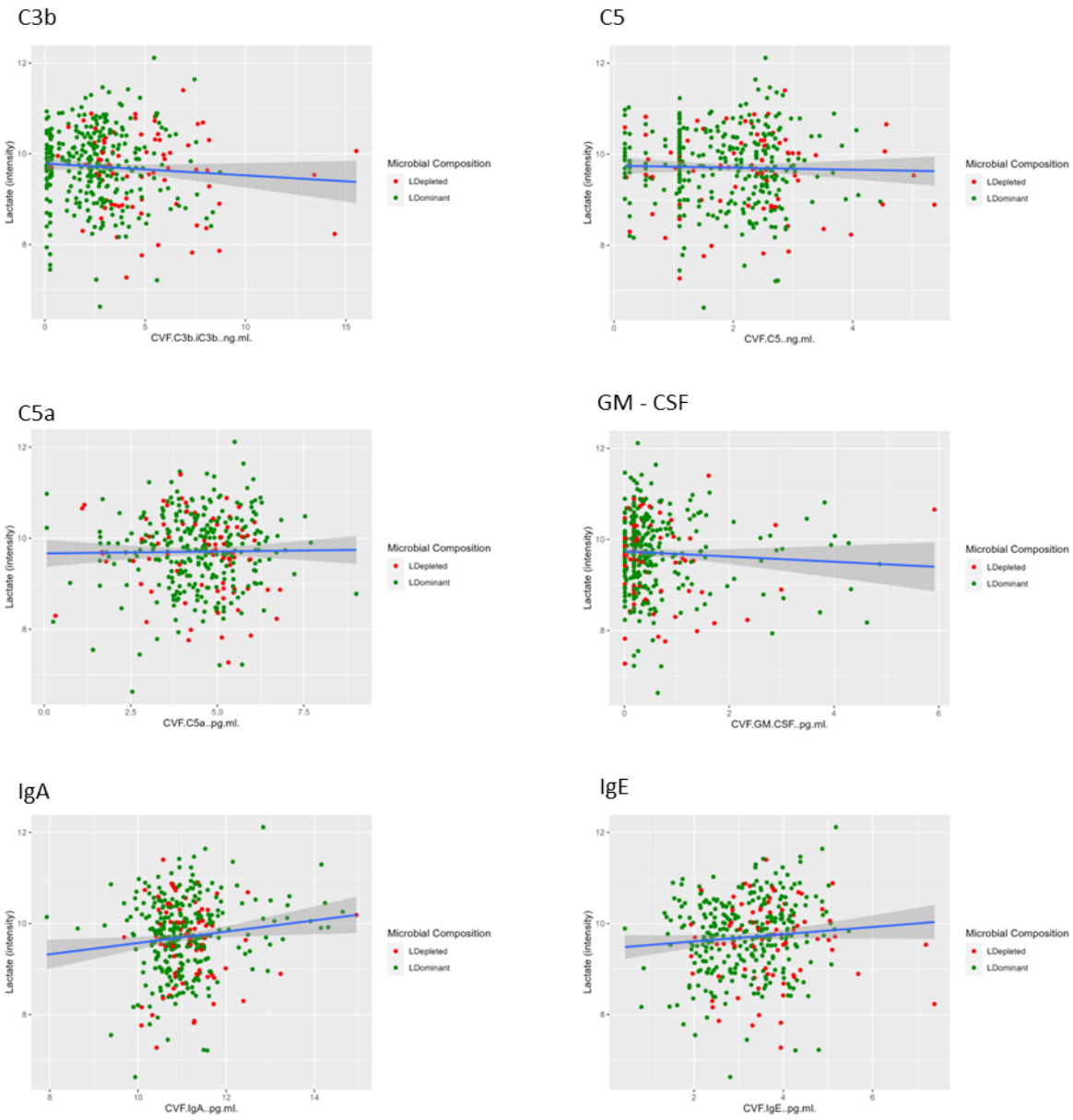
VMET2 DESI-MS (+)

	Df	**Sum0fSqs**	**R2**	**F**	**Pr(>F)**
CST	6	99768.230.	0.022	2.516.	0.001***
GestationalAge	1	16155.440.	0.004	2.444.	0.001***
Ethnicity	7	94423.770.	0.021	2.041.	0.001***
BMI	1	14018.670.	0.003	2.121.	0.001***
Age	1	13275.200.	0.003	2.009.	0.002**
SubjectID	193	2031776.550.	0.453	1.593.	0.001***
Residual	335	2214087.290.	0.494		
Total	544	4483505.140.	1.000		

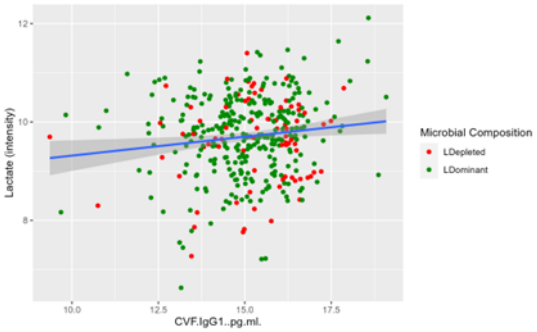
VMET2 DESI-MS (-)

	Df	**Sum0fSqs**	**R2**	**F**	**Pr(>F)**
CST	6	178184.315.	0.049	5.749.	0.001***
GestationalAge	1	14874.142.	0.004	2.879.	0.001***
Ethnicity	7	86027.637.	0.024	2.379.	0.001***
BMI	1	12270.672.	0.003	2.375.	0.003**
Age	1	8373.041.	0.002	1.621.	0.024*
SubjectID	193	1647443.099.	0.452	1.652.	0.001***
Residual	329	1699571.382.	0.466		
Total	538	3646744.289.	1.000		

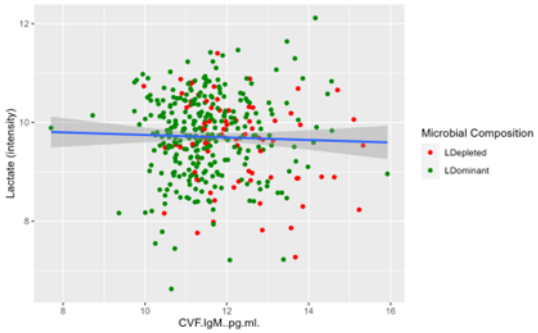
Appendix Figure 1 - Scatter plots with a fitted line representing a linear model between immune markers and lactate levels. The x-axis represents the immune marker values, while the y-axis represents the corresponding lactate levels. Each data point in the plot represents an individual observation. The fitted line visually represents the relationship between the immune markers and lactate levels as modelled by the linear regression.



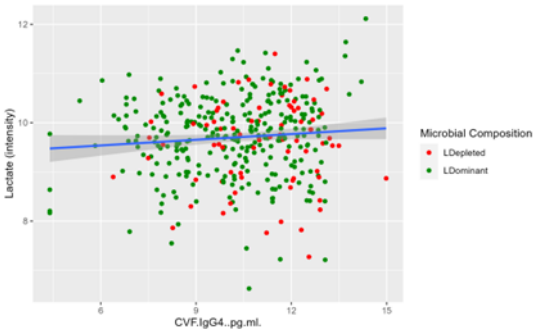
IgG1



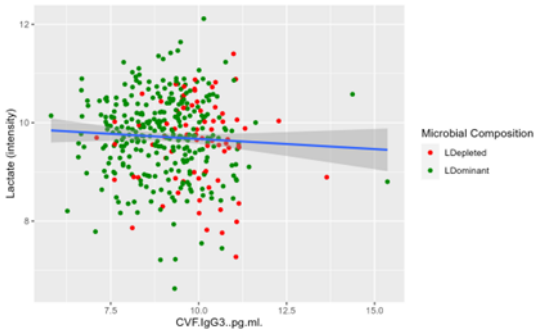
IgG2



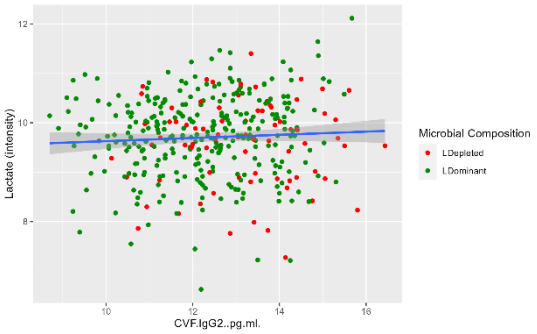
IgG3



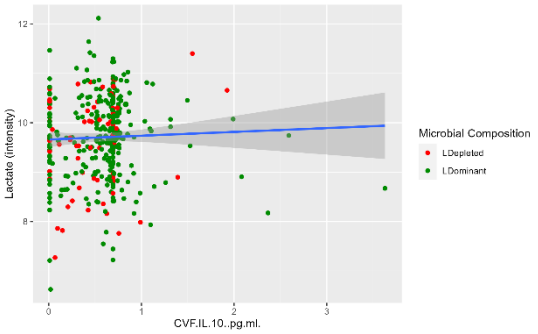
IgG4



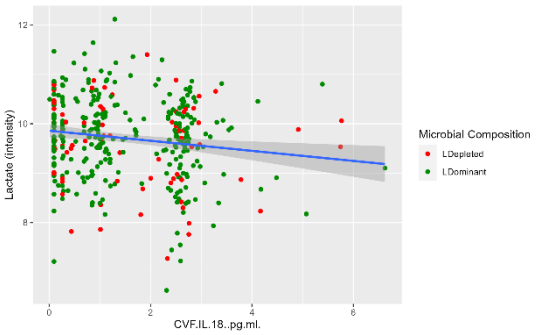
IgM



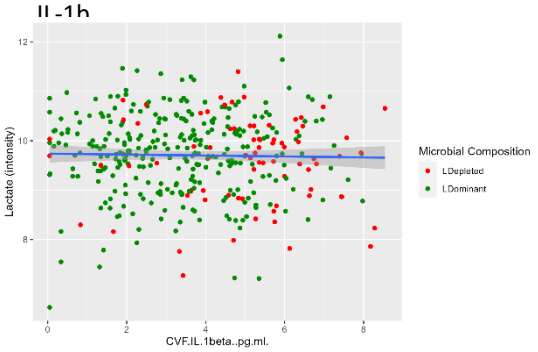
IL-10



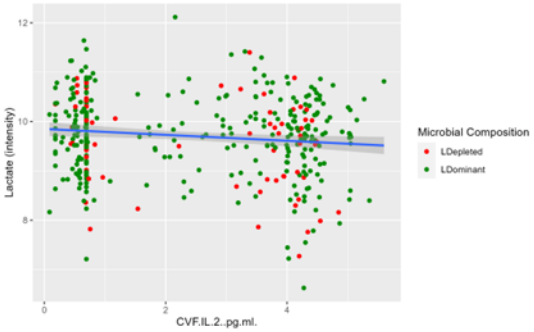
IL-18



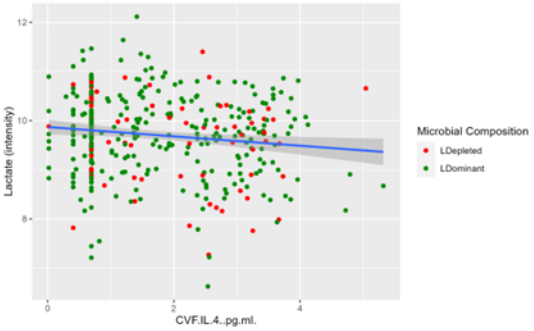
IL-1b



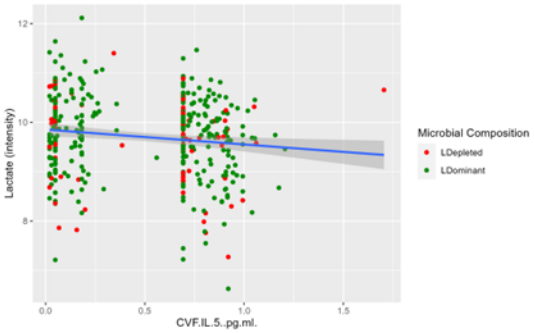
IL-2



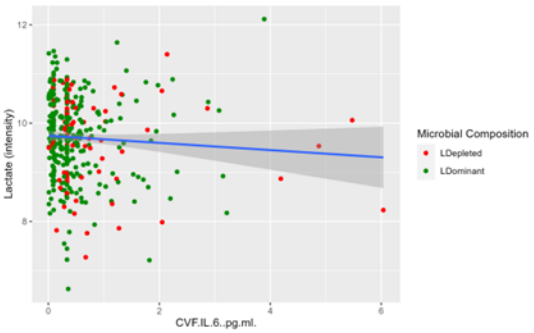
IL-4



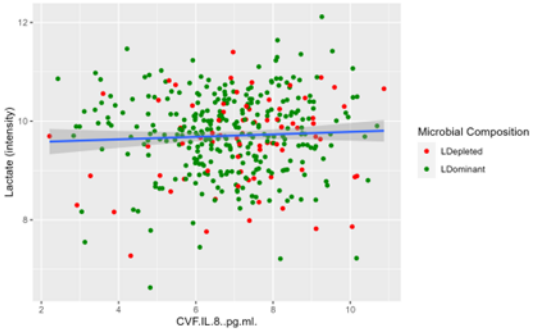
IL-5



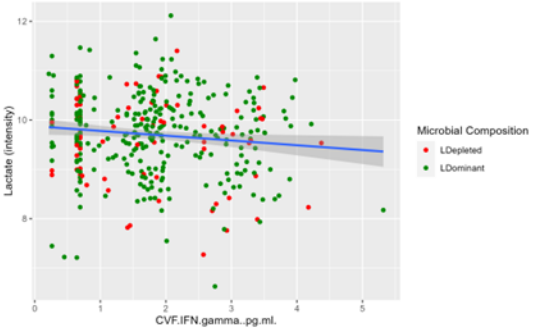
IL-6



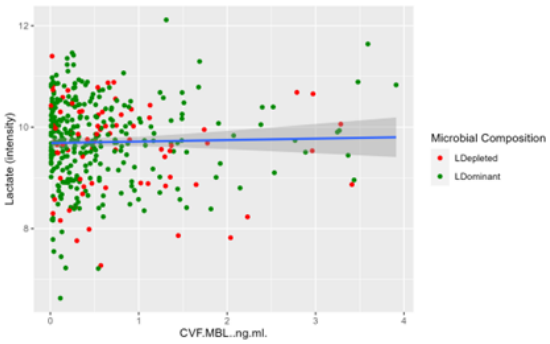
IL-8



INF-y



MBL



TNF-a

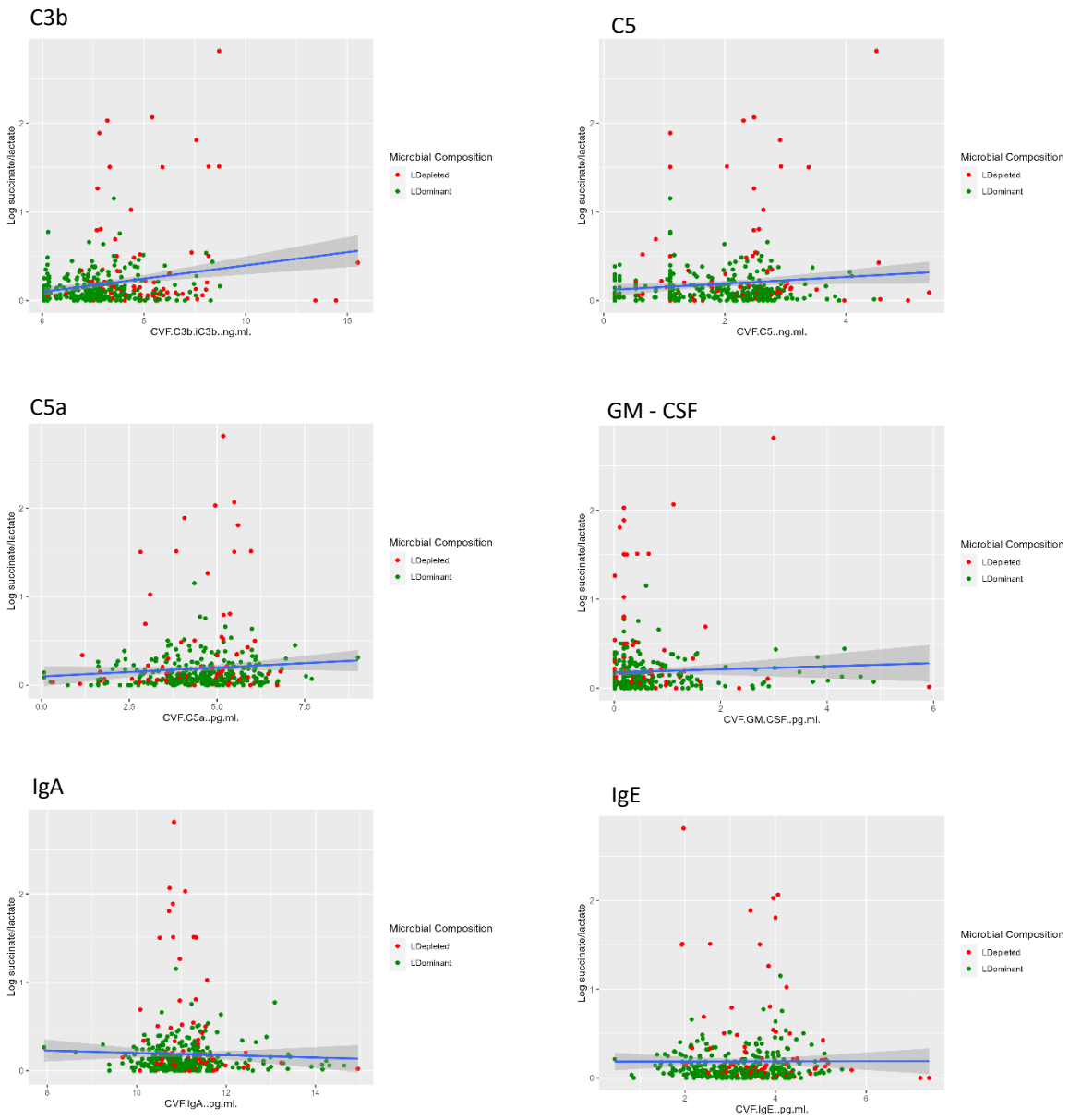


Appendix table 3 - The table provides a summary of the linear regression models between immune markers and lactate levels. It reports the values of R-squared, adjusted R-squared, and the corresponding p-values for each model.

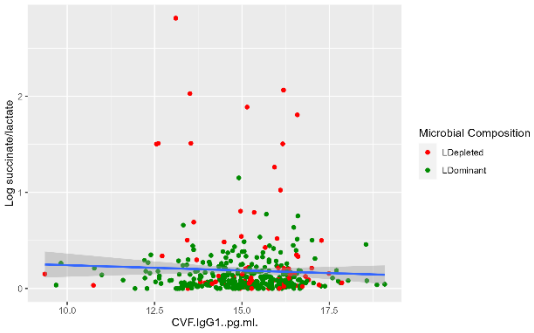
Lactate only

Immune marker	R^2	R^2 adj	p-value
C3b	5.19E-03	2.52E-03	1.64E-01
C5	6.52E-04	-2.02E-03	6.22E-01
C5a	1.74E-04	-2.50E-03	7.99E-01
GM-CSF	3.40E-03	7.35E-04	2.59E-01
IgA	1.55E-02	1.29E-02	1.57E-02
IgE	8.28E-03	5.63E-03	7.80E-02
IgG1	1.70E-02	1.44E-02	1.13E-02
IgG2	3.38E-03	7.17E-04	2.61E-01
IgG3	3.72E-03	1.05E-03	2.38E-01
IgG4	8.01E-03	5.35E-03	8.32E-02
IgM	1.15E-03	-1.52E-03	5.11E-01
IL-10	1.29E-03	-1.39E-03	4.88E-01
IL-18	2.20E-02	1.94E-02	3.92E-03
IL-1β	4.85E-04	-2.19E-03	6.71E-01
IL-2	1.57E-02	1.31E-02	1.50E-02
IL-4	1.80E-02	1.53E-02	9.30E-03
IL-5	1.77E-02	1.51E-02	9.73E-03
IL-6	4.34E-03	1.68E-03	2.02E-01
IL-8	2.51E-03	-1.59E-04	3.33E-01
IFN-γ	1.36E-02	1.10E-02	2.38E-02
MBL	6.56E-04	-2.02E-03	6.21E-01
TNF-α	2.36E-02	2.10E-02	2.81E-03

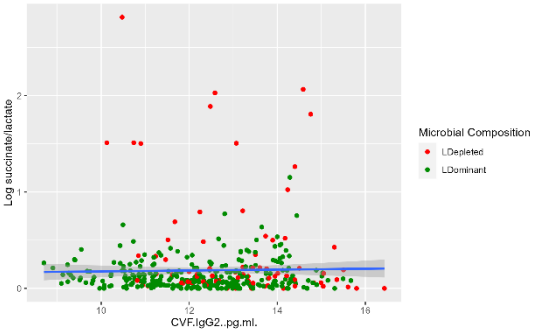
Appendix Figure 2 - Scatter plots with a fitted line representing a linear model between immune markers and log ratio succinate/lactate levels. The x-axis represents the immune marker values, while the y-axis represents the corresponding lactate levels. Each data point in the plot represents an individual observation. The fitted line visually represents the relationship between the immune markers and lactate levels as modelled by the linear regression.



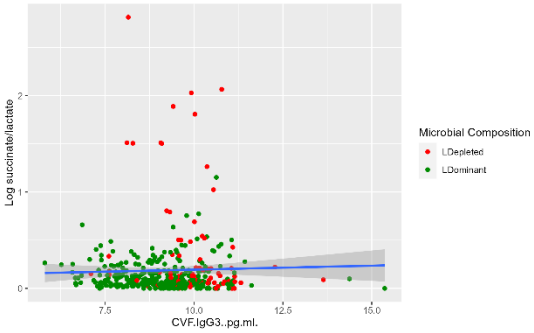
IgG1



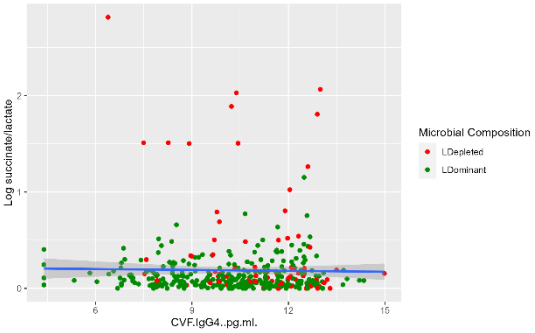
IgG2



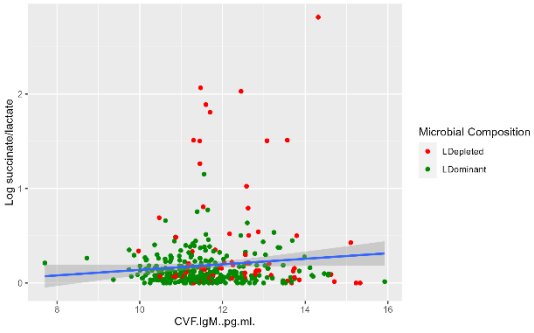
IgG3



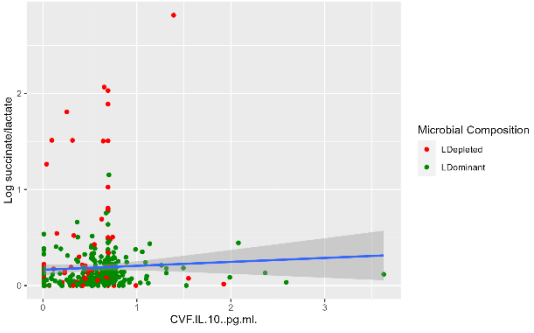
IgG4



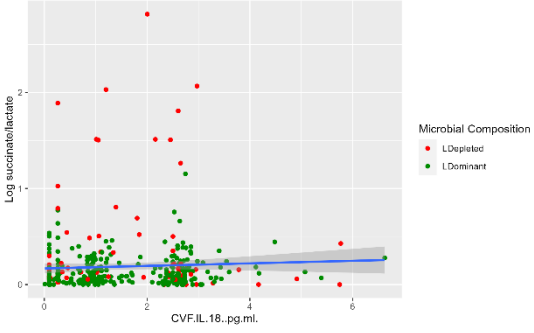
IgM



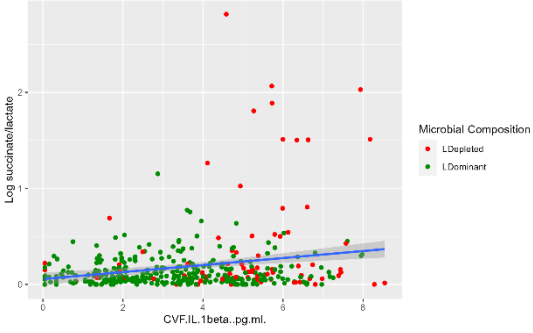
IL-10



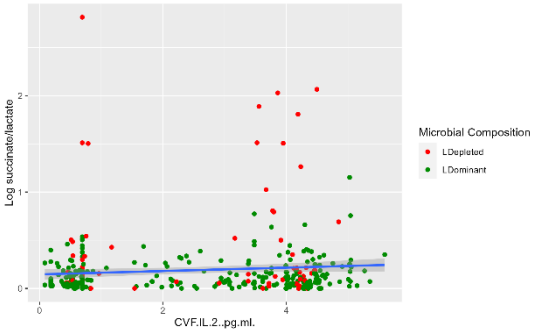
IL-18



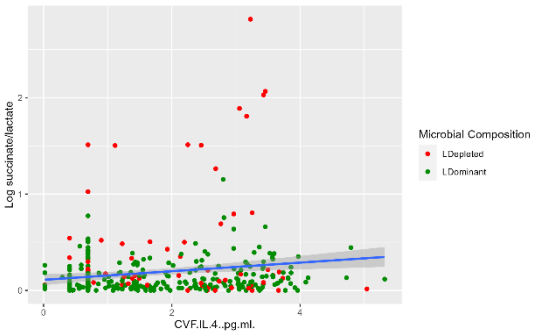
IL-1b



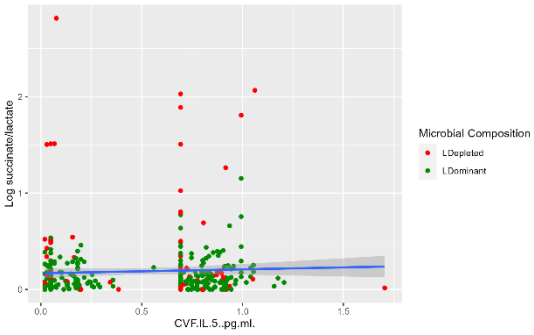
IL-2



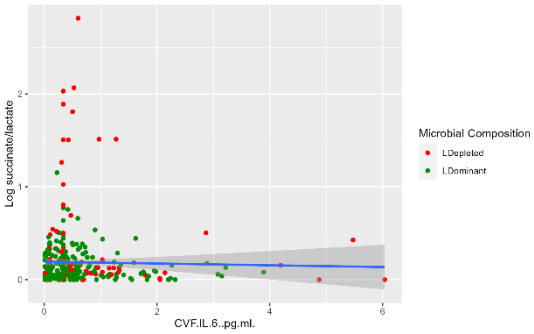
IL-4



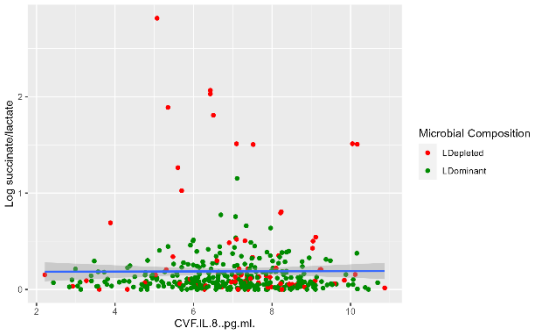
IL-5



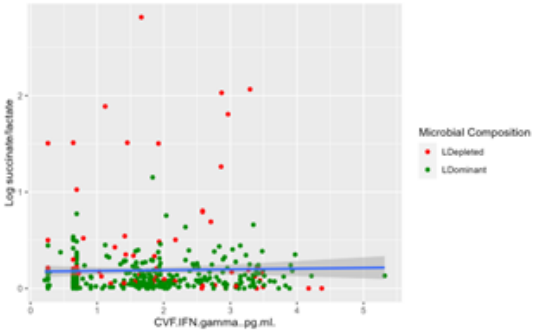
IL-6



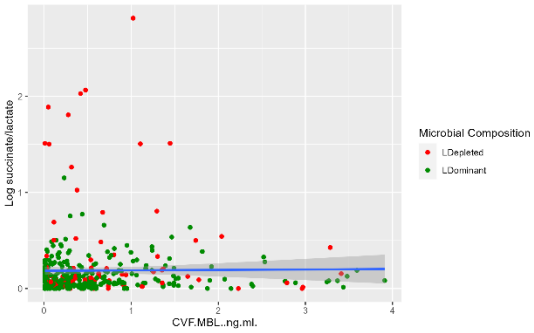
IL-8



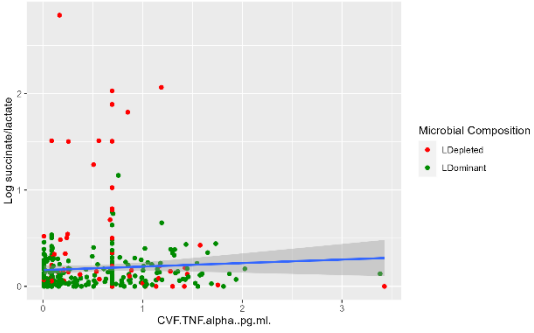
INF-y



MBL



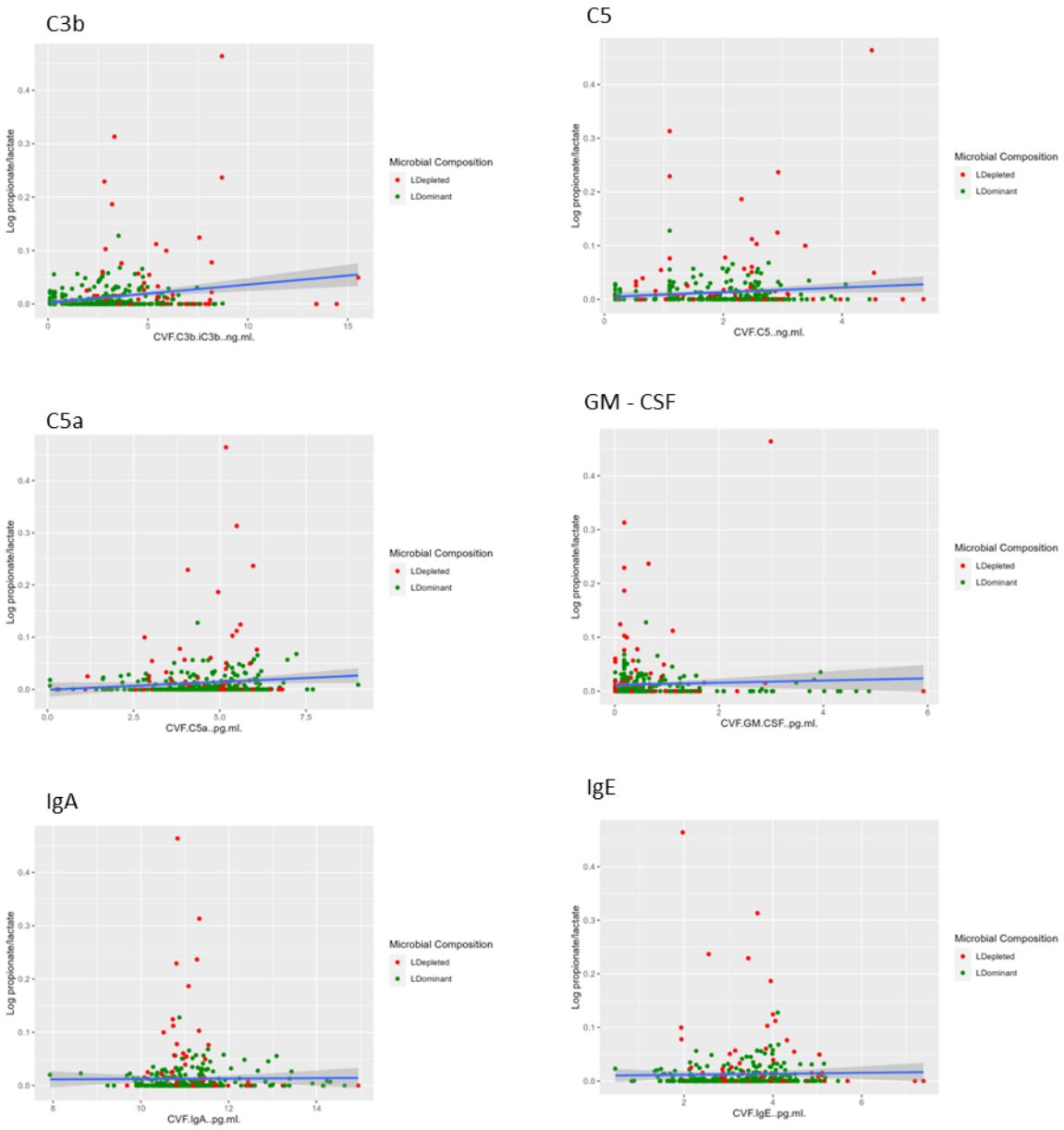
TNF-a



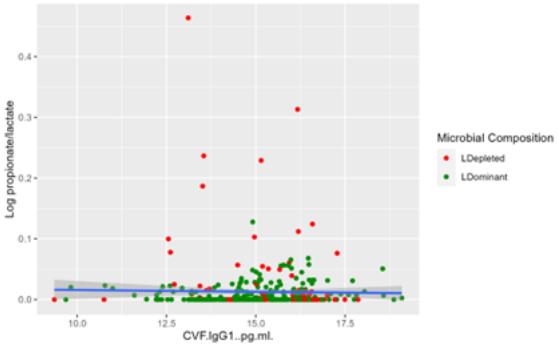
Appendix table 4 - The table provides a summary of the linear regression models between immune markers and the ratio between succinate and lactate levels. It reports the values of R-squared, adjusted R-squared, and the corresponding p-values for each model.

Succinate/Lactate ratio			
Immune marker	**R^2**	**R^2 adj**	**p-value**
C3b	4.18E-02	3.93E-02	6.61E-05
C5	2.06E-02	1.80E-02	5.33E-03
C5a	4.04E-03	1.37E-03	2.19E-01
GM-CSF	7.78E-03	5.12E-03	8.77E-02
IgA	9.17E-04	-1.76E-03	5.58E-01
IgE	5.17E-04	-2.16E-03	6.60E-01
IgG1	4.38E-03	1.72E-03	2.01E-01
IgG2	3.85E-05	-2.64E-03	9.05E-01
IgG3	2.26E-04	-2.45E-03	7.72E-01
IgG4	2.43E-03	-2.34E-04	3.40E-01
IgM	1.63E-02	1.36E-02	1.34E-02
IL-10	5.62E-03	2.97E-03	1.47E-01
IL-18	1.78E-03	-8.90E-04	4.15E-01
IL-1β	3.04E-02	2.78E-02	6.81E-04
IL-2	2.09E-03	-5.78E-04	3.77E-01
IL-4	2.19E-02	1.93E-02	4.02E-03
IL-5	1.90E-04	-2.48E-03	7.90E-01
IL-6	6.98E-05	-2.60E-03	8.72E-01
IL-8	9.54E-04	-1.72E-03	5.51E-01
IFN-γ	7.58E-04	-1.91E-03	5.95E-01
MBL	6.40E-05	-2.61E-03	8.77E-01
TNF-α	6.92E-04	-1.98E-03	6.11E-01

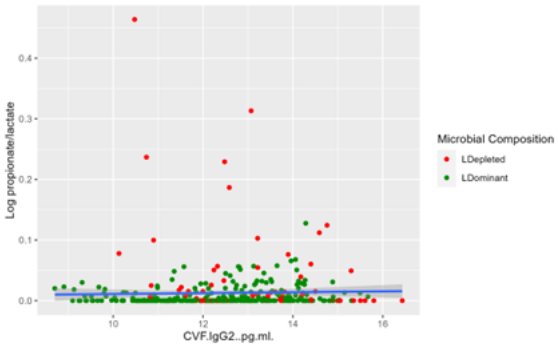
Appendix Figure 3 - Scatter plots with a fitted line representing a linear model between immune markers and log ratio propionate/lactate levels. The x-axis represents the immune marker values, while the y-axis represents the corresponding lactate levels. Each data point in the plot represents an individual observation. The fitted line visually represents the relationship between the immune markers and lactate levels as modelled by the linear regression.



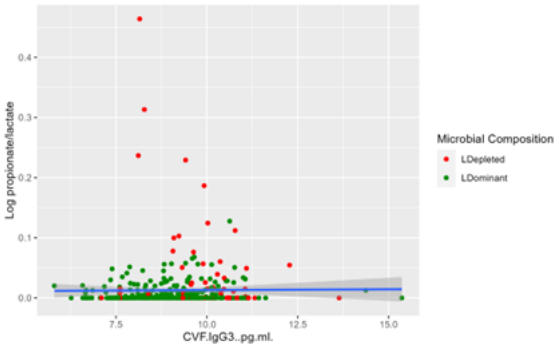
IgG1



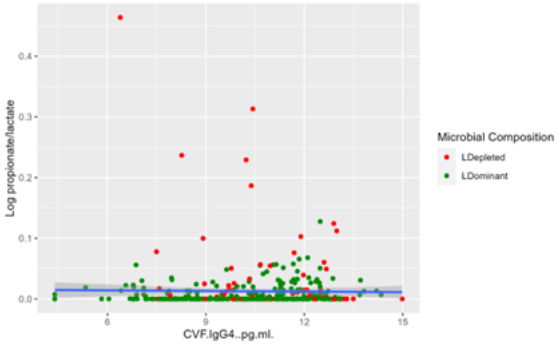
IgG2



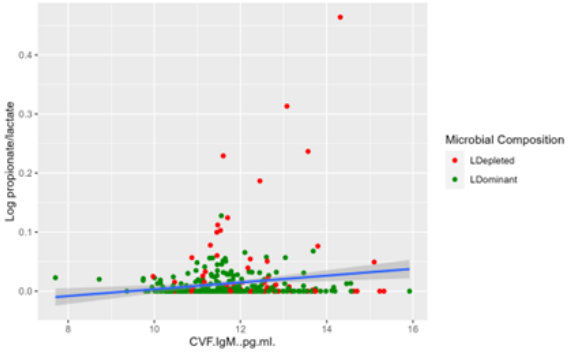
IgG3



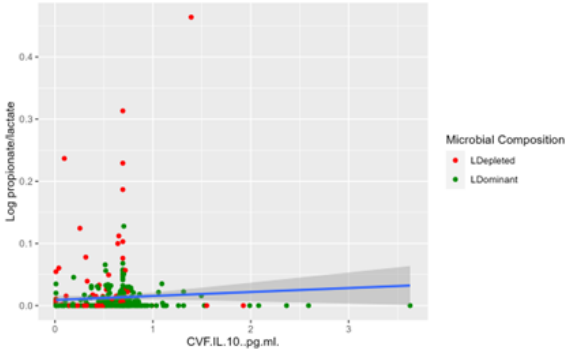
IgG4



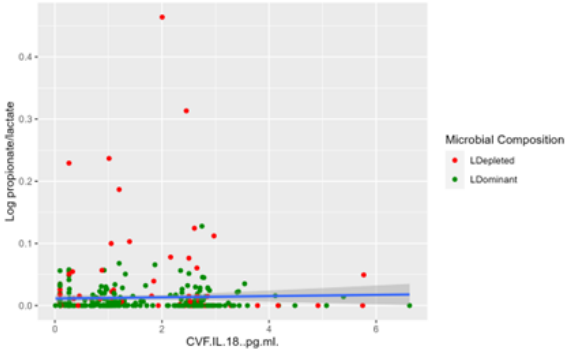
IgM



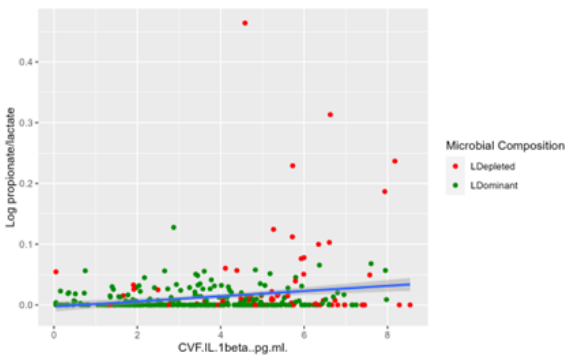
IL-10



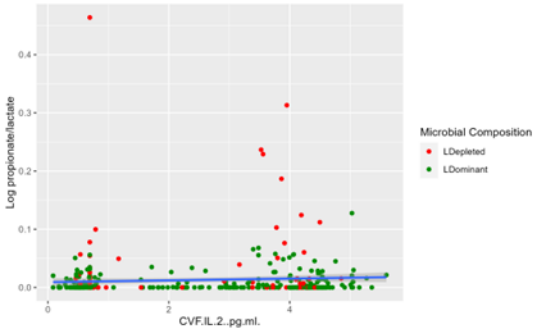
IL-18



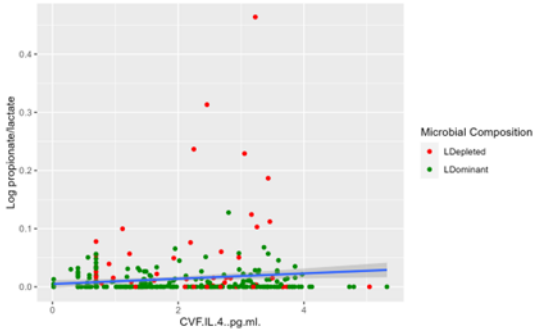
IL-1b



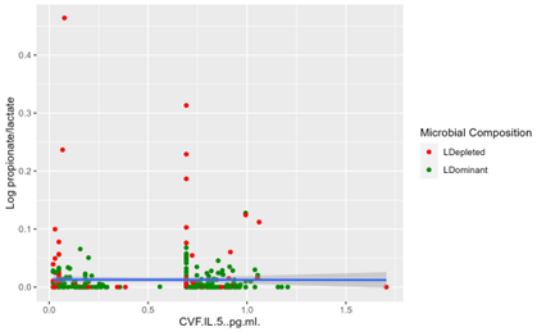
IL-2



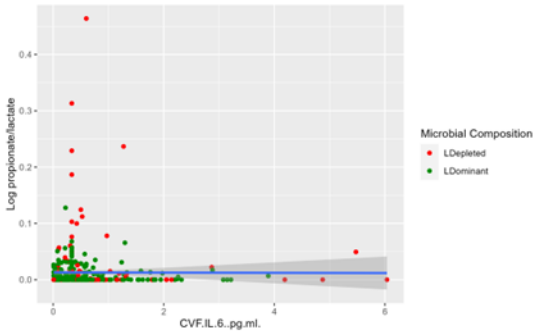
IL-4



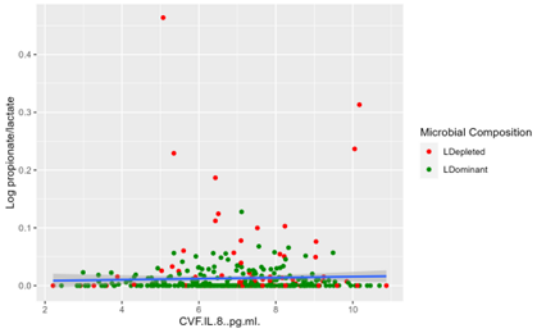
IL-5



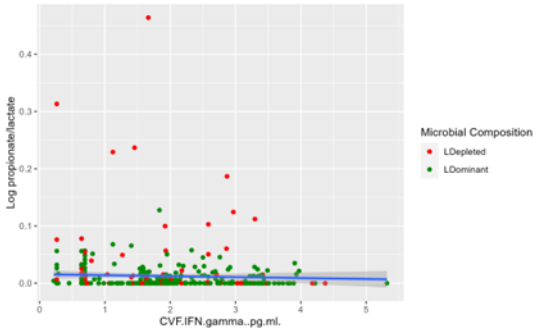
IL-6



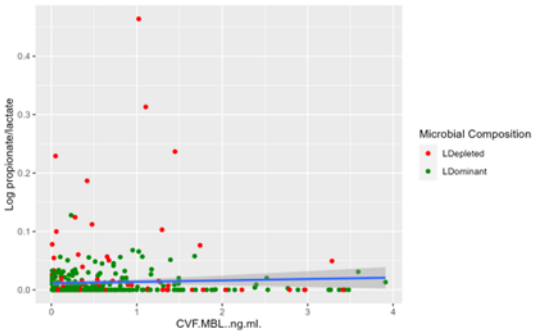
IL-8



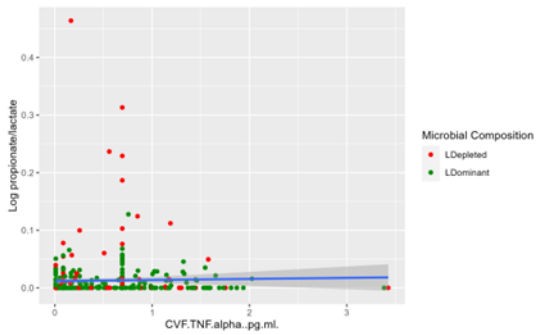
INF-y



MBL



TNF-a

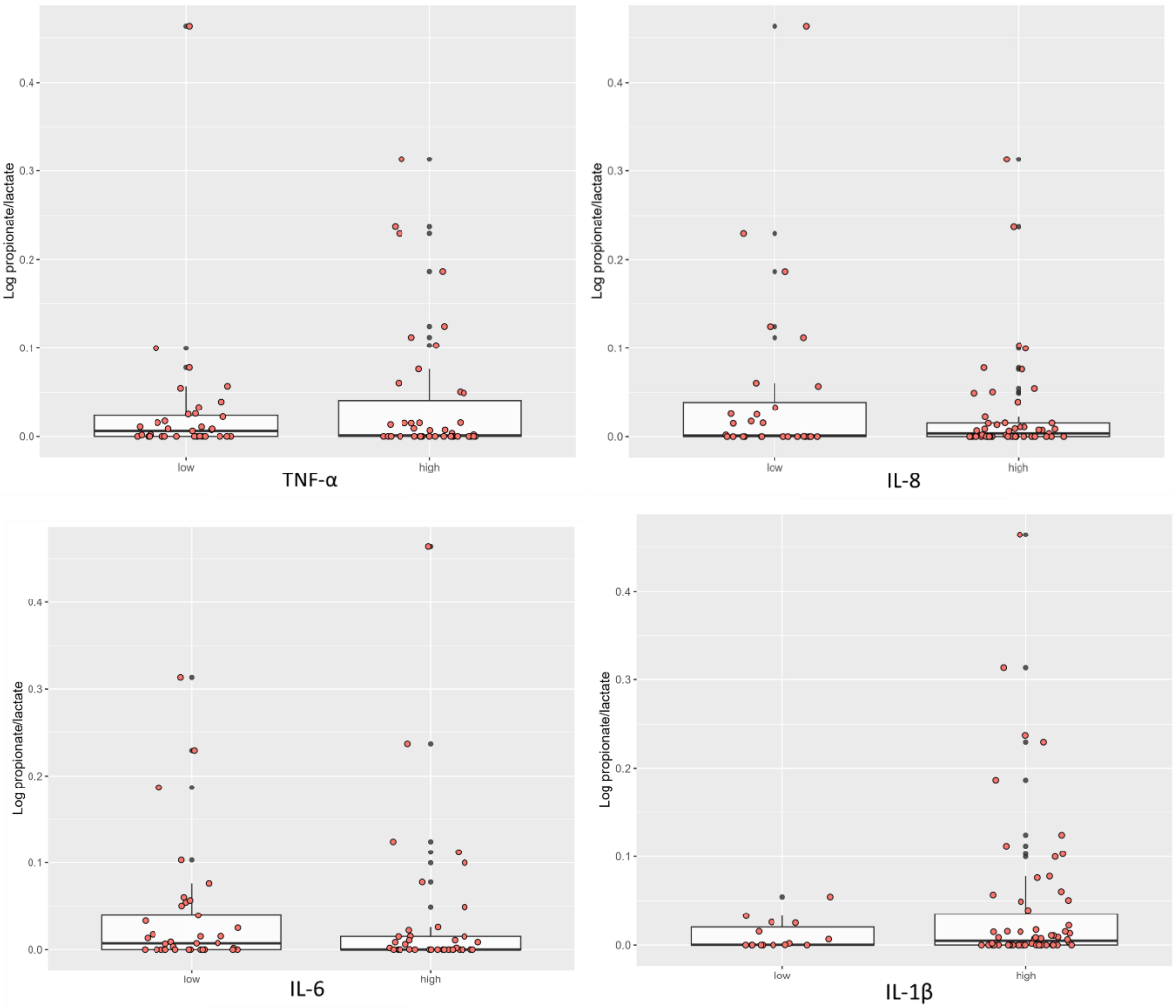


Appendix table 5 - The table provides a summary of the linear regression models between immune markers and the ratio between propionate and lactate levels. It reports the values of R-squared, adjusted R-squared, and the corresponding p-values for each model.

Propionate/Lactate ratio

Immune marker	R^2	R^2 adj	p-value
C3b	3.82E-02	3.57E-02	1.39E-04
C5	1.21E-02	9.49E-03	3.28E-02
C5a	9.29E-03	6.64E-03	6.19E-02
GM-CSF	3.05E-03	3.86E-04	2.85E-01
IgA	3.37E-05	-2.64E-03	9.11E-01
IgE	1.40E-04	-2.53E-03	8.19E-01
IgG1	7.77E-04	-1.90E-03	5.90E-01
IgG2	2.81E-04	-2.39E-03	7.46E-01
IgG3	6.18E-07	-2.67E-03	9.88E-01
IgG4	8.26E-04	-1.85E-03	5.79E-01
IgM	2.77E-02	2.51E-02	1.20E-03
IL-10	4.83E-03	2.17E-03	1.79E-01
IL-18	9.19E-04	-1.75E-03	5.58E-01
IL-1β	4.00E-02	3.74E-02	9.52E-05
IL-2	3.45E-03	7.88E-04	2.56E-01
IL-4	1.81E-02	1.54E-02	9.09E-03
IL-5	9.55E-05	-2.58E-03	8.50E-01
IL-6	5.89E-06	-2.67E-03	9.63E-01
IL-8	1.04E-03	-1.63E-03	5.32E-01
IFN-γ	1.71E-03	-9.56E-04	4.24E-01
MBL	2.10E-03	-5.69E-04	3.76E-01
TNF-α	3.71E-04	-2.30E-03	7.10E-01

Appendix Figure 5 - Box and whiskers plot showing different levels of propionate/lactate ratio in high or low classes of immune markers (IL-1β, IL-8, TNF-α, IL-6) All box and whisker plots are drawn with the lower, horizontal interior line, and upper bounds of the box representing the 25th percentile, median and 75th percentile values, and whiskers extending from the lower or upper box bonds to the position of the most extreme data point within ± 1.5 × IQR (interquartile range).



Appendix Figure 6 - PCA plot generated from metabolic data acquired in both negative (top two plots) and positive ion modes (plots below). The data points in the plot are color-coded based on CST and Study. The cumulative variation explained by the PCA model is approximately 25%.



Appendix Table – 6 PERMANOVA analysis on VMET and VMET2 white patients considering only first sample available (<26 weeks). Estimation of variance explained by risk factors = 0.028. The proportion of metabolome variance explained (R2) per covariate term (rows) is reported for all metabolic profiling datasets, alongside the degrees of freedom (Df), sum-of-squares (SumOfSqs), PERMANOVA Pseudo F-statistic (F),and P-value (Pr(>F)). The "Residual" term represents the remaining unexplained variance. Significance levels for the F-test P-value are highlighted as follows: "***" P-value < 0.001; "**" P-value < 0.01; "*" P-value < 0.05. 0.001 is the smallest numerical estimate

VMET and VMET2 (-)

	Df	**Sum0fSqs**	**R2**	**F Pr(>F)**	
Risk class	2	14909.531.	0.028	3.761.	0.001***
CST	6	38903.975.	0.073	3.271.	0.001***
Gestational Age	1	3703.177.	0.007	1.868.	0.024
BMI	1	2563.333.	0.005	1.293.	0.155
Study	1	37405.590	0.070	18.870	0.001***
Residual	221	438087.838	0.818		
Total	232	535573.444	1.000		

Appendix Table 7 – List of features in negative and positive ion mode found to be significantly different in women with different previous pregnancy history.

m/z negative ion mode	p-value
79.9578	3.06E-02
92.9285	2.25E-02
93.0355	2.78E-02
110.0250	4.58E-02
115.0041	3.07E-02
119.9470	1.91E-02
121.0299	4.93E-03
123.0457	2.40E-02
145.0510	8.54E-03
151.0264	2.56E-02
151.0404	7.44E-03
165.0413	1.80E-02
183.0127	4.17E-02
200.1400	2.31E-02
201.0382	1.33E-02
233.1548	4.40E-03
236.0921	9.81E-03
255.2330	1.88E-02
256.2363	3.85E-02
256.2673	2.04E-03
266.2429	4.49E-02
279.1636	2.41E-03
280.8692	2.93E-02
283.2643	1.03E-02
284.2676	1.56E-02
284.4732	1.01E-02
285.2711	1.04E-02
297.1530	1.83E-02
299.1490	2.73E-02
307.0349	2.33E-02
311.1687	2.07E-02
311.2955	3.48E-02
313.1646	2.05E-02
330.2030	4.73E-02
337.0775	4.17E-02
339.3269	2.22E-03
340.6534	1.32E-03
341.1090	1.69E-02
367.3580	4.17E-03
368.3615	4.29E-03
377.0855	2.96E-03
379.0821	2.18E-02
381.2319	2.29E-02
384.3556	2.09E-02
391.2605	7.89E-03

m/z negative ion mode	p-value
421.2269	5.07E-04
435.2533	1.39E-02
462.7775	1.31E-02
462.9832	8.01E-03
463.7705	9.22E-03
467.6435	6.58E-03
467.8523	9.02E-03
468.8637	7.87E-03
470.4214	3.40E-05
473.0113	4.10E-04

m/z positive ion mode	p-value
105.0334	2.76E-02
110.0712	2.71E-02
120.0808	4.84E-02
132.1019	3.45E-02
133.0971	2.75E-02
135.0028	7.38E-04
149.0596	9.22E-03
150.9764	9.81E-03
171.0990	2.01E-02
199.1688	4.05E-03
229.1534	2.82E-02
240.2684	2.05E-02
298.3466	3.19E-02
310.2068	4.30E-02
322.2717	1.61E-02
325.1770	3.70E-03
331.2096	1.94E-03
337.1044	4.31E-02
338.1079	2.62E-02
339.1103	2.55E-02
351.2492	4.90E-02
353.0788	7.43E-03
353.2662	2.17E-02
354.0816	3.98E-02
365.1056	1.05E-02
366.1401	2.25E-03
368.4250	5.88E-03
369.4285	2.70E-02
377.2660	5.97E-05
380.2838	1.28E-02
381.0803	3.41E-03
381.1300	2.80E-02
381.2968	2.12E-02
393.2991	2.51E-02
409.2798	5.61E-03
425.2143	2.48E-02
443.3345	1.05E-02
483.3651	6.80E-03
487.3594	2.44E-02
516.3082	3.85E-02
517.3109	1.01E-02
521.3445	2.63E-02
524.2730	1.48E-02
527.3916	1.77E-02
529.4950	2.22E-03
534.3570	2.43E-02
543.3849	3.24E-02

m/z positive ion mode	p-value
546.3559	4.12E-02
547.3596	3.90E-02
559.3608	4.18E-02
588.4600	3.33E-02
590.4810	1.61E-02
591.2581	9.18E-03
641.5111	1.13E-02
641.6680	3.89E-02
704.4997	2.79E-02
717.5934	1.30E-02
719.4719	3.02E-02
720.4749	1.83E-02
724.5255	2.43E-02
725.5565	3.95E-02
733.5610	4.71E-02
735.4707	2.72E-02
741.5318	3.98E-02
755.5438	2.97E-03
756.5416	1.17E-02
757.5425	4.17E-03
770.5110	3.47E-02
771.5143	7.76E-03
778.5405	1.60E-03
783.5724	2.43E-02
796.5266	2.05E-02
797.5299	1.96E-03
805.5559	9.39E-03
808.5821	3.02E-02
822.5403	1.31E-04
824.5559	2.11E-02
825.5606	1.03E-02
826.5741	5.82E-03
827.5773	3.38E-02
832.5882	3.69E-02

Appendix Table 8 List of features in negative and positive ion mode found to be significantly different in women that will subsequently experience a cervical shortening or not.

m/z negative ion mode	p-value
93.0355	2.37E-02
107.0506	4.03E-02
121.0299	1.50E-02
145.0510	4.87E-02
150.0561	1.89E-02
157.1237	1.17E-02
165.0413	8.79E-03
185.0569	3.16E-02
187.0073	2.93E-02
200.1738	2.98E-02
225.1858	3.07E-02
233.1548	2.86E-02
272.2311	2.69E-02
353.2331	3.48E-02
368.3615	3.01E-02
412.2809	2.78E-02
428.3022	1.37E-02
462.2998	4.54E-02
484.3665	2.04E-02
497.2949	2.83E-02

m/z positive ion mode	p-value
149.0232	3.50E-03
250.1776	4.72E-02
261.1310	2.99E-02
286.3101	1.32E-02
326.3777	4.60E-02
328.3845	1.72E-03
351.2492	3.54E-02
368.4250	4.89E-02
381.2968	4.36E-02
414.2696	2.62E-02
451.3399	5.73E-03
453.2094	3.12E-02
469.3281	3.97E-02
479.3695	3.81E-02
480.3712	4.12E-02
481.2612	7.81E-03
487.2308	3.61E-02
487.3594	3.31E-02
502.2896	4.50E-02
503.2045	4.80E-02
507.4000	3.11E-02
509.4085	2.15E-02
517.3109	4.42E-02
522.3553	4.46E-02
541.2616	5.60E-03
547.2327	6.79E-03
549.4210	4.37E-02
550.6298	1.35E-02
565.3830	2.75E-02
594.4194	1.92E-02
822.7561	1.43E-02

Appendix Table 9 – Table presenting the Lower Limit of Quantification (LLOQ) and Upper Limit of Quantification (ULOQ) values for a set of 13 steroids as part of a UPLC-MS targeted assay. The LLOQ represents the lowest concentration at which the assay can reliably detect and quantify the steroids, while the ULOQ represents the highest concentration within the assay's linear dynamic range.

	LLOQ (ng/ml)	ULOQ (ng/ml)
E3	0.5	250
AL	0.05	250
HC	0.05	250
21-DEOXY-HC	0.05	250
COR	0.05	250
11-DEOXY-HC	0.1	250
E2	1	250
TES	0.05	100
E1	0.05	250
AND	0.01	100
11-DEOXY-COR	0.05	250
17-OH-PROG	0.05	250
PROG	0.005	25

Appendix Table 10 - Statistically significant metabolite features found associated with microbial composition in the FLIP patient cohort. The features reported are shared with VMET and VMET2 studies: ppm mass error, putative annotation, MS/MS, and sub class of metabolic compounds are reported.

FLIP m/z features	VMET2	VMET	Theoretical m/z	Elevated in	MS/MS	ppm FLIP vs theoretical	ppm FLIP vs VMET2	Metabolite Annotation	Formula	Adduct	Sub Class
117.0565	117.0562	117.0561	117.0557	LDEPL	NO	6.41	2.14	Hydroxyvaleric acid	C5H10O3	M-H-	Fatty acids and conjugates
131.0708	131.0717	131.0717	131.0714	LDEPL	YES	-4.68	-6.97	Hydroxycaproic acid	C6H12O3	M-H-	Fatty acids and conjugates
132.0285	132.0306	132.0306	132.0302	LDOM	YES	-12.73	-15.76	Aspartic acid	C4H7NO4	M-H-	Amino acids, peptides, analogues
173.0920	173.0934	173.0934	173.0932	LDOM	YES	-6.94	-8.10	N2-Acetyl-ornithine	C7H14N2O3	M-H-	Amino acids, peptides, and analogues
201.1207	201.1246	201.1243	201.1245	LDOM	YES	-18.88	-19.38	Leucyl-Alanine	C9H18N2O3	M-H-	Amino acids, peptides, and analogues
203.1020	203.1039	203.1038	203.1037	LDOM	YES	-8.26	-9.25	Valyl-Serine	C8H16N2O4	M-H-	Amino acids, peptides, and analogues
217.1173	217.1195	217.1196	217.1194	LDOM	YES	-9.53	-9.99	Leucyl-Serine	C9H18N2O4	M-H-	Amino acids, peptides, and analogues

221.0108	221.0124	221.0123	221.0125	LDOM	YES	-7.61	-7.16	Carboxysulfanyl-pentanedioic acid	C7H10O6S	M-H-	
231.1330	231.1351	231.1352	231.135	LDOM	YES	-8.67	-9.10	Leucyl-Threonine	C10H20N2O4	M-H-	Amino acids, peptides, and analogues
244.1303	244.1303	244.1302	244.1303	LDOM	YES	0.17	0.17	Leucyl-glycyl-glycine	C10H19N3O4	M-H-	Amino acids, peptides, and analogues
245.0445	245.0432	245.0434	245.0432	LDOM	YES	5.16	5.16	Phosphatidyl glycerol	C6H15O8P	M-H-	
245.1148	245.1143	245.1145	245.1143	LDOM	YES	1.94	1.94	Valyl-Glutamine	C10H18N2O5	M-H-	Amino acids, peptides, and analogues
251.0800	251.0773	251.0773	251.0772	LDOM	NO	11.28	10.88	Methoxymethyl methyl-glucopyranosiduronate	C9H16O8	M-H-	
258.1483	258.146	258.1459	258.1459	LDOM	YES	9.42	9.03	Leucyl-Glutamine	C11H21N3O4	M-H-	Amino acids, peptides, and analogues
259.1294	259.13	259.1299	259.1299	LDOM	YES	-1.80	-2.19	Glutamyl-Leucine	C11H20N2O5	M-H-	Amino acids, peptides, and analogues
280.2346	280.2364	280.2362	280.2364	LDEPL	YES	-6.34	-6.34	Isotope Linoleic acid (C18:2)	(13C)C17H32O2	M-H-	Fatty acids and conjugates
282.2512	282.2519	282.2518	282.252	LDEPL	NO	-2.87	-2.51	Isotope Oleic acid (18:1)	(13C)C17H34O2	M-H-	Fatty acids and conjugates
295.2292	295.2279	295.2277	295.2279	LDEPL	NO	4.56	4.56	HODE	C18H32O3	M-H-	Lineolic acids and derivatives
309.2798	309.2799	309.2798	309.2799	LDEPL	NO	-0.46	-0.46	Phytenic acid (C20:1)	C20H38O2	M-H-	Fatty acids and conjugates

311.2245	311.222	311.2225	311.2228	LDEPL	NO	5.56	8.13	Dihydroxy-octadecadienoic acid	C18H32O4	M-H-	Fatty acids and conjugates
311.2934	311.2955	311.2955	311.2956	LDEPL	YES	-7.15	-6.83	Eicosanoic acid, Arachidic Acid (C20:0)	C20H40O2	M-H-	Fatty acids and conjugates
337.3128	337.3113	337.3108	337.3112	LDEPL	NO	4.61	4.31	Docosenoic acid (C22:1)	C22H42O2	M-H-	Fatty acids and conjugates
339.3271	339.3268	339.327	339.3269	LDEPL	NO	0.61	0.90	Docosanoic acid (C22:0)	C22H44O2	M-H-	Fatty acids and conjugates
342.2492	342.2471	342.2472	342.2472	LDOM	YES	5.87	6.16	N-palmitoyl-cysteine	C19H35NO2S	M-H-	Amino acids, peptides, and analogues
365.3459	365.3425	365.3423	365.3425	LDEPL	NO	9.39	9.39	Nervonic acid (C24:1)	C24H46O2	M-H-	Fatty acids and conjugates
366.3509	366.3458	366.347	366.3459	LDEPL	NO	13.63	13.90	Isotope Nervonic acid (C24:1)	(13C)C23H46O2	M-H-	Fatty acids and conjugates
367.3568	367.3581	367.358	367.3582	LDEPL	YES	-3.73	-3.46	Lignoceric acid (C24:0)	C24H48O2	M-H-	Fatty acids and conjugates
368.2635	368.2629	368.2625	368.2629	LDOM	NO	1.73	1.73	N-palmitoyl methionine	C21H41NO3S	M-H2O-H-	Amino acids, peptides, and analogues
370.2817	370.2784	370.2784	370.2785	LDOM	NO	8.58	8.85	Octadecenylamino-sulfanylpropanoic acid	C21H41NO2S	M-H-	
393.3760	393.3738	393.3741	393.3738	LDEPL	YES	5.59	5.59	Ximenic acid (C26:1)	C26H50O2	M-H-	Fatty acids and conjugates

399.2685	399.2684	399.2688	399.2687	LDOM	YES	-0.45	0.31	Oxosulfanyl-oxazolidinyl-octadecanamide	C21H40N2O3S	M-H-	
400.2756	400.2721	400.2726	400.2721	LDOM	NO	8.86	8.86	Isotope Oxosulfanyl-oxazolidinyl-octadecanamide	(13C)C20H40N2O3S	M-H-	
413.2867	413.2838	413.2847	413.2843	LDOM	NO	5.89	7.10	Methyl-hexadecylcarbamothioylamino-oxobutanoate	C22H42N2O3S	M-H-	
427.2994	427.2999	427.2992	427.3	LDOM	NO	-1.33	-1.10	Ethyl-hexadecylcarbamothioylamino-oxobutanoate	C23H44N2O3S	M-H-	
438.3002	438.299	438.2985	438.299	LDOM	YES	2.84	2.84	LysoPE(O-16:0)	C21H46NO6P	M-H-	Glycerophosphoethanolamines
439.3034	439.3019	439.3015	439.3024	LDOM	YES	2.19	3.33	Isotope LysoPE(O-16:0)	(13C)C20H46NO6P	M-H-	Glycerophosphoethanolamines
450.2685	450.2628	450.2627	450.2626	LDOM	NO	13.15	12.70	LysoPE(16:1)	C21H42NO7P	M-H-	Glycerophosphoethanolamines
455.3317	455.3309	455.3313	455.3313	LDOM	NO	0.89	1.77	Amino-sulfanylheptenyl-amino-oxooctadecanoic acid	C25H48N2O3S	M-H-	
456.3389	456.3342	456.335	456.3347	LDOM	YES	9.28	10.38	Isotope Amino-sulfanylheptenyl-amino-oxooctadecanoic acid	(13C)C24H48N2O3S	M-H-	
476.2815	476.2784	476.2776	476.278	LDOM	YES	7.43	6.59	Lyso-PE(18:2)	C23H44NO7P	M-H-	Glycerophosphoethanolamines
478.2944	478.2937	478.2951	478.2939	LDOM	YES	0.97	1.39	Lyso-PE(18:1)	C23H46NO7P	M-H-	Glycerophosphoethanolamines

483.2762	483.2727	483.2733	483.2728	LDOM	YES	7.13	7.33	Lyso-PG(16:0)	C22H45O9P	M-H-	Glycerophosphoglycerols
483.3705	483.3623	483.3622	483.3626	LDOM	NO	16.30	16.92	Octadecanyloxy-propoxy-thiazolyl-propanamide	C27H52N2O3S	M-H-	
489.3136	489.3148	489.3148	489.3156	LDOM	NO	-4.15	-2.52	Dimethoxyphenyl-hexadecylcarbamothioyl-propenamide	C28H46N2O3S	M-H-	
500.2818	500.2785	500.2781	500.2783	LDOM	YES	6.92	6.52	Lyso-PE(20:4)	C25H44NO7P	M-H-	Glycerophosphoethanolamines
760.5141	760.5141	760.5139	760.5134	LDOM	YES	0.97	0.05	PS(16:0/18:1)	C40H76NO10P	M-H-	Glycerophosphoserines
771.5173	771.5204	771.5165	771.5182	LDOM	YES	-1.15	-4.00	PG(18:1/18:2)	C42H77O10P	M-H-	Glycerophosphoglycerols
788.5518	788.5451	788.5455	788.5447	LDOM	YES	8.99	8.48	PS(18:0/18:1)	C42H80NO10P	M-H-	Glycerophosphoserines