

**Expanding the Applications of Rapid Evaporative Ionisation
Mass Spectrometry (REIMS): Microbial Detection, Diagnostics
and Metabolic Profiling**

A thesis submitted by:

Adam Burke

For the degree of Doctor of Philosophy (PhD)

2019

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-No Derivatives 4.0 International Licence (CC BY-NC-ND).

Under this licence, you may copy and redistribute the material in any medium or format on the condition that; you credit the author, do not use it for commercial purposes and do not distribute modified versions of the work.

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.

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.



Declaration of Originality

I confirm that I am the sole author of this thesis and the work presented throughout is my own. Any research experiments, papers, books or other articles used to aid in conducting this project have been referenced and acknowledged fully.

Abstract

The swift diagnosis of infection and identification of the causative microorganism is imperative to improving patient outcomes. Traditional identification techniques rely on microbial culturing, which is time consuming, labour intensive and subject to interpretation. Molecular methods are expensive and require extensive sample preparation. Commercial MALDI-ToF systems have become integrated into clinical microbiology laboratory workflows due to their ease of use, low per-sample cost, and ability to reduce sample turnaround times by one day. Although these systems have revolutionised clinical labs, they still largely rely on isolation and pre-culture of microorganisms prior to analysis. Some attempts have been made to develop “direct from sample analysis”, however protocols include extensive sample preparation steps and provide unreliable results in the case of poly-microbial samples.

The ambient Rapid Evaporative Ionisation Mass Spectrometry (REIMS) phospholipid based profiling technique has been previously shown to accurately classify isolated clinical microorganisms from agar media. REIMS offers an advantage over existing methods in that real time analysis can be conducted on intact cells with no sample preparation required. This project expands on previous work through the development of a methodology allowing liquid sample analysis, and its direct application to clinical specimens without the requirement of pre-culture.

In the analysis of clinical urine samples, culture negative and culture positive specimens were discriminated successfully, and species-specific biomarkers were identified. Species-specific biomarkers were also detected in synthetic mixed cultures, indicating the potential for REIMS in poly-microbial sample analysis. Additionally, the technique was applied to bovine milk in the investigation of mastitis, and was able to discern mastitis positive and negative milk samples. REIMS was shown to be a versatile culture free microbial identification system with applications outside of the human clinical laboratory, and offers an alternative to existing methods in the veterinary field, which also rely on microbial culture for non-ambiguous species identification.

Acknowledgements

Thank you to Prof. Zoltan Takats for providing me with the opportunity to undertake this project and for his supervision, Prof. Julian Marchesi for his supervision, and Waters Corporation for funding support.

This work is dedicated to my parents, Mick and Denise, who have encouraged, supported and believed in me throughout my life. A huge thanks to them, Lew, Lauren, Taylor and Zack for helping me stay grounded and for patiently sticking by me over the last four years.

I want to thank Kate for her moral support and daily pushes, Frankie for her guidance, Simon for his help with planning and data, and Emmanuelle, Dipa, Anna and all the guys in my group for listening to me grumble and helping me keep my sanity! Thank you also to Monica and the NHS staff in the microbiology lab at Charing Cross Hospital for providing me with samples and allowing me to infiltrate your workspace.

Last, but definitely not least, thanks to my friends back home in Liverpool for providing an escape when I need it, and for doing what friends do best.

Table of Contents

Abstract	iv
Acknowledgements.....	v
Figure List.....	ix
Table List	xiii
Abbreviations List	xiv
Bacterial Species Abbreviations	xvi
1 Background and Literature Review	- 17 -
1.1 Structure.....	- 17 -
1.2 Phylogeny and the Microbiota	- 20 -
1.3 Pathogens	- 22 -
1.4 Urinary Tract Infections.....	- 24 -
1.5 Traditional Identification Techniques	- 31 -
1.6 Genomic Methods	- 35 -
1.7 Clinical Mass Spectrometry	- 37 -
1.8 Ambient Ionisation Mass Spectrometry	- 39 -
1.8.1 Desorption Electrospray Ionisation	- 40 -
1.8.2 Direct Analysis in Real Time	- 41 -
1.8.3 Paper Spray Ionisation	- 42 -
1.8.4 Acoustic Mist Ionisation	- 44 -
1.9 Rapid Evaporative Ionisation Mass Spectrometry	- 45 -
1.10 Non-Human Diagnostic Microbiology	- 49 -
1.10.1 Bovine Mastitis	- 50 -
1.11 MS Statistical Methods	- 53 -
1.11.1 Pre-Processing	- 53 -
1.11.2 Univariate Analysis.....	- 54 -
1.11.3 Multivariate Analysis.....	- 55 -
1.12 Hypothesis/Aims	- 59 -
1.13 Objectives:.....	- 59 -
2 Materials and Methods.....	- 60 -
2.1 Health and Safety	- 60 -
2.2 Ethical Approval	- 60 -
2.3 REIMS Analysis	- 60 -
2.3.1 Bipolar REIMS.....	- 62 -
2.3.2 Monopolar Handpiece REIMS.....	- 63 -
2.3.3 Automated REIMS.....	- 65 -

2.4	Interface Configuration.....	- 69 -
2.5	Cleaning	- 70 -
2.6	Culturing	- 71 -
2.7	Cell line probes experiment.....	- 72 -
2.7.1	Probe Optimisation:.....	- 72 -
2.7.2	Method Comparison	- 73 -
2.8	Mixed Culture Experiment	- 74 -
2.8.1	Bacterial Culture and Identification	- 74 -
2.8.2	Rapid Evaporative Ionisation Mass Spectrometry Analysis	- 75 -
2.8.3	Data Processing and Statistical Analysis	- 75 -
2.8.4	Generation of <i>in silico</i> Mixed Culture Mass Spectra.....	- 76 -
2.9	Milk experiments.....	- 76 -
2.9.1	Cell Line Probe	- 76 -
2.9.2	Laser optimisation	- 77 -
2.9.3	LA-REIMS.....	- 77 -
2.10	Urine Analysis.....	- 78 -
2.10.1	Spiked Urine Method Development (pellet forceps)	- 78 -
2.10.2	LA-REIMS Method Development	- 79 -
2.10.3	Clinical Urine I: 96 Well Plate Pellet Analysis	- 80 -
2.10.4	Clinical Urine II: Microcentrifuge Tube Pellet Analysis	- 80 -
2.10.5	Tandem MS Analysis.....	- 81 -
2.11	AMI-MS	- 81 -
2.12	Data Analysis Workflow:	- 82 -
2.12.1	Pre-processing	- 82 -
2.12.2	Offline Model Builder	- 82 -
2.12.3	MetaboAnalyst	- 83 -
2.12.4	Spectral Richness Tool.....	- 84 -
2.12.5	Spectrum Quality Analyser Tool.....	- 84 -
2.12.6	Biomarker Identification	- 85 -
3	Optimisation and Setup	- 86 -
3.1	Cell Line Probes	- 86 -
3.1.1	Method Comparison	- 92 -
3.2	Laser Optimisation	- 97 -
3.3	Summary and Conclusions	- 99 -
4	Mixed culture analysis.....	- 100 -
4.1	Species-Specific Biomarkers	- 101 -
4.2	Discussion	- 104 -

5	Urine analysis	- 109 -
5.1	Bipolar Forceps Pellet Analysis.....	- 109 -
5.2	LA-REIMS method development	- 113 -
5.3	LA-REIMS Optimisation	- 120 -
5.4	Clinical Urine 1.....	- 123 -
5.5	Clinical Urine 2	- 133 -
5.6	Biomarker identification	- 141 -
5.7	Discussion	- 145 -
5.7.1	Bipolar forceps:	- 145 -
5.7.2	LA-REIMS Method Development	- 146 -
5.7.3	Clinical Urine Experiments.....	- 148 -
5.7.4	Conclusions.....	- 154 -
6	AMI-MS	- 156 -
6.1	Discussion	- 163 -
6.1.1	Conclusions.....	- 165 -
7	Bovine Mastitis	- 166 -
7.1	Discussion	- 175 -
7.1.1	Cell Line Probe REIMS.....	- 175 -
7.1.2	LA-REIMS.....	- 176 -
7.1.3	Conclusions.....	- 177 -
8	General Discussion.....	- 180 -
8.1	REIMS as a Microbial Analysis Technique	- 180 -
8.2	Application to Specific Sample Types	- 182 -
8.2.1	Competing Technology.....	- 182 -
8.2.2	Non-Human Applications.....	- 184 -
8.3	Strengths and Limitations	- 185 -
8.4	Further REIMS Development Work.....	- 187 -
8.5	Conclusions.....	- 188 -
	Publications List	- 190 -
	Conference Presentations	- 191 -
9	References.....	- 193 -
10	Supplementary	- 206 -

Figure List

Figure 1. Simplified representation of Gram-positive and Gram-negative bacterial cell walls	19 -
Figure 2. Causative agents of complicated and uncomplicated urinary tract infections	25 -
Figure 3. The pathogenesis of urinary tract infections.....	28 -
Figure 4. Typical workflow for identification of bacteria in clinical microbiology laboratories using MALDI-ToF systems.....	38 -
Figure 5. Representation of Acoustic Mist Ionisation technique	45 -
Figure 6. Bi-polar electrosurgical forceps REIMS.....	46 -
Figure 7. Automated monopolar REIMS	48 -
Figure 8. Monopolar handpiece REIMS (iKnife)	64 -
Figure 9. Automated Laser-Assisted REIMS.....	67 -
Figure 10. Laser beam spot diameter representation	69 -
Figure 11. Bovine milk lipid layer analysis representation.....	77 -
Figure 12. Modified cell line probes obtained from Waters Research Centre, Hungary	87 -
Figure 13. Signal quality, determined by signal:noise scores of spectral bins, of each modified cell line probe.....	88 -
Figure 14. Example mass spectra of the phospholipid 600-1000 m/z range generated by the different cell line probe types at 8 W	89 -
Figure 15. Representation of yellow long (25 mm length) probe during sample analysis	90 -
Figure 16. Spectral quality, determined by signal:noise ratio scores, of each power level examined for the yellow long (25 mm electrode length) probe	91 -
Figure 17. Mass spectra comparison of an identical <i>Klebsiella pneumonia</i> isolate analysed with monopolar iKnife, bipolar forceps and bipolar cell line probe REIMS.....	93 -
Figure 18. Shannon diversity plot and mean spectral quality scores of different REIMS analysis methods for 20 clinical bacterial isolates (a), and for a single isolate of <i>Klebsiella pneumoniae</i> (b)	94 -
Figure 19. Partial least squares discriminant analysis of Gram positive only (a) and Gram negative only (b) bacterial species replicates for cell line probe REIMS pellet analysis	96 -
Figure 20. Mean combined spectral quality scores, of the 600-1000 m/z region, of each bovine milk sample replicate analysed at increasing laser supply powers (2.5 mm lens distance from sample surface)	97 -
Figure 21. Mass spectra of the 600-1000 m/z range produced from LA-REIMS with a lens distance of 2.5 mm from bovine milk lipid layer surface.....	98 -
Figure 22. 2-dimensional unsupervised principal component analysis of the 600-900 m/z mass range of pure culture isolates	101 -
Figure 23. Comparison of <i>in silico</i> and <i>in situ</i> mixed culture spectral data.....	102 -
Figure 24. Mass spectra comparison of pure isolates <i>Staphylococcus aureus</i> (a) and <i>Pseudomonas aeruginosa</i> (b) along with a mixed culture mass spectra of the 2 isolates in a 1:1 ratio (c).....	103 -

Figure 25. Top ranked species-specific biomarker, based on FDR corrected <i>P</i> value, for each bacterial species within mixed cultures (<i>Escherichia coli</i> , <i>Staphylococcus aureus</i> , <i>Enterococcus faecalis</i> and <i>Pseudomonas aeruginosa</i>)	104 -
Figure 26. PCA (a) and PLS-DA (b) of all bacterial species spiked into culture negative healthy donor urine at dilutions < 10x	109 -
Figure 27. Random forest analysis of healthy, culture negative donor urine spiked with four clinical bacterial species	110 -
Figure 28. Volcano plot of log ₂ fold change (x-axis) and -log ₁₀ FDR corrected <i>P</i> value scores (y-axis) for <i>Staphylococcus aureus</i> (SAUR) spiked urine samples	111 -
Figure 29. Significant biomarkers in <i>S. aureus</i> , determined by FDR corrected <i>P</i> value	112 -
Figure 30. Nitrocellulose membrane (a-b) and filter paper (c-d) LA-REIMS analysis of <i>Escherichia coli</i> (ECOL), <i>Enterococcus faecalis</i> (EFAL), <i>Pseudomonas aeruginosa</i> (PAER) and <i>Staphylococcus aureus</i> (SAUR) spiked into culture negative healthy donor urine at dilutions < 10x	115 -
Figure 31. Nitrocellulose membrane (a) and filter paper (b) LA-REIMS random forest analysis of culture negative donor urine spiked with four clinical bacterial species	116 -
Figure 32. Volcano plots of log ₂ fold change (x-axis) and -log ₁₀ FDR corrected <i>P</i> value scores (y-axis) for <i>Staphylococcus aureus</i> (SAUR) spiked urine samples analysed with nitrocellulose membrane (a) and filter paper (b) LA-REIMS	117 -
Figure 33. Mean combined spectral richness score box plot (a) and quality scores (b) of clinical urine samples analysed with nitrocellulose membrane and filter paper LA-REIMS	118 -
Figure 34. Presence of <i>Staphylococcus aureus</i> significant biomarkers, identified in bipolar forceps REIMS of urine pellets, and in nitrocellulose membrane (a) and filter paper (b) LA-REIMS, across reducing sample concentrations	119 -
Figure 35. Mass spectra of identical urine pellets, in the 600-1000 <i>m/z</i> range, of continuous (a, b) and pulsed (c-f) laser firing modes at varying wattages	121 -
Figure 36. Mean combined spectral richness score box plot (a) and mean combined spectral quality scores (b) of urine pellets analysed with LA-REIMS, in pulsed and continuous laser firing modes across a 1-3 W power range	122 -
Figure 37. PCA (a,c) and PLS-DA (b,d) plots, with species and culture negative/positive groupings, of clinical urine pellet LA-REIMS analysis, 600-1000 <i>m/z</i> range	125 -
Figure 38. PCA (a) and PLS-DA (b) plots, with <i>Escherichia coli</i> (ECOL) verses culture negative grouping, of clinical urine pellet LA-REIMS analysis, 600-1000 <i>m/z</i> range	126 -
Figure 39. Univariate ROC curves of top two most important discriminatory features for <i>Escherichia coli</i> (ECOL) classification, identified through a multivariate random forest selection algorithm ...	127 -
Figure 40. Mass spectra comparison of a culture positive clinical urine sample containing <i>Escherichia coli</i> (a) and a culture negative sample spiked with a pure <i>E. coli</i> isolate (b)	128 -

Figure 41. PCA (a) and PLS-DA (b) plots, with <i>Escherichia coli</i> (ECOL) verses culture negative grouping, of clinical urine pellet LA-REIMS analysis, 600-1000 <i>m/z</i> range, white blood cell positive samples only	129 -
Figure 42. Clinical urine pellet LA-REIMS random forest analysis, 600-1000 <i>m/z</i> range, white blood cell positive samples only	130 -
Figure 43. Volcano plot of log ₂ fold change (<i>x</i> -axis) and -log ₁₀ FDR corrected <i>P</i> value scores (<i>y</i> -axis) for culture positive <i>Escherichia coli</i> (ECOL) containing and culture negative clinical urine sample pellets analysed with LA-REIMS, white blood cell positive samples only	130 -
Figure 44. Univariate ROC curve of model prediction performance (a) through multivariate random forest selection algorithm, and ROC curve permutation test (b)	132 -
Figure 45. PCA (a,c) and PLS-DA (b,d) plots, with <i>Escherichia coli</i> (ECOL)/culture negative grouping, of clinical urine pellet LA-REIMS analysis from microcentrifuge tubes, 600-1000 <i>m/z</i> range	134 -
Figure 46. Clinical urine pellet LA-REIMS, from microcentrifuge tubes, random forest analysis, 600-1000 <i>m/z</i> range, white blood cell positive samples only	136 -
Figure 47. Volcano plot of log ₂ fold change (<i>x</i> -axis) and -log ₁₀ FDR corrected <i>P</i> value scores (<i>y</i> -axis) for culture positive <i>Escherichia coli</i> (ECOL) containing and culture negative (NEG) clinical urine sample pellets analysed from microcentrifuge tubes with LA-REIMS, white blood cell positive samples only.....	136 -
Figure 48. Univariate ROC curves of the top two most important discriminatory features in <i>Escherichia coli</i> (ECOL) classification, identified through multivariate random forest selection algorithm -	138 -
Figure 49. Univariate ROC curve of model prediction performance through multivariate random forest selection algorithm (a), and ROC curve permutation test (b).....	140 -
Figure 50. Mass bin 747.45 <i>m/z</i> (PG 16:0/18:1) (Accurate mass 747.52 <i>m/z</i> [M-H] ⁻) biomarker product ion spectra from an <i>Escherichia coli</i> positive clinical urine sample (a) and Metlin experimental database hit METID40844 (b)	144 -
Figure 51. Example spectral profiles of filter paper (a,c) and pellet (b,d) LA-REIMS analysis of a <i>Klebsiella oxytoca</i> isolate cultured with a D-Serine carbon source	158 -
Figure 52. Example AMI-MS spectral profiles of the metabolite 50-400 <i>m/z</i> region (a) and phospholipid 600-1000 <i>m/z</i> region (b) of a <i>Klebsiella oxytoca</i> isolate cultured with a D-Serine carbon source-	159-
Figure 53. Shannon diversity plot and mean combined spectral quality scores of AMI-MS, filter paper and pellet LA-REIMS analysis of a <i>Klebsiella oxytoca</i> isolate of the 50-400 (a) and 600-1000 <i>m/z</i> (b) ranges	160 -
Figure 54. Mirrored spectral comparison of the 50-400 <i>m/z</i> metabolite region (a) and 600-1000 <i>m/z</i> phospholipid region (b) of a single isolate of <i>Klebsiella oxytoca</i> cultured with a D-Serine carbon source, analysed with AMI-MS and pellet LA-REIMS	161 -
Figure 55. FDR corrected <i>P</i> values (-log ₁₀) verses mass bin scatter plot of the 50-400 <i>m/z</i> (a) and 600-1000 <i>m/z</i> (b) range of pellet LA-REIMS and AMI-MS analysed <i>Klebsiella oxytoca</i> cultured with a range of 192 carbon sources	162 -

Figure 56. Shannon diversity box plot (a) and mean combined quality scores (b) of bovine mastitis sample pellet and layers analysed with cell line probe REIMS, 600-1000 <i>m/z</i> region	167 -
Figure 57. PCA and PLS-DA plots of cell line probe pellet (a,b) and lipid layer (c-f) REIMS analysis of bovine milk 600-1000 <i>m/z</i> range	169 -
Figure 58. Heatmap clustering of top 25 most significant features for mastitis positive verses negative classification, determined through t-test FDR corrected <i>P</i> value rankings, of sample pellets (a) and lipid layer (b) samples analysed through cell line probe REIMS, 600-1000 <i>m/z</i> range.....	170 -
Figure 59. PCA and PLS-DA plots of LA-REIMS lipid layer analysis of bovine milk, 600-1000 (a-d) and 50-500 (e,f) <i>m/z</i> range	171 -
Figure 60. Heatmap clustering of top 25 most significant features for mastitis positive verses negative classification, determined through t-test FDR corrected <i>P</i> value rankings, of lipid layer samples analysed through LA-REIMS, 600-1000 <i>m/z</i> range	172 -
Figure 61. Genera specific biomarkers of mastitis positive bovine milk samples, identified through ANOVA, analysed via lipid layer cell line probe (a) and LA-REIMS (b)	173 -
Figure 62. Three component PLS-DA plot comparisons of cell line probe REIMS (a,c) and LA-REIMS (b,d) analysis of bovine mastitis milk samples.....	174 -

Table List

Table 1. Instrument parameters for REIMS analysis	- 61 -
Table 2. Clinical bacterial isolates and associated culture conditions (CBA: Columbia Blood Agar, CHOC: Chocolate Agar) utilised in cell line probe optimisation for REIMS.....	- 72 -
Table 3. Significant biomarker detection rates in culture negative urine samples spiked with <i>Staphylococcus aureus</i> (SAUR), <i>Pseudomonas aeruginosa</i> (PAER), <i>Escherichia coli</i> (ECOL) and <i>Enterococcus faecalis</i> (EFAL)	- 112 -
Table 4. Clinical Urine 1 overall unlabelled sample prediction matrix.....	- 131 -
Table 5. Clinical Urine 2 overall unlabelled sample prediction matrix	- 139 -
Table 6. Significant biomarker databases matches for <i>Escherichia coli</i> , analysed from clinical urine samples	- 142 -

Abbreviations List

°C	Degrees Celsius
AMI	Acoustic Mist Ionisation
AMR	Antimicrobial Drug Resistance
ANOVA	Analysis of Variance
AUC	Area Under Curve
CBA	Columbia Blood Agar
CF	Cystic Fibrosis
CFU	Colony Forming Unit
CHOC	Chocolate Agar
CMT	California Mastitis Test
cmm	Cells per cubic millimetre
CO ₂	Carbon Dioxide
CSF	Cerebrospinal Fluid
CSV	Comma Separated Values
DART	Direct Analysis in Real Time
DBS	Dried Blood Spot
DESI	Desorption Electrospray Ionisation
DNA	Deoxyribonucleic Acid
DUS	Dried Urine Spot
ELISA	Enzyme Linked Immunosorbent Assay
ESI	Electrospray Ionisation
FAB	Fast Atom Bombardment
FAME	Fatty Acid Methyl Ester
FDR	False Discovery Rate
G	Gram
<i>g</i>	Gravitational Force
GC	Gas Chromatography
GP	General Practitioner
HCCA	α-Cyano-4-Hydroxycinnamic Acid
HPLC	High Performance Liquid Chromatography
IBD	Inflammatory Bowel Disease
ICHNT	Imperial College Healthcare NHS Trust
L	Litre
LA-REIMS	Laser Assisted Rapid Evaporative Ionisation Mass Spectrometry
LC	Liquid Chromatography
Leu-enk	Leucine Enkephalin
LPS	Lipopolysaccharide

LSD	Least Significant Difference
LTP	Low Temperature Plasma Ionisation
M	Molar
m/z	Mass to Charge Ratio
MALDI	Matrix Assisted Laser Desorption Ionisation
MCCV	Monte-Carlo Cross Validation
Min	Minutes
mL	Millilitre
mM	Millimolar
mm	Millimetre
MRM	Multiple Reaction Monitoring (Target Enhancement)
MRSA	Methicillin Resistant <i>Staphylococcus aureus</i>
MS	Mass Spectrometry
ms	Milliseconds
MS/MS	Tandem Mass Spectrometry
MSSA	Methicillin Sensitive <i>Staphylococcus aureus</i>
NEG	Negative
nG	Nanogram
NHS	National Health Service
nL	Nanolitre
OMB	Offline Model Builder
OOB	Out of Bag (error)
P	Probability (value)
PA	Phosphatidic acid
PC	Phosphatidylcholines
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PE	Phosphatidylethanolamine
PEEK	Polyetheretherketone
PG	Phosphatidylglycerol
PI	Phosphatidylinositol
PLS-DA	Partial Least Squares Discriminant Analysis
POS	Positive
ppm	Parts per Million
pps	Pulses per Second
PRESS	Predicted Residual Sum of Squares
PTFE	Polytetrafluoroethylene
QMMS	Quality Milk Management Services

RBC	Red Blood Cell
REIMS	Rapid Evaporative Ionisation Mass Spectrometry
RF	Random Forest
RMS	Root Mean Square
ROC	Receiver Operator Characteristic
rRNA	Ribosomal Ribonucleic Acid
SCC	Somatic Cell Count
Sec	Seconds
SS	Starting Sample
STSS	Streptococcal Toxic Shock Syndrome
TIC	Total Ion Current
ToF	Time of Flight (mass analyser)
UK	United Kingdom
UKAS	United Kingdom Accreditation Service
UPEC	Uropathogenic <i>E. coli</i>
US	United States
UTI	Urinary Tract Infection
UV	Ultraviolet
VIP	Variable Importance in Projection
WBC	White Blood Cell
WRC	Waters Research Centre
µG	Microgram
µL	Microlitre
µm	Micrometre

Bacterial Species Abbreviations

ABAU	<i>Acinetobacter baumannii</i>	MLUT	<i>Micrococcus luteus</i>
BFRA	<i>Bacteroides fragilis</i>	MORG	<i>Morganella morganii</i>
CALB	<i>Candida albicans</i>	PAER	<i>Pseudomonas aeruginosa</i>
CKOS	<i>Citrobacter koseri</i>	PMIR	<i>Proteus mirabilis</i>
ECLO	<i>Enterococcus cloacae</i>	SAUR	<i>Staphylococcus aureus</i>
ECOL	<i>Escherichia coli</i>	SEPI	<i>Staphylococcus epidermidis</i>
EFAL	<i>Enterococcus faecalis</i>	SMAL	<i>Stenotrophomonas maltophilia</i>
EFAM	<i>Enterococcus faecium</i>	SORA	<i>Streptococcus oralis</i>
KOXY	<i>Klebsiella oxytoca</i>	SPNE /	<i>Streptococcus pneumoniae</i> /
KPNE	<i>Klebsiella pneumoniae</i>	BSTP	Group B <i>Streptococcus</i>

1 Background and Literature Review

1.1 Structure

Bacterial cell morphology falls into the three main groups of bacillus, coccus and spiral, with some overlap and variations, such as with coccobacillus like cells exhibited by *Haemophilus influenzae* (1). Bacterial cells possess a cell membrane and cell wall, known as the cell envelope. The cell wall is composed of a mesh like layer of peptidoglycan, consisting of chains of polysaccharide cross linked by flexible peptide bridges laying outside of the cell membrane (2). Peptidoglycan structure, conferred by peptide amino acid sequences, is variable and differs between bacterial species, and is responsible for forming distinct cell shapes which bacteria can direct by preferentially synthesising new peptidoglycan in specific locations, for example on lateral walls which leads to a bacilli (3). The peptidoglycan wall is essential in withstanding turgor pressure and conferring rigidity to the cell, and degradation or disruption of synthesis leads to osmotically vulnerable cells and lysis (2). β -lactam antibiotics exploit this by binding penicillin binding proteins on new peptidoglycan chains, inhibiting peptide bridge formation (4). Cell growth and division is also dependent on peptidoglycan synthesis, with synthesis occurring on division planes and septa, which form part of daughter cells in coccus morphology, or along lateral walls to elongate cells into a bacilli shape as mentioned previously (3, 5). The exception to this is with the mollicutes, which are a class of bacteria possessing no cell walls and are the most simple of bacteria, or any self-replicating cell in nature, possessing a single cell membrane, ribosomes and double stranded circular DNA as their only organelles (6). Rather than relying on a peptidoglycan wall for structural rigidity, members of the mollicutes, for example *Mycoplasma* spp., possess a unique, internal, protein based cytoskeleton, which also confers cell shape and motility functions (6).

Bacterial cell walls are classified into two types based on crystal-violet dye retention, which is utilised in a staining technique coined the Gram stain and discovered by Hans Christian Gram in 1884 (Figure 1) (7, 8). Gram-positive bacteria possess a thick layer of peptidoglycan (20-80 nm) on the outside of the cell membrane, which heavily retains crystal-violet dye causing cells to appear purple when viewed through a microscope.

Gram-negative cells possess a periplasmic space, created by an outer and inner membrane, in which there is a single layer of peptidoglycan of 1-7 nm thickness linked by lipoproteins to the lipid A outer membrane (9). Crystal-violet is not retained by the thinner peptidoglycan layer and is instead washed away by alcohol, but the cells are counter stained by a second dye causing them to appear pink when viewed through a microscope. Gram-positive bacteria have a higher susceptibility to antibiotics, due to the mode of action of many antibiotics being to inhibit synthesis of the peptidoglycan layer (10). Gram staining is historically the primary step in classification and identification of bacteria, although not all bacterial cells stain according to Gram-positive or negative attributes. Acid fast bacteria, such as *Mycobacterium* spp., possess an outer mycolic acid layer of their cell walls, which results in resistance to the stain wash and causes cells to appear Gram-intermediate and potentially be misclassified as Gram-positive (11). A different staining technique is utilised in such cases, but requires a trained microbiologist to recognise. Some bacteria, such as *Clostridium* spp., have variable peptidoglycan thickness throughout different growth stages, causing cells to stain different shades and appear Gram-variable, again requiring a trained microbiologist to recognise (12). The mollicutes, including *Mycoplasma* spp. do not possess peptidoglycan at all and so do not retain primary or counter stain dyes, instead remaining colourless (13).

Bacterial cell membranes are comprised of proteins and lipids, which are primarily amphiphilic glycerophospholipids forming a bilayer, and is responsible for facilitating the active transport and diffusion of small molecules across the membrane and into or out of the cytoplasm. Glycerophospholipids are comprised of two fatty acid chains, a variable head group, phosphate group and glycerol moiety, with some of the commonly observed classes in bacteria being phosphatidylglycerols (PGs), phosphatidylethanolamines (PEs), phosphatidic acids (PAs), cardiolipins (CLs) and phosphatidylcholines (PCs) (14). The fatty acid chains vary in length from 14-20 carbons containing either *iso* or *anteiso* branches or a single *cis* double bond and importantly, bacteria can alter fatty acid composition to direct changes in membrane viscosity, which in turn alters the permeability of the membrane to hydrophobic molecules as well as effecting protein-protein interactions and active transport mechanisms (15).

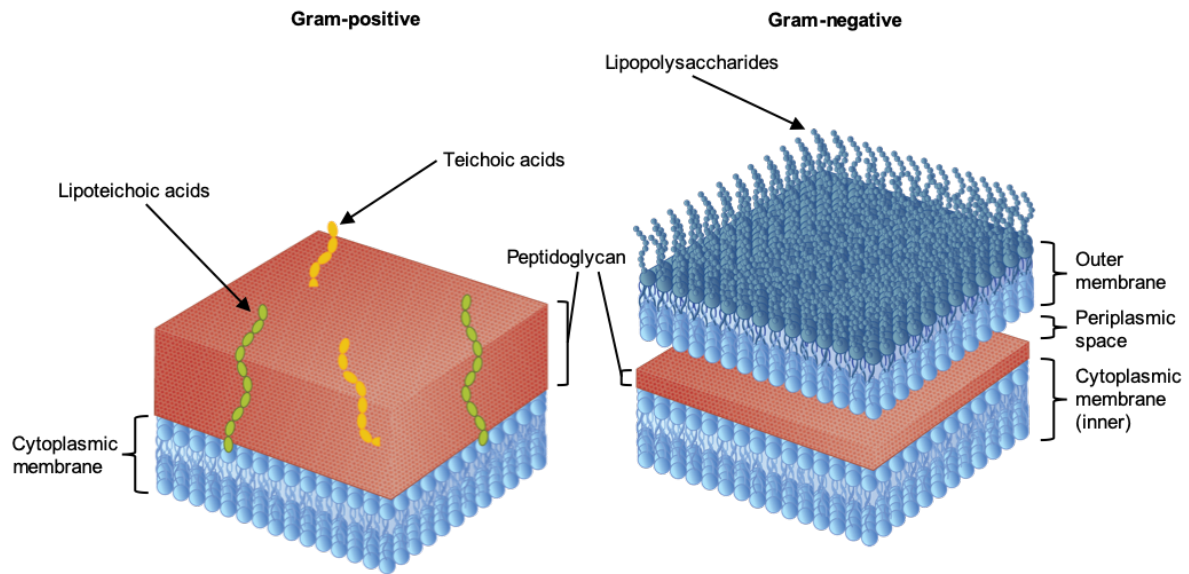


Figure 1. Simplified representation of Gram-positive and Gram-negative bacterial cell walls. Gram-positive bacteria possess a single cytoplasmic membrane and a thick peptidoglycan layer. Gram-negative bacteria possess an outer and inner membrane between which is a periplasmic space containing a thin peptidoglycan layer. Membrane proteins are not shown.

The ability of bacterial cells to initiate these biophysical changes to the membrane through new phospholipid synthesis, as well as in existing phospholipid structure within the bilayer, is critical to their survival during environmental stresses, such as in changes to pH, temperature, osmotic pressure and salinity (16). Saturated, straight chain fatty acids cluster more tightly together within the membrane than *cis* double bonded unsaturated chains possessing a kink, leading to reduced membrane permeability and an increased phase transition temperature (16). By increasing the synthesis and incorporation of lower melting point *anteiso*-branched chain fatty acids into the membrane, cells can adapt to growth at lower temperatures by maintaining a homeostatic fluidity, as observed in *Bacillus* spp. and *Listeria* spp. in response to cold shock treatment (17, 18). *Listeria* spp. also control the proportion of *anteiso*-branched fatty acids in response to changes in pH, increasing the incorporation into the membrane at >8.0 pH to increase membrane fluidity and resist surfactant and alkali related damages (19). This function is important to the role of *Listeria monocytogenes*, for example, as a food borne pathogen, which can survive and grow at low temperatures occurring during refrigeration of food products and pH stresses observed during

decontamination and processing (20). Fatty acid profile changes have also been observed in response to pH stress by *Escherichia coli*, *Pseudomonas* spp., *Salmonella* spp. and *Streptococcus mutans* (21). Bacterial cells may also alter biophysical membrane properties via modification of surface charge achieved through lysinylation of the glycerol head group of acidic PGs, decreasing the overall negatively charged bilayer and conferring resistance of the membrane to cationic peptides, defensin incorporation and neutrophil killing (15). Biophysical membrane alterations respondent to the environment therefore allows bacterial cells to rapidly adapt, subvert the innate immune system, conserve energy, and maintain membrane homeostasis, allowing energy to be directed toward other cell functions such as growth, which is essential in pathogen survival within a host as well as in free living environmental bacteria.

1.2 Phylogeny and the Microbiota

Bacterial taxonomy was historically established by the observed laboratory culture phenotype, including Gram stain status, growth conditions and cell shape. This method was superseded by the adoption of molecular techniques, which utilises gene sequencing to assign bacteria taxonomic rankings under the three-domain classification system. Environmental sequencing of the highly conserved 16S rRNA gene has enabled the classification of uncultured clades and phyla, which have been previously unknown due to culture limitations, revealing an enormous increase in diversity and the realisation that the vast majority of life in the Bacteria domain may be undiscovered, with only up to 1% of soil dwelling microorganisms readily culturable using standard techniques (22). New advances in whole genome sequencing are continuing to improve bacterial taxonomy by providing higher taxonomic resolution than the 16S rRNA gene sequencing approach, with methods such as average nucleotide identity providing insights into genome comparisons not previously possible (23).

Sequencing of extracted DNA from human samples has revealed over three million genes of microbial origin, and around 100 trillion microbes making up the human microbiota, the majority of which is from the gastrointestinal tract (24). The Firmicutes and Bacteroidetes make up 90% the phyla within the gut, which is intimately involved

in nutrition, metabolism, immune function and physiology (25). In the metabolic sense, bacteria synthesise vitamins and ferment non-digestible substrates into short chain fatty acids, which humans can utilise (25). In particular, propionate regulates gluconeogenesis and satiety, and a reduction of short chain fatty acids produced in the gut is linked with an increased insulin resistance and obesity rate (24). Colonisation of the gut occurs at birth through the ingestion of vaginal secretions containing an inoculum of the mother's microbiota, with caesarean section delivered infants demonstrating a reduced microbial load compared to vaginal births, along with reduced immunological functions (26). Commensal bacteria in the gut exhibit a competitive exclusion effect by attaching to epithelial cells and occluding pathogens from doing so, as well as by competing for food and affording a protective barrier to disease to the host (27).

The gut microbiota effects the innate and adaptive immune system and regulates immune homeostasis of the host, with dysregulation sometimes leading to autoimmune disease (28). Interactions of commensal microorganisms with gut macrophages produces a tolerogenic response, in which no inflammatory cytokines are produced (29). CD4⁺ T cell maturation is modulated by the microbiota, and in studies performed on germ-free mice, a significant reduction of CD4⁺ T cells is observed, and several commensal bacterial species have been linked with specific CD4⁺ T cell pathways (30-32). Since the gut microbiota has significant impact on immune homeostasis, the composition of gut microbiota therefore has a range of effects on health and disease, with some species conformations linked to autoimmune and metabolic disease. In inflammatory bowel disease (IBD) patients, an increase in Proteobacteria numbers and a reduction in Firmicutes and Bacteroidetes is observed, responsible for inducing an increased inflammatory immune response (33). Antibiotic and probiotic therapy may lead to remission of IBD, indicating the role of the microbiota in the disease, although resistance to therapy usually emerges relatively rapidly (34). The mechanism for the development of IBDs, such as Crohn's disease and ulcerative colitis, is postulated to be due to a reduction in beneficial, immune regulating commensals leading to a disruption in immune homeostasis rather than single or multiple disease causing agents, but specific taxonomic increases in *Fustobacterium* spp. and *E. coli* are characteristically

observed (35). In some trials, a faecal microbiota transplant from healthy individuals into IBD patients resulted in remission of disease, demonstrating that restoration of a healthy microbiota can reverse inflammation, and also has applications in other gastrointestinal syndromes such as recurrent *Clostridium difficile* associated diarrhoea (36, 37).

1.3 Pathogens

Pathogens are agents of disease and have a wide economic impact on human health and agriculture. In humans, opportunistic bacteria living on the skin, within the gastrointestinal tract or in the environment may gain entry through cuts or abrasions, perforations of the bowel, disruptions in the commensal microbiota or by a compromised immune system (38). The immune system can be compromised by a wide range of factors, such as aging, immunosuppressant drugs, fatigue, infectious or non-infectious disease, malnutrition, pregnancy or for genetic reasons (39, 40). Primary pathogens, such as *Bacillus anthracis* and *Vibrio cholerae*, do not require host defences to be weakened through wounds or a compromised immune system, as they have evolved specialised and complex mechanisms for subverting host defences and are capable of causing infection and disease in otherwise healthy individuals (41). Opportunistic pathogens include common skin commensals such as *Streptococcus* and *Staphylococcus* spp., environmental and skin dwelling *Pseudomonas aeruginosa* and gut commensal *C. difficile*.

Streptococcus pyogenes can cause skin (cellulitis, impetigo) and throat (pharyngitis) infections, pneumoniae, sepsis and streptococcal toxic shock syndrome (STSS), with >500 000 deaths associated with infection by the species each year, reaching a morbidity of around 12% (42, 43). Some strains possess a hyaluronic acid capsule which enables mucoid colony formation, adherence to the epidermis and confers resistance to phagocytosis, and produce hyaluronidase to facilitate breakdown of host hyaluronic acid, increasing permeability of host mucosa and enabling colonisation and spreading of bacterial cells (44, 45). A surface antigen, known as M protein, also helps in avoiding phagocyte ingestion by binding to host proteins and inhibiting opsonisation (46). STSS

occurs when a superficial infection of the epidermis progresses to the point of necrotising fasciitis, or from direct introduction through compromised skin such as in a burn or cut, often with rapid onset (47). The M protein antigen is responsible for promoting the massive immune response characteristic of toxic shock (48). Once released from the surface of bacterial cells, M antigen is capable of binding to fibrinogen in blood plasma, forming aggregates which then bind to granular leukocytes and elicit an oxidative burst following binding to endothelial cells (49). The endothelium is subsequently damaged, leading to vascular leakage, thrombophilia, hypotension and organ damage (50).

Staphylococcus spp. inhabit the skin and nasal passages of healthy people, but can cause skin infections, endocarditis, localised or systemic infections from the insertion of contaminated medical devices and toxic shock syndrome (51). *Staphylococcus aureus* is persistently carried in 20% of healthy nasal passages as a commensal, and intermittently in 30% (52). Carriage in the nasal passage is important in transmission of bacterial cells via hands and onto surfaces, through airborne transmission to surfaces, or through direct contact with other people (53, 54). Once introduced to the nasal cavity, the bacterial cells are able to avoid host immune responses and adhere to mucosa through a variety of methods, including resistance mechanisms to antimicrobial peptides present in nasal secretions, production of protein A which binds to and deactivates immunoglobulins, and surface protein G binding to the epithelium, along with other adhesins (54-56). *S. aureus* is of particular concern in nosocomial infections, due to the emergence of resistance to methicillin and vancomycin, which is also increasing in community acquired infections as well (57-59). *S. aureus* is a leading cause of sepsis when introduced to the blood stream through insertion of medical devices or intravenous lines, and is associated with a high mortality rate of up to 39% in some populations (60). *Staphylococcus epidermidis* is likewise strongly associated with infections arising from medical devices, due to the formation of biofilms on the surface of intravenous lines, prosthetics and catheters (61). Biofilms are formed following cell adhesion to surfaces and subsequent growth and cell to cell adhesion, mediated by polysaccharide intercellular adhesin, leading to amassing of cells in layers (62). A pericellular matrix of polysaccharides surrounding bacterial cells mediates adhesion

both to surfaces, such as plastics used in medicine, and to other cells, as well as conferring protection from host immune defences and antimicrobial drugs by acting as a physical barrier (63). The formation of protective and persistent biofilms means that *S. epidermidis* is the leading cause of infections originating from indwelling medical devices, likely originating from contamination of device surfaces from patient or medical staffs' skin prior to insertion (64). Infection can be localised to the device, cause endocarditis originating from prosthetic heart valves and can also cause sepsis and other secondary infections following cell detachment from mature biofilms (65). Patients with endocarditis with *S. epidermidis* as a causative agent have a high risk (38%) of developing intracardiac abscesses, with a mortality rate of 24% (66). *S. aureus* can also cause endocarditis, and is the causative agent in up to 35% of all cases, with a mortality rate of >55% (67, 68).

1.4 Urinary Tract Infections

The most common community and nosocomial acquired infections are those of the urinary tract, with women, immunocompromised and catheterised patients at particularly high risk to acquiring infection (69). Approximately one third of women under the age of 24 will encounter a urinary tract infection (UTI) requiring treatment with antimicrobial drugs, and up to half of women will encounter the same within their lifetime (70-72). UTIs impose a significant public health and economic burden and have potential to cause distress and lower quality of life to patients, costing the US health system \$3.5 billion per year (73). UTIs, classified as cystitis if affecting the lower urinary tract or pyelonephritis if affecting the upper urinary tract, are uncomplicated in patients considered healthy with no abnormalities of the urinary tract, and are usually mono-microbial, with the most common causative agent being *E. coli* (Figure 1) (74, 75). Complicated UTIs arise in patients who are immunocompromised, catheterised, pregnant, possess a neurological or structural abnormality, have a neurological disease, urinary obstruction or renal failure (76, 77). Catheter associated UTIs are responsible for around 70-80% of all reported complicated UTIs, and are much more likely to be poly-microbial in nature than uncomplicated infections, with *E. coli*, *Klebsiella pneumoniae*, *Staphylococcus saprophyticus*, *Proteus mirabilis*, *Enterococcus faecalis*, *P.*

aeruginosa, group B *Streptococcus* and *S. aureus* typical infectious agents (Figure 1) (78-80).

Since bacteria are present and can be cultured from the urine of asymptomatic, healthy people, a positive UTI is usually defined as $>10^5$ CFU/mL, although lesser culture levels are accepted as positive if a known urinary pathogen is cultured (81, 82). This is complicated by patients often presenting with symptoms indicative of a UTI, but with a negative urine culture result returned in as much as 20% of women (83). For this reason, a UTI diagnosis is provided based on a combination of a positive double urine culture of the same causative organism in asymptomatic patients, a single positive urine culture in symptomatic or catheterised patients performed by a trained biomedical scientist and analysis of urinary tract symptoms by a physician (84, 85).

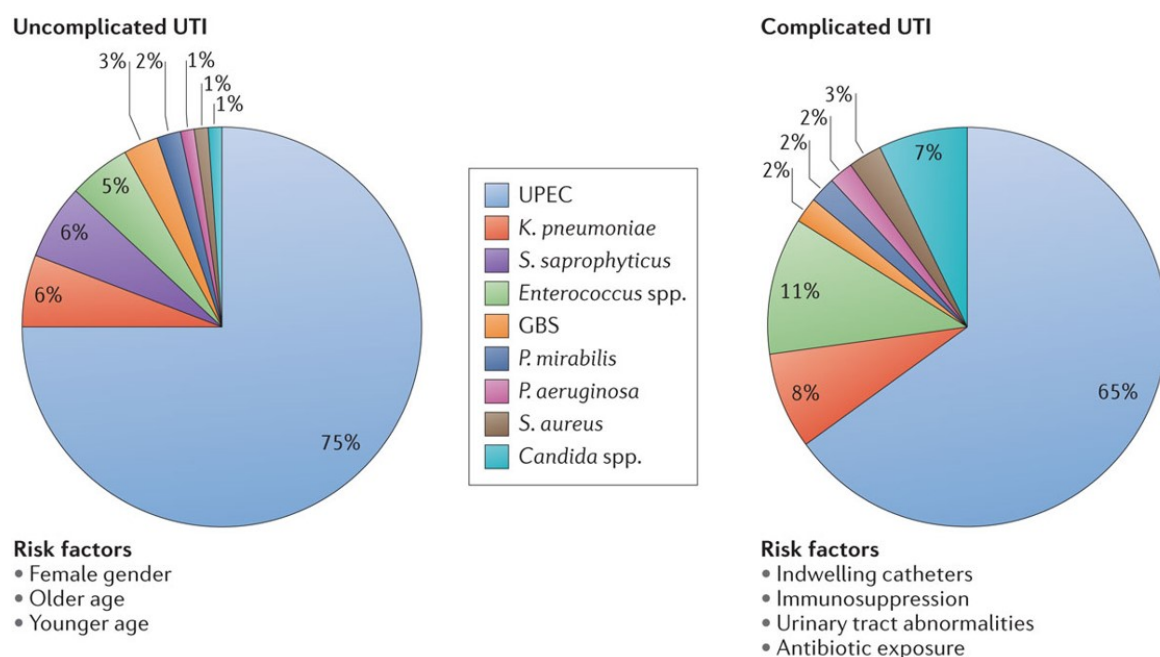


Figure 2. Causative agents of complicated and uncomplicated urinary tract infections. Urinary Tract Infections may be uncomplicated if no urinary tract abnormalities are present, or complicated if a structural abnormality is present or the patient is immunocompromised or catheterised. The most common causative agent for both UTI classifications is UPEC, with complicated UTIs having a much higher risk of a poly-microbial infection. Reproduced with permission from Springer Nature Publications © 2015 (70) (Supplementary 20).

Complications of UTIs can be potentially serious for certain demographics such as pregnant women, the elderly, children and immunocompromised patients, with catheter associated UTIs in particular being a common source of secondary acquired septicaemia, and accompanying a higher risk of mortality and morbidity (70, 86).

Uncomplicated cystitis is normally self-limiting and resolves without treatment, although symptomatic UTIs are ordinarily empirically treated with antibiotics to hasten resolution of symptoms and to reduce the risk of complications developing, such as secondary infections (87). Empirical antimicrobial therapy is utilised due to the length of time required for a definitive culture identification and antimicrobial susceptibility testing to be completed by clinical microbiology laboratories (88, 89). Although usually adequate for outpatients and community acquired UTIs, patients with indwelling catheters or other complications empirically treated with broad spectrum antibiotics see little to no improvement in clinical outcomes until a targeted treatment is able to be offered, due to the prevalence of multi-drug resistant pathogens in these circumstances (90, 91). Asymptomatic infections of the lower urinary tract are usually not treated with antimicrobial drugs, unless a complicated asymptomatic infection is diagnosed as these often progress to symptomatic cases (92). Treatment is avoided in these asymptomatic cases as spontaneous resolution typically occurs, and unnecessary antimicrobial therapy leads to negative drug side-effects for patients, increased emergence of antimicrobial resistance and human microbiome disruptions (93). Gastrointestinal microbiome disruption potentially leads to acute or chronic disease, such as antibiotic associated diarrhoea and pseudomembranous colitis as a result of opportunistic overgrowth of multi-drug resistant *C. difficile* (94, 95).

The majority of both complicated and uncomplicated UTIs are caused by *E. coli*, responsible for around 80% of all diagnosed UTIs in women, who are the most at risk demographic (96). Uropathogenic *E. coli* (UPEC) are unique strains within the species that have evolved adaptations to aid in host immune evasion, such as through the production of adhesins mediating attachment and invasion of uroepithelial cells, and through the formation of biofilms (Figure 3) (97). *K. pneumoniae* and *S. saprophyticus* also bind directly to the uroepithelium, and *P. aeruginosa*, *Pr. mirabilis* and

Enterococcus spp. are capable of creating biofilms to avoid host immune response (98-100). Adhesion to the uroepithelium by UPEC occurs when adhesin organelles, including P pili, type 1 pili, S pili and Dr family adhesins, bind to host cells. Adhesion to specific host cell residues or glycolipids present in the ascending urinary tract and kidney facilitate cystitis progression to the upper tract, leading to pyelonephritis (Figure 2) (101).

Invasion of uroepithelium cells follows adhesion of type 1 pili to superficial facet cells of the bladder lumen, and leads to persistent and recurring UTIs, which host defences such as bladder cell exfoliation, urine pH, osmolarity, iron scavenging lactoferrin, anti-adherence factors and bulk flow are unable to completely overcome if invasion and subsequent refuge occurs (102-107). Host immune responses to UTIs involves neutrophil, mast cell and macrophage recruitment to infection sites, as well as nitric oxide and defensin antimicrobial factor production by uroepithelial cells (108-110). Despite this response, UPEC can persist within the bladder lumen and within uroepithelial cell cytoplasm or vacuoles for weeks, and lead to increased susceptibility to future infections even if successfully cleared as a result of damage to the uroepithelium via inflammatory response and cell exfoliation (105, 106, 111, 112). Rapid identification of UPEC is therefore important in reducing the risk of recurrent infections.

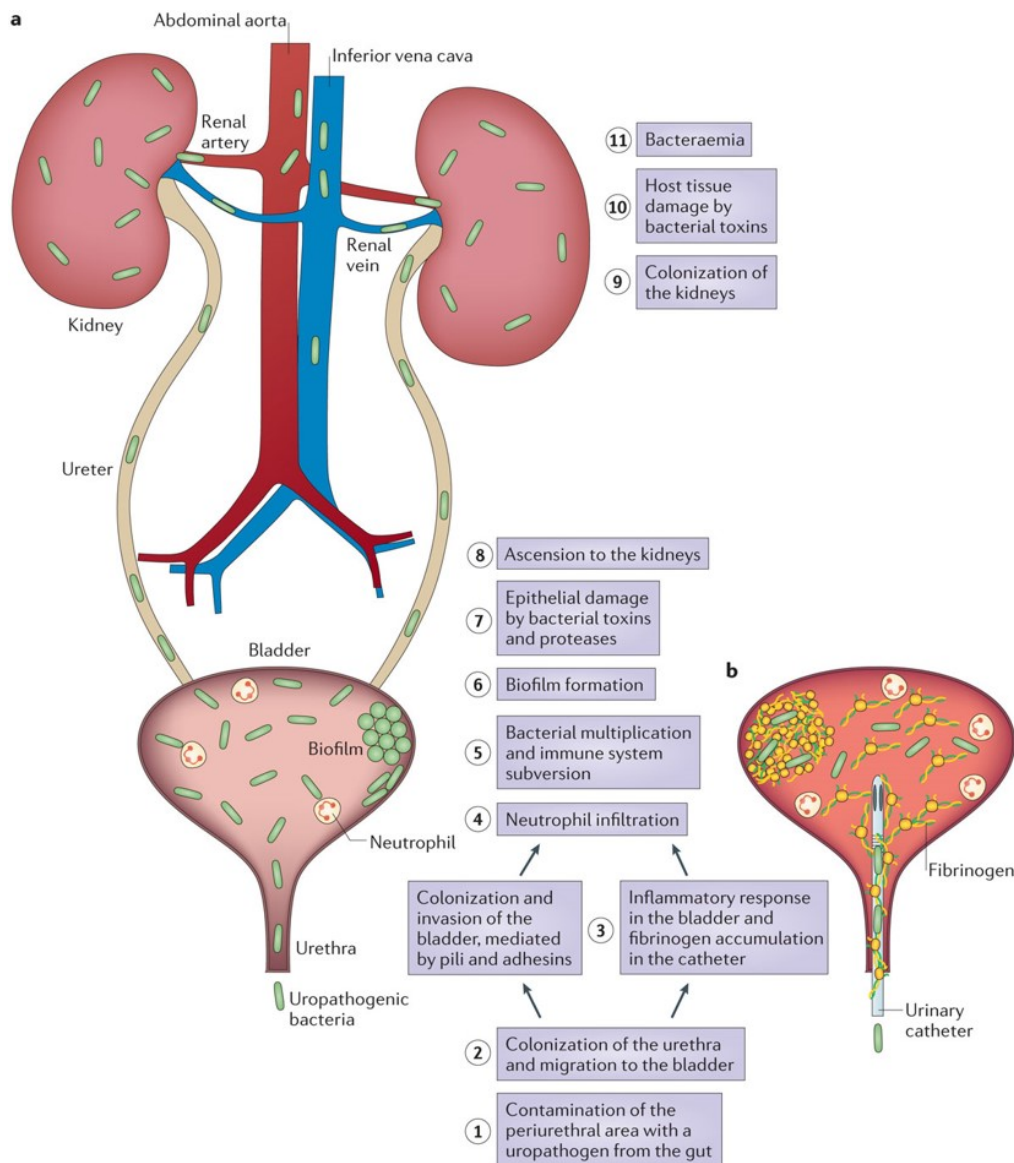


Figure 3. The pathogenesis of urinary tract infections is mediated by adhesins and virulence factors allowing colonisation, invasion of uroepithelial cells and evasion of host immune responses. Contamination of the urethra, followed by proliferation, biofilm formation and migration up the ureter to the kidneys results in pyelonephritis or bacteraemia if left untreated (a). Complicated UTIs arise from catheterisation due to the bladder being directly compromised, with fibrinogen mediated biofilm formation on catheters forming a reservoir for continued colonisation of the bladder lumen (b). Reproduced with permission from Springer Nature Publications © 2015 (70) (Supplementary 20).

UTIs may be acquired from migration of Enterobacteriaceae from the gastrointestinal tract into the urethra, since the majority of uropathogens are from this family (80). Transmission of uropathogens between people can occur during sexual activity, from

contaminated water, or food from poorly washed hands or the environment (113). Since significant CFU/mL values of >100 000 are possible within the bladder lumen, it is not surprising that contamination of other parts of the body occurs following urination (114). The higher incidence of UTIs in women compared to men can be explained by anatomical differences in the length of the urethra, and risk of acquiring a UTI for women is increased following sexual activity, but greatly reduced when a condom is used, supporting a sexual transmission route (113, 115). Additionally, the incidence of UTIs is greatest amongst women of the 18-29 age group; an age range of typically increased sexual activity (113). UPEC strains isolated from the urethra of women's sexual partners have been found to be twice as likely to be the causative agent of a UTI than *E. coli* strains isolated from gastrointestinal tract, supporting the theory that uropathogens are introduced to the urethra from either the vaginal cavity, skin or directly from the urethra of the sexual partner (116). Evidence for transmission through non-sexual contact with people or with a contaminated environment is shown by both nosocomial epidemics of strains possessing unusual but identical resistance profiles, and local epidemics of unusual resistance profiles caused by travel of a few people to countries from which the resistance profiles are widespread (117, 118).

Antimicrobial resistance is increasing in urinary pathogens due to a wide array of resistance mechanisms (119, 120). UPEC and *K. pneumoniae* have been found to encode plasmid acquired extended-spectrum β -lactamases, conferring resistance to β -lactam antibiotics including cephalosporins, penicillins and aztreonam (121, 122). The incidence of extended-spectrum β -lactamase producing UPEC strains was once confined to nosocomial UTIs, but has spread to community acquired UTIs since 2001, with 24% of a resistant subset analysed in the UK being found to originate from the community (123). Resistance spreads relatively rapidly through horizontal gene transfer methods of transduction, conjugation or transformation in local populations and geographical areas, with resistant isolates at high incidence in hospitals, nursing homes and in developing countries where antibiotic use is widespread (124, 125). Other species of the Enterobacteriaceae family have encoded AmpC enzymes capable of conferring resistance to cephamycin, as well as cephalosporins and β -lactamase inhibitors which are often co-formulated with β -lactam antibiotics to increase their efficacy, and also confer

resistance to carbapenems in *K. pneumoniae* lacking a particular outer-membrane protein (120, 121, 126). New developments such as improved β -lactamase inhibitors targeting carbapenemase produced by *K. pneumoniae* offer newer more effective antimicrobials and new co-formulations are combating the emergence of resistance (127). However, the swift diagnosis of UTIs is still imperative to good clinical outcomes for patients. Failure to identify an infection and prescribe appropriate, targeted treatment can lead to recurrent and persistent infections, antimicrobial resistance, ascending infections to the upper urinary tract and complications including secondary infections.

The diagnosis of UTIs is not straight forward, and different guidelines exist depending on patient demographics due to the variable risk factors associated with distinct groups. In most cases, presentation of one or more symptoms such as dysuria (pain when urinating), nocturia (night-time urination) or cloudy urine is met with testing via biochemical test strips and follow up culturing if positive (128). Any symptoms of pyelonephritis including kidney pain, nausea, fever and myalgia (muscle pain) are met with immediate empirical antibiotic treatment and follow up culturing. In infants and pregnant women, culturing and subsequent testing is usually recommended with presentation of any symptoms due to increased risk factors. For people over the age of 65 the range of symptoms is greater and may include incontinence, delirium, suprapubic pain and haematuria, with antimicrobial therapy recommended only after culturing and sensitivity testing is carried out due to the increased risk of resistance, unless symptoms are severe. Although new, alternative treatments to antimicrobial drugs, such as vaccines and small molecules targeting bacterial toxins and adhesion factors are being developed to aid in persistent and recurring infections, and in antimicrobial resistance, diagnosis still relies on isolating and identifying the uropathogen or symptoms (80, 129). Therefore, reducing turnaround time for this step can have a significant impact for treatment quality, and potentially reduce the incidence of empirical antimicrobial therapy which applies pressure for the emergence of resistance.

1.5 Traditional Identification Techniques

Traditionally, clinical microbiology laboratories have relied on techniques such as culturing, microscopy, biochemical tests, and culture independent methods, such as toxin or antigen detection, for the diagnosis of infection. Culturing is still extensively used in diagnostic laboratories, as it is a requirement of most subsequent tests to have an isolated pure colony of the microorganism of interest. However, it is labour intensive and often the rate-limiting step of laboratory workflows. Organisms can be slow growing and fastidious, and as in the case of *Mycobacterium* spp., can take several weeks or even months to obtain adequate culture, resulting in empirical treatment rather than the administration of specific antimicrobial drugs targeted to strain susceptibilities (130). Microbial speciation was often based upon biochemical tests, which were often commercially developed and sold as systems such as the API® ID strip and VITEK® 2 system (bioMérieux inc., Marcy-l'Étoile, France). These systems can still be found in some clinical laboratories as they aid in identification, although are subjective and still rely upon prior isolation of the organism, with results typically not received in less than 18 hours. In the case of the VITEK® 2, algorithms based on analysing fluorescence changes resulting from metabolic alterations enable the system to provide Gram-negative bacteria identification in around three hours. In addition, it is capable of performing susceptibility testing and determining minimum inhibitory concentrations (MICs) via monitoring of bacterial growth (131). However, in one study the system struggled to distinguish *Acinetobacter* isolates and failed to deliver MICs for any antimicrobial drug other than co-trimoxazole in the case of *Stenotrophomonas maltophilia* (131).

Immunoassays, such as ELISA or latex agglutination tests, are useful in detecting specific microbial antigens and are therefore performed when a specific causative agent is suspected, for example *Streptococcus pneumoniae* in the case of pulmonary infections. These tests therefore offer faster turnaround times than techniques requiring culturing in such instances. For suspected bacterial meningitis *H. influenzae*, *S. pneumoniae* and *Neisseria meningitidis* can be detected in cerebrospinal fluid (CSF) with a rapid latex fixation test. Here, latex beads coated with antibody specific to the pathogen are mixed

with a patient's CSF sample resulting in agglutination or 'clumping' of carrier beads if the antigen is present, however disadvantages include false positives due to antigen cross reactivity and low sensitivity (130).

More clinical microbiology laboratories are beginning to employ automation in sample processing. Clinical urine samples were traditionally first processed by manual microscopy or biochemical test strips but are now typically processed using automated microscopy or flow cytometry methods; such as with the sediMAX (Menarini Diagnostics, Berkshire, UK), iQ200 (Beckman Coulter, California, US) or Sysmex urinalysis platform (Sysmex, Kobe, Japan). These high throughput automated/semi-automated platforms enable biomedical scientists to quickly screen samples as culture negative or positive (132). Automated methods are increasingly important in laboratory workflows, as up to 80% of urine specimens received are culture negative and require no follow up tests (133). Although more streamlined than manual microscopy methods, the sediMAX workflow is still laborious and labour intensive, requiring at least three dedicated staff members to complete urinalysis. Received samples are aliquoted into tubes and loaded into the instrument, which dispenses the sample into single-use cuvettes and performs centrifugation. Digital images are taken of samples across the whole field of view, at different focusing planes, and an artificial neural network analyses recorded images to identify urine sediment particles such as bacteria, yeasts, erythrocytes, leukocytes, red blood cells, epithelial cells, urinary casts and crystals (134). Following detection of bacteria or yeasts, a sample is flagged as culture positive and highlighted in a report for follow-up work by a biomedical scientist. Importantly, yeasts and red blood cells cannot be differentiated, leading to samples containing either being manually inspected and a culture positive/negative classification manually assigned (135). These samples are not flagged by the instrument, so staff must look for samples containing red blood cells or yeasts and perform a manual inspection on either to confirm the correct particle assignment. Some samples are flagged as "unsure" by the system software, such as those containing urinary crystals which have a variable morphology. Such samples are also manually inspected by a biomedical scientist, diverting resources from other clinical duties (136). In the ICHNT clinical microbiology laboratory, culture plates are prepared during sample processing and are inspected

following an incubation period to confirm sediMAX culture negative results, as well as to follow up culture positive samples with identification and antimicrobial susceptibility testing. In theory, the sediMAX system reduces culturing requirements in the clinical laboratory. In practice, however, culturing is still performed on each urine sample received, as data obtained from the system can be ambiguous, unreliable or require follow-up culturing for clinical reasons. The system is therefore only useful at reducing microscopy workload on the clinical staff, as those samples which are returned as culture negative with a high degree of certainty by the machine do not have to be reviewed manually.

Automated fluorescence flow cytometry is provided by the Sysmex UF-1000i, which analyses side scatter and forward scatter light from a red semiconductor laser, along with fluorescence, to determine particle granularity and size (137). With this system, a minimum sample volume of 1 mL is required for analysis, compared with 2 mL for the sediMAX (138). Loaded samples are suspended in a sheath fluid and pass through a gating mechanism to separate bacteria and other particles into two separate laminar flow cells, where fluorochrome staining of particles, and proprietary fluorochrome dye staining of bacterial nucleic acid occurs (138). Bacteria staining in a separate chamber is an improvement over older models of the system and reduces red blood cell interference whilst increasing bacterial detection rates (137). Stained particles and bacteria are ejected through a nozzle in a tight jet and irradiated with the laser, causing fluorescence and distinct light scattering patterns dependent on particle size and shape, which the instrument uses to produce scattergrams for storage or interpretation by laboratory operators (137). An automated identification of red blood cells, white blood cells, erythrocytes, urinary casts, bacteria, yeasts and epithelial cells is provided, with cell counts calculated for bacteria, red blood cells, white blood cells, casts and epithelial cells (137). Laboratory operators set parameters for what should be deemed culture positive or negative, taking into account both the bacterial and white blood cell counts. Although the gold standard of a culture positive urine sample is $\geq 10^5$ CFU/mL, samples will be typically flagged for follow up culturing if a threshold level of 39 bacterial cells per μL , corresponding to 10^4 CFU/mL, and a white blood cell count of $>100/\text{mL}$ is reported (133, 139). In this way, the system can reduce the culturing load of negative

samples for the laboratory, with some studies demonstrating a reduction of 28% (133). Reporting of false negatives is dependent on the threshold levels set by the instrument operator, but some studies have reported high false positives of white blood cell counts, which are misclassified epithelial cells, by the instrument (138, 140). Since white blood cell counts are included in determination of culture positive or negative samples, this may contribute to the generation of false positives by the instrument as well, although false positives are of a lesser clinical concern as they would be screened out after culture follow-up of the sample. False negatives are of clinical concern, however, and the instrument has demonstrated unacceptable levels of 12.5% false negative reporting in some studies, with 37.5% of those clinically significant, leading to the conclusion that the system is not suitable for use in clinical laboratories (133, 138, 141). A recent development improves on the UF-1000i with the UF-5000 system, which utilises a different laser setup and analyses depolarised side scattered light in addition to the scatter patterns analysed in the older system, with the aim of improving bacterial determination and also assigning a Gram-negative or positive classification (142). The new system has shown a false negative reporting rate of 0.6% and a reduction in culturing workload of 55.5% in the single study returned in a literature search conducted within a clinical laboratory (143).

The iQ200 system and laboratory workflow is similar to that of the sediMAX. Rather than recording whole field microscopy images of urine sediment, the iQ200 system captures images of urine particles suspended between two laminar layers of suspending fluid as the uncentrifuged sample is passed through a flow cell (144). A neural network analyses digital images by considering particle size, shape, contrast and texture to detect bacteria, yeasts, urinary crystals, leukocytes, erythrocytes, epithelial cells, red blood cells and urinary casts (145). Samples are tagged as culture positive or negative based on identification of leukocytes, bacteria and “all small particles” indicators, or flagged for follow up work by a biomedical scientist if returned with an “artefact” flag (146). Importantly, false negatives are sometimes reported by the instrument, which flags a sample as culture negative using the three particle indicators, whilst manual culturing reveals an opposite culture positive result (147). In consequence, clinical microbiology laboratories using the iQ200 will often opt to culture samples when aliquoting for input

to the instrument, as with those utilising the sediMAX, and several studies of the systems indicate that sensitivity and specificity as a screening tool is inadequate to negate the culture requirement (145, 147). Samples flagged as culture negative with a high degree of certainty, however, could potentially reduce the culturing and follow up work load of a diagnostic laboratory by up to 35%, conferring significant cost savings and freeing up staff resources for other tasks if software recognition is improved (148).

Automated systems can all be of use as a preliminary screen for culture negative or positive urine samples, and to reduce the workload of clinical laboratories in culturing negative samples. However, they have drawbacks in both the cost of the instruments and in reporting of false negatives and positives. Reporting of false negatives is of clinical concern, as an incorrect reporting of a non-UTI would delay treatment and potentially effect clinical outcomes, particularly for at risk patients. All of the systems discussed still require a culture step following identification of culture positive samples, and often clinical laboratories will choose to prepare all samples for culture anyway. A problem with all automated analysers is that they record and report on dead bacterial cells as well as living, unlike with the gold standard of manual culturing (149). In addition, although the gold standard of a culture positive urine sample is typically $\geq 10^5$ CFU/mL, this value can change between clinical laboratories utilising automated systems to reduce recording of false negatives, making comparison between systems and laboratories difficult (142). There exists a need for better automated analysis, if not to replace the culturing step completely than at least to reduce the incidence of false negatives further.

1.6 Genomic Methods

Polymerase chain reaction (PCR) is a commonly used method for microbial diagnostics. Traditional methods may rely on identifying amplicons through gel electrophoresis or gene probe hybridisation methods. However, real time PCR is more commonly used as no post PCR analysis is required thereby reducing turnaround times and limiting contamination. Furthermore, the technique can often be used for quantification purposes which can be important for some infections, such as *Chlamydia trachomatis*

and *Neisseria gonorrhoeae* (150-152). Targeted primers or labelled gene probes which hybridise to specific complementary DNA sequences can be used in the rapid detection of microorganisms within clinical samples or for culture confirmation, with results obtainable within between two to four hours. After amplification of the target gene by PCR, hybridisation of the specific DNA sequence with the gene probe produces a detectable signal. Commercial tests are obtainable for *Neisseria gonorrhoeae*, *Chlamydia trachomatis* and Group A streptococci, but have relatively high limits of detection and therefore have limited use in cases where gene copy numbers are low (130).

The most common genomic method used in bacterial identification outside of routine clinical labs is 16S rRNA gene sequencing, as the gene is ubiquitous throughout divergent phyla, has well-conserved regions for primers, and contains genus- or species-level hypervariable regions (153). As such, it is well suited for classifying microorganisms which are difficult to grow in culture and those which have ambiguous biochemical test results. It has been used successfully with next generation sequencing platforms, including Illumina MiSeq and Roche 454 in the analysis of mixed-community microbiome samples (154). In these cases, DNA is extracted from microbial cells typically using a commercial extraction kit of which several varieties are available. Kits may enzymatically or mechanically lyse cells and bind DNA to a matrix. Bound DNA can be eluted and specific primer sets used so as the 16S rRNA gene region is targeted. PCR is performed to amplify the regions of interest and in the case of Illumina 16S rRNA gene sequencing library preparation, sequence adapters and barcodes are added. The samples are run through the sequencer which uses a fluorescent imaging system to determine the sequence. Datasets can be extremely large due to the millions of short reads generated and therefore analysis can be complex and time consuming, with commonly used bioinformatics tools for classifying bacterial taxonomy often only classifying to genus level due to the uniform nature of the 16S rRNA gene resulting in little evident variation between species (155). Although molecular techniques are useful, they are infrequently used in clinical settings due to extensive preparation requirements, lack of expertise, complicated analysis and costs. Additionally, studies analysing the microbiome through a 16S rRNA gene sequencing approach are not truly quantitative and often over-represent some taxa at the expense of others due to

differences in gene copy numbers between species (156). In such studies it is necessary to correct for gene copy numbers through the use of bioinformatic tools which improve relative abundance estimates. A novel method of quickly identifying species and subspecies of the microbiota within tissue samples could therefore prove revolutionary to clinical microbiology and drastically improve disease outcomes of patients.

1.7 Clinical Mass Spectrometry

Commercially available mass spectrometry (MS) systems used in clinical laboratories are based on matrix-assisted laser desorption ionisation-time of flight (MALDI-ToF) technologies; which analyse a microbial colony mixed with a proton supplying matrix via irradiation with a UV laser (157). Subsequent analysis of ionised proteins, up to 100 kDa, and spectral comparison to a reference spectra database allows classification to genus or species level, and reduces sample turnaround time by approximately one day (157, 158). Incorporation of MALDI systems to clinical microbiology laboratories improves patient outcomes, reduces hospital stay periods and improves time until correct antibiotic action through more rapid sample turnaround (159, 160). Economic impact of MALDI systems is also positive, with reduced waste, culture media and labour costs, resulting in 2-5 fold cost savings when compared to traditional culture methods and, translating to around \$94,500 per year for a clinical laboratory handling 105,000 samples (160). Although the initial cost of MALDI systems is high, annual savings allow laboratories to offset this in approximately three years, explaining the widespread adoption of the technology (159).

These technologies however, still rely upon the isolation of a pure microbial culture (Figure 4). This means clinical samples such as urine and sputum undergo significant processing and incubation periods; which is time consuming and labour-intensive. Although MALDI-ToF has, in many ways, revolutionised the workflow of clinical microbiology laboratories, there is still further room for improvement for mass spectrometry in terms of sample preparation, turnaround times, and functional classifications of clinically significant microorganisms. No information on clinically relevant phenotypes such as antibiotic-resistance, virulence or serotypes is provided

which makes additional conventional testing necessary. Other limitations of the technology include spectral interference due to the presence of endospores, for example in *Clostridium* spp., leading to poor classification results and requiring young cultures to be used in such instances. Microorganisms possessing capsules, such as *Streptococcus* spp., *Haemophilus influenzae* and *K. pneumoniae* are often difficult to lyse thereby resulting in poor extraction yields and classification due to weak spectral quality (161).

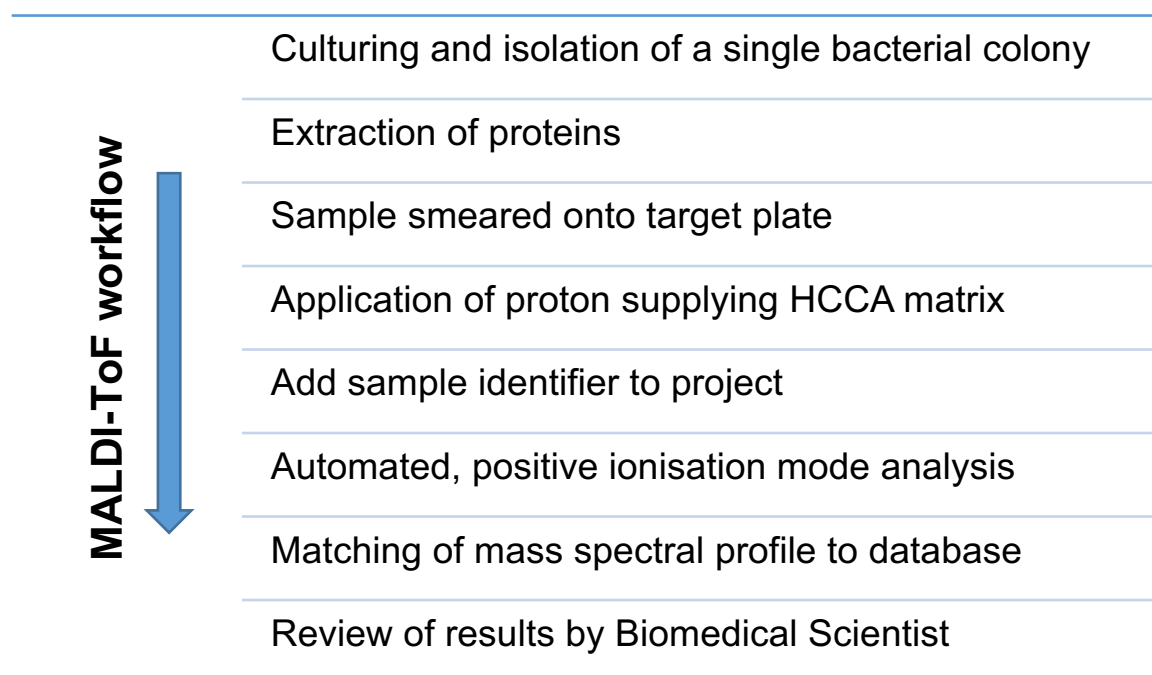


Figure 4. Typical workflow for identification of bacteria in clinical microbiology laboratories using MALDI-ToF systems.

MALDI-TOF has some potential in detecting bacteria directly from urine samples containing a single disease-causing organism, however some sample preparation is required. Unreliable classifications are reported in samples containing more than one microorganism, making identification of causative agents in complicated UTIs problematic since they are much more likely to be poly-microbial than uncomplicated UTIs (162). Preparation includes centrifugation of culture positive samples containing $>10^5$ CFU/mL to eliminate leukocytes and form a bacterial pellet. Proteins from within the pellet must be extracted using acetonitrile and formic acid prior to analysis on the

MS. Accurate bacterial classification is found in around 90% of single species samples, while an unreliable result in around 8% (162). In poly-microbial samples 45% are reported with an unreliable classification (162, 163).

The Bruker MALDI SepsiTyper kit™ has been developed for use in identification of bacteria directly from blood culture broths and cerebrospinal fluid samples. A sample preparation time of between 10 and 30 minutes is required, involving multiple buffer suspension and centrifugation steps, followed by extraction with acetonitrile and formic acid (164, 165). Although designed for direct from blood culture analysis, the kit is mainly used where a single causative microorganism is present (166). 100% and 70% correct species classification is reported from blood culture samples for Gram-negative and Gram-positive bacteria, respectively. However, the kit does not classify poly-microbial samples successfully, often only identifying one species in samples containing multiple causative agents (166, 167). In addition, Gram-positive organisms and *Streptococcus* spp. are poorly resolved when using the Bruker system for direct from sample blood analysis (168).

1.8 Ambient Ionisation Mass Spectrometry

Traditional mass spectrometry techniques require ions to be generated within the instrument vacuum. Ambient ionisation, however, allows ions to be generated in normal atmospheric conditions and thus samples can be analysed in their native form, without any extensive sample preparation, extractions or other manipulation. As yet, no ambient techniques are used in clinical microbiology laboratories, although several methods have been developed for research purposes and are available commercially. The advantages of ambient ionisation techniques in microbial analysis over alternatives such as MALDI are that ionisation is not reliant on a proton supplying matrix, thereby reduced sample preparation time as bacteria can be analysed directly from a petri dish if desired. Samples can be freely accessed during analysis with ambient techniques, meaning that they can be manipulated or moved around (169). Some ambient ionisation methods used for bacterial analysis are desorption electrospray ionisation (DESI), direct

analysis in real time (DART), low-temperature plasma ionisation (LTP), and paper spray ionisation.

1.8.1 Desorption Electrospray Ionisation

DESI was introduced in 2004 and was the first described ambient ionisation technique (170). The method involves firing an electrospray of charged solvent droplets at a solid-phase sample from a short distance, which pneumatically desorbs ions from the sample surface and introduces them to the mass spectrometer inlet capillary (169). Both large biomolecules and small organic molecules can be ionised in this way. Reproducible but distinct mass spectra have been generated in the analysis of bacterial species using DESI, with differentiation between species possible via analysis of spectral profiles containing phospholipids, fatty acids and lipopeptides (171). Several sample preparation methods have been utilised for bacterial DESI analysis, including analysis of intact biofilms on membranes, analysis of washed cells on glass slides and analysis of cells suspensions on glass slides (171-173). Early work on cells isolated onto PTFE showed unique spectral fingerprints for *E. coli* and *P. aeruginosa* (174). Negative ion mode analysis of *Bacillus subtilis* biofilms produces clusters of lipopeptide ions around 1000 m/z and phospholipids in the 600-800 m/z region, which are highly species-specific and comprise of cell envelope lipids along with secreted surfactins (173). Intact biofilm analysis suggests DESI may be used as an *in vivo* bacterial classification technique. In another study, classification of seven distinct clinical species was possible by analysing mass spectra of the 50-500 m/z region containing fatty acids, with a sample analysis time of less than two minutes (172). Spectra produced for each species possessed a unique fatty acid pattern, which allowed good species class separation in principal component analysis, even with Gram-positive and negative species grouped together. However, batch effects were observed between different analysis days, likely due to growth media nutrient concentration fluctuations on batches made on different days, or ambient instrument condition changes such as temperature and air pressure.

The identification of fatty acid biomarkers may allow biomarker based species classification with DESI, rather than relying on a mass spectral profile recognition based classification method (172). Thin film DESI imaging is a method in which an imprint of

a culture from agar media is taken using various solid hydrophilic membranes, for example mixed cellulose ester filter membranes. Cultures can be imaged from an imprint without any further sample preparation, and analysis reveals the spatial spread of metabolites and secreted products, providing insights into metabolic exchanges and relationships between species in microbial communities (175). The method has been utilised to investigate the communication dynamics and metabolic exchange of secondary metabolites, with a spatial resolution of 250 μm , between *B. subtilis* and *Streptomyces coelicolor*, which both alter their secreted defensive metabolites in response to secondary metabolites available from each other (176). The disadvantage of this technique is that the pneumatic action of the solvent spray requires samples to be analysed from solid membranes, as analysis from native agar media results in destruction of the soft surface.

A recent development in DESI replaces the pneumatic carrier gas with a dual capillary setup, in which no pneumatic force is applied to the sample surface. In nanoDESI, a nanolitre size liquid solvent bridge is formed between two capillaries and rests on the sample surface, desorbing analytes as it rolls around (177). Desorbed ions travel up the second capillary and are aspirated into the mass spectrometer. This approach allows non-destructive imaging of a sample, and has been successfully used in imaging of live metabolite production directly from cultures on agar media (178). The main drawback of DESI is the wide array of experimental parameters to be optimised for each experiment, as although these allow the technique to be highly dynamic and adaptable to specific samples and methodologies, there is no consensus on an overall optimal methodology for microbial analysis, with solvent composition, spray angle, nebulising gas pressure and nanoDESI geometries differing between groups.

1.8.2 Direct Analysis in Real Time

DART mass spectrometry shares some similarities with DESI in that it was introduced shortly after the DESI technique and achieves ionisation of a sample by desorbing analytes from sample surfaces by applying an ionising gas or aerosol, under ambient conditions. DART ionisation sources produce a corona discharge of helium gas between two electrodes containing ions (179). The heated plasma is then released from the

ionisation source and forms reagent ions through a cascade of gas-phase reactions with ambient air, which are both directed at a sample in solid, liquid or gaseous form (180). Both plasma and reagent ions colliding with a sample surface produce ionised analyte particles, which enter the mass spectrometer inlet across a short reaction zone (181).

Like DESI, the DART technique has been applied to the analysis of bacteria. Fatty acid methyl ester (FAME) profiles are useful biochemical biomarkers for bacterial classification and are analysed via gas-chromatography mass spectrometry (GC-MS) in reference laboratories for taxonomy research (182). Although not routinely used in clinical microbiology laboratories due to the expertise required, sample preparation time and the wide availability of newer, purpose built, commercialised mass spectrometry identification systems, analysis of FAME signatures with mass spectrometry can still provide a reliable classification method (183). FAME profiling with DART avoids the lengthy chromatographic separation step and can provide signatures in the 150-400 m/z range from whole cells of *S. pyogenes*, *E. coli* and *C. burnetti* in a much shorter time frame of four minutes, compared to several hours for GC-MS (184). Differences in FAME profiles of the three isolates were observed in the mass spectral, intensity, and temporal domains, allowing differentiation. Strain level differentiation may also be possible with DART, as FAME profile differences have also been observed in three strains of *C. burnetti* (184). DART has also been used to directly analyse fatty acids of bacteria without methylation required in the production of FAMEs. This approach allowed the classification of five Gram-negative and five Gram-positive clinical isolates of different species, based on their unique fatty acid spectral profiles, with 100% prediction accuracy (185). Disadvantages of the DART technique mainly lie in the lack of research done into applications for microbiology. More work and optimisation needs to be completed in order for a robust methodology to be realised.

1.8.3 Paper Spray Ionisation

Paper spray ionisation is achieved by applying a high electrical potential to porous media, such as paper or a wooden toothpick, producing gas phase ions from the surface which are introduced to a mass spectrometer for analysis. The paper or other material

is shaped in such a way as to have a sharp tip, on which a sample is applied and moistened with a solvent/water mixture. A voltage of 4500 is applied across the material and a plume of ions is emitted from the tip (186). The simple methodology of the technique has proven highly adaptable, with a wide range of media including pen nibs, glass fibre filters, leaves and chromatography paper used to generate ionisation (186). The solvent flow rate used to moisten the porous media effects how ions are generated; in high flow rates and low electrical currents, multiple pulsating and Taylor cone (distortion of liquid into a cone shaped jet under an electrical field) sprays are generated emitting variable plume droplet sizes, whilst with low flow rates and high electrical currents, a single Taylor cone spray and corona discharge are generated along with small plume droplet sizes, which has shown better ionisation of non-polar molecules than the first mode (187). Both emission modes are ran sequentially during sample analysis, with the second mode occurring when solvent supply is near depletion. Zero volt paper spray ionisation is also possible by utilising the native mass spectrometer instrument's vacuum as the pneumatic force acting on the sample, although with a two orders of magnitude lower signal intensity when compared to standard paper spray methods (188). In this case, paper is held 0.2-0.5 mm away from the mass spectrometer inlet capillary, from which the suction causes release of ion containing solvent droplets from the sample. Ions are believed to be generated by analyte existing in solution phase ion pairs within the solvent mixture, which disassociate under aerodynamic forces breaking up solvent droplets within the mass spectrometer and undergo multiple evaporation and Coulombic fission events (188).

A method using chromatography paper in zero volt paper ionisation has been used in bacterial analysis, with spectra highly comparable to those obtained with standard paper spray, albeit with reduced intensity as expected (189). In this method, single colonies were placed onto the paper and moistened with ethanol to begin droplet formation. Observed lipid profiles from *C. farmeri*, *C. freundii*, *P. aeruginosa*, *Acinetobacter baumannii*, *S. aureus*, *S. epidermidis*, *B. subtilis* and *B. thuringiensis* were significantly different, and clear separation of each species was achieved under principal component analysis, as well as groupings of the Gram-positive and negative species (189). Although signal intensity can be increased by up to three orders of magnitude

when applying 220 V to the paper, the simplicity of the 0 V method along with high signal to noise ratio of the produced spectra lend preference to its use in future experiments involving bacterial analysis. As with DART, little work has been published on paper spray analysis of bacteria, and although it shows promise for adoption in clinical microbiology laboratories, more optimisation experiments are required. Optimisation should include the testing of different porous medias, as these effect ionisation rates, along with solvent mixture ratios (190).

1.8.4 Acoustic Mist Ionisation

AMI-MS is a new technique which samples from 96 or 384 microtiter plates by firing an ultrasonic wave from beneath each well sequentially. Liquid samples are required for this process, and acoustic droplet ejection is utilised via an acoustic transducer to produce a mist of droplets of femtolitre volume (Figure 5). A negative voltage applied to a collection capillary imparts an opposite, positive voltage to the surface of the liquid, thereby producing a positively charged droplet mist. Charge is transferred to analytes within droplets as high capillary temperatures of 180°C act to desolvate the sample, leading to the production of analyte ions. The process is rapid, with analysis speeds of three samples per second achieved through the use of an automated stage which moves plate wells over the inducer. A proof of principle study using AMI demonstrated that ionisation of peptides, cell metabolites and small drug molecules is possible with this system, and the technique has been applied in pharmacological high throughput screening of compound libraries (191, 192). Although not yet applied to analysis of bacteria, the technique is of interest due to its simplicity and rapidity.

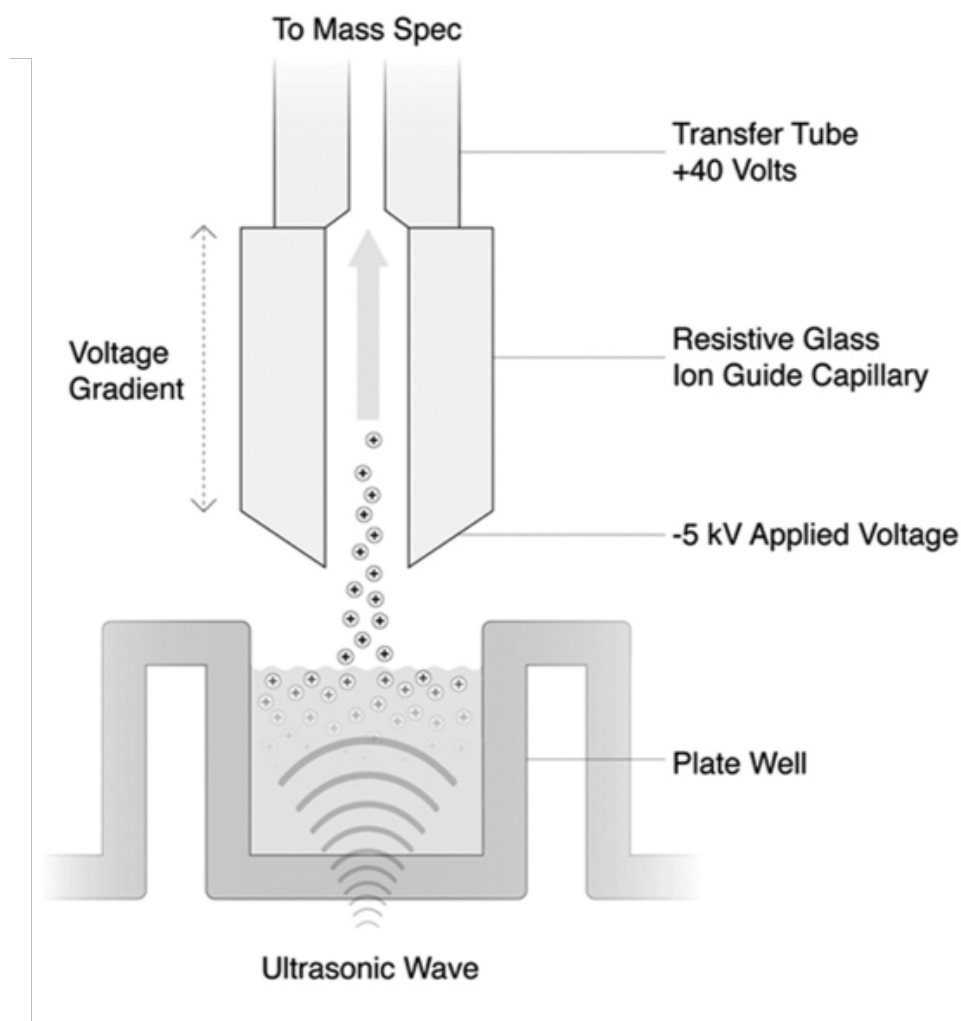


Figure 5. Representation of Acoustic Mist Ionisation technique. Charged fluid droplets are sampled in ambient conditions by applying an ultrasonic wave to a liquid sample contained within a plate well. A voltage applied to a capillary imparts a charge to the sample surface, attracts droplets and acts as an ion guide to transport ions directly to the inlet of the mass spectrometer. The technique is high throughput and is capable of analysing three samples per second. Reproduced with permission from SAGE Publications, © 2016 (191) (Supplementary 21).

1.9 Rapid Evaporative Ionisation Mass Spectrometry

REIMS is an ambient ionisation technique which is performed on intact samples in near real time, with no pre sample processing or extraction required. Application of a radiofrequency electrical current (100kHz – 4MHz) rapidly heats biomass which is vaporised and introduced to a mass spectrometer via a transfer line utilising the native instrument vacuum. The vapour contains lipid coated droplets, intracellular and

extracellular metabolites, which are generated by the disintegration of cells via joule heating. A heated collision surface present in the REIMS interface further breaks apart and vaporises these droplets, leading to production of gas phase ions, which can be used in biomarker or spectral profile matching analysis. The electrical current is applied to the biomass using either handheld bi-polar or automated mono-polar probes. The *in vivo* iKnife surgical device has been used to successfully differentiate between malignant and healthy tissue during surgery using the REIMS technique (193). For bacterial analysis, the REIMS classification mechanism is based on molecules present in the cell wall and cell membrane of microbes, as acquired mass spectra display high signal intensity in the fatty acid and lipid mass regions. Analytes mainly consist of phosphatidylethanolamines, phosphatidylglycerols and phosphatidic acids in the case of bacteria, and phosphatidylethanolamines, phosphatidic acids and phosphatidylinositols in the case of fungi (158, 194). Culture media does not affect species classification and thus isolates can be grown in their optimal conditions without affecting results (158).

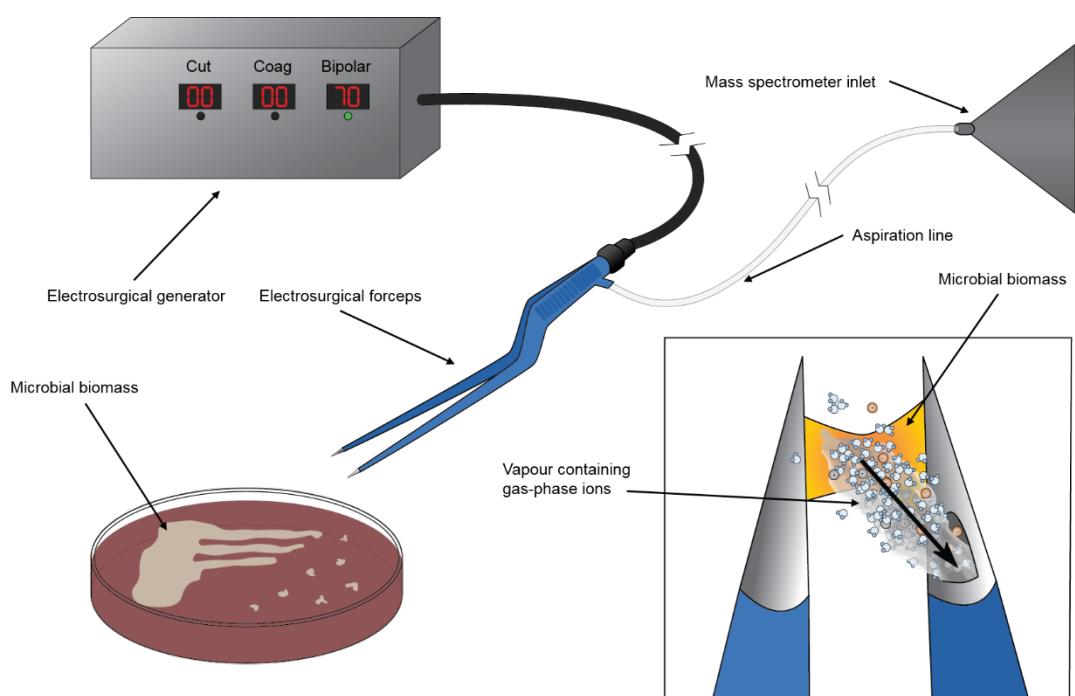


Figure 6. Bi-polar electrosurgical forceps REIMS. The forceps serve as a REIMS hand-piece to vaporise microbial biomass. Vapour is channelled through the embedded fluid line to the mass spectrometer inlet using the native vacuum of the instrument.

In bi-polar probe REIMS, electrosurgical forceps are used as a sampling probe and the operator manually picks up bacterial colonies from an agar plate before activating the radiofrequency electrical current and closing the tips of the forceps together. Upon contact, the biomass heats rapidly and the vapour is channelled to the mass spectrometer as mentioned previously, utilising the native vacuum of the instrument as suction (158, 194) (Figure 6). The primary benefit of this technique over existing clinical MS systems is that it is conducted on intact cells, requiring no sample preparation or extraction prior to analysis. A high-throughput, automated REIMS method has also been developed which utilises monopolar probes (195). In this method, plate handling and colony picking is capable of being performed automatically by a bespoke Tecan Freedom EVO robotic system coupled with Pickolo SciRobotics software (Figure 7). Operators place petri dishes into a rack and the system places them individually onto an illuminated stage, which also acts as the return electrode for the monopolar probe electrode on the robot arm. A camera takes a picture of the culture media which is displayed on screen. The software can automatically select colonies for analysis, although manual colony selection is currently utilised to ensure optimal data quality. Once colonies have been selected, a robotic arm picks up an electrode probe and descends to the colony where the radiofrequency electrical current is activated and biomass is rapidly heated and vaporised. The vapour is channelled to the mass spectrometer, using the native vacuum of the instrument, through a connecting PTFE tube, as in the bi-polar method described above. The system is also capable of picking a small amount of biomass into a 96-well plate which can be removed once full for subsequent genomic analysis. This simplified workflow removes user input, and thus variability and inconsistency in technique when using the bipolar method. As such, it is well suited to the workflow of clinical laboratories, as it is capable of analysing one isolate with five replicate acquisitions every two minutes.

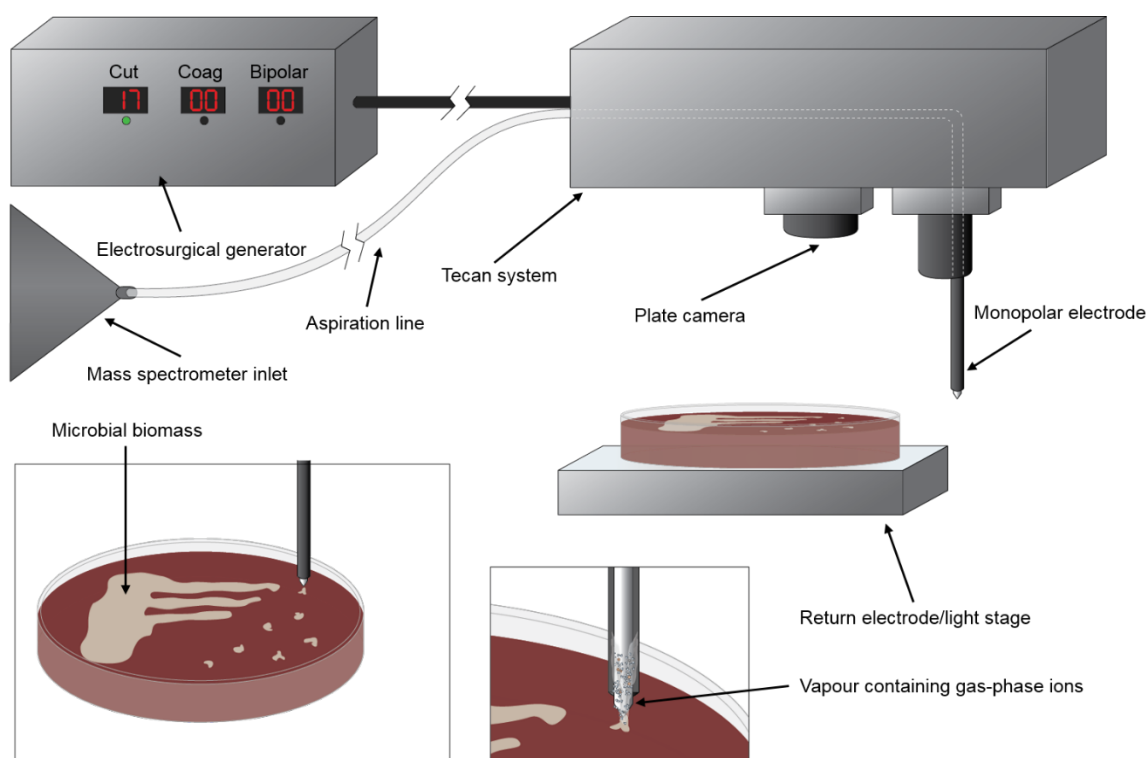


Figure 7. Automated monopolar REIMS. A modified Tecan liquid handling system acts as an automated REIMS high-throughput analysis platform. Colonies can be selected automatically using the Pickolo software, analysed using mono-polar probes and picked into a 96-well plate for subsequent genomic analysis.

Both bipolar forceps and automated monopolar REIMS have proven successful in the generation of distinct spectral profiles of 25 clinical species of bacteria, as well as clinically significant yeasts and filamentous fungi, where 96.3% and 93.9% accuracy was achieved in species classification of bacterial species through Random Forest decision making trees, for the bipolar and monopolar methods respectively (195). For pathogenic yeasts of the *Candida* genus, 96% and 100% species classification has been achieved for the two methods (196). Unpublished data obtained from analysis of 45 clinically significant bacterial species suggests a classification accuracy to species level of >93%, genus level of >95% and Gram stain level of >99%. In addition, an analysis methodology for filamentous fungi has been developed, with the *Trichophyton*, *Scopulariopsis* and *Mucor* genera showing distinct spectral profiles and good separation under principal component analysis. Although bacterial differentiation with REIMS based on the unique

spectral fingerprint signatures produced by each species is useful, the identification of species-specific biomarkers would expand applications to speciation of mixed cultures and *in vivo* tissue diagnostics. The 50-2500 *m/z* region is acquired during REIMS analysis, and the optimal region for classification of both bacteria and fungi has been identified as 600-900 *m/z*, where phospholipids are observed (195). Fatty acid analysis may provide even higher classification accuracies, achieved through fragmentation of parent phospholipids and examination of daughter ions, although this work is in the early stages.

1.10 Non-Human Diagnostic Microbiology

Diagnostic microbiology extends beyond the human clinic and into veterinary applications. In particular, infectious diseases of livestock are a major economic burden and their diagnosis and control is critical to animal welfare, food supply security and agronomic health. The estimated worldwide cattle population is 1.3 billion, reared in conditions ranging from small sustenance farms to large, corporate, intensive ranches (197). Under such intense husbandry conditions, infectious disease control is essential to prevent production losses and maintain food supply chains. In addition, zoonotic diseases are a health concern to humans, and require the implementation of swift control measures following identification. The most obvious of these are foodborne illnesses resulting from faecal matter contamination of livestock meat intended for human consumption with *E. coli*, *Salmonella* spp., *Listeria* spp. and *Campylobacter* spp, but tuberculosis, brucellosis, chlamydiosis and bubonic plague also have a possible zoonotic origin. Bovine tuberculosis is caused by pulmonary infection by *Mycobacterium bovis*, which is capable of causing infections in humans through consumption of unpasteurised milk products, direct contact with infected livestock or through air transmission (198). The burden of zoonotic tuberculosis disease in humans is believed to be under-reported, due to the difficulty in differentiating between *M. tuberculosis* and *M. bovis* using non-molecular methods and the slow growth rate of both species, where up to six weeks is required to achieve culture (199, 200). The diagnostic procedure in livestock is a tuberculin sensitivity skin test, with follow up culture if a positive result is reported, and bovine tuberculosis is a notifiable disease (201). The rapid identification of causative agents of disease is critical in controlling

spread, however the economic reality of animal agriculture is that diagnosis is not carried out unless legislation requires reporting or a farmer faces significant monetary losses, leading to the use of broad spectrum antimicrobials and contributing to the emergence of resistance.

1.10.1 Bovine Mastitis

Bovine mastitis is a bacterial infection of mammary glands causing inflammation, pain and economic losses to the farmer due to milk production losses, milk discard, veterinary treatment, labour increases and early culling (202). Mastitis is the most common and most economically significant bacterial infection of dairy cows, with an estimated cost of £300 million per year to dairy farmers in the UK, and \$2 billion per year in the US (203). Mastitis can be caused by wide array of bacteria and yeasts, but 80% of cases are a result of infection by *E. coli*, *S. aureus*, *Streptococcus uberis*, *Streptococcus dysgalactiae* or *Streptococcus agalactiae*, with infection presenting as either clinical, sub-clinical or chronic (204). Sub-clinical manifestations are caused by intracellular, pathogenic *Sta. aureus*, *Str. dysgalactiae* or *Str. agalactiae* and are associated with a higher economic cost than clinical manifestations as they account for the majority of mastitis cases, present no visible indications, often go undiagnosed, can reduce milk production by up to 25%, and cause an increase in the somatic cell count (SCC) of milk, which is used as a marker of milk quality affecting farmers' payments (203-205). Clinical infections are typically caused by environmental opportunistic *E. coli* or *S. uberis*, present with visible inflammation of the udder, abnormalities of milk and systemic symptoms such as fever, increased heart rate and behavioural changes, results in between 6-11% loss in milk production depending on the lactation stage of the cow, causes decreased lactation duration and increases culling rates (203-205).

Selective breeding has led to cows which produce significantly more milk than is needed to rear a calf, resulting in metabolic stresses, and increased milk production is correlated with a higher incidence rate of mastitis (206). Modern farming practices also increase the rate of mastitis through trauma caused by automatic milking machines and increased cow density leading to easier transmission of bacteria between cows. Infection arises when bacteria gain entry to the teat canal, usually following lactation, when the

canal remains open for around two hours and the obstructive keratin plug has been dispelled (203). Cows are also naturally more susceptible to mammary gland infection prior to parturition, as the teat canal dilates and encounters reduced defence mechanisms (207). Entry of the teat canal and subsequent subversion of immune defences leads to colonisation of the mammary gland. Recruitment of leukocytes to the infection site causes inflammation and tissue damage, resulting in epithelial cell and leukocyte influx into secreted milk, which is detected as increased SCCs during milk quality testing (203).

Due to the substantial economic costs of mastitis, diagnosis must be swift to minimise financial ramifications for the farmer and maintain welfare standards for the cow. Selective and differential culturing techniques provide a definitive identification of the causative agent and leads to a confirmed diagnosis, however these are not preferred due to costs, the turnaround time of two to seven days to achieve a result, and the requirement of sending samples to a laboratory when quicker “cow-side” tests are available (208). SCCs are used as a marker in several cow-side tests to determine if an infection is present, such as the California Mastitis Test (CMT), portachek assay, Fossomatic SCC test, and DeLaval cell counter (DeLaval, Cardiff, UK) (209). The CMT utilises a surfactant to disrupt the cell membrane of SCCs, causing release of nucleic acids which subsequently aggregate to form a matrix, increasing the viscosity of the milk sample. Although swift and extremely cost effective, the test is of low sensitivity and is prone to human error, being difficult to interpret at lower SCC levels. Portachek assays enable determination of a SSC through an enzymatic reaction and are also cost effective, but have the drawback of low sensitivity. Cell counting methods, including the fossomatic SCC test and DeLaval counter, offer higher sensitivity and rapid results through staining and optical fluorescence of DNA, although with a higher expense and expertise required. Proteomic analysis has enabled immunoassay approaches, such as ELISA, which target protein biomarkers within mastitis positive milk using antibodies, although specific antibodies for the target bacterial or immune defence proteins are required. The drawback of such methods is that although 80% of mastitis is caused by five species of bacteria, over 120 causative agents have been identified, meaning that the high specificity of ELISA can be a hindrance and potentially provide false negative

results (203). Immunoassays must also be carried out within a laboratory and so cannot be utilised as a rapid cow side test.

MALDI-ToF has been utilised in culture-free detection of mastitis causing bacteria in reconstituted milk samples which were spiked with the five most common causative agents of infection, and incubated to produce a range of CFU/mL concentrations in order to determine detection limits (210). In this, bacterial ribosomal proteins were extracted from cell pellets obtained by centrifugation of milk samples. Cell lysis, protein extraction and washing was performed by using a commercially available extraction kit intended for blood sample processing. Correct species identification by MALDI analysis was dependent on bacterial cell loads within the sample, with $>10^6$ - $>10^8$ CFU/mL required for *Sta. aureus*, *E. coli*, *Str. agalactiae*, *Str. dysgalactiae* and *Str. uberis* (208). Although MALDI based tests have potential in offering a rapid, culture free identification methodology, they have yet to be tested on true mastitis milk samples.

Although a preliminary positive identification of mastitis can be performed by one of the rapid cow side test methods, identification of the causative agent can only be followed up by other methods, such as culturing, in a central laboratory. Causative agent identification is required as treatment options and control measures depend on environmental or pathogenic sources of infection (208). Some steps have been taken toward reducing turnaround times for laboratories with the investigation into culture-free MALDI-MS methods, but relatively high CFU/mL are required for detection. Real time PCR has also been utilised through the commercial availability of specific kits, such as the PathoProof Mastitis Complete-12 Kit (Fisher Scientific, Loughborough, UK), and offers a rapid identification method without any culture requirement (211). These kits have shown comparable diagnostic sensitivity and specificity to culture methods in detecting *S. aureus*, but reduced sensitivity and specificity for other *Staphylococcus* species (211). PCR methods are also prone to contamination, require targeted primer optimisation and detect non-viable cells, which may be of no clinical relevance.

Therefore there still exists a need for a rapid, high throughput, preferably cow side test which can be performed on milk samples, or linked directly to automatic milking

machines for continuous sampling. Existing MALDI-MS methods could not be applied in this way due to the protein extraction protocols required, favouring an ambient ionisation technique. Successful integration of a test meeting these requirements would improve animal welfare, increase food security and reduce economic losses of the farmer and industry on the whole.

1.11 MS Statistical Methods

1.11.1 Pre-Processing

The data analysis methods discussed here are particularly relevant to time of flight (ToF) MS due to the inherent nature of the technology, which relies on ion drift times to discern mass and so is more susceptible to changing environmental conditions than other mass analysers. ToF MS is the only mass analyser technology utilised in this thesis.

Analysis via mass spectrometry generates mass spectra, containing large amounts of data on m/z values and associated intensities at specific scan times. Before any statistical analysis of data can be completed, data matrices must be generated from mass spectra and corrected for background noise and mass drift to allow similarity comparisons between datasets. This is generally done to an m/z range of interest to reduce processing time and computational power, as spectra are usually acquired over a greater range than is necessary. Pre-processing scripts and toolboxes may be run in R, MatLab or proprietary software to perform these functions, or via cloud-based tool suites.

Under standard operating conditions and modes, background noise is generally present at all m/z values at a low intensity, with chemical, electrical or random origins. Noise is generated due to high sensitivity of instruments and the non-inert atmosphere and materials used to operate and maintain them, and can interfere with analysis if analyte peaks of interest are of low intensity. This noise is reduced or removed via background subtraction algorithms.

Mass drift occurs due to changes in instrument conditions throughout the course of an experiment, such as ambient temperature, flight tube temperature, humidity, pressure

or voltage fluctuations. Mass drift correction can be performed by aligning a specific peak to a known reference compound infused into the mass spectrometer continuously during sample analysis. The known m/z value of the reference peak is used to adjust analyte peaks and align m/z values to the shift observed in the lock-mass compound. Similar or identical samples analysed under the same experimental parameters often demonstrate different intensity ranges due to the changes in instrument conditions mentioned above, and discrepancies between sample preparation. Normalisation of acquired spectra is therefore required to compare datasets of similar samples analysed at different time points, and is achieved using total ion current (TIC), mean or median fold normalisation techniques which bring the two dataset intensity ranges into alignment. Data transformation and scaling may also be employed during pre-processing, such as log transformation to reduce heteroscedasticity and minimise skew, and pareto scaling to reduce dominance or relative importance of large fold changes of variables compared to smaller ones. Following completion of pre-processing, binning is performed at defined m/z intervals to couple aggregate intensities and m/z values together in order to reduce data dimensionality. The binned m/z value and intensity serve as representatives of the bin, which can be exported as a simple table for further data analysis.

1.11.2 Univariate Analysis

The identification of significant biomarkers is commonly performed using parametric univariate analysis techniques such as the t-test and analysis of variance (ANOVA). Parametric tests carry assumptions, generally being that data is approximately normally distributed and homoscedastic, which if violated lead to increased Type I or II errors (212). If these assumptions are not met, non-parametric tests are used as they make no assumptions about population data, however are considered less statistically powerful. Data transformation techniques can be employed to increase the normality of data distributions so that parametric tests may be carried out.

In the case of two sample classes, a t-test can be carried out to determine if the sample classes are statistically different to each other. ANOVA is performed on datasets with multiple sample classes, and examines the variance within means across sample classes.

For mass spectrometry data, this means variance in means within mass bins are examined and a calculation on their statistical difference is performed, with a P value assigned to each feature. The P value tests the null hypothesis that data from all classes has identical means, and provides a probability that the observed difference between sample classes has occurred by chance, with a generally accepted threshold of $P \leq 0.05$. A P value of 0.05 indicates a false positive rate of 5%, and therefore classes with P values at or below this value are counted as statistically different as the observed differences are unlikely to have occurred by chance, and the null hypothesis of identical means is rejected. Post-hoc tests, such as Fishers LSD, reveal the sample classes to which the statistically difference bins occur and account for false positives by providing a significance level, with only those bins above the least significant difference level counted.

1.11.3 Multivariate Analysis

Multivariate analysis techniques are utilised in mass spectrometric data to examine similarities and differences in high dimensionality datasets across several features simultaneously, and are useful in datasets in which discrimination based on single biomarkers is not sufficient. In bacterial analysis, many features are shared throughout different species as a relatively high degree of biological similarity is shared between cell envelopes. Multivariate analysis is well suited to identifying correlations between samples and discerning features with high discriminatory power. Multivariate techniques can be split into supervised or unsupervised approaches, where supervised analysis refers to tests which consider user defined groupings and unsupervised refers to those which search for data variances without considering any pre-defined user data groupings (or class labels).

1.11.3.1 Principal Component Analysis (PCA)

PCA is an unsupervised technique commonly used to visualise class differences by mapping data onto a plot following orthogonal linear transformation. Data variance is displayed on coordinate axes as principal components, with the greatest variance assigned to principal component one, second greatest to component two, etcetera. The

number of components computed is selected by the user based on the amount of variance that is wished to be explained. PCA reduces a datasets dimensionality by removing low importance variables and grouping variables with high covariance, which makes it ideal for classification algorithm incorporation.

Data output from two-dimensional PCA is graphically represented as a scatter/score plot, with components one and two as the axes, allowing visualisation of general data trends and groupings. The percentage of variance the components are accountable for is also displayed, with one and two responsible for the most variance, resulting in them customarily forming the axes of the plot. Two points which appear close together on a PCA score plot indicates a high degree of similarity, or a low degree of variance, between the data. A loading plot is also generated along with a score plot, which displays the contribution of data points to variance within a component. In mass spectrometry data, data points on the score plot represent a mass spectrum, and data points on the loading plot represent spectral features which contribute to the separating power of the component. By considering the score and loading plot together, it is possible to discern how strongly particular variables influence the principal components.

PCA is a non-bias technique due to its unsupervised nature, which is positive when looking for general data trends but also has the potential to analyse variance arising from batch effects, growth media differences and experimental factors, rather than biological differences between samples. This enables the technique to be used as a data auditing tool for the purpose of ruling out bias caused by batch effects or other experimental parameters. PCA is often used as the first step in multivariate analysis due to its dimensionality reduction properties, before progressing on to discriminant analysis. Strong data groupings observed in a PCA model from batch effects, for example from samples analysed on different days, indicate poor quality data which should be reanalysed and not subjected to further statistical analysis. Retention of only the highest variance responsible components eliminates unnecessary data of low importance, such as that of background noise, improving the performance of subsequent tests. PCA has been established as an appropriate test for mass spectral profiling, and REIMS data in particular, due to its high complexity and dimensionality (158, 194-196, 213).

1.11.3.2 Discriminant Analysis

Discriminant analysis techniques are supervised methods in that they take into account user assigned data groupings when searching for variable variance within datasets. This lends a certain amount of bias to the outputs, although careful consideration of cross validation and performance scores enables techniques such as partial least squares discriminant analysis (PLS-DA) to be powerful exploratory, feature selection and classification model tools (214). PLS-DA can be considered as the supervised counterpart to PCA, in that it reduces dimensionality of complex datasets whilst preserving data covariance in the data and associated labelling, and maximises class separation by rotating PCA axes. Multivariate linear regression techniques are utilised for class discrimination, and the algorithm can also be utilised for important feature selection and classification. Features responsible for high intra-class variance can be identified through variable importance in projection (VIP), which assigns scores to features based on their influence in the model and presents them in ranked order for user review (215). VIP scores indicate which features contribute most to observed variation between classes.

Although useful, PLS-DA can be prone to model overfitting, so consideration of cross validation scores is essential. All PLS-DA plots are presented in this thesis with corresponding bar charts of R^2 and Q^2 scores, or a summary score. R^2 , or the coefficient of determination, is an indication of how well the model fits the data, or how well regression predictions fit real data points. The score is a percentage of the response variable variation, with high scores indicating a better fit of the model to the dataset. The coefficient of model prediction (Q^2) is an estimate of the predictive ability of the model and is calculated via cross validation. In each cross validation, predicted data are compared with original data and the sum of squared errors is calculated. The prediction error is summed over all samples (Predicted Residual Sum of Squares or PRESS), with good predictions possessing a low PRESS or high Q^2 score. There is no firm Q^2 score threshold for a robust model, as model performance is dependent on other factors such as R^2 and external validation, although a Q^2 of > 0.5 is generally indicative of good predictive ability (214). The optimal number of components computed in the model can be discerned by comparing Q^2 scores from different model builds. It is possible to

have negative Q^2 , meaning that the model is not at all predictive or is overfitted. Permutation tests are also performed for model quality assessment, in which new PLS-DA models are created with permuted class labels and the distribution of cross validation scores compared to provide an overall model significance value. In datasets with many variables and small sample sizes, such as with mass spectrometric data, discriminant analyses can find differences between sample classes by chance, but the features responsible for discrimination will not be applicable to other datasets. Permutation tests give an indication of PLS-DA reliability in prediction accuracy and assign associated P value based on the frequency of class separation from permuted data repeats compared to the original dataset.

1.11.3.3 Random Forest

Random forest (RF) classification algorithms use a supervised approach in growing hundreds of decision-making trees to populate a forest of strong selecting features. The principle of the algorithm is based on Condorcet's jury theorem, in that each decision tree is weak in classification but marginally better than $P = 0.5$, however a forest of these trees together make a strong classification model (216). RF is a non-linear machine learning method which generates hierarchical trees, nodes of variable/threshold pairs and leaves of sample classes. The root of the tree consists of the initial sample set and the first decision node, from which each possible variable and threshold pair are tested to generate a sample grouping, with the supervised label used to determine the best combination. This process is repeated at each tree node until leaves of classes are generated. To classify an unknown sample, the variable/threshold pair of the sample is compared to each tree node and follows down the classification route until a leaf, or sample class, is reached. Diversity is introduced to the model through bootstrapping, in which random subsets of sample classes are selected to build decision making trees, usually with the range of 400-1000 trees in a forest. The advantages of RF over linear discriminant techniques are that it is not prone to overfitting and features important in classification can be determined through the native RF algorithm rather than running separate tests. Throughout the experimental section of this paper multiple univariate and multivariate tests will be utilised and compared, and their performance scores considered, in order to arrive at a robust conclusion.

1.12 Hypothesis/Aims

It is proposed that the scope of applications of REIMS can be expanded to include clinical sample types such as urine, and that the system is capable of identifying species-specific biomarkers in poly-microbial samples. Secondly, REIMS can be used as a culture-free identification method for non-human clinical samples with applications in the veterinary field.

1.13 Objectives:

- i. REIMS will be used to firstly assess if individual bacterial species can be detected in poly-microbial samples. Bacterial isolates manually mixed *in vitro* will be examined to determine if observed spectra are additive or suppressive. This will give an indication if analysis of true poly-microbial samples will be viable.
- ii. REIMS methodology and instrumentation will be developed and optimised to allow analysis of liquid samples.
- iii. REIMS will be carried out on clinical urine samples, and the biomarker database will be utilised to identify microbial components of the samples. The microbial composition of clinical samples will be determined using the existing workflow of an NHS clinical microbiology laboratory, including conventional culturing techniques, biochemical tests and MALDI-ToF in order to validate REIMS analysis.
- iv. Biomarker analysis will allow the beginning of a repository of species-specific mass spectral features.
- v. The system will be compared with Acoustic Mist Ionisation Mass Spectrometry (AMI-MS), a novel technique still in the development stage, to determine the strengths and weaknesses of each approach.
- vi. Potential of the optimised system in the analysis of non-human veterinary samples will be assessed so further commercial applications of the technique can be realised.

2 Materials and Methods

2.1 Health and Safety

Analysis using REIMS was approved by the Imperial College Health and Safety team after consideration of a completed risk assessment for work involving biological material (Form Bio1, project assessment I.D.: IC-4987). All microorganisms and project samples were treated as Hazard Group II materials and handled within containment level 2 facilities. REIMS analysis was performed within a custom-made class 2 biological safety cabinet with the associated safety procedures followed to minimise risk of potential pathogen exposure. Decontamination and emergency procedures relating to spillages or exposure were in place within the laboratory and were followed at all times. Occupational health monitoring was also in place and an assessment of the project was conducted along with a vaccination programme.

2.2 Ethical Approval

Analysis of urine samples was approved by the Tissue Management Committee of the Imperial College Healthcare Tissue Bank (application number R15079: Development of a Direct from Sample Diagnostic Test for the Identification of Urinary Tract Infections).

2.3 REIMS Analysis

Mass spectrometric analysis was performed on a Waters Xevo G2-XS QToF instrument (Waters, Wilmslow, UK) fitted with a custom REIMS atmospheric interface ion source (Waters Research Centre, Budapest, Hungary). All analysis was carried out in negative ion sensitivity mode across a mass range of 50-2500 m/z (Table 1). Experiments were aimed to be completed in one day where possible, to prevent extensive mass drift. Before beginning each experiment, the instrument was calibrated using a 5 mM sodium formate solution (Sigma-Aldrich, Dorset, UK). The solution was received as a 4 M liquid stock and diluted to the working concentration using HPLC grade 2-propanol (ROMIL Cambridge, UK) and HPLC grade water (VWR, Leicestershire, UK), and stored at 4°C for up to one month. Solution was drawn into a 5 mL glass HPLC syringe (Hamilton, Bonaduz, Switzerland), connected to the instrument inlet capillary using HPLC PEEK tubing (Sigma-Aldrich, Dorset, UK) and infused at a rate of 0.25 mL/min for up to 10

seconds via a syringe pump (Scientific Instrument Services, New Jersey, US). Calibration was performed across the 50-2500 m/z range using the manual Intellistart method through MassLynx, with pass parameters of <3 ppm RMS residual mass and >34 reference peaks detected. To account for mass drift, a leucine-enkephalin lock mass compound was infused at a rate of 0.2 mL/min throughout the duration of each experiment through a HPLC pump (Knauer, Northampton, UK). The leu-enk stock solution was made using HPLC grade 2-propanol (ROMIL Cambridge, UK) and stored at -20°C. A 10 nG/mL working solution was made fresh before the start of each experiment and stored in an insulated cool box during pumping to prevent degradation of the compound, and discarded at the end of each day.

Table 1. Instrument parameters for REIMS analysis. A mass range of 50-2500 m/z is acquired in negative ion mode to capture spectral profiles covering a wide range of interest.

Scan time	1000 ms
Scan mode	Sensitivity
Mass analyser	Time of Flight
Ion mode	Negative
Mass range	50-2500 m/z
Sampling cone	80 V
Source offset	50 V
Source temperature	Off
Collision surface current	2.8 A

Generator	Erbe
Mode	Cut
Power	9 W

REIMS analysis was carried out using bipolar forceps, monopolar handpiece, monopolar probes or laser ablation within a custom-made Class II biosafety cabinet (Bigneat, Waterlooville, UK) to contain microbial vapour. Both bipolar forceps and monopolar handpiece (iKnife) REIMS are manual, user operated methods, with monopolar probe and laser assisted REIMS semi-automated via the use of a custom-made liquid handling robot (Tecan Group Ltd., Männedorf, Switzerland & Waters Research Centre, Budapest, Hungary). The Class II biosafety cabinet was specifically made to accommodate Tecan instrumentation and laser apparatus safely, with a door interlock system capable of cutting off power to the laser should the cabinet door be opened. Ports for REIMS transfer tubing and a germicidal UV lamp were fitted, along with a light stage for viewing agar plates via an inbuilt camera, and a return electrode to enable monopolar REIMS.

2.3.1 Bipolar REIMS

Electrosurgical forceps with an inbuilt aspiration line (Erbe Medical Ltd., Leeds, UK) were connected to an electrosurgical generator (Erbe Medical Ltd., Leeds, UK) set to bipolar mode with a power rating of 70 W. 1.5 M PTFE transfer tubing of 1/8in. outer diameter, 1/16in. inner diameter (Sigma-Aldrich, Dorset, UK) was connected to the forceps aspiration line outlet and inlet capillary of the mass spectrometer. The forceps were used to pick up approximately 5 μ L microbial biomass, either directly from agar media or from a pellet, before activating the power supply using a foot pedal and closing the tips to make light contact with the biomass (Figure 6). Analysis of liquid culture was also possible by disconnecting the transfer tubing and holding it at the mouth of the microcentrifuge tube (Eppendorf, Stevenage, UK) or Falcon (Fisher Scientific, Loughborough, UK) tube before submerging the forceps into the liquid culture and activating the power supply. Power was increased to 80-90 W with this method to produce sufficient vapour for analysis. Contact of the forceps tips and biomass produces joule heating and vaporisation, before the vapour is aspirated into the aspiration line through the native vacuum of the mass spectrometer. Power was supplied to the forceps for between 1 and 5 seconds depending on the volume of vapour produced. In the case of mucoid forming bacteria such as *P. aeruginosa* or when analysing liquid cultures,

longer heating times were required to adequately vaporise the sample. Mass spectra for three separate vaporisations were recorded for each sample.

Cleaning of the forcep tips was required between analyses of each isolate due to charring and build-up of biomaterial on the tip surface, resulting in reduced vapour formation. This was performed by wiping residual biomass with a sheet of absorbent tissue (Kimberly-Clark, Kent, UK) soaked in 70% v/v ethanol solution (VWR, Leicestershire, UK) until charring was removed. The aspiration line inlet was prone to clogging with residual biomass if the tips of the forceps were not carefully cleaned after each acquisition, or if the biomass was forced upwards toward the inlet upon pinching the tips together. To overcome this, the transfer tubing was disconnected from the forceps and a 10 mL disposable syringe (Fisher Scientific, Loughborough, UK) filled with 2-propanol (ROMIL Cambridge, UK) before connecting to the aspiration line outlet and plunging the solvent through. One to three attempts were typically necessary to completely clear the aspiration inlet and line of debris. The forceps were either dried by again using the syringe to plunge through air, or by connecting to a nitrogen source and flushing. Rarely, the PTFE transfer tubing or mass spectrometer inlet capillary could also accumulate deposits if the aspiration inlet was covered with biomass, which were cleaned by removing and flushing with 2-propanol (ROMIL Cambridge, UK) or by removing and replacing the tubing in the case of heavy deposition. The tubing and capillary were air or nitrogen dried, as described previously, before connection to the instrument.

2.3.2 Monopolar Handpiece REIMS

A modified, short blade, electrosurgical scalpel (Medtronic, Budapest, Hungary), or iKnife, with embedded transfer tubing was connected to an electrosurgical generator (Erbe Medical Ltd., Leeds, UK) set to monopolar “cut” mode. Power supply to the iKnife was typically 11 W but was adjusted between 8-12 W based on the sample type being analysed, with isolates prone to excessive charring and splattering analysed at a lower power. As with bipolar REIMS, PTFE transfer tubing was connected to the hand piece and the inlet capillary of the mass spectrometer (Figure 8). Power was activated by pressing a button on the hand piece and the blade was lightly touched to the surface of

biomass before moving it along the culture for 2-5 seconds. Produced vapour was drawn into the transfer tubing by the native vacuum of the instrument. After three distinct vapour acquisitions, the blade was cleaned by wiping on absorbent tissue soaked in 70% v/v ethanol solution. In case of heavy charring or deposition, the blade was removed from the hand piece and submerged in 100% 2-propanol before placing in an ultrasonic bath for 10 minutes.

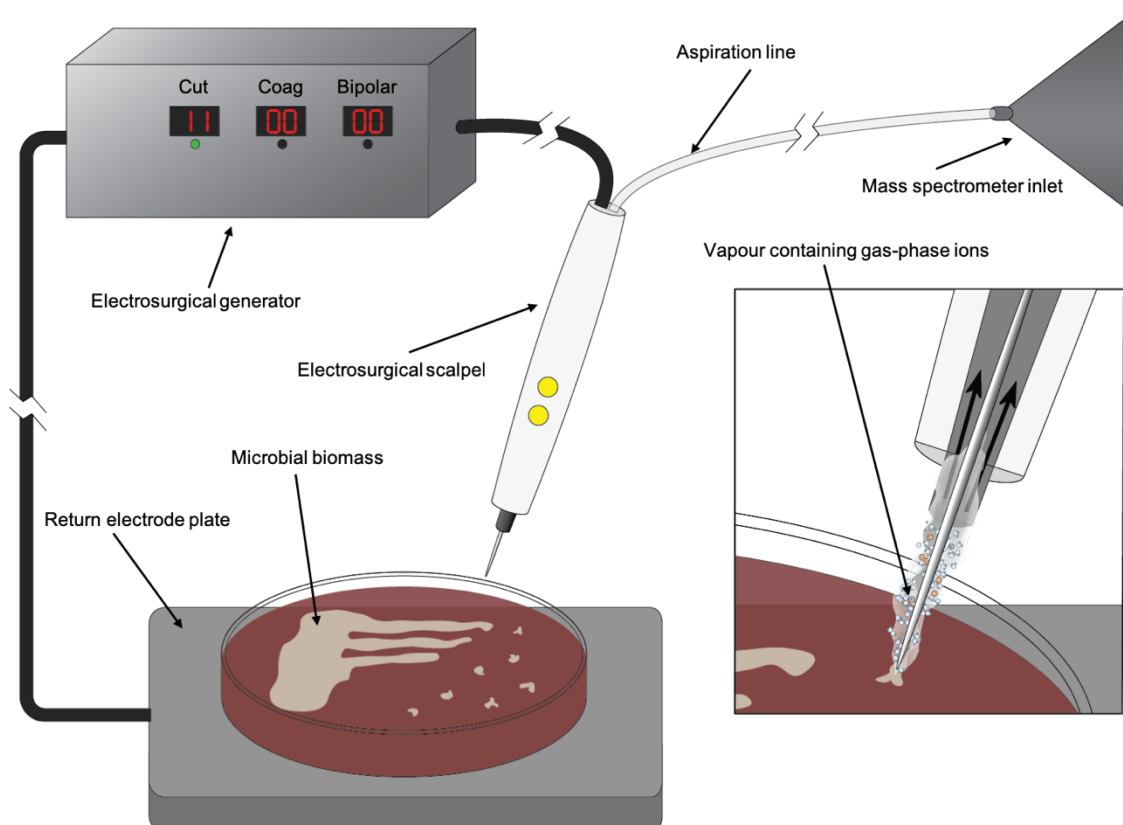


Figure 8. Monopolar handpiece REIMS (iKnife). A modified electro-surgical scalpel is connected to an electro-surgical generator set to cut mode, 11 W. A culture plate is placed on a return electrode and the handpiece is brought down to make contact with microbial biomass. An activation button is pressed and the biomass is vaporised through joule heating. The resultant vapour, containing gas phase ions, is channelled through the handpiece to the mass spectrometer inlet capillary.

2.3.3 Automated REIMS

2.3.3.1 Electric Diathermy

Scripts for running of the automated platform were written in Freedom EVOware software (Tecan Group Ltd., Männedorf, Switzerland) according to the requirements of each experiment. Parameters altered with scripts included three-dimensional positioning of the robotic arm for different sample types, automation of MassLynx acquisition windows, filename generation, probe power supply time and acquisition time. The reusable, purpose-built, conductive probes (Waters Research Centre, Budapest, Hungary) were loaded into a 96 well rack and placed within the Tecan instrument. The sample to be analysed was placed onto the light stage and the relevant script started through EVOware following Tecan initialisation (Figure 7). Power for the monopolar probes was supplied by an Erbe electrosurgical generator, as with the hand-piece methods, connected to the Tecan instrument through a connection adapter (Waters Research Centre, Budapest, Hungary). Power activation occurred through prompting by an automated script which also raised and lowered probes in a circular pattern to avoid vaporisation inhibition via charred biomass build up. Spectra were acquired for 70 seconds for each sample, with three replicates achieved during this time. A three second delay was inputted between each replicate to allow for vapour to completely clear from the transfer line. After each acquisition replicate, used tips were ejected into a collection box by the robotic arm. Tips were collected and placed in a cleaning rack which, when full, was submerged into 100% 2-propanol and allowed to soak for up to 10 minutes. The rack was removed from the solvent bath and air dried within a ducted cabinet. Dry tips were reloaded into the instrument rack on a continuous basis for ongoing analysis.

2.3.3.2 Laser-Assisted

For laser-assisted REIMS, the monopolar probe arm of the Tecan instrument was modified to hold the laser lens assembly (Figure 9). A Class 4 CO₂ laser of 10.6 μ M wavelength (Omniguide, Massachusetts, US) was retrofitted into the Tecan housing using a fibre optic beam delivery system routed through the monopolar probe arm. The Class II biological safety cabinet was also modified to contain diffuse reflections by

fitting a louvre cover over the viewing panel, as well as an interlock system which cuts off power to the laser if opened. The Omniguide module was connected to a helium gas cylinder, for cooling purposes, with a supply pressure of 2 bar during operation.

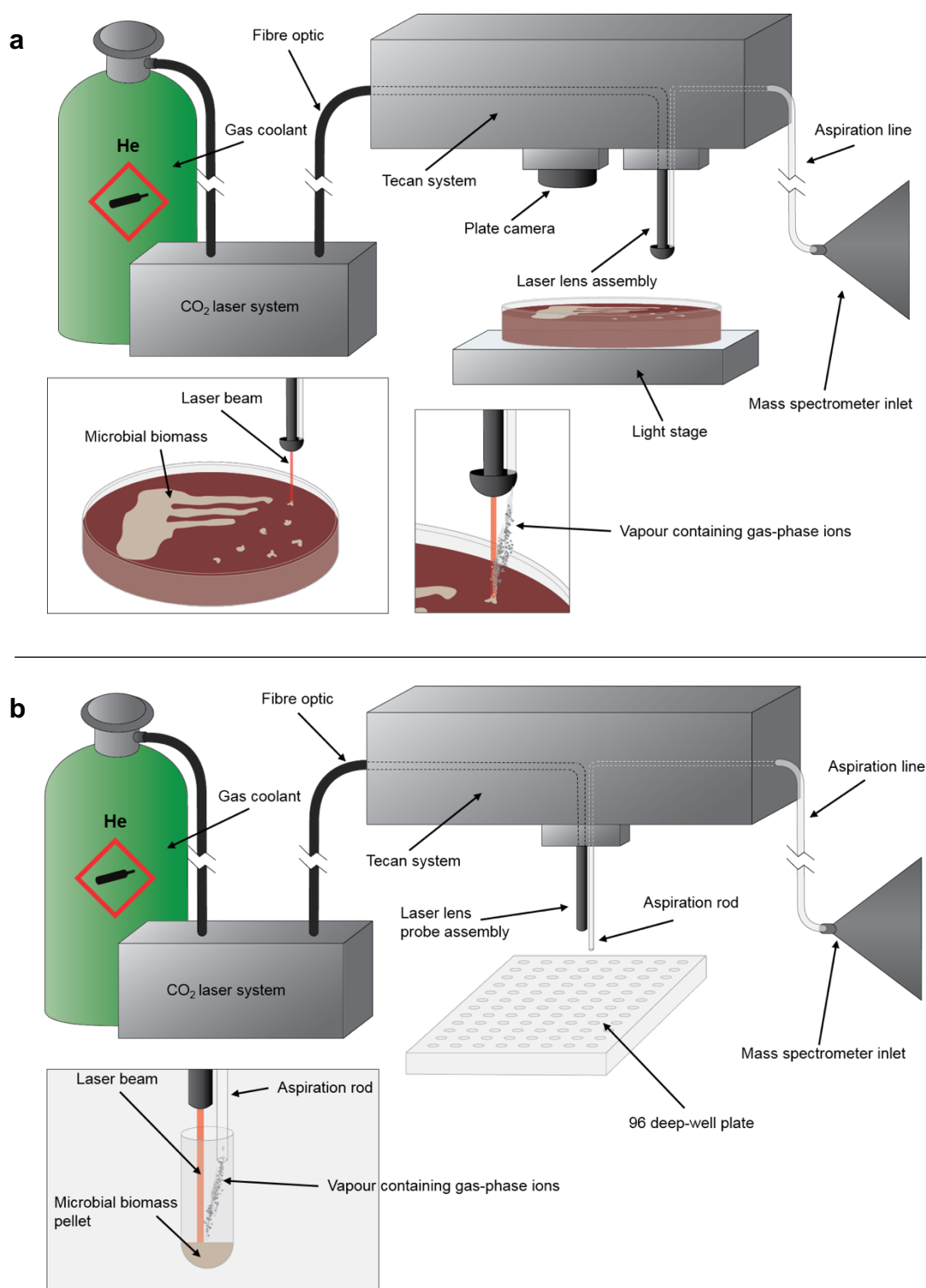


Figure 9. Automated Laser-Assisted REIMS (culture plate setup (a) and 96 well plate setup (b)). A modified Tecan liquid handling instrument was fitted with a CO₂ fibre optic laser system. A robotic arm holds the laser lens, which descends to a user selected colony once a camera records an image of a culture plate (a), or an aspiration rod descends into wells of a deep-well plate to analyse cell pellets (b). The aspiration rod allows vapour to be captured within the well.

Laser parameters were altered depending on sample type and format, with the “Z-max” value set to equate to a 2-2.5 mm distance from the sample surface, resulting in a beam diameter of 485-520 μM (Figure 10). Scripts were written to accommodate analysis from agar media, 24 well plates, 96 deep well plates or racked 2 mL microcentrifuge tubes. In particular, the “Z-max” value, referring to the depth setting of the probe, was checked and adjusted before each experiment to ensure the laser lens was approximately 2 mm from the sample surface. Laser power was typically set at 1.5-2.5 W for solid sample analysis (plate cultures and cell pellets) or 2.5-3.5 W for liquid sample analysis (liquid surface and droplets). Energy delivery mode was set to pulsed, with the superpulse option switched on, resulting in a beam gated pulse width of 500 ms and a secondary pulse rate of 375 pps (superpulse mode).

The continuous beam mode was used in optimisation experiments for liquid cultures, with the superpulse mode activated. EVOware scripts were written to automatically guide the laser fibre optic, activate the beam and start the acquisition in MassLynx. A wait period of 3 seconds was inputted to allow vapour dispersal between sample replicates as with the monopolar probe method. Cleaning procedures for LA-REIMS are greatly simplified thanks to no contact being made with samples. To clear splattering or deposit build up from heavy vapour generation on the laser lens assembly, the plastic lens cover was removed after each experiment and submerged in 2-propanol before placing in an ultrasonic bath for 10 minutes. As with the other methods, transfer tubing was replaced or flushed with 2-propanol to clean. The instrument inlet capillary and T-piece were also submerged in a 1:1 ratio of 2-propanol (ROMIL, Cambridge, UK) and HPLC grade water (VWR, Leicestershire, UK) after each experiment to remove residue.

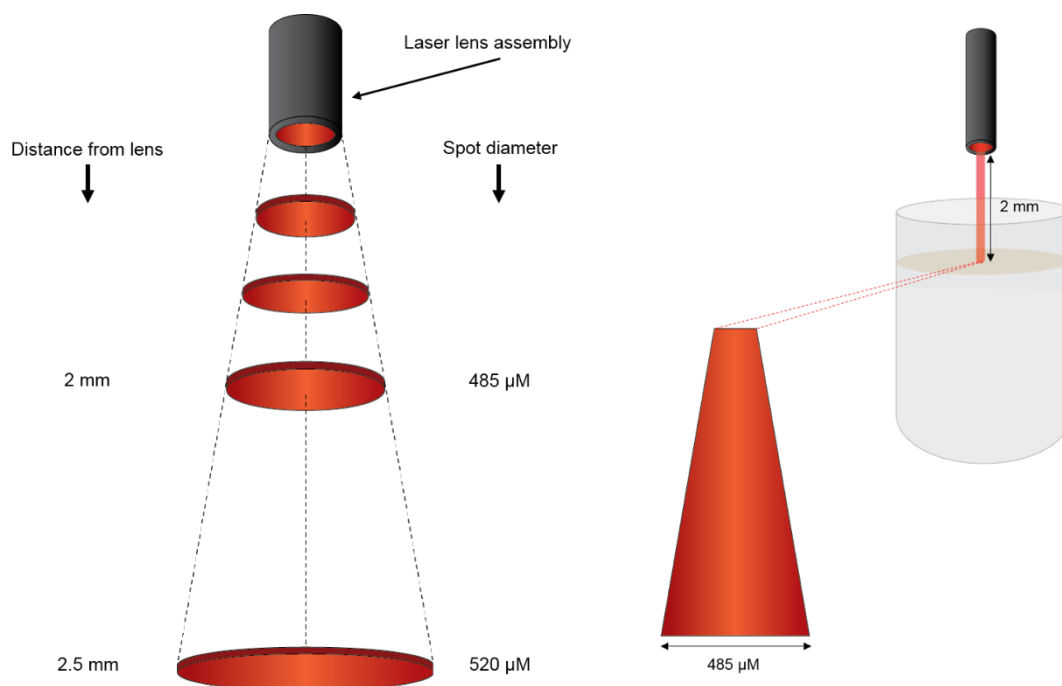


Figure 10. Laser beam spot diameter representation. At a distance of 2 mm from sample surface, laser beam diameter is equal to 485 μm . The beam diverges with increasing distance away from the sample.

2.4 Interface Configuration

Two generations of the REIMS atmospheric ion source interface were utilised in the completion of this study. Version one of the interface was utilised in handheld bipolar and monopolar REIMS, with version two used for automated monopolar probe and LA-REIMS analysis. V1 comprised of a vacuum chamber housing enclosing a kanthal coil collision surface, to which an external bench power supply (RS Components, Northants, UK) was connected via external connection points and set to deliver 2.1 A. A collared inlet capillary was attached to the interface to maintain a 2mm capillary outlet distance from the collision surface. A custom PTFE T-piece (Waters Research Centre, Budapest, Hungary) was attached to the external section of the inlet capillary, and PTFE transfer tubing leading to the REIMS handpiece was attached via luer connection to the top inlet. PEEK tubing was connected to the parallel inlet for infusion of 2-propanol matrix solvent. V2 was optimised to include an isolated collision surface chamber, allowing removal of the inlet capillary and collision surface for cleaning without venting the instrument vacuum. An inbuilt connection to the instrument interface source power

supply allowed the collision surface to be powered by the internal instrument electronics, controlled by a button scripted into the MassLynx tune page graphical user interface. The T-piece was also modified to be smaller and crafted from stainless steel rather than PTFE, to allow for more vigorous solvent cleaning and lessen the chance of solvent leeching from the plastic.

2.5 Cleaning

As part of a general maintenance routine, the instrument was vented once per week and the REIMS interface removed to allow for cleaning. After removing source housing, the capillary was removed and placed into a measuring cylinder which was filled with a 1:1 ratio of 2-propanol (ROMIL, Cambridge, UK) and HPLC grade water (VWR, Leicestershire, UK). The StepWave ion guide was also removed using the StepWave removal tool, similarly placed into a measuring cylinder and immersed in the solvent mixture described above. Both measuring cylinders were placed in an ultrasonic bath and sonicated for 10 minutes to clean. The housing, collision surface and StepWave surroundings were also wiped clean using cotton swabs or Kimwipes (VWR, Leicestershire, UK). Once the StepWave and interface were replaced the instrument was pumped to vacuum again for use the following day. Each day after calibration, a reduced cleaning protocol was followed to remove remnants of sodium formate solution from the interface, by subjecting the T-piece and inlet capillary to the same procedure described above. This can be done without venting the instrument as the V.2 REIMS source has an isolating valve, allowing the external components to be removed without effecting the instrument vacuum. Laser lens housing, bipolar forceps or monopolar probes was also cleaned daily following each experiment by sonicating or wiping with cleaning solution as required.

In the case of bipolar forceps, soaking overnight followed by sonicating in the cleaning solution was carried out in the case of heavy biomass deposit buildup around the forcep tips or aspiration inlet. The instrument vacuum pump (Edwards, West Sussex, UK) was fitted with an all in one exhaust filter assembly (Scientific Instrument Services, New Jersey, US) since extraction to the outside air was not possible within the lab. A monthly

cleaning protocol of the filter assembly was carried out in which the housing was washed and the charcoal trapping exhaust filter component replaced. The coalescing filter element was replaced every six months.

2.6 Culturing

An ongoing collaboration of the wider MicrobeID project with the Imperial College Healthcare NHS Trust microbiology laboratory (ICHNT) at Charing Cross Hospital was utilised for the majority of isolates analysed in this project. All pure bacterial isolates were initially collected by registered Biomedical Scientists within the ICHNT microbiology laboratory during routine clinical microbiology work, mostly from blood cultures. Collected isolates routinely enter -80°C Microbank (Pro-Lab Diagnostics, Wirral, UK) stores. Isolates from these stores were purity plated onto Columbia Blood Agar (CBA) media (Oxoid, Basingstoke, UK) to confirm absence of contaminants. All isolates were previously identified using the conventional laboratory workflow, including MALDI-ToF MS analysis using a Bruker Microflex LT instrument (Bruker Daltonics, Coventry, UK), according to the manufacturer's instructions and UKAS accredited standard operating procedures of the clinical laboratory. A second Microbank store, for exclusive use by the MicrobeID team, was populated with these validated isolates and stored in the MicrobeID laboratory at -80°C. Isolates used in experiments were obtained from the second store and cultured aerobically, anaerobically or at 5% CO₂ depending on the specific isolate, by dipping a flexible, disposable, sterile 1 µL culture loop (VWR, Leicestershire, UK) into Microbank glycerol broth before streaking on columbia blood agar (CBA) or chocolate (CHOC) growth media, followed by incubation at 37°C for 24 hours (Table 2). Culture manipulation was performed within a Class II biological safety cabinet within the MicrobeID laboratory.

Table 2. Clinical bacterial isolates and associated culture conditions (CBA: Columbia Blood Agar, CHOC: Chocolate Agar) utilised in cell line probe optimisation for REIMS. A wide range of clinical isolates was chosen to ascertain optimal conditions for analysis in real clinical settings.

Organism	Culture conditions	Growth Media	Organism	Culture conditions	Growth Media
<i>Acinetobacter baumannii</i>	37°C, O ₂	CBA	<i>Morganella morganii</i>	37°C, O ₂	CBA
<i>Candida albicans</i>			<i>Proteus mirabilis</i>		
<i>Citrobacter koseri</i>			<i>Pseudomonas aeruginosa</i>		
<i>Enterobacter cloacae</i>			<i>Staphylococcus aureus</i>		
<i>Enterococcus faecalis</i>			<i>Staphylococcus epidermidis</i>		
<i>Enterococcus faecium</i>			<i>Stenotrophomonas maltophilia</i>		
<i>Escherichia coli</i>			<i>Streptococcus pneumoniae</i>	37°C, 5% CO ₂	CHOC
<i>Klebsiella oxytoca</i>			<i>Streptococcus oralis</i>		
<i>Klebsiella pneumoniae</i>			<i>Haemophilus influenzae</i>	37°C, anO ₂	CBA
<i>Micrococcus luteus</i>			<i>Bacteroides fragilis</i>		

2.7 Cell line probes experiment

2.7.1 Probe Optimisation:

Modified cell line probes were obtained from the Waters Research Centre, Hungary, for the purpose of optimising the REIMS setup for liquid culture analysis. A total of 14

probes were received with a range of steel alloy electrode materials, geometry and electrode spacer configurations (Supplementary Table 2). A stock solution of Brain Heart Infusion broth was made by dissolving 37 g of dehydrated media (Oxoid, Basingstoke, UK) into 1 L of purified water (TripleRed, Bucks, UK) before decanting 200 mL into 250 mL Duran bottles (VWR, Leicestershire, UK). An *E. coli* isolate was inoculated into one bottle using a 10 µL culture loop. The broth was placed on to a shaking platform within a 37°C aerobic incubator overnight. The culture was agitated before 1 mL was extracted using a pipette and added to each well of 2x96 deep-well plates (Eppendorf, Stevenage, UK). The plates were sealed with an aluminium plate seal (VWR, Leicestershire, UK) and placed within a sealed plate bucket before centrifugation (Eppendorf 5810R) at 3486 x g, 4°C for 10 minutes to obtain a cell pellet. The supernatant was discarded into a Duran bottle for autoclaving. Each cell line probe was connected to an electrosurgical generator (Erbe Medical Ltd., Leeds, UK) set to bipolar mode and used to analyse each cell pellet in replicates of three, with heating powers 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 W. Analysis was performed by activating the generator foot pedal and lowering the probe electrode into the cell pellet within the plate wells until vapour was produced.

The optimal probe and power level were selected by comparing mass spectra for the different power levels and examining signal:noise ratio scores along with spectral richness.

2.7.2 Method Comparison

A secondary experiment was conducted to compare the optimal cell line probe to bipolar forceps and monopolar shortblade handheld probe (iKnife) REIMS. In this, 20 clinically significant microbial isolates, in five replicates, were streaked from cryovial (ProLab Diagnostics, Wirral, UK) stocks onto Columbia Blood agar plates (Oxoid, Basingstoke, UK), or Chocolate agar plates in the case of *H. influenzae*, and cultured overnight (Table 2). Isolates requiring anaerobic conditions were placed into a 2.5 L anaerobic culture jar along with an AnaeroPack atmospheric oxygen absorbing sachet (Mitsubishi Gas Chemical, Tokyo, Japan), before being placed into the incubator. Cultures were checked for purity the following day and disposable 10 µL culture loops

were used to inoculate 10 mL sterile Brain Heart Infusion broth vials (Oxoid, Basingstoke, UK) with the pure isolates. Biomass was analysed directly from the remaining agar plates using handheld monopolar and bipolar REIMS. The vials were placed into racks on a shaking platform within an incubator overnight, or in an anaerobic culture jar as before. Liquid cultures were agitated before 1 mL was pipetted into 96 deep-well plates (Eppendorf, Stevenage, UK). The plates were sealed with an aluminium plate seal (VWR, Leicestershire, UK) and placed within a sealed plate bucket before centrifugation (Eppendorf 5810R) at $3486 \times g$, 4°C for 10 minutes to obtain a cell pellet. The supernatant was discarded into a Duran bottle for autoclaving. Cell pellets were analysed using both the optimal probe from the previous experiment, and with bipolar forceps. For bipolar forceps REIMS, the PTFE transfer tubing was disconnected from the forcep aspiration outlet and held at the top of each well during analysis to prevent biomass entering and clogging the line. The optimal analysis method was investigated by comparing speciation accuracy for each of the REIMS types, along with signal:noise ratio scores and spectral richness.

2.8 Mixed Culture Experiment

2.8.1 Bacterial Culture and Identification

Three isolates, collected from different clinical patients, of four clinically relevant bacterial species (*E. coli*, *S. aureus*, *P. aeruginosa* and *E. faecalis*) were collected from the Imperial College Healthcare NHS Trust (ICHNT) microbiology laboratory. Isolates were chosen based on their clinical significance and high occurrence rates in the laboratory, and were previously identified using the conventional laboratory workflow including MALDI-ToF MS analysis using a Bruker Microflex LT instrument (Bruker Daltonics, Coventry, UK), according to the manufacturer's instructions and UKAS accredited standard operating procedures. The 12 isolates were aerobically cultured for 24 hours on standard Columbia Blood agar (CBA) media (Oxoid, Basingstoke, UK) at 37°C and mixed using sterile 1 µL culture loops on fresh CBA plates. Mixtures were prepared in all possible distinct combinations of two species, in ratios of 1:1, 1:2, 2:1, 3:1 and 1:3. All work was performed in a containment level 2 laboratory with the associated safety procedures followed to minimise risk of potential pathogen exposure.

2.8.2 Rapid Evaporative Ionisation Mass Spectrometry Analysis

The MS instrument was calibrated prior to analysis using standard 5 mM sodium formate solution recommended by Waters, over 50 to 2500 m/z mass range. Each sub-experiment was performed on a single day. For the mixed culture sub-experiment sample mixtures, alongside pure cultures of each bacterial species, were analysed directly from agar plates using a Waters Xevo G2-XS QToF (Waters Corporation) instrument in negative ion sensitivity mode and the automated REIMS platform as previously described (195). Negative ion mode spectra were acquired in continuous mode at one scan per second for one minute over a 50 to 2500 m/z mass range, in five technical replicates. A 2-propanol solvent containing 10 nG/mL leu-enk lock-mass solution was infused at a rate of 0.2 mL/min to allow for correction of mass drift during analysis, aid ionisation and prevent build-up of vapour residue within the REIMS interface and inlet capillary of the instrument. Analysis was performed within a Class II Biosafety Cabinet to prevent inhalation of the bacterial vapour and contamination of samples from environmental microorganisms.

2.8.3 Data Processing and Statistical Analysis

Mixed culture data was imported in raw format to custom Waters Offline Model Builder (OMB) software (version: 1.1.29.0) and processed to perform mass drift correction, background subtraction, total ion current normalisation and binning to 0.1 Da. Using a mass range of 600 to 900 m/z data matrices were produced for each ratio and exported for further analysis. Univariate analysis of variance (ANOVA) was performed by editing the OMB generated data matrices of mass spectral bins to a standard CSV file using a custom R (R Project) script and uploading to the cloud based MetaboAnalyst 3.0 (217) suite of statistical analysis tools before selecting the one-way ANOVA tool. Only those mass bins with a false discovery corrected $P < 0.05$ were used in subsequent analysis. Two-dimensional unsupervised principal component analysis was also performed on the 600-900 m/z mass range using the PCA tool within the chemometrics analysis path, after log transformation and Pareto scaling of data.

2.8.4 Generation of *in silico* Mixed Culture Mass Spectra

Pure isolates of the microbial cultures were also analysed and used to produce *in silico* models of predicted mixed culture spectra to provide a comparison to the *in situ* mixed culture spectra acquired by the high-throughput REIMS platform. To achieve this, data matrices containing mass bins of pure samples were produced using Waters OMB software as before and imported to R. Synthetic matrices were created by multiplying each mass bin intensity value by the intended ratio of the sample within the culture mixture, according to the following formula:

$$a \cdot A / b \cdot B$$

Where a=ratio value of sample 1, A=sample 1 mass bin, b=ratio value of sample 2, B=sample 2 mass bin.

The synthetic mixture mass bins were plotted along with those produced *in situ* to examine the correlation.

2.9 Milk experiments

2.9.1 Cell Line Probe

A collaboration with Quality Milk Management Services (QMMS, Draycott, Cheddar) was set up to determine if REIMS could be an effective, rapid test for bovine mastitis screening. A total of 141 bovine mastitis positive milk samples were obtained from QMMS, with causative microorganisms pre-identified through the standard laboratory workflow of biochemical and culture-based tests. A total of 15 culture negative, non-mastitis, bulk tank milk samples were also obtained for analysis. Samples were frozen and stored at -80°C within 24 hours of collection by QMMS. Prior to REIMS analysis, samples were thawed at ambient room temperature and a 2 mL aliquot added to 24 well plates before centrifugation at 4°C, 3486 x *g* for five minutes (maximum *g* of centrifuge with plate rotor attachment). The resulting surface lipid/fat layer was analysed directly from the plates using cell line probe REIMS at 8 W and standard instrument parameters (Figure 11). Raw data was exported from MassLynx into OMB software for preprocessing.

A data matrix of mass bins was generated in OMB and uploaded to MetaboAnalyst 3.0 (217) for principal component analysis and ANOVA.

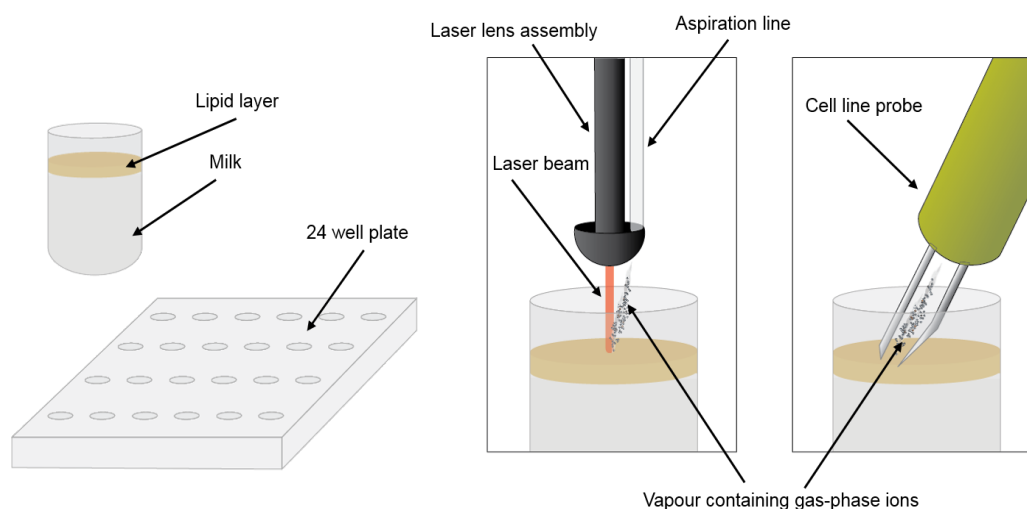


Figure 11. Bovine milk lipid layer analysis representation. Milk samples, obtained from Quality Milk Management Services (QMMS), were aliquoted into 24 well plates before centrifugation. The resulting lipid surface layer was analysed with both LA-REIMS and cell line probe REIMS in order to determine the capability of the technology in discerning mastitis positive/negative milk samples.

2.9.2 Laser optimisation

Whole milk was obtained from a supermarket for the purpose of laser optimisation only, before analysis of QMMS samples on the LA-REIMS platform. Aliquots of 2 mL were pipetted into 2x 24 well plates and centrifuged at 4°C, 3486 x *g* for five minutes to obtain a lipid surface layer. The lipid layer was analysed directly from the plates using the laser power outputs 1, 1.5, 2, 2.5, 3 and 3.5 W. Z-max values of 1820 and 1815 were tested, corresponding to 2.0 and 2.5 mm lens distances from the sample surface. Signal:noise ratio scores were generated for each power level and lens distance combination and optimal parameters for analysis were selected.

2.9.3 LA-REIMS

A second 2 mL aliquot of the milk samples obtained from QMMS were thawed and processed as in the cell line probe analysis experiment. Optimal laser power and lens parameters determined from laser optimisation using supermarket whole milk were

used in the analysis of the lipid layer of the samples (Figure 11). Data analysis was performed in the same way as in the cell line probe analysis experiment, to directly compare and contrast the two methods.

2.10 Urine Analysis

2.10.1 Spiked Urine Method Development (pellet forceps)

2.10.1.1 Spiked Urine Sample Preparation

The same isolates obtained from ICHNT and used in the mixed culture experiment (section 2.8) were again aerobically cultured for 24 hours on CBA media at 37°C. 500 mL urine was obtained from a single healthy donor. Standard 10 µL disposable culture loops (VWR, Leicestershire, UK) were used to place 30 µL of biomass from each pure bacterial species into a 5 mL microcentrifuge tube containing donor urine. The starting solution was diluted 2x, 5x, 10x, 100x, 1000x and 10,000x using the donor urine and vortexed to mix. Additional serial dilutions up to $\times 10^6$ were performed for plate counting. Sample tubes were placed in a centrifuge (Eppendorf 5810R) set to 4°C and centrifuged at 4500 $\times g$ for 10 minutes. The supernatant was removed from each tube leaving a pellet for REIMS analysis. 100 µL of each dilution factor was pipetted onto CBA plates at the time of dilution and incubated aerobically overnight at 37°C. Culture plates were removed from the incubator and colonies counted on the appropriate dilution media to determine CFU/mL of each of the serial dilutions.

2.10.1.2 REIMS Analysis

Microcentrifuge tubes were placed in racks within the Class II Biosafety Cabinet and cell pellets analysed using the handheld bipolar REIMS method described previously (158). Forcep power was set to 70 W and a 10 nG/mL leu-enk lock-mass solution was infused throughout analysis as before.

2.10.1.3 Data Analysis

Raw data files were batch processed within MassLynx v4.1 (Waters, Wilmslow, UK) for lock-mass correction to 554.26 m/z and background subtraction. Species-specific biomarkers identified from ANOVA in the mixed culture sub-experiment were

manually interrogated by opening each sample spectra and searching the mass regions of interest. Peak intensity of biomarkers were recorded in a spreadsheet for each sample and data visualisation completed in Microsoft Excel 2013. Biomarker intensity for each dilution analysed was correlated to CFU/mL values to ascertain the limit of detection of the biomarkers looked at within each sample. Blank urine sample data was also interrogated in the same manner to determine at what concentration the biomarkers became masked by the background spectra of urine.

2.10.2 LA-REIMS Method Development

Isolates utilised in this experiment were identical to the previous spiked urine experiment, and were cultured, prepared and stocks diluted in the same manner, however the samples were not centrifuged, and made in duplicate. In addition, a greater range of concentrations were created, up to 10^7 . 500 mL of healthy urine was obtained from the same donor for serial dilution creation. Additional culture plates were prepared, as before, for CFU/mL enumeration the following day. Pre-cut nitrocellulose membrane filters of 0.45 μ M pore size, 13 mm diameter (Whatman, Kent, UK) and grade 595 qualitative filter paper of 4-7 μ M pore size, 125 mm diameter (Whatman, Kent, UK) were placed into 24 well plates (Corning, Flintshire, UK). 40 μ L of the spiked urine serial dilutions was pipetted onto the membrane and filter paper surfaces. Plates were placed briefly onto a vortex mixer (VWR, Leicestershire, UK) to disperse the liquid bead throughout the media.

Sample plates were loaded into the Tecan instrument and a specialised script, adapted for 24 well plate analysis, ran. For power optimisation, the additional sample plate was analysed with power setting 1-4.5 W, increasing in 0.5 W increments, for the two media types. Continuous and pulsed firing mode, along with super pulse mode, was selected and the laser lens assembly was set to 2 mm distance from the media surface.

Plate counts were performed the following day on the appropriate dilution media to determine CFU/mL of the samples analysed.

2.10.3 Clinical Urine I: 96 Well Plate Pellet Analysis

A total of 560 clinical urine samples were obtained from the ICHNT microbiology laboratory over the course of one week, following progression through the standard lab workflow, which includes initial urine sediment examination screening (sediMAX, Menarini Diagnostics, Winnersh, UK), manual microscopy and differential culturing followed by identification via MALDI-ToF system if required. A mixture of GP referral and ward acquired samples were obtained, including culture negative, culture positive, no significant growth and healthy donor samples. Samples were determined to be culture positive, negative or “no significant growth” by examining NHS laboratory reports available for each sample. Samples reported as “no significant growth”, or those containing mixed growth, were excluded from the dataset to remove ambiguity on infection status.

Samples were aliquoted into 96 deep well plates and analysed using the LA-REIMS setup, with a custom made deep well plate aspiration arm and focusing lens (Waters Research Centre, Budapest, Hungary) on the day they were collected. A 2 mL volume of each sample was pipetted into plate wells before centrifugation at $3486 \times g$ for 10 minutes. Sample supernatant was discarded and the remaining pellet was dried as much as possible by removing residual liquid with a micropipette. Pellet plates were analysed with a custom Tecan script, in both continuous and pulsed firing modes as well as the super pulse option, at 1.5-2 W.

2.10.4 Clinical Urine II: Microcentrifuge Tube Pellet Analysis

A second batch of 288 clinical urine samples was collected in an identical manner as before, with the same exclusion rules and distribution. Samples were aliquoted into 2 mL microcentrifuge tubes and centrifuged at $21\,500 \times g$ for 10 minutes, before the supernatant was discarded. Tubes were briefly vortexed to position the pellet in the centre of the tube, and were placed in a rack within the Tecan instrument. LA-REIMS analysis was performed as previously.

2.10.5 Tandem MS Analysis

MS/MS was carried out on significant biomarkers identified from clinical urine samples containing *E. coli*. Biomarkers were determined as significant based on univariate analysis techniques outlined in the MetaboAnalyst workflow below. Importance in class variation between groups in partial least squares discriminant analysis and random forest analysis were also used to determine significant biomarkers. To carry out MS/MS, urine pellets from the previous experiment were thawed and analysed as before, with LA-REIMS. Target biomarker m/z was entered into quadrupole mass filter settings and a collision energy of 35 eV was selected. Optimal collision energy was determined by adjusting the value until the relative intensity of the precursor ion was approximately 1/3 of the highest intensity product ion. Argon gas was used as the supply to the collision chamber. MS/MS spectra were acquired in triplicate for three samples containing *E. coli* as the causative agent, for the top five significant biomarkers. Full instrument parameters are available in supplementary materials (Supplementary Table 1).

2.11 AMI-MS

A method comparison study was carried out to determine how LA-REIMS compares with other MS methods in small metabolite detection. Phenotype microarray plates, containing proprietary media in 96 well format supplemented with 192 different carbon sources (BIOLOG, California, US) (Supplementary 15), were utilised to culture a single isolate of *K. oxytoca*.

For LA-REIMS, a 40 μ L aliquot of each sample was spotted onto grade 595 qualitative filter paper of 4-7 μ m pore size (Whatman, Kent, UK) and placed within 24 well plates (Corning, Flintshire, UK). Samples were analysed at 4.5 W, continuous firing mode with super pulse activated. An additional LA-REIMS analysis was carried out, where samples were extracted from the plates and centrifuged in 2 mL microcentrifuge tubes at 21 500 $\times g$ for 10 minutes, and supernatant discarded. Pellets were analysed at 3.5 W, pulsed fire with super pulse mode activated.

Samples were prepared for AMI-MS (AstraZeneca, Macclesfield, UK) by aliquoting 3 μL of each whole sample into 50 μL 100% methanol (VWR, Leicestershire, UK) within a 384 well plate. Samples were analysed on the ultrasonic platform with a 0.5 second pulse time. A second batch of samples suspended in 50 μL water/0.1% acetic acid was also analysed. Non-fragmentation target enhancement (ToF-MRM) was enabled with precursor ions at 150, 300, 500 and 750 m/z to improve ion sensitivity in the four distinct mass ranges. Following acquisition, the four mass ranges were combined for each sample using a custom script, taking higher intensity ions over low intensity ions for overlapping mass regions.

For all methods, the 50-2500 m/z range was acquired, although 50-300 m/z was inspected with particular interest for analysis of small metabolites.

2.12 Data Analysis Workflow:

Data for each experiment was generated as raw files by MassLynx software. Raw files were transferred to an external hard drive at the end of each day for processing.

2.12.1 Pre-processing

For manual inspection of spectra, raw files were pre-processed in MassLynx using the accurate mass measure function from the tool page. Input files were loaded and automatic peak detection parameters set to automatic, with lock mass correction enabled for the leu-enk compound at 554.26 m/z averaged over 10 scans, with a mass window of ± 0.3 . Processed raw files were saved with a “_Processed” file name suffix and opened within MassLynx for inspection.

2.12.2 Offline Model Builder

Offline Model Builder (OMB) is a compilation of custom data processing scripts and graphical user interface written by the Waters Research Centre. Raw data was imported to the programme and metadata labels added for each file, imported from a Microsoft Excel spreadsheet. Files were selected and total ion current (TIC) chromatogram inspected for each sample. Depending on sample format during analysis, TIC sections

were selected and labelled individually or left as wholly selected if samples were saved as distinct files. Pre-processing of the samples by lock mass correction, using the same parameters as with MassLynx preprocessing, and background subtraction was carried out before exporting the data matrix. Mass binning was typically carried out at 0.1Da, over the 600-1000 m/z lipid range, although the full 50-2500 m/z range was utilised for some optimisation experiments and a 1Da bin size was required for the spectrum quality analyser tool.

2.12.3 MetaboAnalyst

Data processing, normalisation and statistical analysis was performed in the cloud based analysis tool suite MetaboAnalyst 3.0 (217). Before data matrices were uploaded to the site, CSV files exported from OMB were opened in Microsoft Excel and manually formatted by adding a mass bin header row and sample number column before being resaved. The formatted CSV files were uploaded into the statistical analysis MetaboAnalyst pipeline. Interquartile range data filtering was carried out on sample sets containing high similarities, such as in clinical urine containing many host features, or in mastitis milk samples with highly similar lipid profiles. Log transformation and Pareto scaling were also carried out in most cases to reduce the relative importance of large fold changes within the dataset.

In multivariate analysis, unsupervised Principal Component Analysis (PCA) was used as an initial test on all datasets, together with Partial Least Squares – Discriminant Analysis (PLS-DA) as a supervised approach, for distinguishing of discrete groupings within the data. For these, principal components one and two were plotted along with 95% confidence regions and Q2 scores recorded. Cluster analysis heatmaps of the top 25 most significant features, from t-test, were generated using MetaboAnalyst to assess overall feature profile similarities. Univariate analysis techniques, including Fold Change Analysis, t-tests and Analysis of Variance (ANOVA), along with Fisher's Least Significant Difference post-hoc test, were used in significant biomarker discovery investigations. Spectral mass bins with a false discovery rate (FDR) corrected P value of < 0.05 were accepted as significant biomarkers. Box plots were generated for each significant spectral feature to aid in group overlap consideration. For analysis of two

groups only, volcano plots displaying fold change against FDR corrected *P* value from t-tests were utilised to aid in significant feature selection. Box plots were also generated in this case. ANOVA and t-tests data was exported to Microsoft Excel in tabular form for further manual inspection, processing and creation of charts.

The biomarker analysis tool suite within MetaboAnalyst was utilised for exploratory multivariate receiver operator characteristic (ROC) analysis of significant biomarkers. Monte-Carlo cross validation (MCCV) was used to create ROC curves, with biomarker importance assessed by 2/3 of the sample set in repeated sub-sampling. The top 100 most important biomarkers were utilised in a random forest classification model validated with 1/3 of the sample subset left out previously. Features were ranked according to random forest decrease in accuracy. Non-parametric re-sampling was utilised to generate 95% confidence intervals.

2.12.4 Spectral Richness Tool

A custom programme script and graphical user interface created in-house was used to examine the spectral abundance after background subtraction, using Shannon-Wiener diversity index. Shannon index scores increase along with evenness and richness of data communities, providing a graphical summary of richness spread. Data matrices exported from OMB as CSV files were inputted to the custom script and a series of box plots generated demonstrating spectral abundance and evenness of each matrix.

2.12.5 Spectrum Quality Analyser Tool

A second custom programme script, created by WRC, was provided to determine spectral quality by generating signal:noise ratio scores for each sample within a dataset. The tool was used to determine and eliminate poor quality data from datasets before model building, or for comparing different experimental methodologies. Raw, unprocessed data matrices were exported from OMB and were opened within the programme interface, and files selected for analysis. Quality scores were generated using the range vs background method, where a distinct mass range and background region were inputted. Typically, the lipid region of 500-1000 *m/z* was used as the range

of interest and a higher mass region of 1600-2000 m/z used for background. The leu-enk lock mass compound bin, at 554.26 m/z , was inputted to the peak exclusion list so as to not interfere with the quality calculation. Quality scores are calculated by taking the quotient of the median intensity of the background peaks and the sum of the most intensive 20 peaks within the signal range. Only those peaks three times the intensity of the background median are used for calculation. If no peaks are found at this intensity level, an error message is displayed for the file and it is excluded from the final matrix. Scores were exported to a CSV file and exported to Microsoft Excel for chart creation.

2.12.6 Biomarker Identification

The accurate mass of biomarkers was calculated by performing an accurate mass measure in MassLynx, and pre-processing through background subtraction and lock-mass correction. Biomarker accurate mass was referenced to the MicrobeID list of known biomarkers. If a match was not found, a tentative identification was conducted by carrying out a literature search on the exact mass of each biomarker. A LipidMaps(218) glycerophospholipid precursor/product ion database search was conducted by entering precursor and product ion m/z values along with their associated intensity. An intensity threshold of five was used, with an ion mass tolerance of ± 0.5 m/z . Database monoisotopic mass hits were manually inspected for total number of product ion matches, mass deviation and glycerophospholipid class. A Metlin MS/MS metabolite database(219) search was performed using the monoisotopic mass of the top two matches from LipidMaps, with a 15ppm tolerance, accounting for common adducts in REIMS. Experimental database hits were preferred over *in silico*. Fragmentation patterns for each likely hit were examined and compared to the MS/MS spectra acquired with REIMS. Biomarker identities were assigned according to both database matches and by referencing MS/MS ion fragmentation patterns.

3 Optimisation and Setup

Early development and optimisation work on REIMS in bacterial analysis has been previously reported (220). Basic experimental parameters from this early work, as well as more recent studies, were utilised throughout this project, with optimisation experiments conducted for new sample types and analysis methods as required. Basic parameters include instrument voltages, nebulising solvent, flow rates, lock-mass compound infusion, interface configuration and instrument ionisation mode. The analysis mass range was also treated as a basic experimental parameter, although a wider range was acquired and included in data analysis pipelines in the analysis of novel samples types, including milk and urine. The initial optimisation and general, early-stage method development for novel analysis techniques are presented in this section. Specific experiments utilising these techniques, and further optimisations, are presented in Sections 4-7.

3.1 Cell Line Probes

The 14 modified cell line probes, obtained from WRC, differed by a range of steel alloy electrode materials, geometry and electrode spacer configurations (Figure 12). Yellow probes consisted of no insulating sheath or electrode spacer, with electrode material of tactile 304 annealed stainless steel. The three yellow probe types differed by electrode lengths, with the short probe measuring 20 mm, medium probe 22.5 mm and long probe 25mm. Blue probe types all consisted of a 23 mm electrode along with a short insulating sheath, differing by electrode material (Figure 12). A long insulating sheath was present on green probes, along with an electrode spacer to prevent complete contact closure of the electrode tips. Electrode length was 24 mm, with varying alloy compositions.

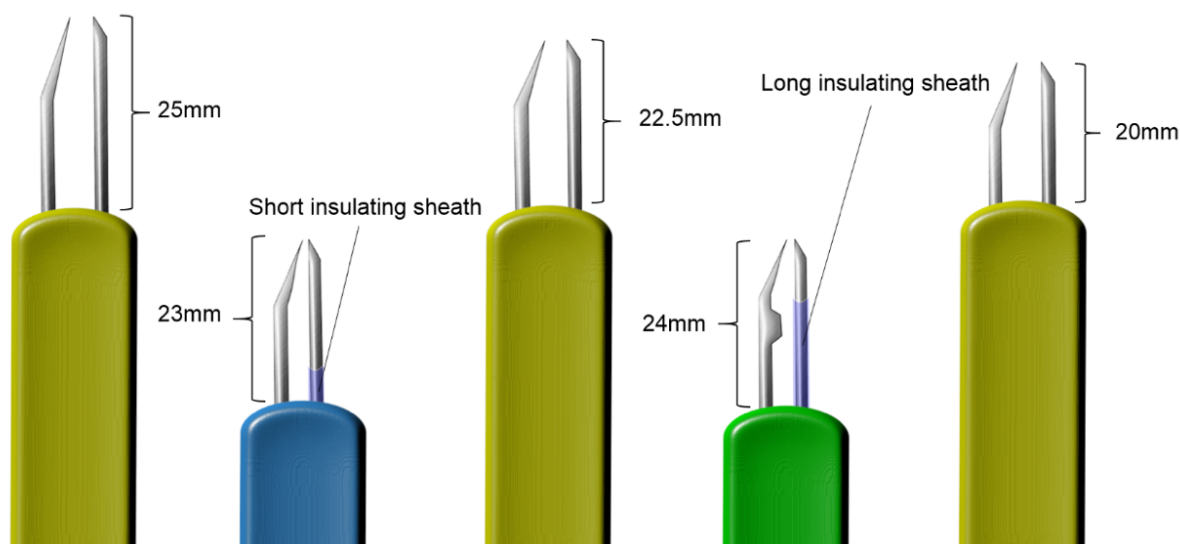


Figure 12. Modified cell line probes obtained from Waters Research Centre, Hungary, consisting of various electrode geometries, alloy compositions, electrode lengths, insulation sheaths and electrode spacers.

Analysis of an *E. coli* cell pellet with each probe at varying wattages revealed that the 25 mm yellow probe, stainless steel (half hard) and stainless steel (full hard) blue probes gave superior signal quality to the other probes (Figure 13). Mass spectral features of the *E. coli* isolate differed significantly between probe types (Figure 14). The presence of the electrode spacer and insulating sheath on green probes prevented adequate contact closure around the biomass for the majority of sample replicates, resulting in generally poor vaporisation of the sample and signal quality. The long electrode length of the 25 mm yellow probe allowed for greater malleability of the electrode tips, ensuring biomass could be more easily manipulated between the contact closure points (Figure 15). High variability of spectral quality between samples was observed, mainly due to the range of different power levels examined, but also due to difficulty in initiating successful vaporisation of some samples, deposit build-up on electrode tips and fluctuating moisture levels within cell pellets (Figure 16).

Amongst the power levels tested, <5 W generally failed to vaporise the sample, or resulted in very poor vaporisation with inadequate volumes of gas phase ions reaching the mass spectrometer. Levels between 5-9 W resulted in high volumes of vapour being

generated and superior spectral richness and quality. A 10 W power level hindered vaporisation by causing significant charring of the sample and high deposit generation on electrode tips, inhibiting current flow through the rest of the biomass.

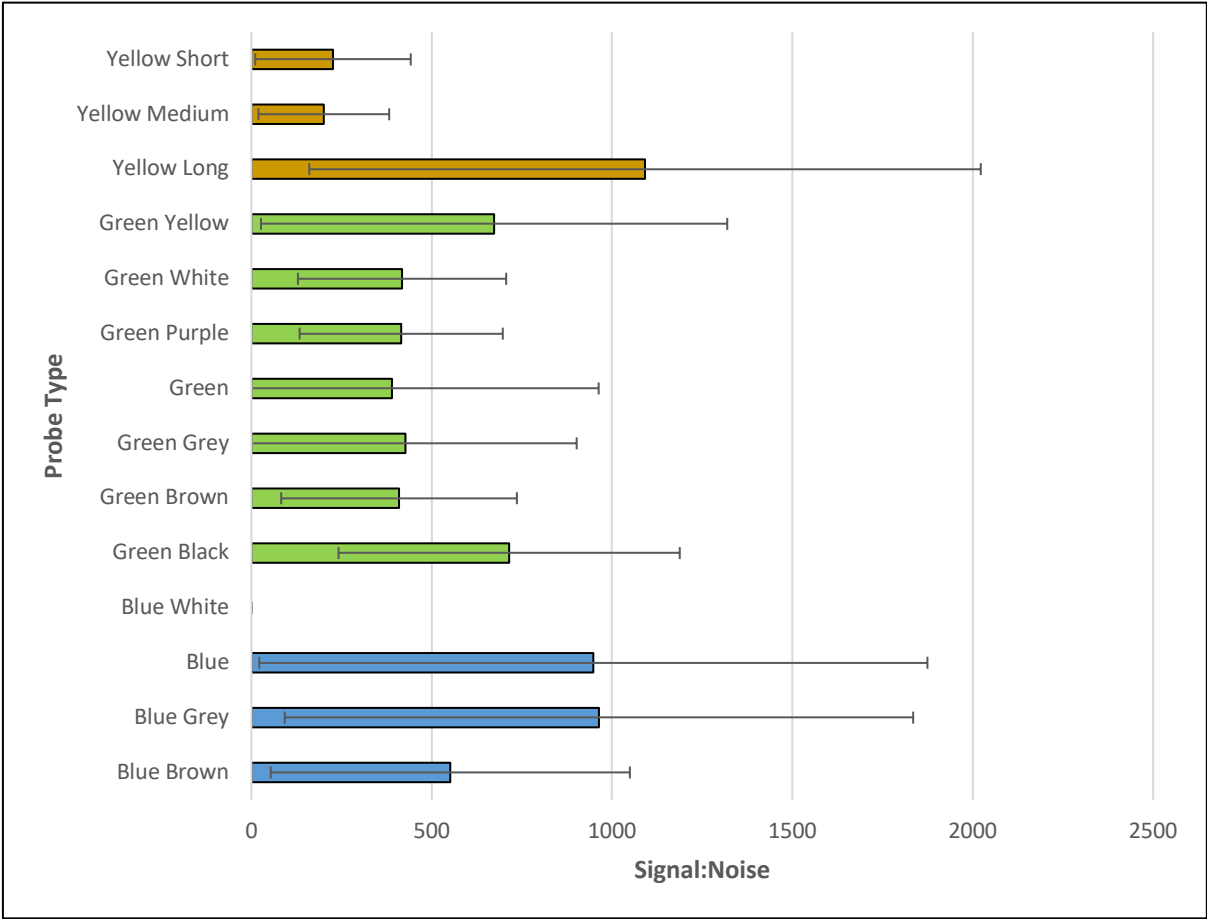


Figure 13. Signal quality, determined by signal:noise scores of spectral bins, of each modified cell line probe. High variability of scores, represented by large standard deviation scores (shown as error bars) for each probe, were observed across all probes due to the range of different power levels investigated.

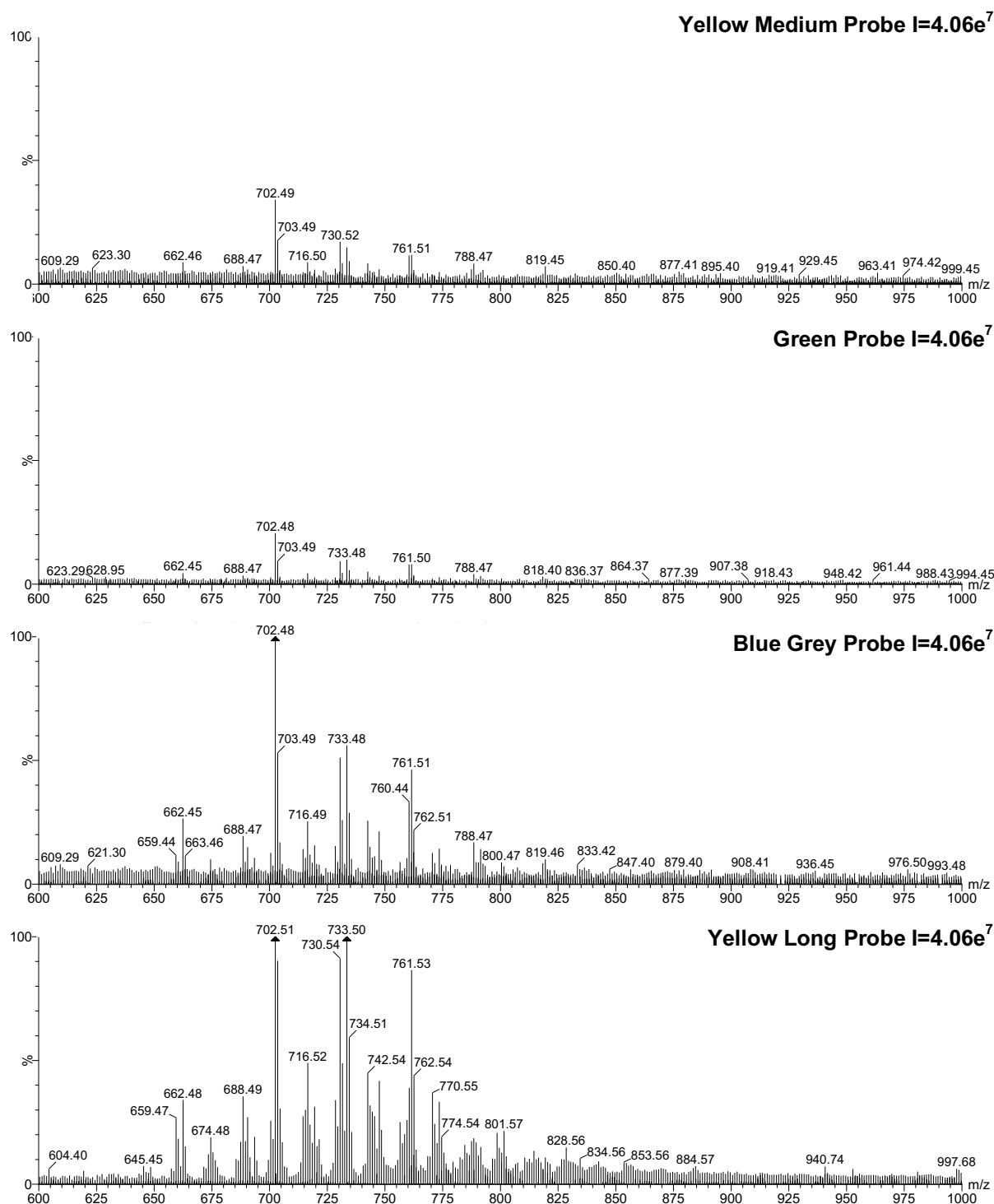


Figure 14. Example mass spectra of the phospholipid 600-1000 m/z range generated by the different cell line probe types at 8 W power supply. Blue-grey (stainless steel (full hard)) and yellow long (25 mm) probes displayed the richest signal. Spectra lockmassed with vertical axis link enabled, range magnification x2.

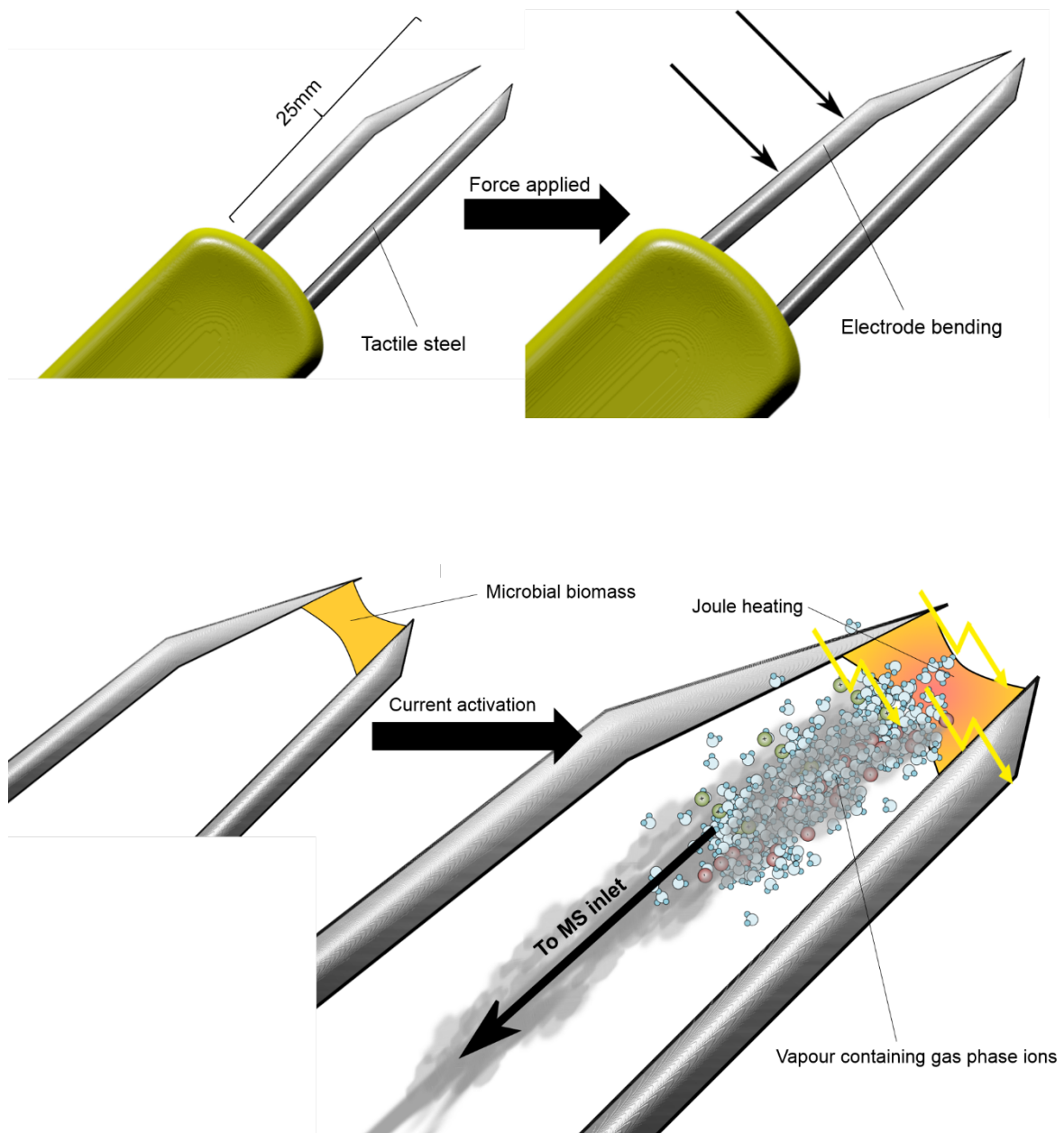


Figure 15. Representation of yellow long (25 mm length) probe during sample analysis. Light force is applied by the user to the bottom electrode to bend the tips closer together. Microbial biomass between the tips is subjected to an electrical current, causing joule heating within the sample and vaporisation. Vapour is channelled to the mass spectrometer using the native vacuum of the instrument.

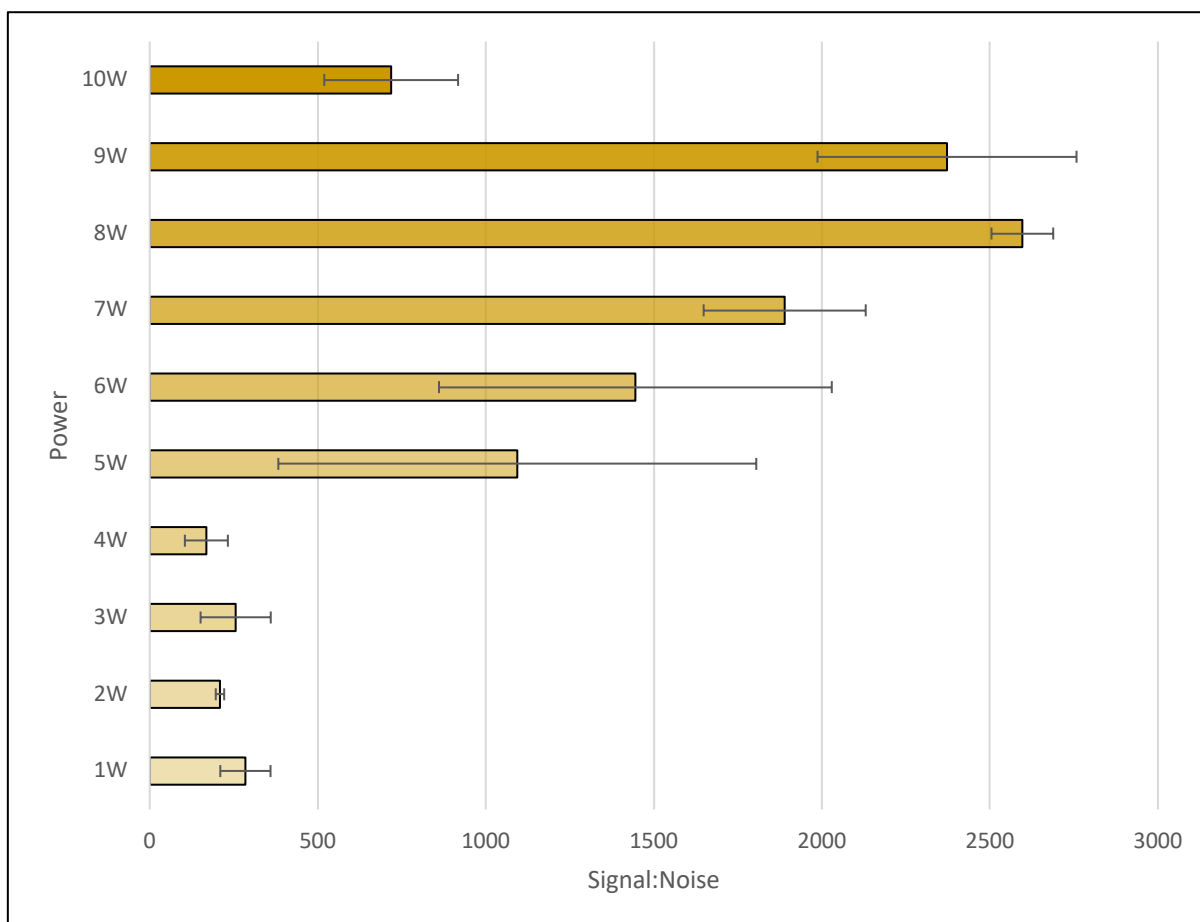


Figure 16. Spectral quality, determined by signal:noise ratio scores, of each power level examined for the yellow long (25 mm electrode length) probe. 8 W was the optimal level for high quality spectra, with a small variation between sample acquisition replicates. Power levels <5 W failed to initialise adequate vaporisation, whilst 10 W caused charring and splattering of the sample. Standard deviation is represented in error bars.

3.1.1 Method Comparison

A method comparison experiment was carried out to compare the spectral richness, quality and speciation efficacy of the optimal cell line probe with the bipolar forceps REIMS on liquid culture samples processed by centrifugation into cell pellets. Results were compared with the established handheld monopolar hand piece (iKnife) and bipolar forceps REIMS methodologies of traditional agar culture plates in order to identify the optimal analysis method for future liquid format samples.

Analysis of cell pellets was not possible with the monopolar hand piece, due to the requirement of the sample being placed onto a return electrode, and the cell pellets being too small. Analysis of culture direct from agar media was not possible with modified cell line probes, as the electrode tips require manipulation to be brought together which is not possible in the absence of a hard surface to press an electrode arm into (such as a microcentrifuge tube containing the cell pellets). Analysis of the cell pellets from liquid cultures was difficult in general, as it was not possible to entirely remove the liquid broth from the pellet resulting in inhibition of sample evaporation. Processed mass spectra from the four analysis techniques, however, were largely comparable in terms of peak numbers in the phospholipid range (Figure 17). Pellet analysis of liquid samples was one-two orders of magnitude lower in signal intensity for bipolar forceps REIMS compared to agar culture analysis, and two-three orders of magnitude lower for modified cell line probe analysis, likely due to the higher moisture content of the cell pellets compared to agar cultures resulting in a higher specific heat capacity. Cell line probe analysis produced mass spectra with a much greater level of background noise than that of bipolar forceps REIMS, due to the lower surface area of the probe electrode compared to the bipolar forceps tips, meaning less biomass could be harvested for vaporisation at each replicate.

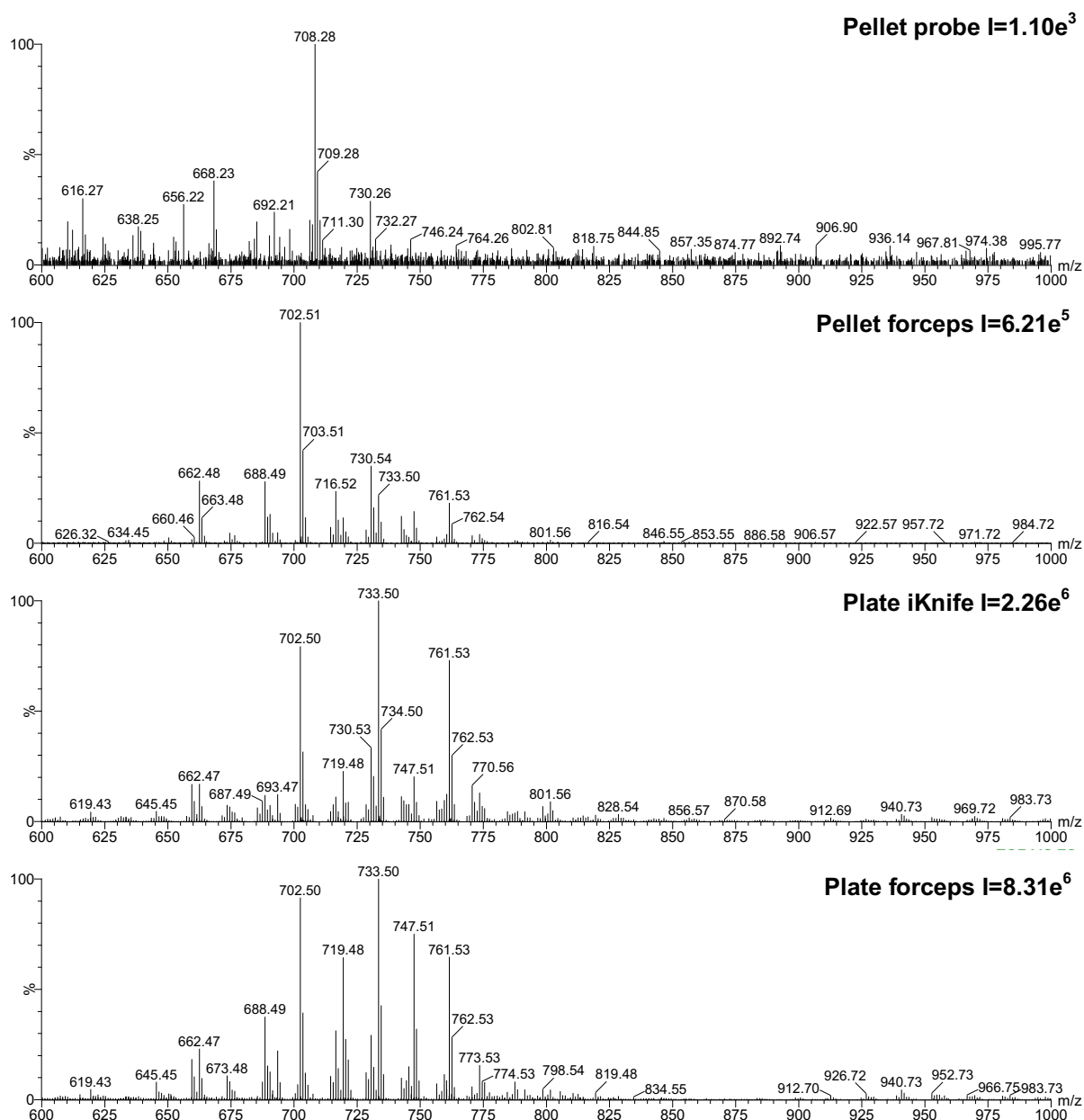


Figure 17. Mass spectra comparison of an identical *Klebsiella pneumoniae* isolate analysed with monopolar iKnife, bipolar forceps and bipolar cell line probe REIMS, as cell pellet from liquid culture and harvested directly from an agar plate. Spectra have been processed for background subtraction and lockmass correction.

Spectral richness of the phospholipid range, represented by Shannon-Wiener diversity index, appeared highest in the cell line probe analysis method, although is shown to be the opposite when only relatively high quality spectra, determined through signal:noise ratio scores, are analysed, as in the case of a *K. pneumoniae* isolate (Figure 18). This indicates the high variability of background noise between samples due to user technique.

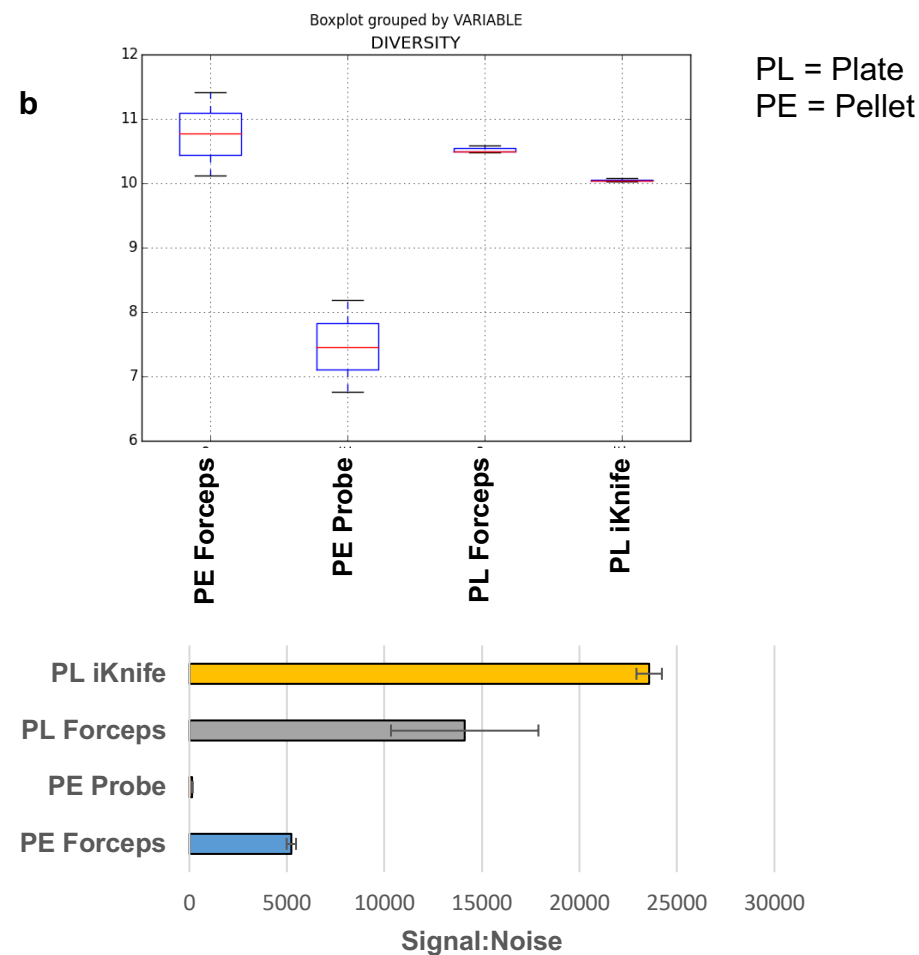
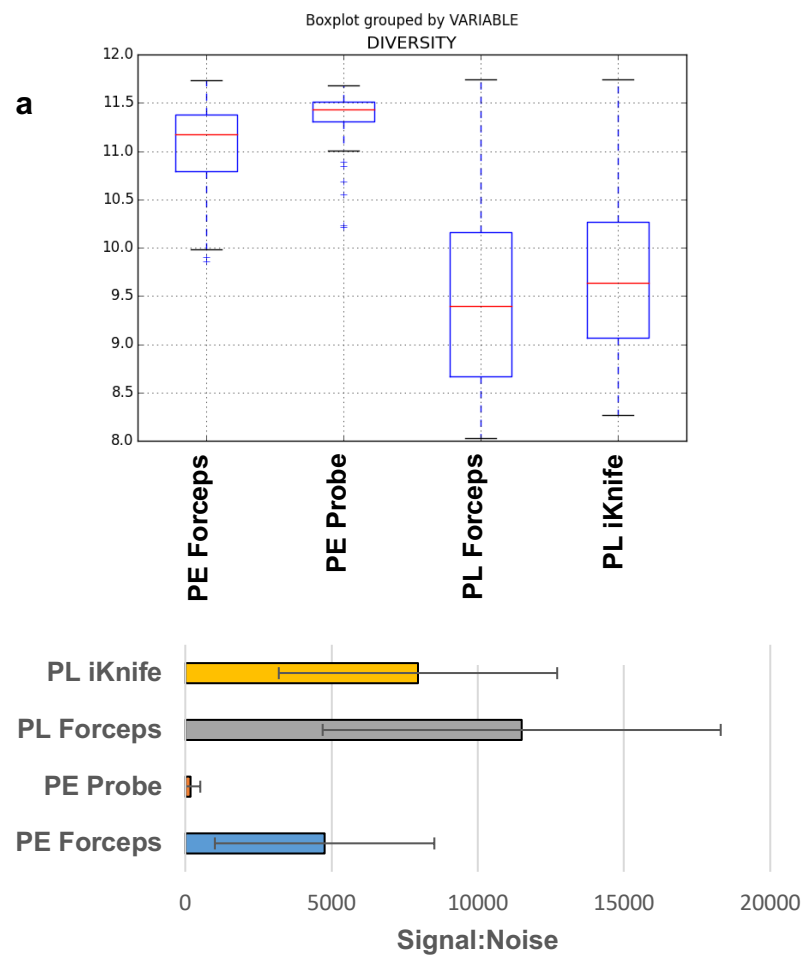
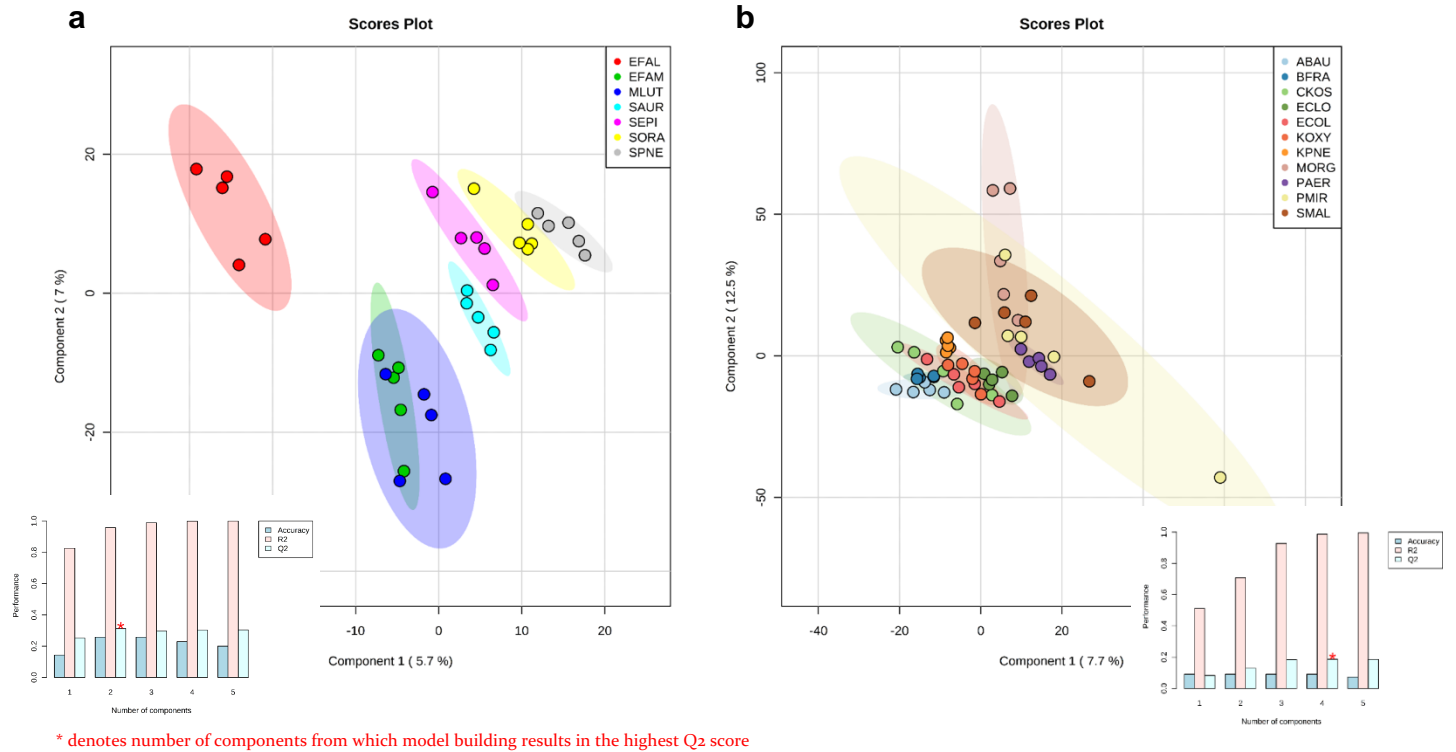


Figure 18. Shannon diversity plot and mean spectral quality scores of different REIMS analysis methods for 20 clinical bacterial isolates (a), and for a single isolate of *Klebsiella pneumoniae* (b). The spectral richness of the phospholipid 600-1000 m/z region, across 20 isolates, appears to be highest in the liquid culture pellet analysed with modified cell line probe REIMS, however this is due to the high level of background noise contained within the spectra, represented by a low quality score. When only considering analysis of the higher spectral quality *K. pneumoniae* isolate, agar plate analysis methods perform better, demonstrating that the low spectral quality for other isolates contributed to the high richness score in the previous test.

Supervised discriminatory analysis of all bacterial species sampled with cell line probe REIMS failed to achieve strong groupings of sample replicates or separation of distinct species (Supplementary 3), however samples grouped according to their Gram status in separate data matrices performed well in grouping and separation (Figure 19). Therefore the larger biomarker changes observed between Gram status may be masking the smaller inter-species changes, causing models including all isolates to perform poorly. A similar result was seen in the bipolar forceps analysis of cell pellets (Figure 19). Models including bipolar forceps and iKnife analysis of agar plates showed good clustering of replicates and strong separation of species in both unsupervised and supervised analysis (Supplementary 4, Supplementary 5). Superior discriminatory performance of samples analysed from agar media is expected as forceps and iKnife probe analysis parameters, such as geometry, supply voltage and instrument conditions, have previously been developed and optimised by both Waters and the MicrobeID group for this purpose. Liquid analysis, however, has not been previously possible with these methods, hence the need for development of an alternative methodology. Cell line probe analysis of pellets is preferred over bipolar forceps analysis due to the observed discriminatory power.

Cell line probe REIMS



Bipolar forceps REIMS

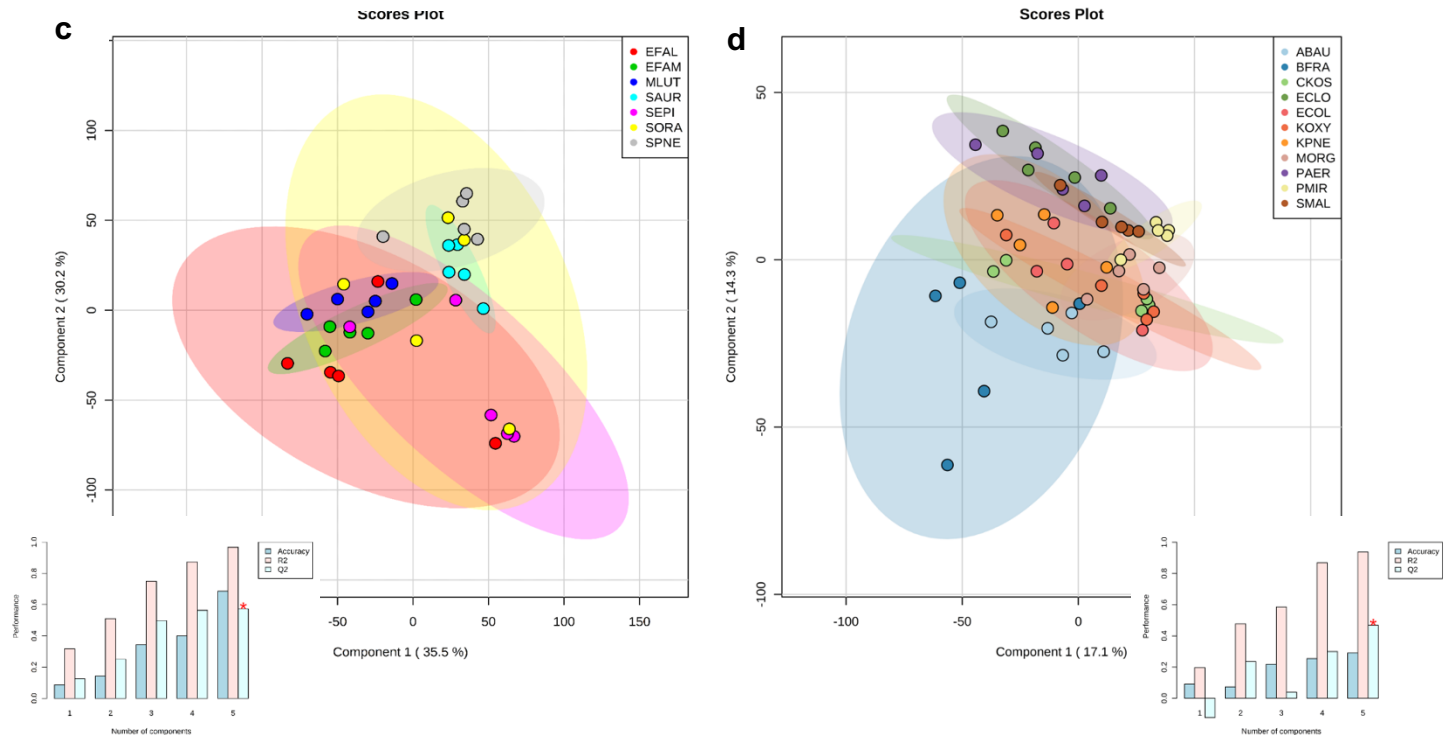


Figure 19. Partial least squares discriminant analysis of Gram-positive only (a) and Gram-negative only (b) bacterial species replicates for cell line probe REIMS pellet analysis. Strong clustering of species replicates is seen with good separation. In bipolar forceps pellet REIMS partial least squares discriminant analysis of Gram-positive only (c) and Gram-negative only (d) isolates some clustering of species replicates is seen, but with poor separation. Overall, cell line probe REIMS (a-b) achieves superior discrimination of species and is the preferred method for analysis of cell pellets from liquid culture.

3.2 Laser Optimisation

A laser system was developed and optimised with WRC alongside optimisation of modified cell line probes. Optimisation was conducted using laser-assisted REIMS to investigate the system potential for direct liquid analysis. Bovine milk was initially utilised for this purpose as a collaboration was set up to investigate the potential of REIMS as a bovine mastitis screening tool. Additionally, the high lipid content of milk offered an ideal test substrate for optimising parameters for phospholipid analysis of bacterial cultures.

Milk lipid layers analysed with a lens Z max value of 1820, or a 2 mm lens distance from the sample surface, resulted in no signal for the majority of samples (Supplementary 6). Spectral quality scores of all combined spectra from the 1815 Z max setting, or 2.5 mm from the sample surface, suggest 3.5 W is the optimal laser power (Figure 20), and resulted in good signal intensity for the majority of laser powers tested (Figure 21). A lens distance of 2 mm from the sample surface corresponds with a laser beam diameter of 485 μm , whilst 2.5 mm from the sample surface corresponds with a beam diameter of 520 μm . A more dispersed beam reaching the sample surface from the 2.5 mm distance setting therefore results in higher vaporisation efficacy. Powers above 3.5 W were not tested to avoid overheating and resultant splattering of the sample. A large range of quality scores was observed between each sample replicate for single power ratings, attributed to slightly different volumes of milk pipetted into each plate well resulting in varying distances between lipid layer surfaces and the laser lens.

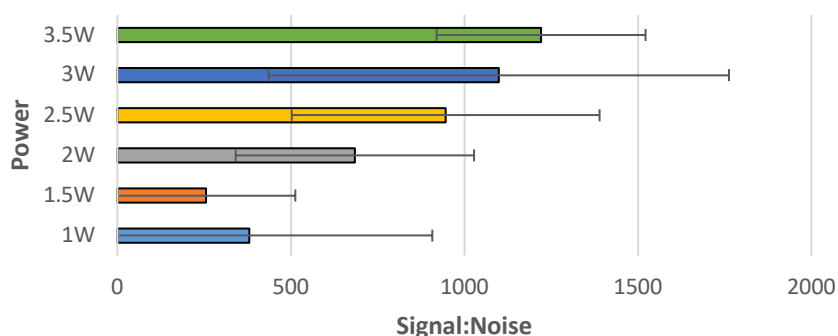


Figure 20. Mean combined spectral quality scores, of the 600-1000 m/z region, of each bovine milk sample replicate analysed at increasing laser supply powers (2.5 mm lens distance from sample surface). Variability in spectral quality between each sample replicate is attributed to slightly different milk volumes in plate wells, resulting in distance differences from the laser lens.

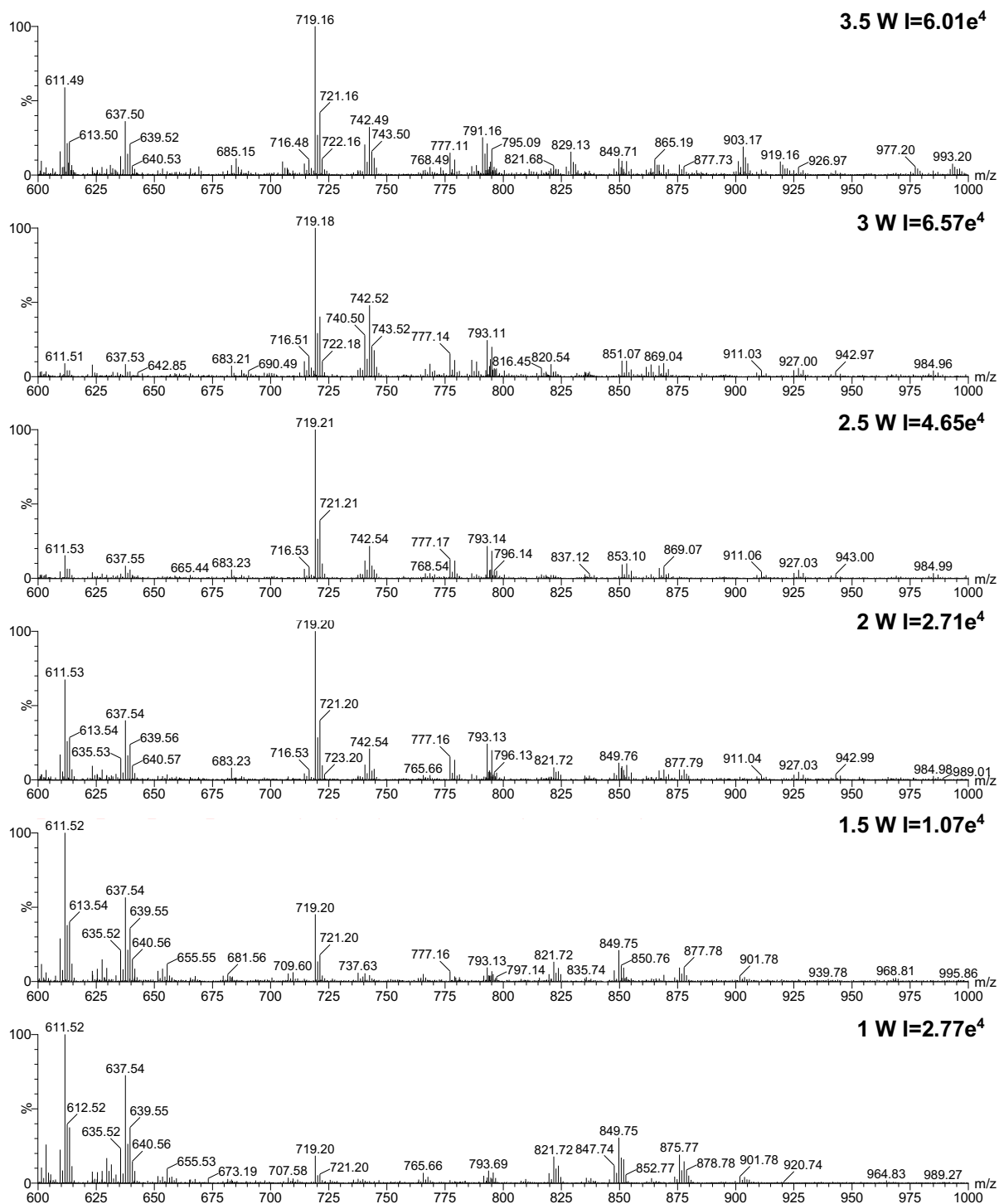


Figure 21. Mass spectra of the 600-1000 m/z range produced from LA-REIMS with a lens distance of 2.5 mm from bovine milk lipid layer surface. Background subtraction and lock mass correction applied. Good signal intensity was achieved in all laser powers tested.

3.3 Summary and Conclusions

Comparison of discriminatory analysis from the four REIMS analysis methods shows that bipolar forceps and iKnife REIMS perform best in speciation with solid agar cultures, with robust model performance scores. Newly developed cell line probe REIMS and bipolar forceps REIMS sampling from liquid culture shows comparable models, although cell line probe REIMS is preferred for pellet analysis due to a simpler user technique and slightly better discriminatory power. The already established and previously optimised bipolar forceps REIMS and iKnife REIMS analysis outperformed the new methodologies in discriminatory analysis, as expected. This early optimisation and development experiment demonstrates that acquisition of mass spectra containing a degree of spectral richness in the 600-1000 m/z phospholipid region has been achieved, therefore determining that analysis of liquid samples, via the centrifugation and pellet sampling methodology investigated here, is feasible. This technique can be further optimised and developed for specific sample types, and is investigated in Sections 5: Urine analysis, and 7: Bovine Mastitis.

Overall LA-REIMS setup and optimisation was conducted as part of the wider MicrobeID group project and is outside the scope of this thesis, however the specific optimisation and method development into liquid sample analysis conducted here shows that vaporisation of milk samples is possible. The automated nature of LA-REIMS analysis and the robotic Tecan setup means that optimisation will be conducted prior to analysis with each sample type. Further LA-REIMS method development is explored in Section 5: Urine analysis. The hypothesis that the applications of REIMS can be expanded to culture free microbial analysis of clinical samples has been supported here through the initial exploration and method comparison of liquid analysis techniques.

4 Mixed culture analysis

REIMS has thus far been utilised in the analysis of pure bacterial cultures, with little investigation into samples containing more than one species of bacteria. This experiment was conducted in order to determine the presence and/or extent of any suppression effects caused to the intensity of species biomarkers in samples containing two distinct bacterial species, in order to serve as a proof of principle in REIMS of analysing poly-microbial clinical samples. The four bacterial isolates used in this investigation were chosen based on their clinical significance, specifically in urinary tract infections, their taxonomic range, and high incidence rate within the clinical microbiology NHS laboratory. Good separation and sample class clustering of each pure species was achieved through unsupervised principal component analysis of the pure samples, as shown in previous studies, demonstrating that REIMS is capable of distinguishing between the pure sample isolates tested based on their phospholipid spectral signal within the 600-900 m/z mass range (Figure 22)(195). REIMS analysis was carried out on bacterial mixtures, created within the laboratory by manually mixing pure isolate biomass in varying ratios. Synthetic mass spectra of culture mixtures were also generated *in silico* for the four bacterial isolates, using the data obtained from analysis of the pure cultures. Correlation analysis between the *in situ* and *in silico* generated m/z bins shows very strong correlations, typically above 0.8, between expected and observed relative intensities (Figure 23).

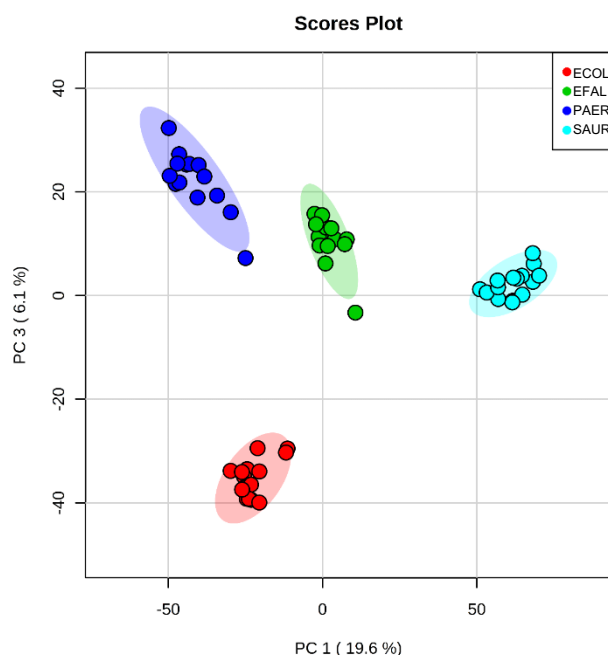


Figure 22. 2-dimensional unsupervised principal component analysis of the 600-900 m/z mass range of pure culture isolates. Data was log transformed and subjected to Pareto scaling before plotting of components. Good separation was achieved of each pure isolate, demonstrating that REIMS can successfully distinguish between the pure cultures analysed based on their phospholipid mass spectral signal.

4.1 Species-Specific Biomarkers

When comparing pure isolate spectra with that of the same isolates in mixed cultures, isolate specific peaks can be consistently and readily seen by eye, even before any specific data analysis is carried out. For example, in the case of *P. aeruginosa* and *S. aureus* culture mixture in a 1:1 ratio, visible *S. aureus* specific peaks include 693.47 and 735.52 m/z , and visible *P. aeruginosa* specific peaks include 673.48 and 716.52 m/z (Figure 24). Univariate ANOVA tests were completed between the three isolates of each of the four bacterial species previously described, using MetaboAnalyst 3.0 statistical tool suite (217). Using a false discovery corrected $P < 0.05$ threshold, a total of 129, 892, 119, and 182 m/z bins were identified as significantly higher in relative intensities for *E. coli*, *E. faecalis*, *P. aeruginosa*, and *S. aureus* respectively, within the m/z range of 600 to 900 at 0.1 Da bin sizes. To determine the utility of these m/z as species-specific biomarkers in controlled bacterial mixes, the relative intensity of the top ten, based on FDR corrected P value, was measured for each relevant bacterial mixture after mass

spectra had undergone background removal, mass drift correction, TIC normalisation, and mass binning to 0.1 Da. Only those features which met these criteria were included in the analysis. Biomarkers identified as species-specific in pure cultures were still able to be detected in the bacterial mixtures (Supplementary 7). The relative intensity of the top ranked biomarkers, based on FDR corrected *P* values, varied throughout isolate concentration in mixtures, and tended to reduce in intensity with reducing concentration of the isolate within a mixture (Figure 25).

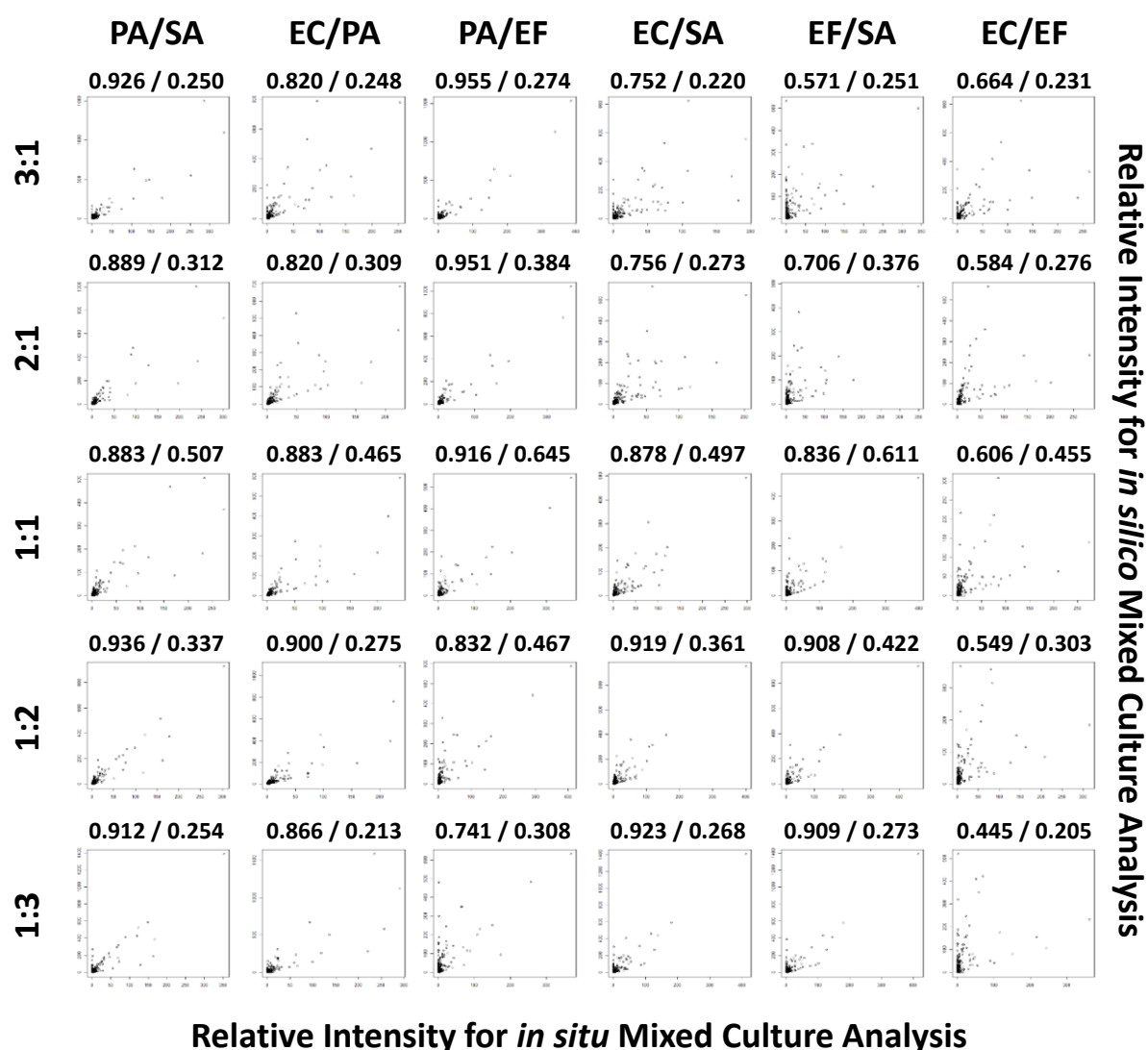


Figure 23. Comparison of *in silico* and *in situ* mixed culture spectral data. High correlations between intensities of actual and predicted mass peaks are observed, but the overall signal intensity for *in situ* mixed cultures is suppressed. Row headers indicate ratios of bacterial mixtures. Left-hand figure above each graph indicates correlations between relative intensities of *in situ* and *in silico* mixtures, and right-hand figure indicates fold change difference between relative intensities of *in silico* and *in situ* mixtures. EC = *E. coli*; PA = *P. aeruginosa*; SA = *S. aureus*; EF = *E. faecalis*.

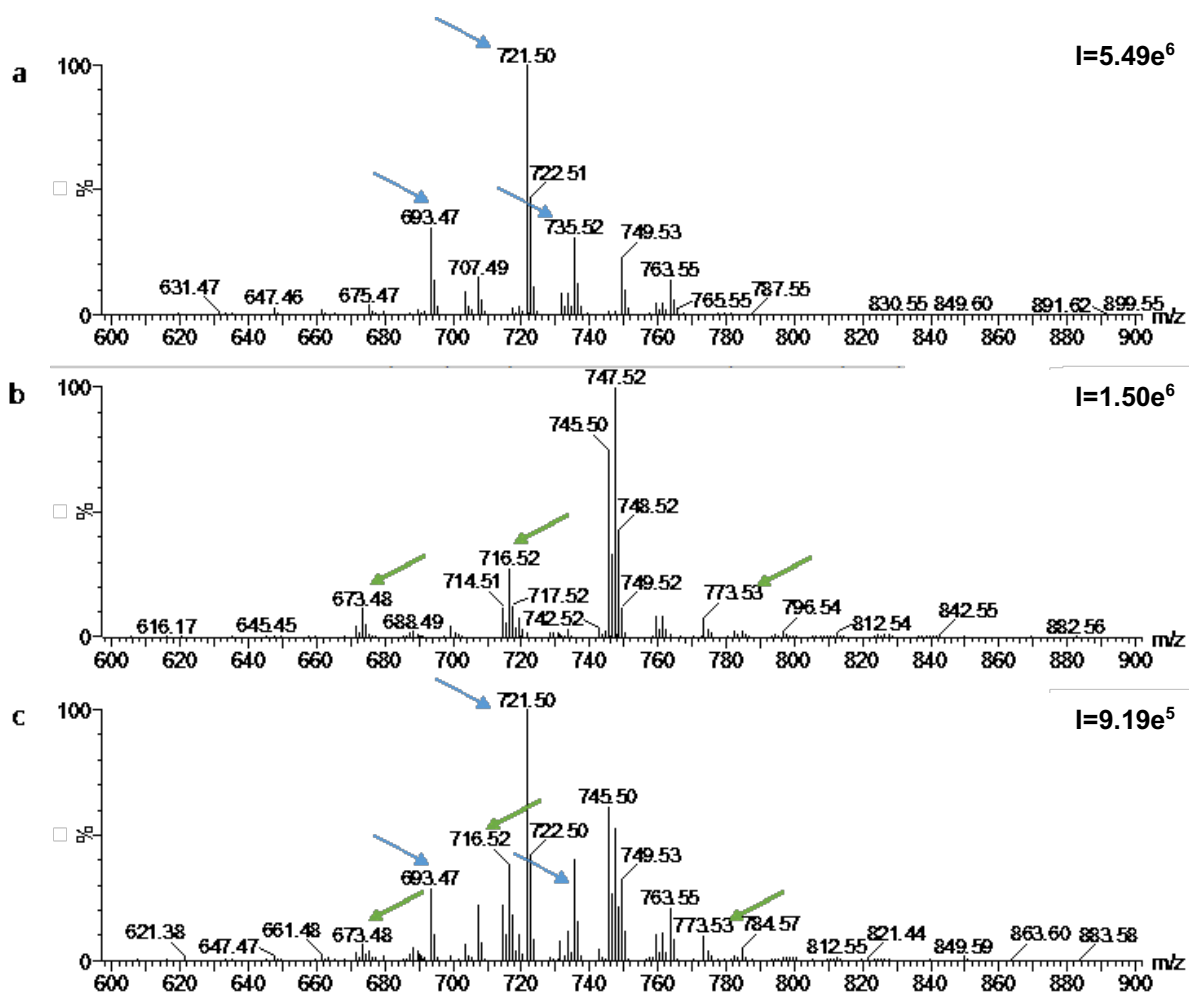


Figure 24. Mass spectra comparison of pure isolates *Staphylococcus aureus* (a) and *Pseudomonas aeruginosa* (b) along with a mixed culture mass spectra of the 2 isolates in a 1:1 ratio (c). *S. aureus* specific peaks (blue arrows) are visible in the mixed culture spectra along with *P. aeruginosa* specific peaks (green arrows), demonstrating that species-specific peaks are not subject to significant suppression effects in the mixtures analysed. Spectra are lock mass corrected and background subtracted.

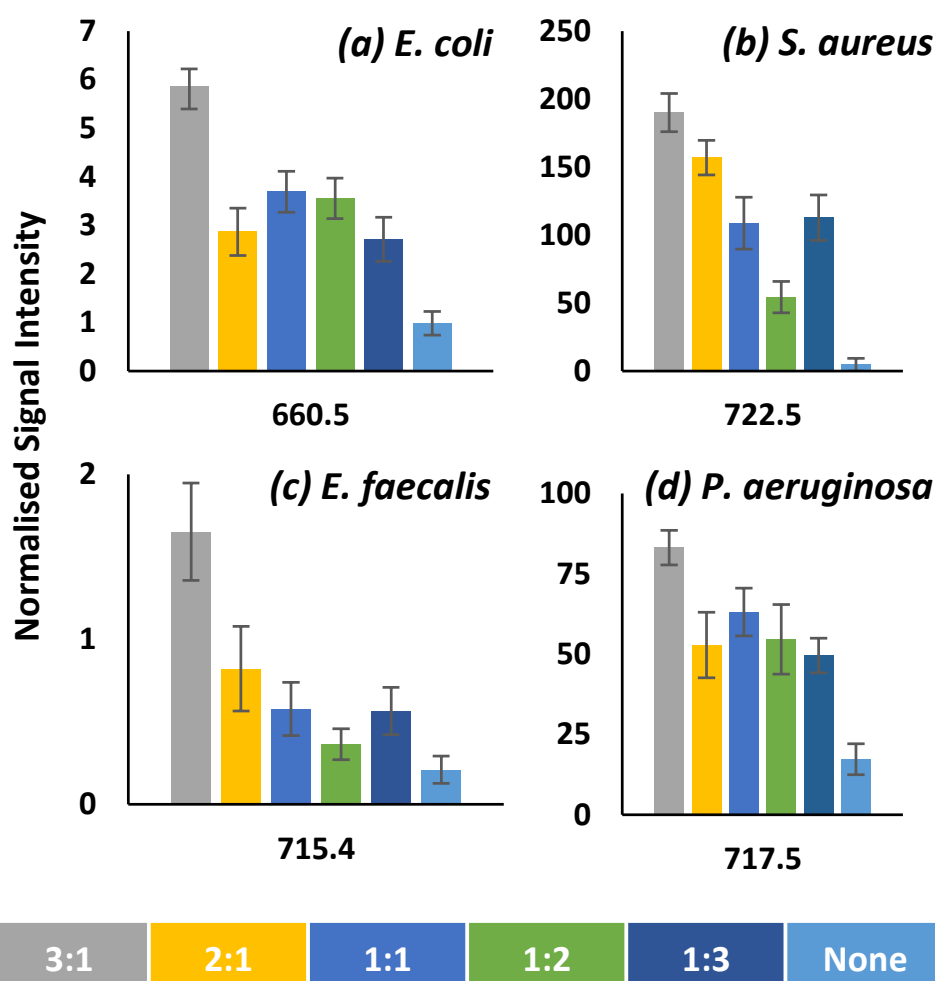


Figure 25. Top ranked species-specific biomarker, based on FDR corrected P value, for each bacterial species within mixed cultures (*Escherichia coli*, *Staphylococcus aureus*, *Enterococcus faecalis* and *Pseudomonas aeruginosa*). Significantly higher intensity levels are seen in all five ratio mixes containing the target species. The top ten ranked species-specific biomarkers are shown in Supplementary 5.

4.2 Discussion

The identification of species-specific biomarkers of four clinically significant bacterial isolates, and their detection in controlled mixed culture ratios with REIMS without any significant intensity suppression effects, indicates REIMS may be capable of analysing clinical samples without pre-culture. This supports the hypothesis that the scope of REIMS can be expanded to clinical samples, which along with potentially possessing more than one causative agent are likely to contain bacterial cells from the native host microbiota. In particular, for a direct analysis method of clinical urine samples to succeed, any analysis technique must be capable of detecting pathogenic biomarkers

above the baseline of biomarkers exhibited by *Lactobacillus*, *Corynebacterium*, *Streptococcus*, *Actinomyces* and *Staphylococcus* spp. which, together with 30 additional bacterial genera, constitute the commonly resident microbiota in the bladder and urine of healthy females (221). Although the levels of bacterial cells in the urinary microbiota are of a low concentration, and almost certainly less than the 10^5 CFU/mL level widely accepted as the threshold of a UTI, sensitive direct urine analysis techniques may be capable of detecting biomarkers from these organisms even at such low levels (222).

Some investigations into the urinary tract microbiota have shown cell loads of up to 10^3 CFU/mL, which is below the threshold of a positive UTI but above the level of clinical significance, as specific uropathogens at levels of 10^2 CFU/mL have been shown to be indicative of a symptomatic bladder infection (223). Therefore, a system designed to analyse urine directly should ideally be able to detect bacterial biomarkers at cell loads of 10^2 CFU/mL and account for the background spectral peaks of the microbiota at this level. The limit of detection for REIMS is investigated in Section 5: Urine analysis.

The hypothesis that REIMS is capable of detecting species-specific biomarkers within poly-microbial samples is also supported through this experiment, as this was achieved in controlled mixed cultures at varying ratios. Strong correlations were observed in the relative intensities of *in silico* and *in situ* culture mixture mass spectra (Figure 23), meaning that the rank of signal intensities for a chosen mass bin is highly similar between the *in silico* and *in situ* mass spectra. However, mixed cultures containing high ratios of *E. faecalis* showed typically lower correlations, at approximately 0.6 to 0.7, compared to the >0.8 observed for other mixtures. This may be due to incompletely homogenised mixtures or sampling from inadequate surface area of mixtures, and would likely be improved with a larger sample set and more homogenous method of mixing. For all mixtures, although the correlation between the relative intensity ranks of each *m/z* bin was high, *in silico* relative intensities were typically two to four fold higher than what was observed in the *in situ* mixed culture mass spectra. However, the high level of correlation between the two mass spectra types suggests that the mixing of bacterial cultures does not lead to the suppression of any particular spectral feature, but rather a general suppression of overall signal intensity. This provides a reasonable

expectation that m/z bins identified through the univariate analysis of pure bacterial cultures as different between species will translate as suitable biomarkers for the detection of that species within a mixed culture.

Pure culture isolates show high spectral richness within the 600-900 m/z region. This is consistent with what has previously been reported to be the optimal region for species level taxonomic classification, of 93.9% accuracy, of a range of 25 clinically significant microorganisms using the automated REIMS platform (195). It has been previously determined that the relative intensity of peaks is also important in species level classification, rather than simply a presence or absence of specific peaks, meaning biomarkers shared between species are not necessarily non-significant for classification (196). The identification of species-specific biomarkers within the controlled mixed cultures analysed here presents a proof-of-principle for direct REIMS analysis of clinical samples, for example urine, blood and sputum. This may be performed without the requirement for the pre-isolation of pure cultures and will be achieved by constructing a database comprising species-specific biomarkers for clinically important microorganisms. Definitive biomarker identifications corresponding to observed m/z values have not been determined in this experiment as the purpose was to determine if analysis of mixed cultures is possible on a proof of principle basis. Biomarker identifications are sought in Section 5.6: Biomarker identification, from clinical samples.

Currently, no commercial microbial diagnostic system capable of reliably analysing poly-microbial samples is in use in clinical microbiology laboratories, and limited applications exist for culture-free analysis directly from clinical samples. The MALDI Sepsityper kit is designed for direct analysis of blood culture broths in the detection of septicaemia and has been used with some success in poly-microbial blood sample classification, although significant sample preparation is required (164, 166). A 1 mL aliquot of a blood sample is taken and lysis buffer is added before centrifugation to obtain a pellet, which is then washed in buffer and re-centrifuged. The pellet is then resuspended in ethanol and water and centrifuged again. A total of five centrifugation steps are required in the sample preparation, including an extraction step performed

with acetonitrile and formic acid. The final solution is spotted onto a MALDI plate along with a matrix and analysed on a MALDI-ToF platform loaded with a MBT Sepsityper Compass IVD module (Bruker Daltonics, Coventry, UK), which takes into account background spectra obtained from blood cells. Identification of poly-microbial samples is based on generating synthetic spectra of pure organisms and combining them to produce a reference “fingerprint” which can be compared with the observed spectra. The system is limited to identifying two species only, and this can only be performed if a reference spectrum is available with a specific confidence score which the module generates for each sample to indicate the quality of spectral fit. In a study using this system to examine 143 poly-microbial blood samples, a correct bi-microbial classification was achieved in 34.3% of samples, with correct classification of only a single organism in 54.5% of samples, and 11.2% of samples returned no classification result (224). In 22 poly-microbial samples containing more than two causative agents, correct classification was achieved for two causative agents in 45.5% of samples, and a classification for one causative agent in 50% of samples and 4.5% of samples returned no classification for any microorganism (224). Swift diagnosis and targeted treatment of infection is imperative to improving clinical outcomes, and since the incidence of poly-microbial septicaemia is growing, a rapid and accurate method of identifying causative agents is required to improve the worse outcomes typically associated with such cases (225, 226). A commercialised REIMS system may offer a viable alternative to MALDI-ToF and the Sepsityper kit if no pre-culturing or extensive sample preparation is required, and if poly-microbial samples can be correctly identified, which this experiment suggests is possible.

Objective i has been met in this study as individual bacterial species have successfully been identified, through the identification of species-specific biomarkers, in mixed cultures, and associated spectra have been found to not be significantly suppressive. This indicates analysis of true, *in vivo*, poly-microbial clinical samples is feasible. Further constructive work would include examining synthetic mixed cultures containing genera prevalent in poly-microbial infections other than those of the urinary tract, for example *E. faecalis*, *K. pneumoniae* and *A. baumannii* in blood, and examining more of the common genera found in the native urinary tract microbiota for suppressive

effects in culture mixtures (227). This work was not carried out at this stage, however, as progression of the investigation to analysis of clinical urine samples was the overall aim. The findings of this study are therefore tested in clinical urine samples in the following chapter (5: Urine analysis).

5 Urine analysis

5.1 Bipolar Forceps Pellet Analysis

The clinical isolates investigated in the previous experiment (4: Mixed culture analysis) were utilised here in the spiking of healthy, culture negative donor urine, with the aim of determining if species-specific biomarkers could be detected over the background spectra produced by urine. In bipolar forceps REIMS analysis, good separation of distinct species was achieved in both unsupervised and supervised analysis (Figure 26). In dilutions > 10x, clustering of species in unsupervised analysis was much weaker due to the lower intensity of species-specific biomarkers in dilute samples, although supervised analysis still yielded separation with a strong Q₂ model score (Supplementary 8). Classification of species was also good through random forest analysis, with a high prediction rate and only two misclassifications (Figure 27). Mean decrease accuracy scores, shown through a random forest feature importance matrix, indicates that several biomarkers are important for classification of each species, although there was little similarity in top significant biomarkers identified through this method than in the previous mixed culture experiment (4: Mixed culture analysis).

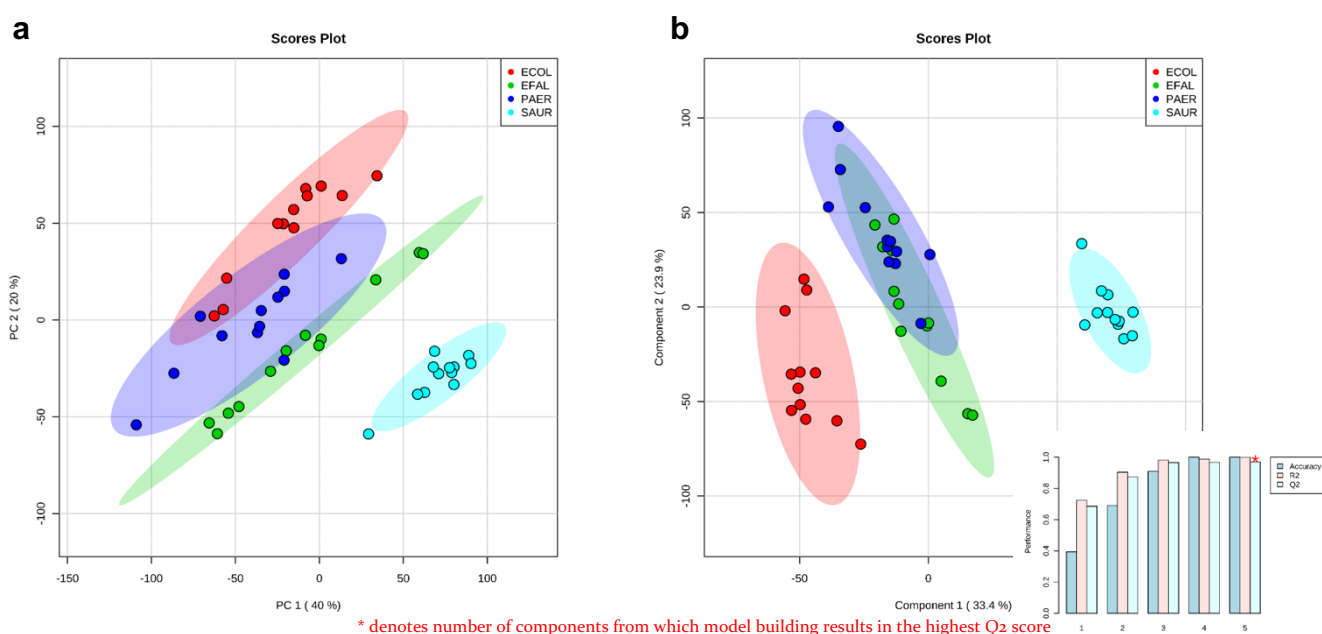


Figure 26. PCA (a) and PLS-DA (b) of all bacterial species spiked into culture negative healthy donor urine at dilutions < 10x. Good species separation was achieved in both supervised and unsupervised analysis.

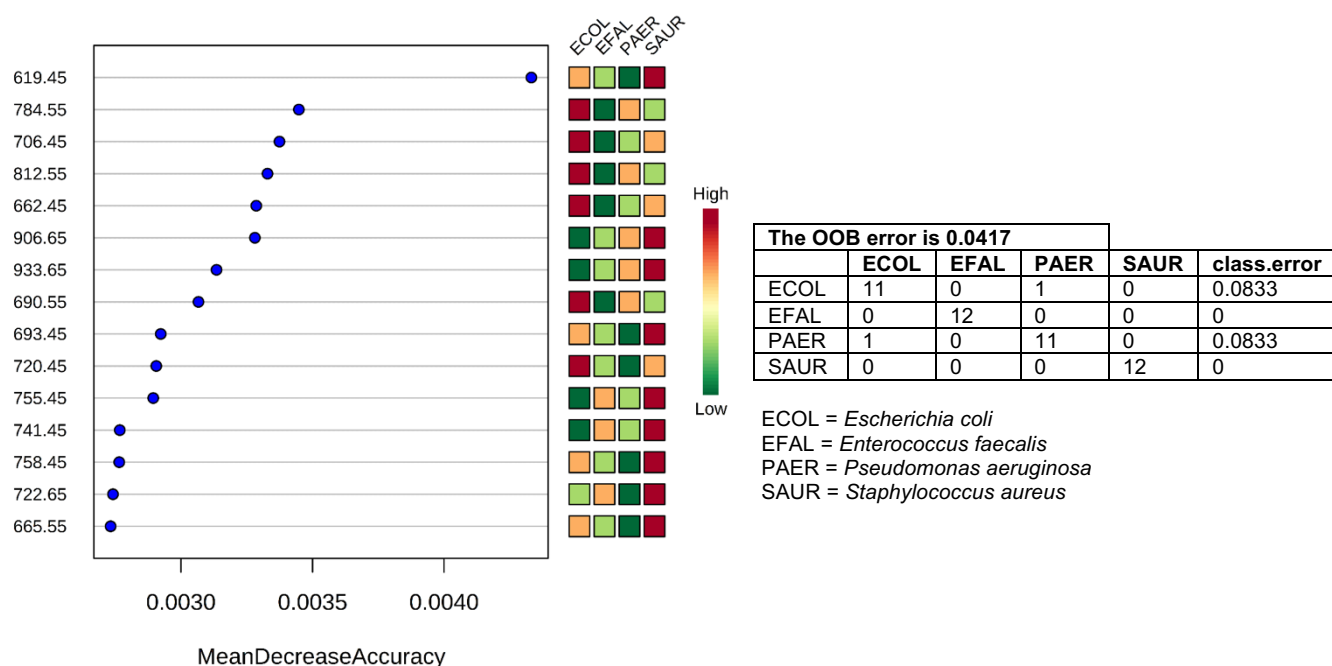


Figure 27. Random forest analysis of healthy, culture negative donor urine spiked with four clinical bacterial species. The model demonstrates a good prediction rate, with a low out-of-bag error and only two sample misclassifications. Biomarkers for each species are displayed in a feature importance matrix, ranked by mean decrease accuracy.

Volcano plots of fold change and t-tests scores identified several features which were significantly raised in bacteria containing samples (Figure 28), including features concurrently identified in pure agar cultures of the samples. Biomarker intensity was also found to reduce in positive correlation with reducing sample concentration within the spiked urine. Significant biomarkers identified in the previous mixed culture experiment were also found during analysis of spiked urine samples here, albeit at different significance levels, suggesting the spectral features are not significantly inhibited in intensity by the native urine spectrum (Figure 29). The limit of biomarker detection, based on visibility of significant biomarkers in mass spectra at the lowest concentration tested, was not reached for *S. aureus*, *P. aeruginosa* and *E. coli*. The limit of detection for *E. faecalis* containing samples was reached at a dilution of 2x. This is due to low intensity of biomarkers detected in *E. faecalis*, even at undiluted concentrations, compared with the other species tested. CFU/mL calculations for the most dilute samples tested were 10^5 for three of the isolates and 10^6 for *E. coli*. This means

the top 10 biomarkers included in analysis are detectable at these levels, and likely at much lower CFU/mL values (Table 3).

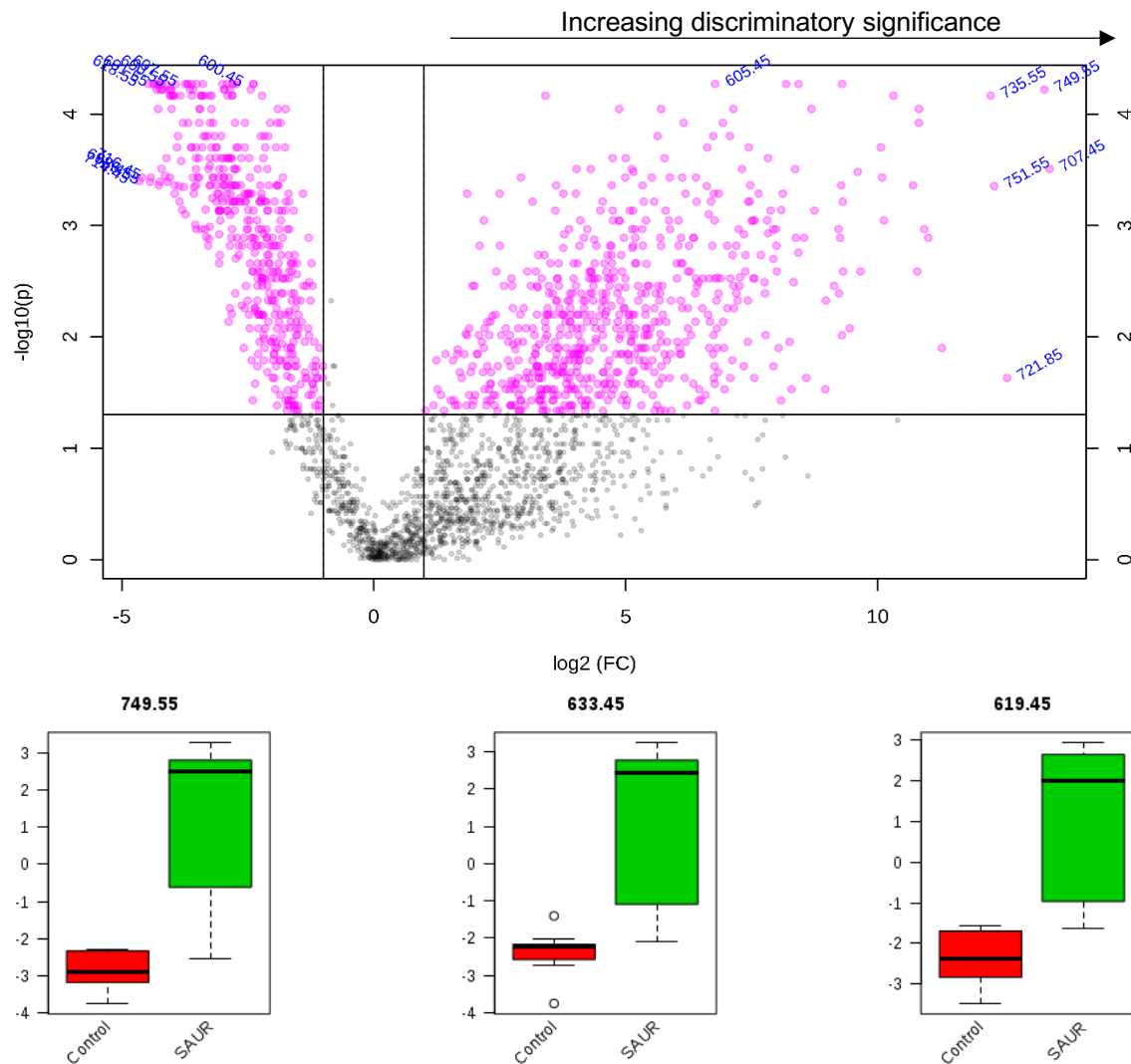


Figure 28. Volcano plot of $\log_2$ fold change (x-axis) and $-\log_{10}$ FDR corrected P value scores (y-axis) for *Staphylococcus aureus* (SAUR) spiked urine samples. Features appearing to the top right of the plot have both high fold change and low P score in *S. aureus* containing samples, suggesting they are species-specific and significant in classification. Box plots of three of the top ranking features show a significant increase in significance in samples containing *S. aureus* verses those which do not. 619.45 m/z is also observed in random forest analysis. Range of P values is large due to all concentrations of spiked samples being included in analysis.

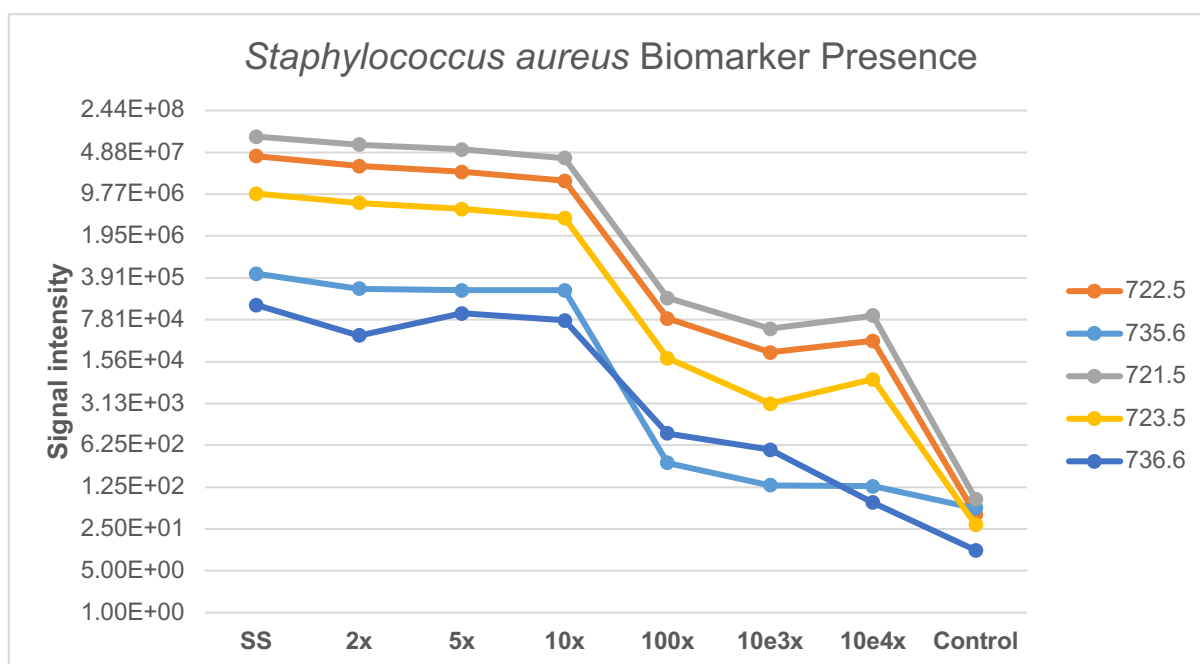


Figure 29. Significant biomarkers in *S. aureus*, determined by FDR corrected *P* value are visible in decreasing intensity across the concentration range, demonstrating that biomarker intensity is likely correlated with microbial cell count. Biomarkers of the mixed culture experiment (4.1-Species-Specific Biomarkers) are identifiable in spiked urine samples with intensity reducing with sample concentration, demonstrating signal intensity is not significantly inhibited by either the native urine spectrum or the other microorganisms present within the mixtures. The x axis displays dilution factors of 10,000x, 1000x, 100x, 10x, 5x, 2x and starting solution (SS).

Table 3. Significant biomarker detection rates in culture negative urine samples spiked with *Staphylococcus aureus* (SAUR), *Pseudomonas aeruginosa* (PAER), *Escherichia coli* (ECOL) and *Enterococcus faecalis* (EFAL). Dilutions of spiked samples were spread onto agar media and cultured to determine CFU/mL of each dilution. The limit of biomarker detection was not reached for *S. aureus*, *P. aeruginosa* or *E. coli*, whilst biomarkers failed to be detected at dilutions greater than 2x in *E. faecalis*. SS = starting solution.

Species	CFU/mL of SS sample	CFU/mL of 10 ⁴ dilution	Biomarkers Detected?
SAUR	3.00E+09	3.00E+05	Yes
PAER	9.90E+09	9.90E+05	Yes
ECOL	3.20E+10	3.20E+06	Some
EFAL	5.10E+09	5.10E+05	No

5.2 LA-REIMS method development

The existing LA-REIMS laser lens, designed for analysis from agar media and 24 well plates (Figure 9, Figure 11), was utilised here to develop an automated, high throughput method for urine analysis since manual analysis with bipolar forceps is time consuming and labour intensive. Analysis from nitrocellulose membranes and filter paper spotted with spiked urine was attempted. Results from this method development experiment were compared with the bipolar forceps pellet analysis experiment (5.1: Bipolar Forceps Pellet Analysis) to determine optimal analysis parameters.

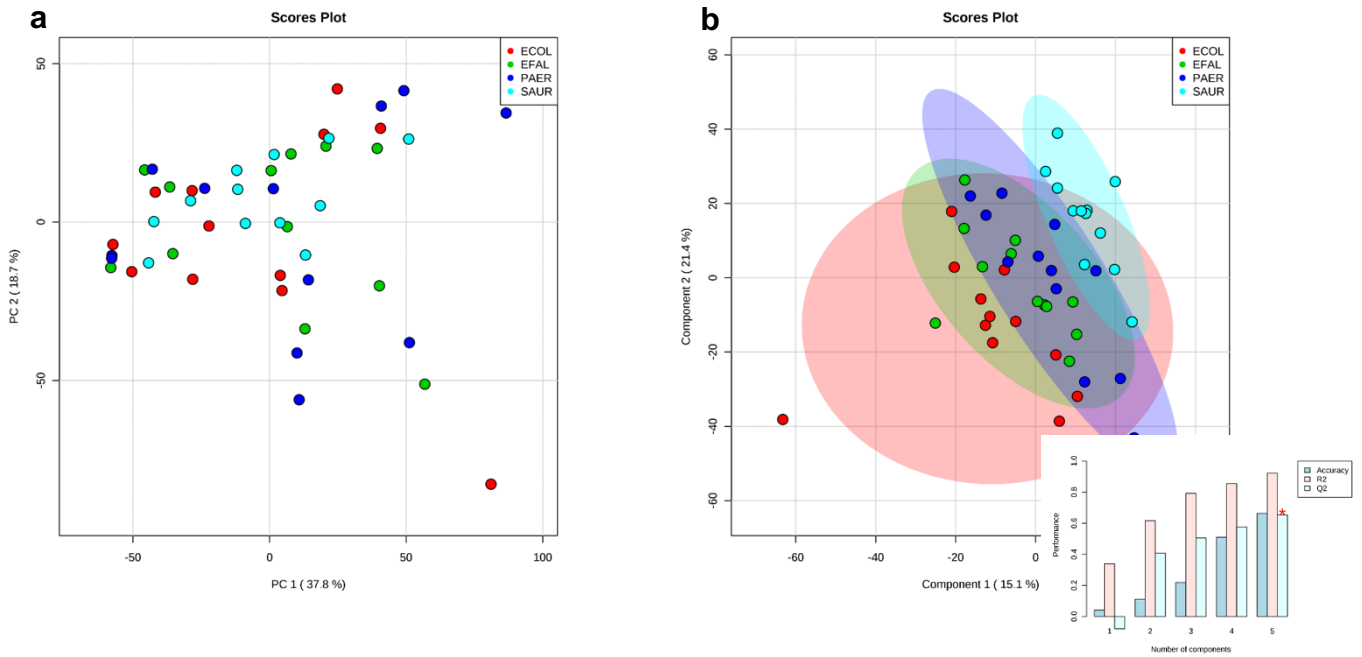
For nitrocellulose membrane analysis of spiked urine samples, poor separation of distinct species was achieved in urine samples spiked with the four bacterial isolates in unsupervised analysis, but good clustering of species was achieved in supervised analysis (Figure 30). Classification through random forest analysis was also poor, with many misclassifications and high out-of-bag error, as well as individual class errors (Figure 31). Both random forest and volcano plots of fold change and t-test scores show some overlap in top-ranking significant biomarkers when compared with the previous bipolar REIMS pellet experiment (5.1: Bipolar Forceps Pellet Analysis), in particular 749.55 and 722.65 m/z are conserved as a significant class defining feature in the two models (Figure 32). In analysis from filter paper, clustering of distinct species was better in both supervised and unsupervised analysis than with nitrocellulose membrane analysis, however separation was weak (Figure 30). Significant biomarkers were again detected through random forest analysis and t-test/fold change volcano plots, however top-ranking features were largely not comparable to the previous bipolar REIMS experiment (Figure 32). Classification with random forest performed similarly to analysis with nitrocellulose membranes, although top ranking significant features differed (Figure 31).

Spectral richness was greater in samples analysed from filter paper, although spectral quality was lower, indicating the presence of a large amount of non-isolate specific background noise (Figure 33). This is consistent with what was observed during analysis, as laser ablation of filter paper produced a significant amount of polymer peaks, likely

from cellulose. The range of quality scores in nitrocellulose membrane analysis was high due to uneven absorption of urine aliquots, resulting in the laser beam periodically firing at dry membrane due to the difficulty in positioning both the membrane, sample and laser analysis point with accuracy. Firing at dry membrane results in poor quality spectra, whilst firing at saturated membrane resulting in higher quality spectra with a better signal:noise ratio.

The limit of detection, based on visibility of significant biomarkers in mass spectra at the lowest concentration tested, was not completely reached for both filter paper and nitrocellulose membrane analysis, however biomarker intensity reduced in positive correlation with sample dilution, as in the bipolar REIMS experiment (Figure 34). The most dilute samples of urine were of $\approx 10^2$ CFU/mL, with some biomarkers still being detected at this level.

Nitrocellulose membrane method



Filter paper method

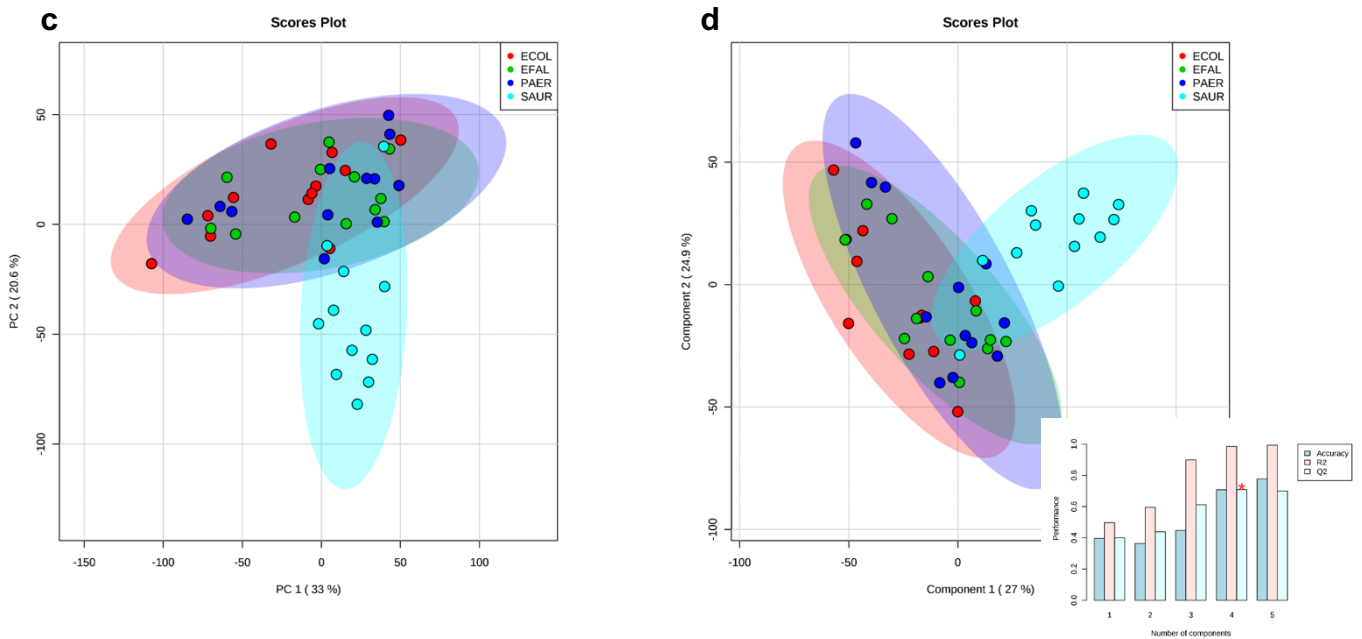
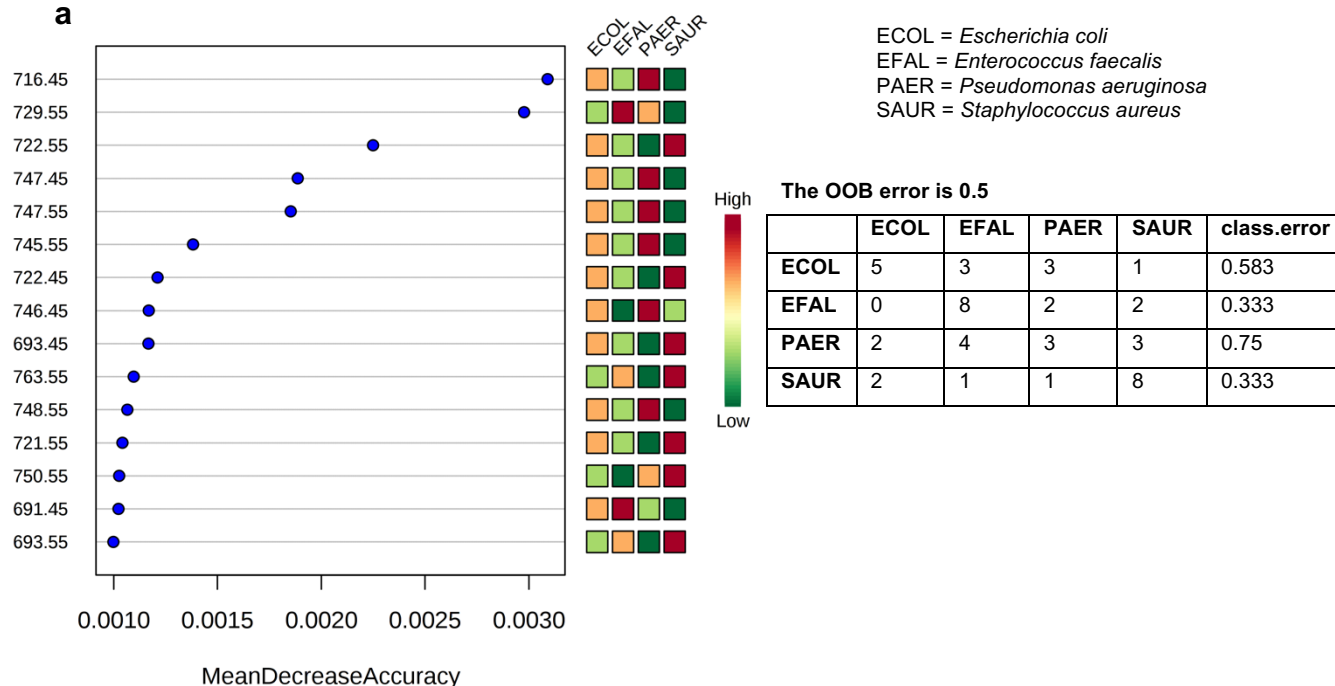


Figure 30. Nitrocellulose membrane (a-b) and filter paper (c-d) LA-REIMS analysis of *Escherichia coli* (ECOL), *Enterococcus faecalis* (EFAL), *Pseudomonas aeruginosa* (PAER) and *Staphylococcus aureus* (SAUR) spiked into culture negative healthy donor urine at dilutions < 10x. Poor species separation was achieved in unsupervised analysis (a) with nitrocellulose membranes, however supervised PLS-DA shows some species clustering. Good clustering of species was achieved in both unsupervised (c) and supervised PLS-DA (d), with some species clustering.

Nitrocellulose membrane method

a



Filter paper method

b

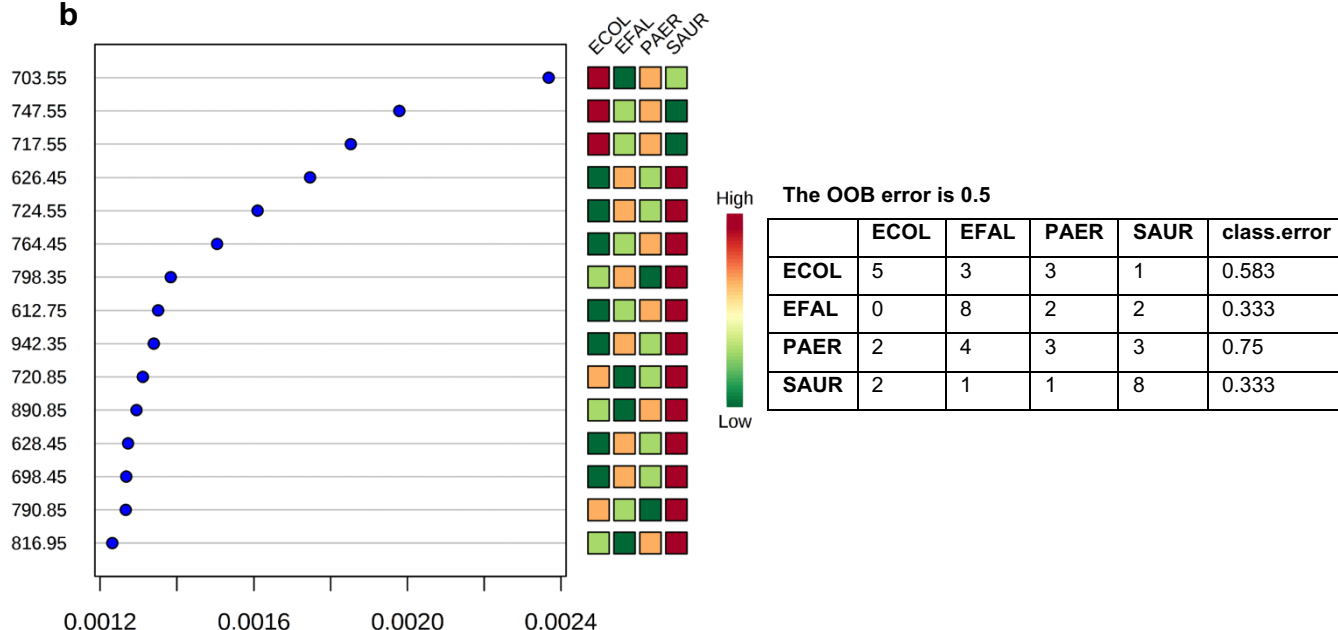


Figure 31. Nitrocellulose membrane (a) and filter paper (b) LA-REIMS random forest analysis of culture negative donor urine spiked with four clinical bacterial species. Misclassification of species is high, represented by a high out-of-bag error and individual class errors. Top-ranked biomarkers for each species are displayed in a feature importance matrix, ranked by mean decrease accuracy. Some significant feature overlap (of top-ranked features) with bipolar REIMS forceps analysis can be seen, with 722.55 *m/z* (+/- 0.10) conserved across both random forest models in *S. aureus*. 747.55 *m/z* is also conserved across both models in *E.coli*.

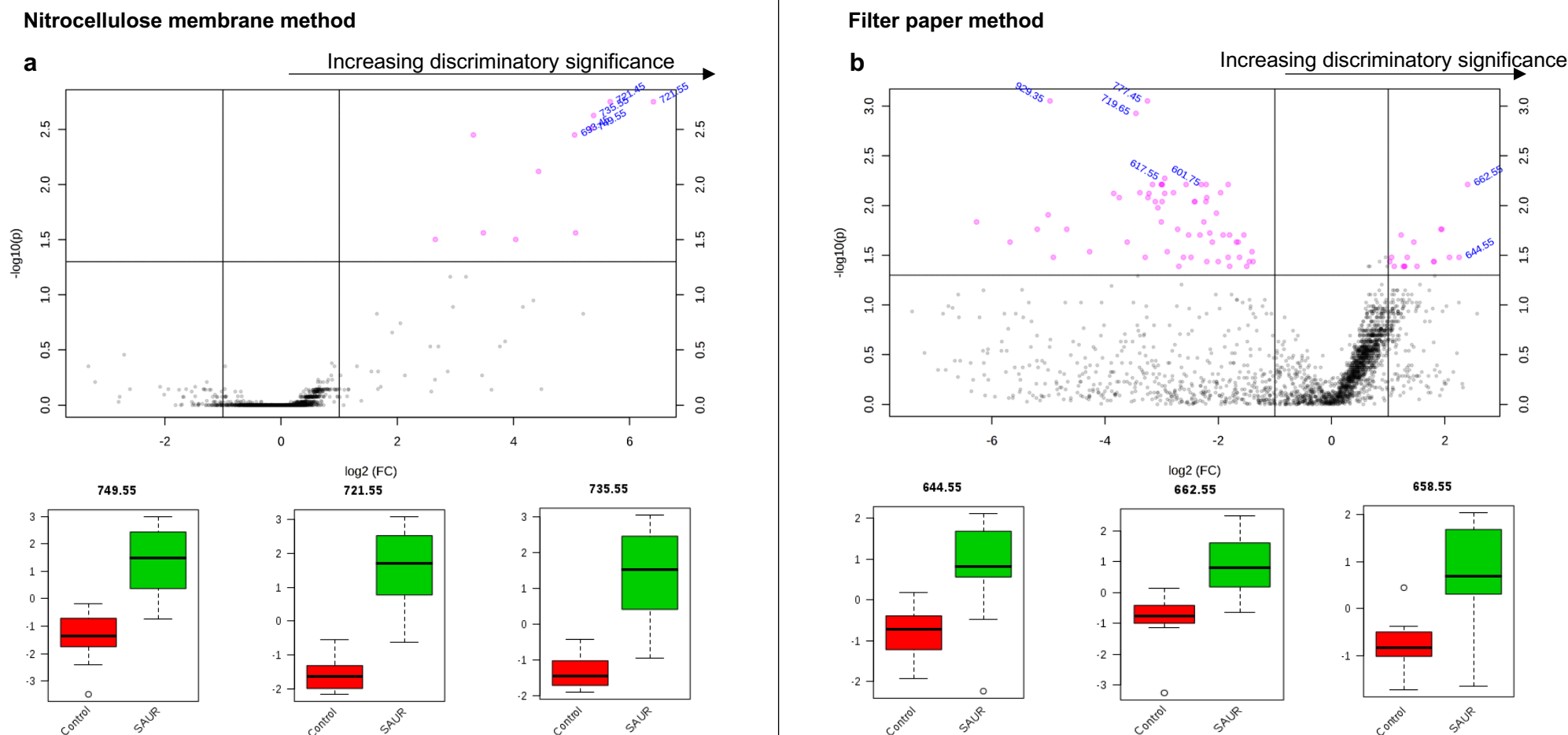


Figure 32. Volcano plots of log₂ fold change (x-axis) and -log₁₀ FDR corrected *P* value scores (y-axis) for *Staphylococcus aureus* (SAUR) spiked urine samples analysed with nitrocellulose membrane (a) and filter paper (b) LA-REIMS. Features appearing to the top right of the plot have both high fold change and low *P* score in *S. aureus* containing samples, indicating they are both species-specific and significant in classification. Although fewer significant features were detected compared to bipolar REIMS pellet analysis, box plots of three of the top ranking features for both methods show a significant increase in significance in samples containing *S. aureus* verses those which do not. 721.55 *m/z* is also observed in top-ranking random forest features, and 662.55/749.55 (+/- 0.1) *m/z* are identified as significant in bipolar REIMS analysis of urine pellets, suggesting they are conserved biomarkers for *S. aureus* across different REIMS methods.

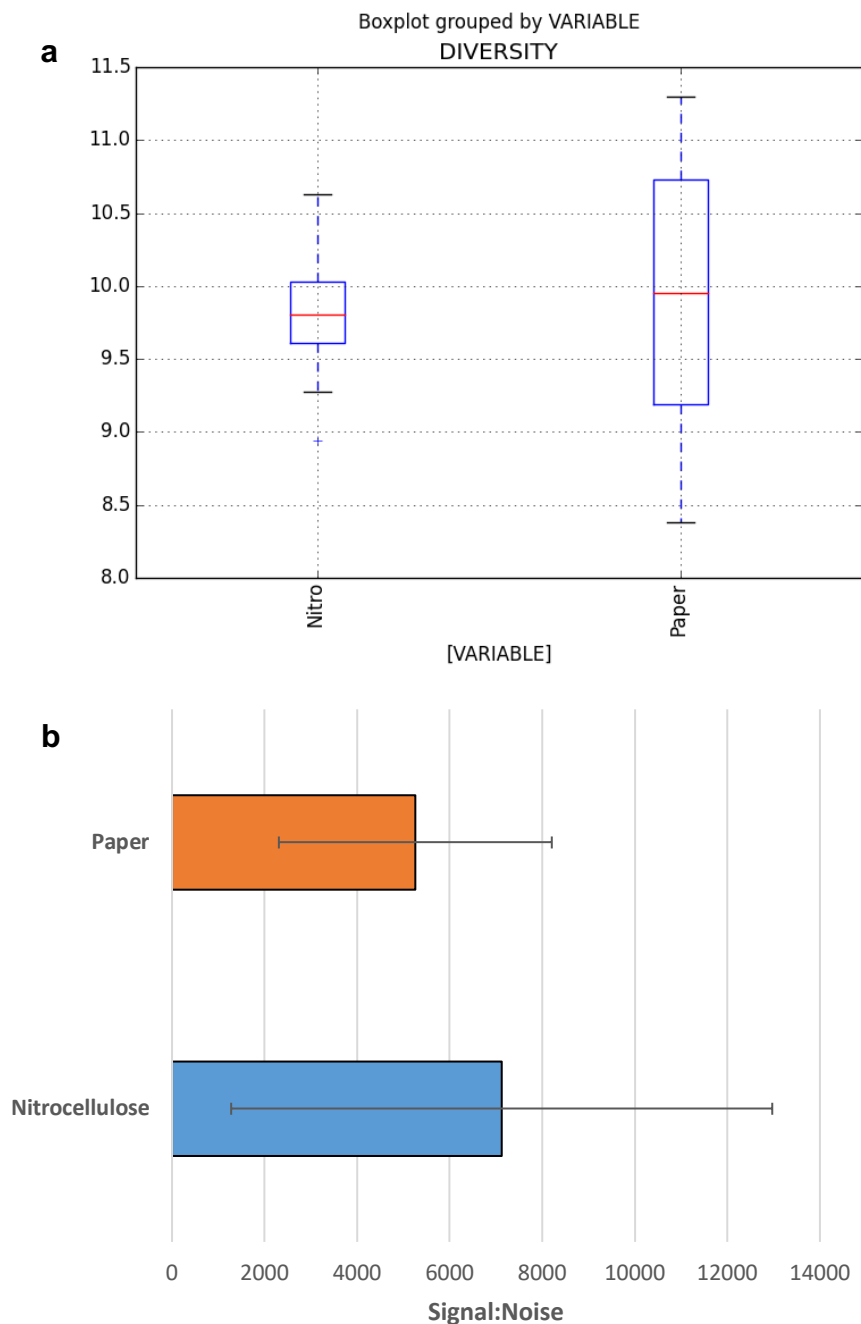


Figure 33. Mean combined spectral richness score box plot (a) and quality scores (b) of clinical urine samples analysed with nitrocellulose membrane and filter paper LA-REIMS. Spectral richness was higher for samples analysed with filter paper REIMS, however with a large range due to background noise generated from burning paper. Spectral quality was generally higher in nitrocellulose membrane LA-REIMS, however a large variation between samples due to noise, uneven sample absorption and difficulty in optimal positioning of the laser beam was observed.

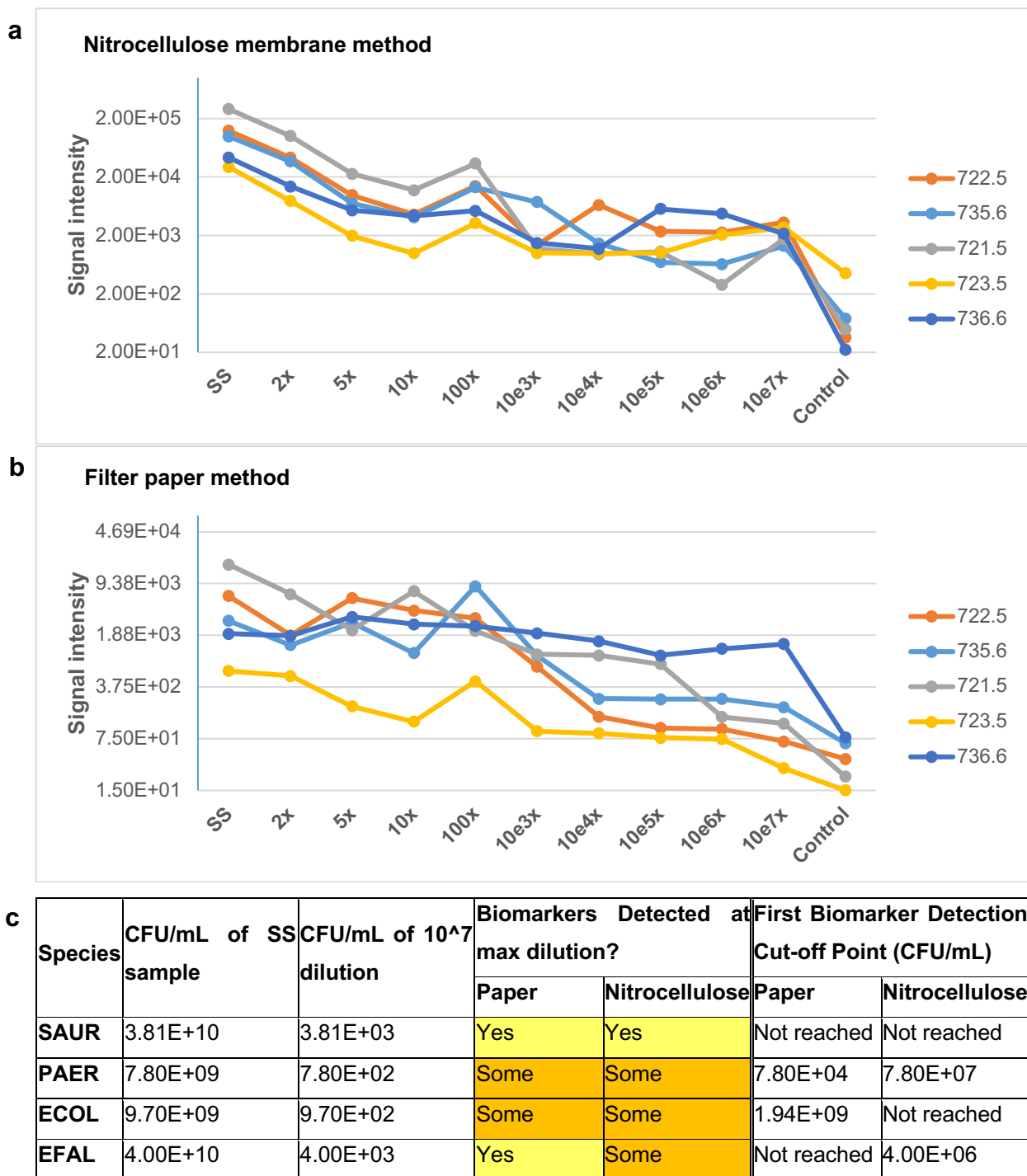


Figure 34. Presence of *Staphylococcus aureus* significant biomarkers, identified in bipolar forceps REIMS of urine pellets, and in nitrocellulose membrane (a) and filter paper (b) LA-REIMS, across reducing sample concentrations. All biomarkers were identified in the two methods, although were not in top-ranking significant feature lists generated by random forest or volcano plots. Biomarkers reduced in intensity in correlation with sample dilution factor. (c) Significant biomarker detection rates in culture negative urine samples spiked with *Staphylococcus aureus* (SAUR), *Pseudomonas aeruginosa* (PAER), *Escherichia coli* (ECOL) and *Enterococcus faecalis* (EFAL). The limit of biomarker detection was not fully reached for any isolate, however some biomarkers failed to be detected sooner in analysis from nitrocellulose membranes compared to filter paper.

5.3 LA-REIMS Optimisation

A modified laser probe for the automated Tecan analysing platform was received from WRC, Hungary, enabling LA-REIMS analysis of deep-well plates which was previously not possible with the Tecan monopolar probe or standard LA-REIMS setup (Figure 7). This experiment aimed to optimise the modified probe for clinical urine pellet analysis from 96 well plates by investigating pulsed and continuous laser emitting modes and a range of laser powers. Analysis revealed spectral richness and quality varied widely across both laser wattage and mode parameters (Figure 35). Although spectral richness appeared higher in continuous mode, this was due to continuous laser fire burning through biomass pellets and sometimes reaching through to the plastic below, producing a mass spectrum rich in noise (Figure 36). Continuous mode also occasionally caused pellets to overheat and splatter. This was difficult to control for even at low beam powers, as biomass pellets were irregularly shaped and positioned with variable thickness. For this reason pulsed mode was preferred, as the pulsing of the beam allows for a lower maximum temperature to be reached, and splattering did not occur at any power level tested. In general, pulsed mode gave higher quality spectra than continuous (Figure 36). A power level of 1.5-2 W in pulsed mode was deemed optimal for biomass pellet analysis, as overall spectral richness at these levels was preferred, although quality was slightly reduced at 2 W due to increased background noise peak intensity.

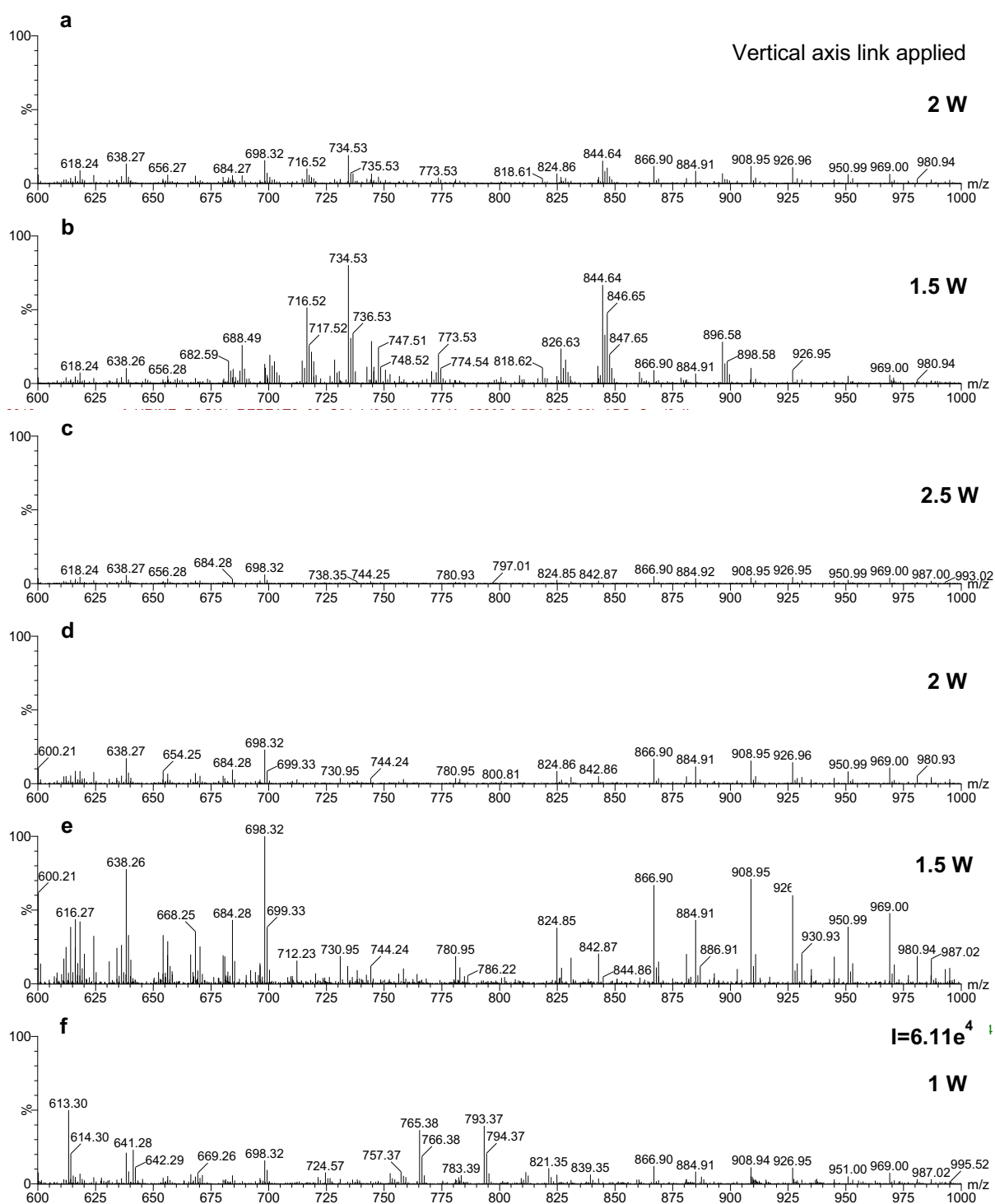


Figure 35. Mass spectra of identical urine pellets, in the 600-1000 m/z range, of continuous (a, b) and pulsed (c-f) laser firing modes at varying wattages. Pre-processing of spectra for background subtraction and lockmass correction has been applied. Spectral feature richness and intensity is highly variable dependant on both laser firing mode and power.

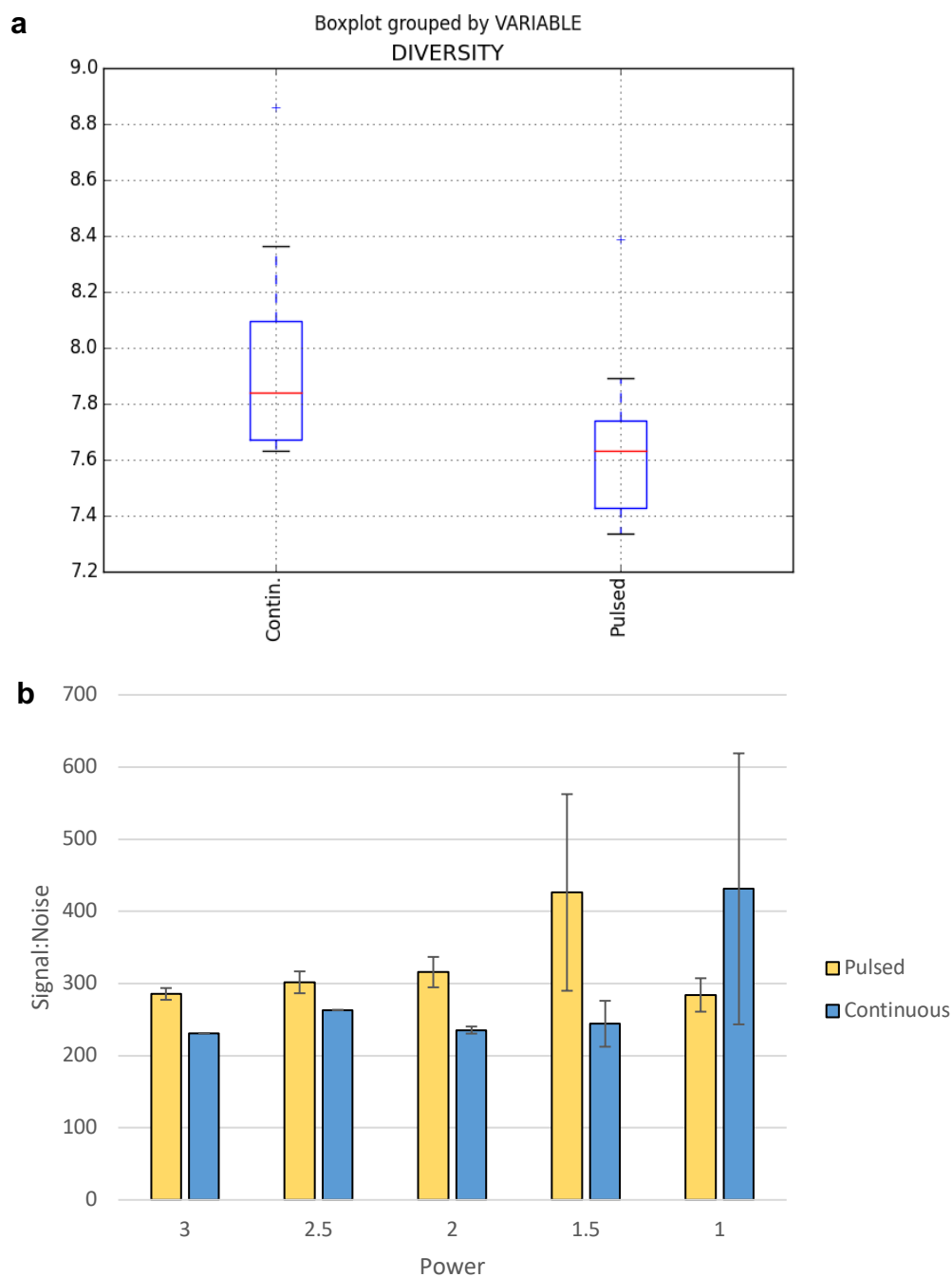


Figure 36. Mean combined spectral richness score box plot (a) and mean combined spectral quality scores (b) of urine pellets analysed with LA-REIMS, in pulsed and continuous laser firing modes across a 1-3 W power range. Spectral richness was consistently higher for samples analysed in continuous firing mode, however this was due to background noise generated via burning through the urine pellet and into the plastic underneath. A laser power of 1 W produced optimal quality spectra for continuous firing mode, whilst a 1.5 W supply power produced optimal quality for pulsed mode. A greater range of quality scores was observed in continuous mode at 1 W than for pulsed mode at 1.5 W, due to splattering and overheating of the sample pellet leading to higher background noise in some samples.

5.4 Clinical Urine 1

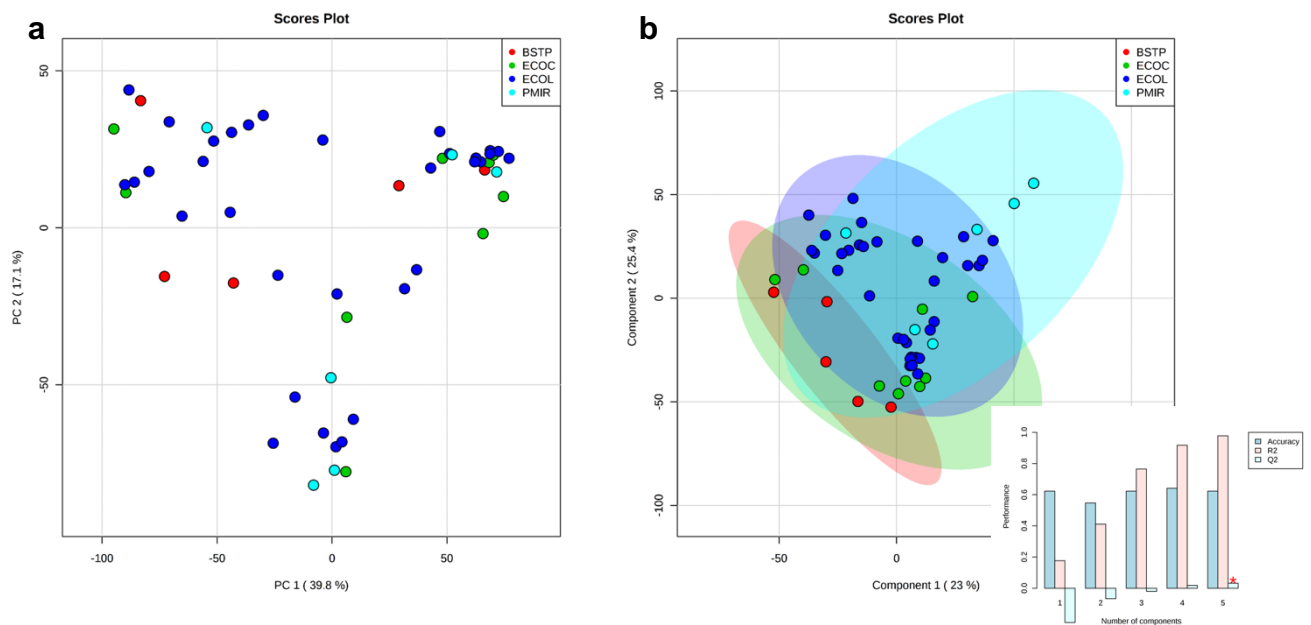
Analysis of urine pellets is preferred over nitrocellulose membrane and filter paper REIMS, due to ease of sample preparation and more consistent biomarker identification. Here, pellet analysis is attempted with LA-REIMS, rather than bipolar forceps, to increase commercial viability through automation and high-throughput capability. Classification and metadata of urine samples obtained from ICHNT clinical microbiology laboratory was not available until after the LA-REIMS experiment had been ran. As a result, some samples were discarded due to there being less than three replicates for a single bacterial species, more than one species present within a single sample, or a non-confident identification result returned. *E. coli* was the most common causative agent in urinary tract infections within the samples obtained (52% of samples), followed by *Enterococcus* spp. (14%) *P. mirabilis* (11%) and *S. agalactiae* (reported as Group B *Streptococcus*) (5%). Samples are not always classified to species level within the ICHNT laboratory as it is not clinically necessary, whilst some are simply reported as “coliforms” or “mixed”. The latter were discarded from the dataset as insufficient metadata in these cases could distort analysis models, classification accuracy and significant biomarker identification. Samples were included in analysis regardless of growth amount in culturing, as long as they were not reported as a possible contaminant or skin commensal from incorrect urine sample collection technique.

Relatively small pellets were produced after plate centrifugation, due to low max centrifuge speeds when the plate adapter is fitted to the rotor. Visible pellets were present in most samples, however in some only a thin biomass coating on the bottom of the plate well was seen after discarding supernatant. This may be overcome in future by switching to analysis from microcentrifuge tubes, which can be centrifuged at higher speeds.

No species grouping or separation was achieved in unsupervised analysis, while supervised analysis showed overfitting and a poor model performance, represented by a low Q₂ score (Figure 37). Samples grouped by culture negative/culture positive status performed better in supervised analysis, although class separation was still relatively

poor (Figure 37). Random forest and volcano plot analysis revealed some biomarkers significant in culture positive identification, but with high class errors observed (Supplementary 9 Supplementary 10). With all culture positive samples other than those containing *E. coli* as the causative agent removed, model performance was much better, with stronger clustering, separation and a higher Q₂ model performance score in supervised PLS-DA analysis (Figure 38). Random forest and t-test/fold change volcano plot analysis identified 35 biomarkers significant in classification of *E. coli* culture positive samples over culture negative samples, although a high class error was observed in classification (Supplementary 11 Supplementary 12).

Species grouping



* denotes number of components from which model building results in the highest Q₂ score

Culture negative/positive grouping

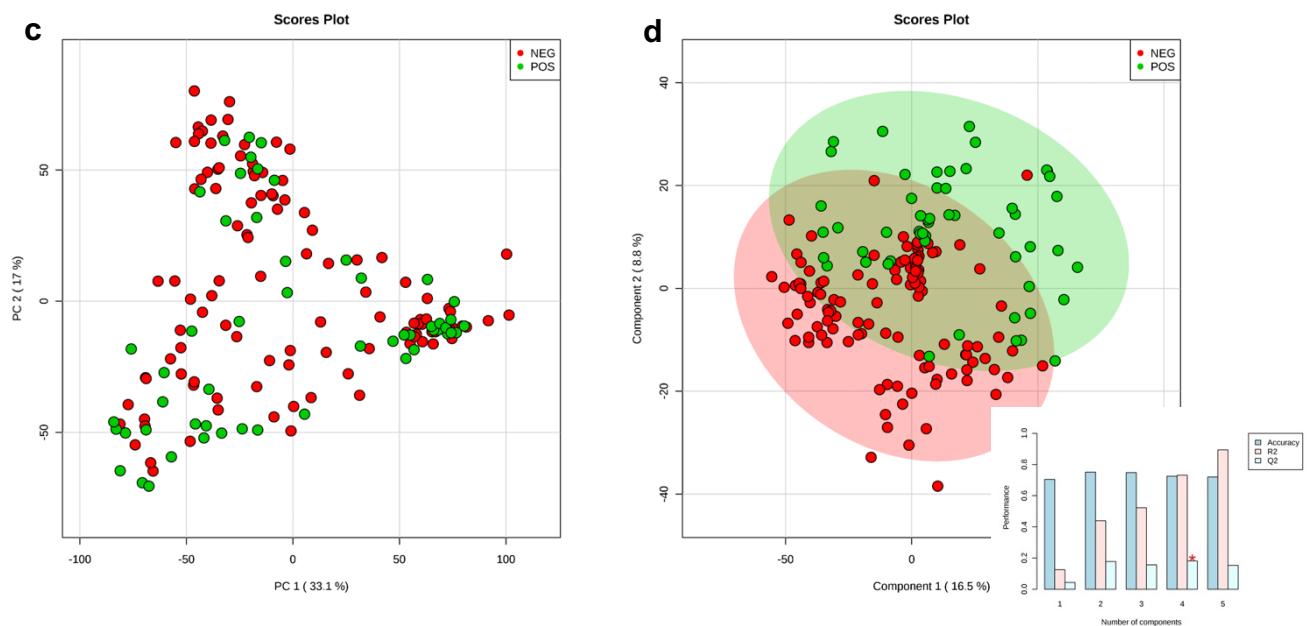
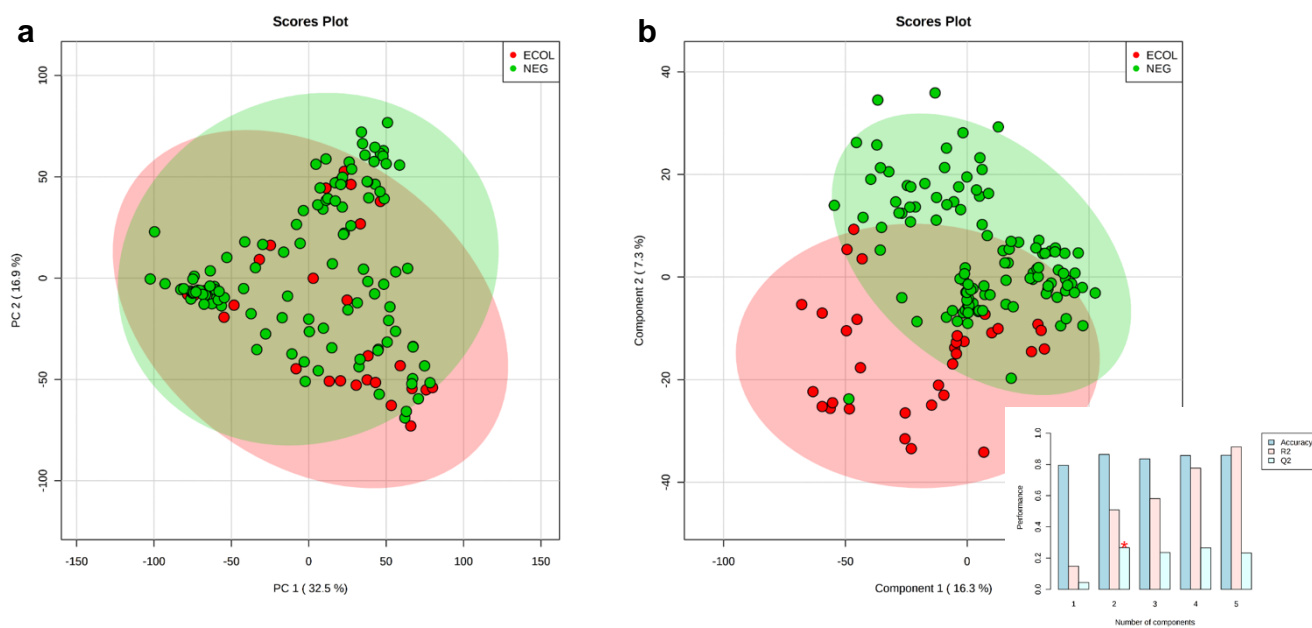


Figure 37. PCA (a,c) and PLS-DA (b,d) plots, with species and culture negative/positive groupings, of clinical urine pellet LA-REIMS analysis, 600-1000 m/z range. Species classification of urine samples was determined through the standard ICHNT microbiology laboratory workflow. No species separation was achieved in unsupervised analysis, however supervised PLS-DA shows some species clustering with a low model Q₂ score, suggesting overfitting. Supervised PLS-DA of culture negative/positive samples (d) shows good class clustering and some separation, with a better Q₂ score than species groupings. BSTP = Group B *Streptococcus*, ECOC = *Enterococcus cloacae*, ECOL = *Escherichia coli*, PMIR = *Proteus mirabilis*.



* denotes number of components from which model building results in the highest Q₂ score

Figure 38. PCA (a) and PLS-DA (b) plots, with *Escherichia coli* (ECOL) versus culture negative grouping, of clinical urine pellet LA-REIMS analysis, 600-1000 m/z range. No class separation was achieved in unsupervised analysis, however supervised PLS-DA shows good class clustering and some separation, with a better Q₂ score than discriminant analysis of culture positive versus negative groupings.

Univariate ROC curve analysis of two of the top significant *E. coli* biomarkers, 688.45 and 780.55 m/z , demonstrates their significance and strong classifying ability, represented by high AUC scores (Figure 39). Importantly, significant biomarkers previously identified in pure *E. coli* isolates are present in culture positive clinical urine specimens containing *E. coli* as the causative agent. Spectral comparison of a healthy, culture negative donor urine sample spiked with an *E. coli* isolate and the clinical urine specimen clearly show the presence of the shared 688.45 m/z biomarker (Figure 40).

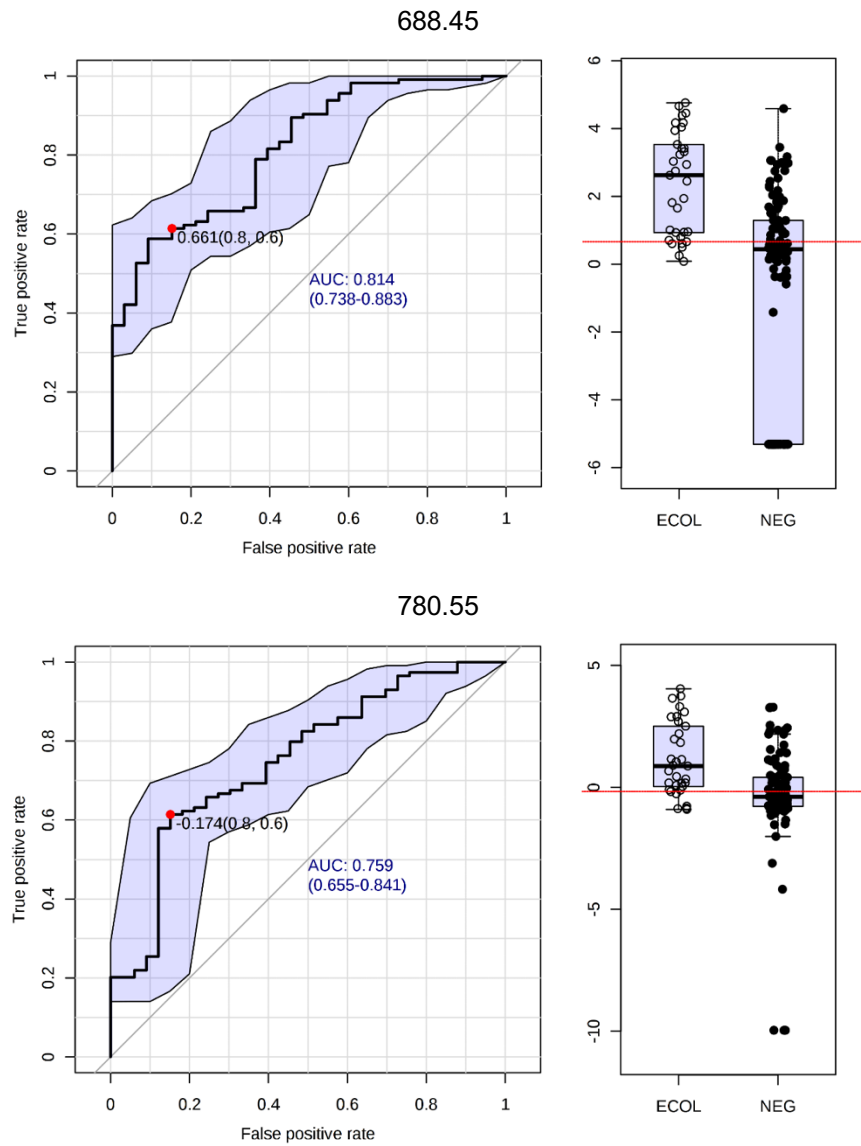


Figure 39. Univariate ROC curves of top two most important discriminatory features for *Escherichia coli* (ECOL) classification, identified through a multivariate random forest selection algorithm. The 95% confidence interval is shown in blue. High area under the curve (AUC) values and a curve approaching the upper left corner of the plot indicate strong performance of the biomarker in discriminating between *E. coli* containing and culture negative clinical urine samples.

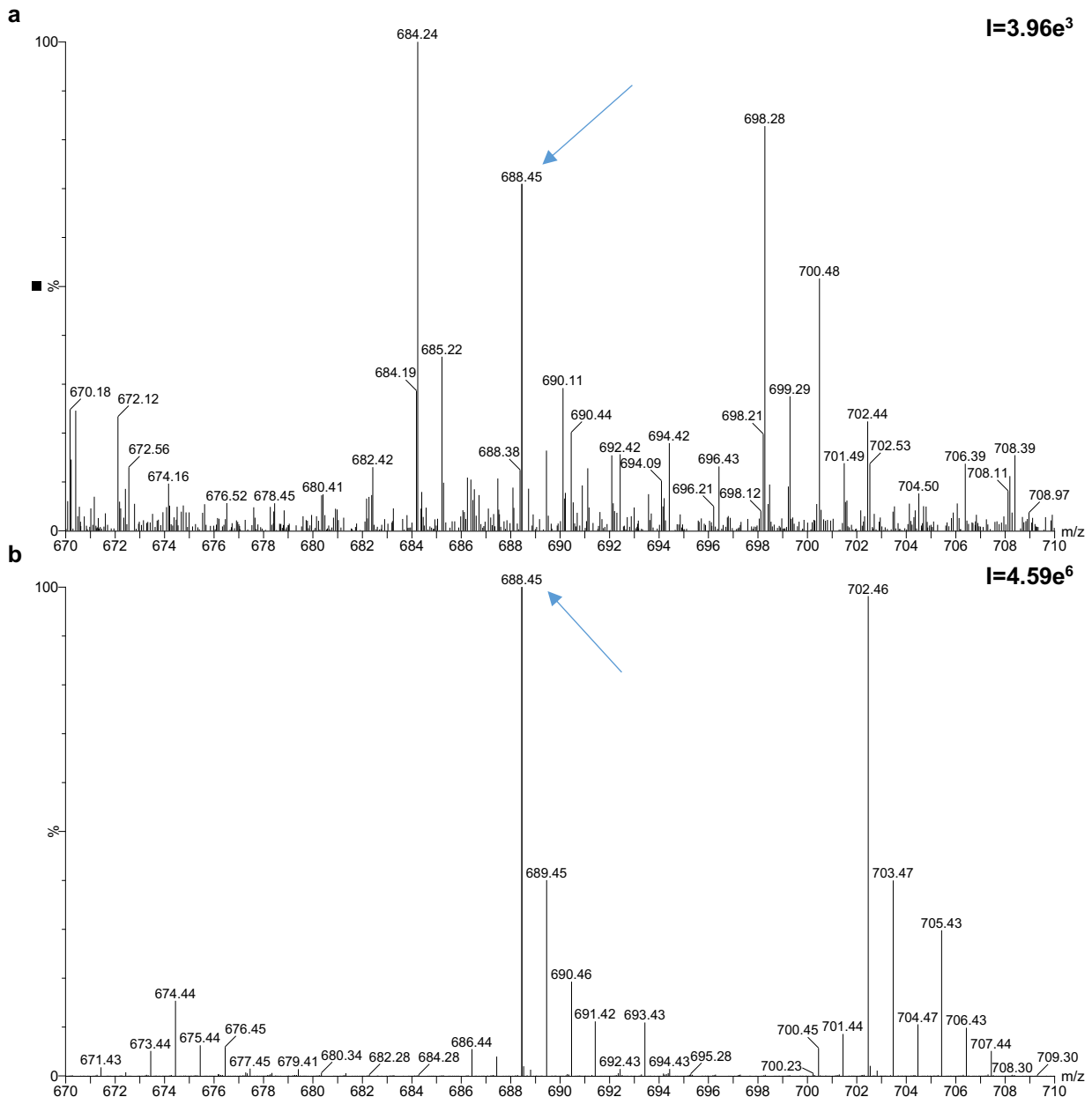
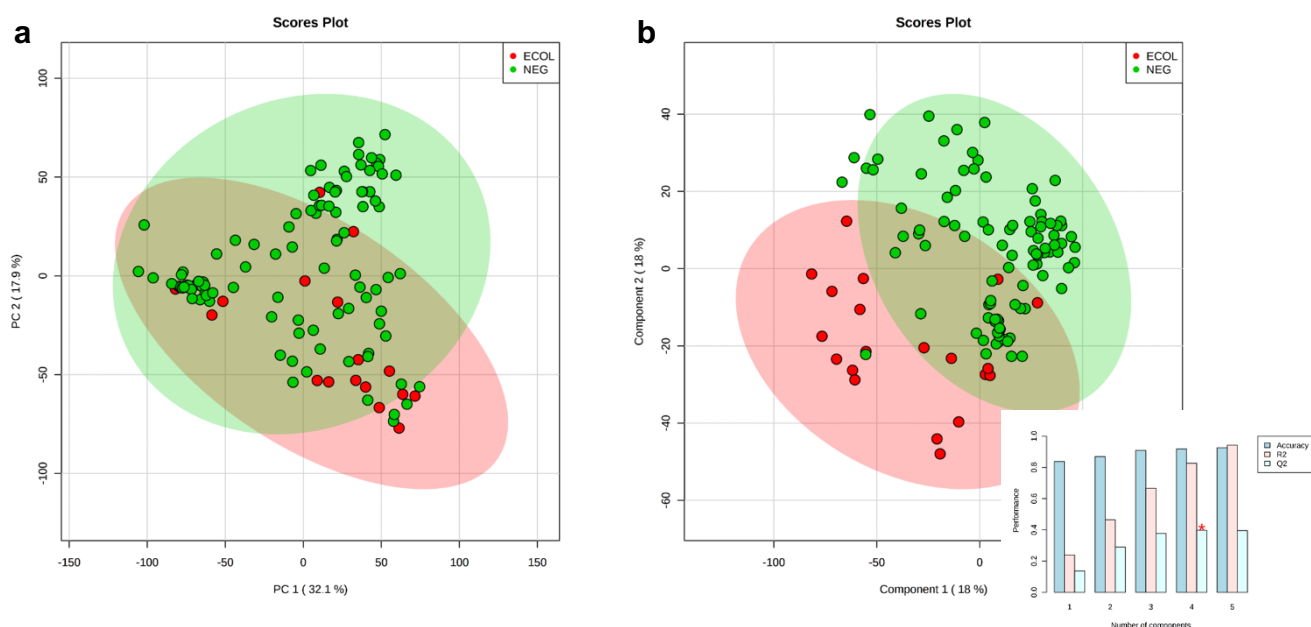


Figure 40. Mass spectra comparison of a culture positive clinical urine sample containing *Escherichia coli* (a) and a culture negative sample spiked with a pure *E. coli* isolate (b). The 688.45 m/z biomarker, identified as significant for classification by both random forest and univariate ROC curve analysis, can be seen in both samples. Background subtraction and lockmass correction applied.

A subset of data was generated based on the white blood cell count (WBC) reported by the ICHNT laboratory, once metadata was obtained. Samples classed as culture positive, but with a WBC <100, were excluded from the subset. Similarly, samples classed as culture negative with a WBC ≥ 100 were also excluded. Data analysis was repeated for the subset.

PCA and PLS-DA models of the new dataset were improved compared to the previous, with better separation and clustering of classes and a higher Q₂ score (Figure 41). Classification through random forest analysis was also improved, with both a better out of bag and class specific error score, although class error for *E. coli* was still poor (Figure 42). A total of 234 biomarkers, significant in *E. coli* classification, were identified through t-test/fold change volcano plot analysis (Figure 43), approximately seven times more than the previous dataset. The top significant biomarkers identified previously were still of top-ranking significance in this dataset, represented in both volcano plot and associated box plots.



* denotes number of components from which model building results in the highest Q₂ score

Figure 41. PCA (a) and PLS-DA (b) plots, with *Escherichia coli* (ECOL) versus culture negative grouping, of clinical urine pellet LA-REIMS analysis, 600-1000 *m/z* range, white blood cell positive samples only. Good class clustering was achieved in unsupervised analysis and supervised PLS-DA shows good separation with a much better Q₂ score than the previous dataset.

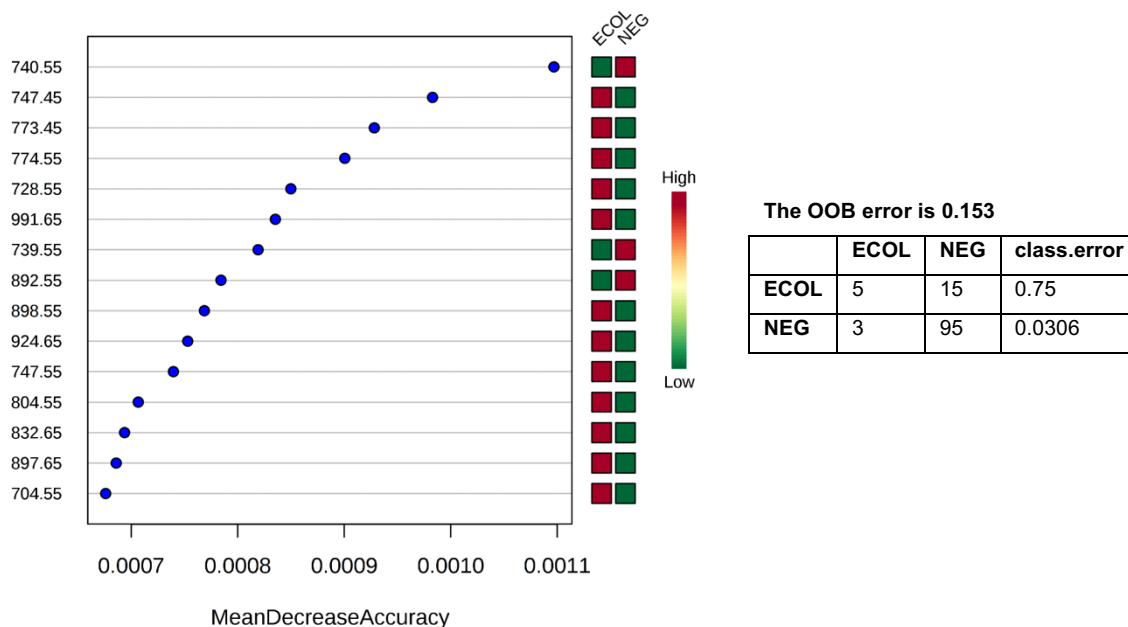


Figure 42. Clinical urine pellet LA-REIMS random forest analysis, 600-1000 m/z range, white blood cell positive samples only. Misclassification of *Escherichia coli* (ECOL) containing samples is high, represented by a high class error. Top-ranked biomarkers for *E. coli* containing urine samples are displayed in a feature importance matrix, ranked by mean decrease accuracy. Both *E. coli* class error and overall out of bag error are improved over the previous dataset.

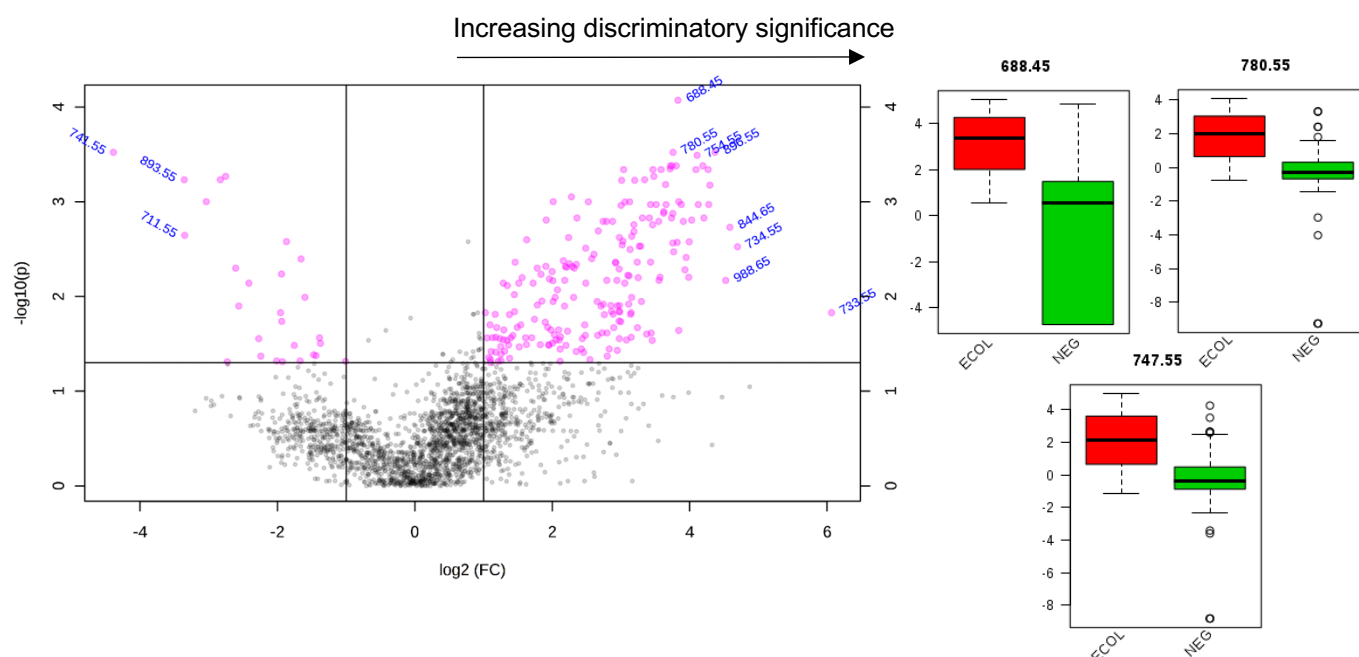


Figure 43. Volcano plot of $\log_2$ fold change (x-axis) and $-\log_{10}$ FDR corrected P value scores (y-axis) for culture positive *Escherichia coli* (ECOL) containing and culture negative clinical urine sample pellets analysed with LA-REIMS, white blood cell positive samples only. Features appearing to the top right of the plot have both high fold change and low P score in *E. coli* containing samples, suggesting they are class specific and significant in classification. Box plots of three top ranking features show a significant increase in significance in *E. coli* containing samples versus culture negative. Features identified as significant were also significant in the previous dataset, along with an additional 199 not previously identified.

To test robustness of the subset model in classifying both culture negative and *E. coli* culture positive samples, a validation model was created by inserting 10 unlabelled samples into a random forest selection algorithm based on the top 15 most significant biomarkers, with 100x cross validations. Top biomarkers were selected based on their significance rank in both volcano plot analysis and random forest. In the ten unlabelled sample predictions, only one *E. coli* sample was misclassified (Table 4), with an overall predictive class accuracy of 73% across all cross validations (Supplementary 13). ROC curve analysis and permutation tests further demonstrate model robustness, with a high AUC value of 0.832 and a narrow 95% confidence interval margin in the ROC curve and low *P* value of < 0.05 over 1000 permutation times (Figure 44).

Table 4. Clinical Urine 1 overall unlabelled sample prediction matrix. 10 samples were removed from the classification model and inserted into a validation model with their class label removed. A random forest prediction algorithm was ran to assign a class to the unlabelled samples. The overall result matrix following 100 cross validations results in only one *Escherichia coli* (ECOL) sample misclassified as culture negative (NEG).

Name	Probability	Predicted Class	Actual
21	0.91667	NEG	NEG
35	0.95333	NEG	NEG
49	0.99	NEG	NEG
64	0.98	NEG	NEG
91	1	NEG	NEG
100	0.65667	NEG	ECOL
106	0.64	ECOL	ECOL
112	0.54667	ECOL	ECOL
114	0.60333	ECOL	ECOL
116	0.78333	ECOL	ECOL

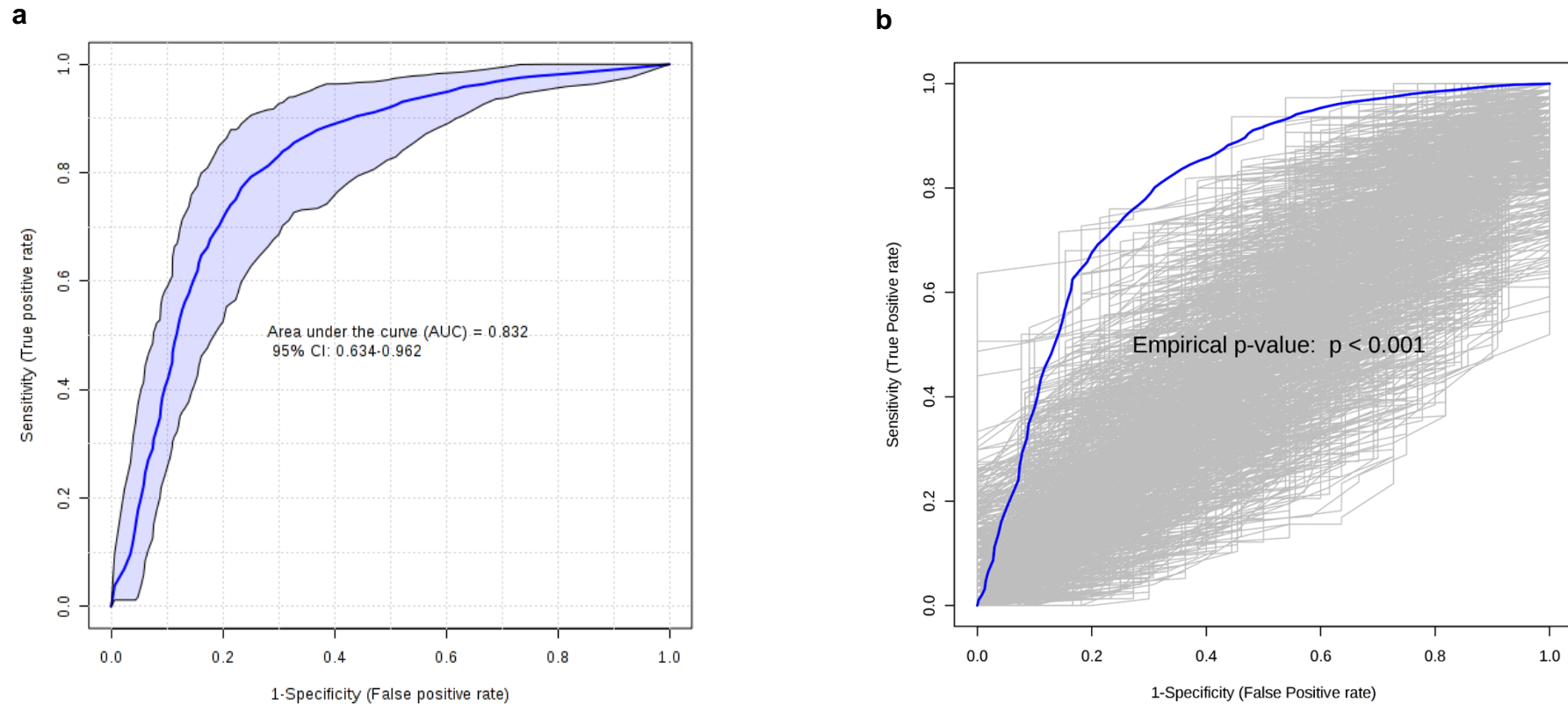


Figure 44. Univariate ROC curve of model prediction performance (a) through multivariate random forest selection algorithm, and ROC curve permutation test (b). 100 cross validations were performed to generate the model, with the 95% confidence interval shown in blue. High area under the curve (AUC) values and a curve approaching the upper left corner of the plot indicate strong performance of the model in classifying 10 unlabelled samples. In the permutation test, a sampling distribution was built by resampling the data used in the multivariate random forest selection algorithm with 1000 permutation times to assess the AUC performance measure. The low P value indicates the AUC values from each permutation are not statistically different, and the model is robust.

5.5 Clinical Urine 2

A second set of clinical urine samples were obtained from the ICHNT laboratory and analysed in microcentrifuge tubes from a rack placed within the automated Tecan system, with an aim to increase classification accuracy and biomarker discovery. This was made possible by the acquisition of a modified laser lens assembly and aspiration rod, which descends into each microcentrifuge tube. Centrifugation of urine samples at a higher $21\,500 \times g$, compared to the previous $3486 \times g$, was made possible by avoiding the use of plate rotor attachments in the centrifuge. The result of this was a greater biomass cell pellet for analysis.

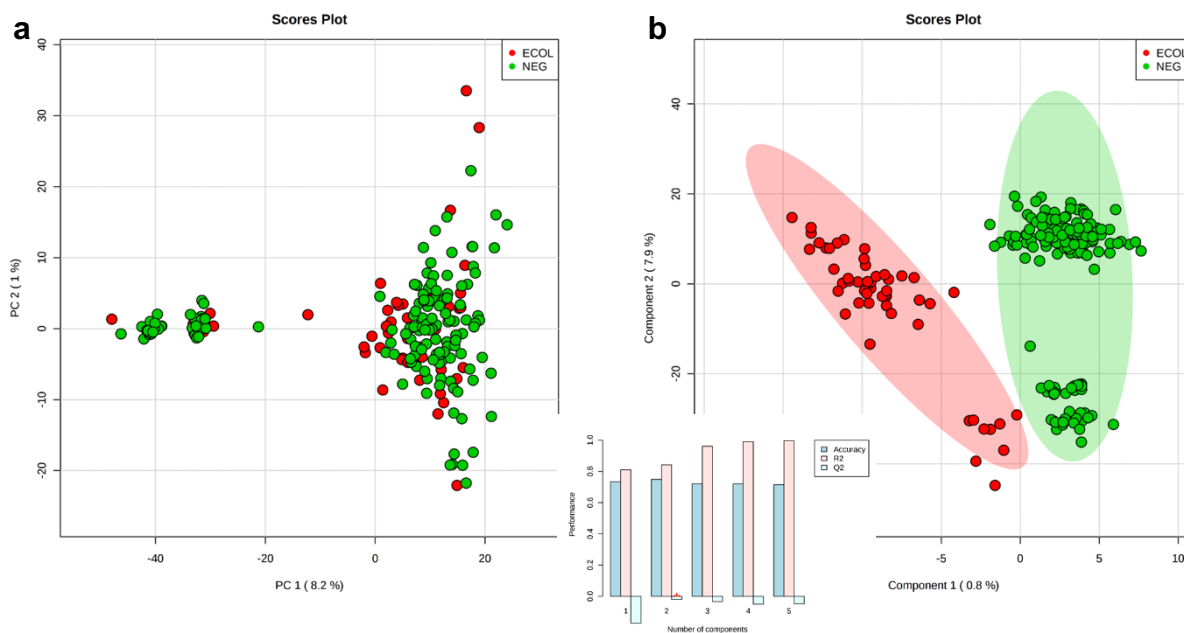
E. coli was again the most common causative agent (67% of samples), followed by *P. aeruginosa*/*S. agalactiae* (6%) and *S. marcescens*/*Enterococcus* spp. (5%). Some samples were discarded from the dataset, as before, if they were reported as “mixed”, had less than three replicates or an ambiguous result to avoid distorting the data model. Due to low replicate numbers of species other than *E. coli* present in the dataset, *E. coli* versus culture negative samples was the focus of statistical analysis and model building.

Although a greater biomass pellet was achieved over the clinical urine 1 experiment, pellets were attached to the wall of the microcentrifuge tubes rather at the bottom, due to the angle of the rotor within the centrifuge. As a result, pellets were not vaporised by the laser unless manually manipulated into position with a sterile pipette tip. This drastically increased sample analysis time.

Unsupervised analysis displayed poor separation and clustering of sample classes. Whilst supervised analysis displayed good clustering and separation, Q2 scores for the model were very low, suggesting overfitting and a poor quality model (Figure 45).

Model improvement was attempted by correcting for white blood cell count within samples, as in the previous clinical urine experiment. Samples reported as culture positive but with a white blood cell count <100 , and samples reported as culture negative but with a WBC of ≥ 100 were removed from the dataset and model building performed again.

Original dataset



* denotes number of components from which model building results in the highest Q2 score

WBC corrected dataset

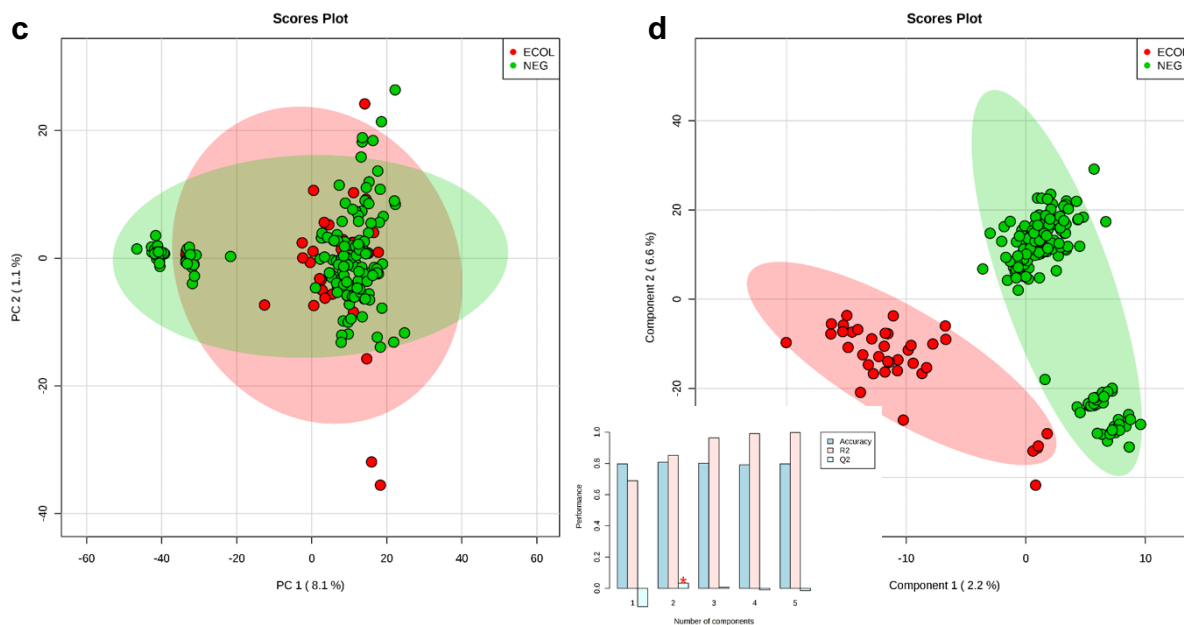


Figure 45. PCA (a,c) and PLS-DA (b,d) plots, with *Escherichia coli* (ECOL)/culture negative grouping, of clinical urine pellet LA-REIMS analysis from microcentrifuge tubes, 600-1000 *m/z* range. No class separation was achieved in unsupervised analysis in the original dataset, however supervised PLS-DA shows good class clustering and some separation, although with a negative Q2 score suggesting overfitting. When samples with a white blood cell (WBC) count <100 were removed, separation through supervised PLS-DA was improved, along with a higher Q2 score than the previous dataset, although still suggesting overfitting.

For the corrected dataset, supervised analysis showed good separation of the two classes with a higher Q2 score than the uncorrected dataset. The model suggests overfitting however, due to the low Q2 score (Figure 45). Poor model performance is further revealed in random forest analysis, where top ranking features were identified in culture negative samples as well as samples containing *E. coli*. No *E. coli* containing samples were identified as such by the prediction model, resulting in a class error of 1, meaning the top ranking features identified were not significant in classification or were not detected in many samples (Figure 46).

Volcano plot analysis of fold change and FDR corrected *P* values revealed only three biomarkers significant in *E. coli* classification, compared to 234 identified in the previous LA-REIMS pellet analysis method directly from 96 well plates (Figure 47). Although the three identified significant biomarkers were also detected in the previous experiment, they ranked poorly compared to other more significant biomarkers, based on their FDR corrected *P* value.

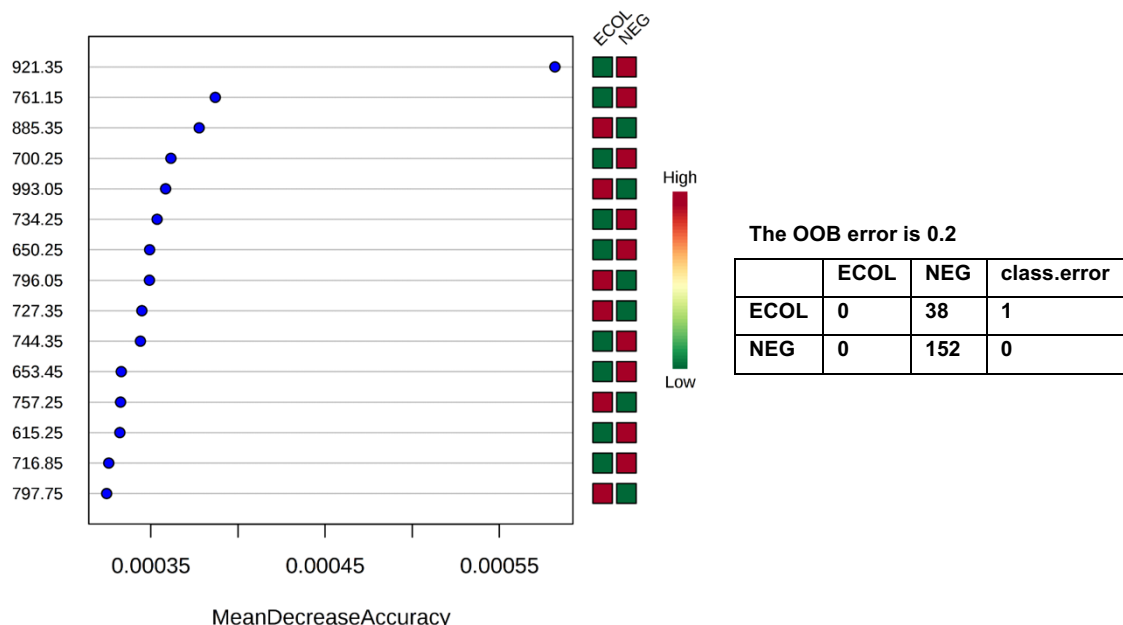


Figure 46. Clinical urine pellet LA-REIMS, from microcentrifuge tubes, random forest analysis, 600-1000 m/z range, white blood cell positive samples only. Misclassification of *Escherichia coli* (ECOL) containing samples is high, represented by a total class error. Top-ranked biomarkers ranked by mean decrease accuracy, split between *E. coli* containing samples and culture negative (NEG) samples.

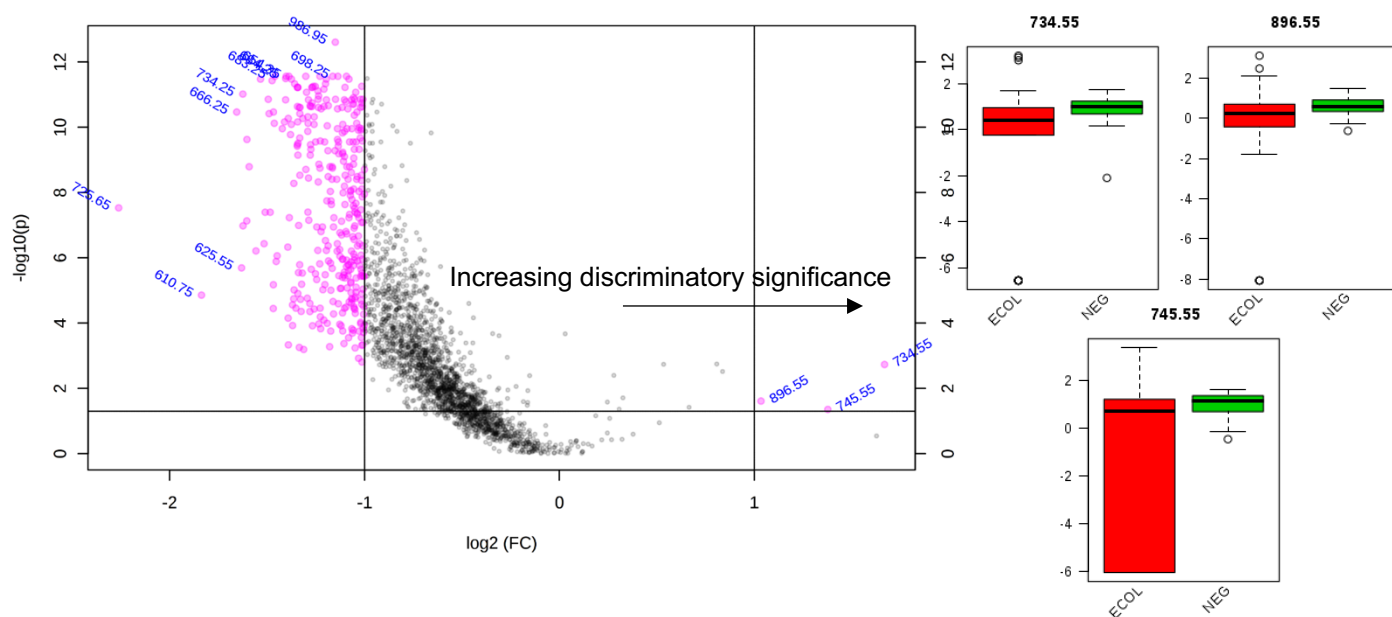


Figure 47. Volcano plot of log₂ fold change (x-axis) and -log₁₀ FDR corrected P value scores (y-axis) for culture positive *Escherichia coli* (ECOL) containing and culture negative (NEG) clinical urine sample pellets analysed from microcentrifuge tubes with LA-REIMS, white blood cell positive samples only. Features appearing to the top right of the plot have both high fold change and low P score in *E. coli* containing samples, suggesting they are class specific and significant in classification. Only three biomarkers were identified as significant in *E. coli*, with very low fold change and FDR corrected P value scores, reflected in box plots which show low, but statistically significant increase in significance of the three biomarkers can be seen in *E. coli* containing samples verses culture negative.

The random forest selection algorithm did not yield the same significant biomarkers identified in volcano plot analysis. Univariate ROC curves of two of the top ranking biomarkers for *E. coli*, from random forest, show poor model robustness through a wide 95% confidence interval margin and low AUC values (Figure 48). Additionally, the random forest tool ran to generate ROC curves yielded a different set of biomarkers than the previous tool. A validation model was created by inserting 10 unlabelled samples into a random forest selection algorithm based on the top ranking 15 biomarkers, with 100x cross validations. In the overall unlabelled sample prediction, five samples were misclassified (Table 5) with an overall predictive class accuracy of 49% across all cross validations (Supplementary 14). Model robustness was further demonstrated to be poor by a low ROC curve analysis AUC value of 0.461 and a wide 95% confidence interval margin and high *P* value of 0.618 over 1000 permutation times (Figure 49).

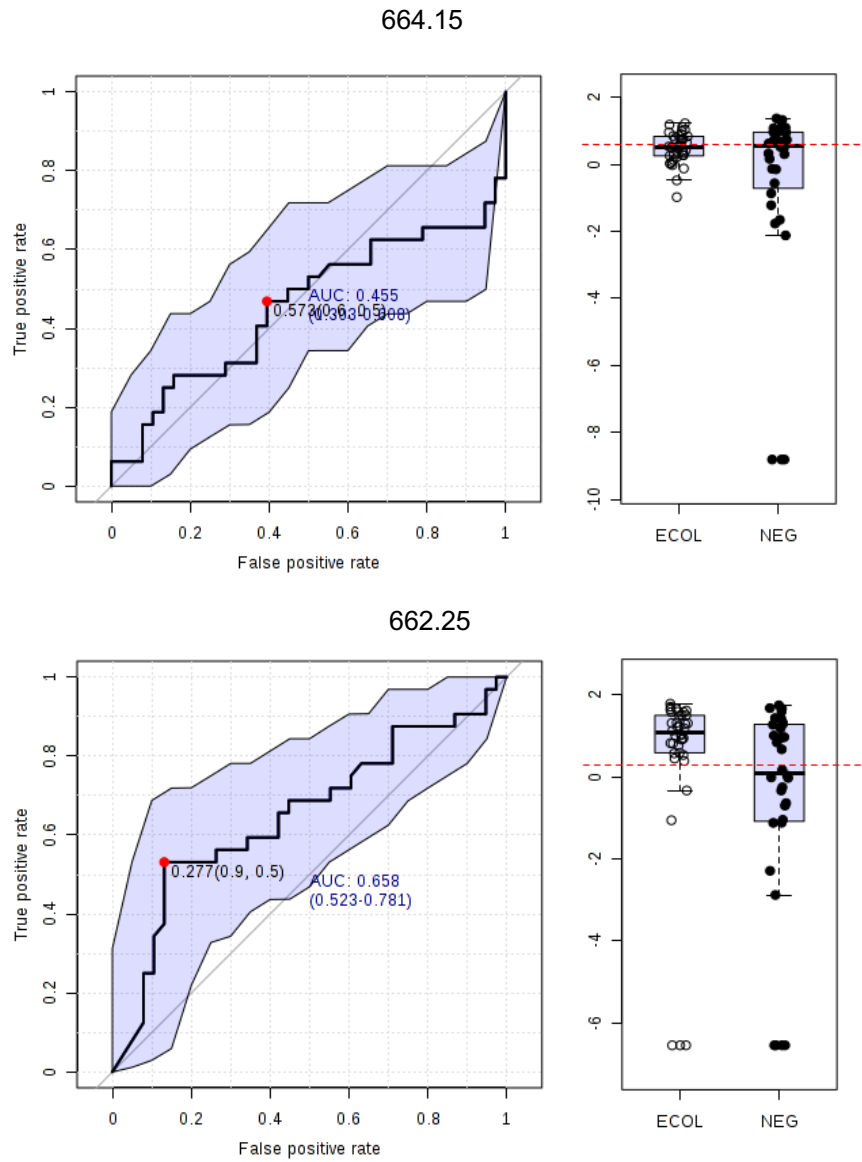


Figure 48. Univariate ROC curves of the top two most important discriminatory features in *Escherichia coli* (ECOL) classification, identified through multivariate random forest selection algorithm. The 95% confidence interval is shown in blue. High area under the curve (AUC) values and a curve approaching the upper left corner of the plot indicate strong performance of the biomarker in discriminating between *E. coli* containing and culture negative clinical urine samples.

Table 5. Clinical Urine 2 overall unlabelled sample prediction matrix. 10 samples were removed from the classification model and inserted into a validation model with their class label removed. A random forest prediction algorithm was ran to assign a class to the unlabelled samples. The overall result matrix following 100 cross validations results in five samples misclassified. ECOL = *Escherichia coli*, NEG = culture negative.

Name	Probability	Predicted Class	Actual
4	0.513	NEG	ECOL
7	0.633	ECOL	ECOL
12	0.527	NEG	ECOL
16	0.607	ECOL	ECOL
28	0.533	ECOL	ECOL
47	0.657	ECOL	NEG
54	0.51	ECOL	NEG
59	0.517	NEG	NEG
62	0.58	NEG	NEG
68	0.707	ECOL	NEG

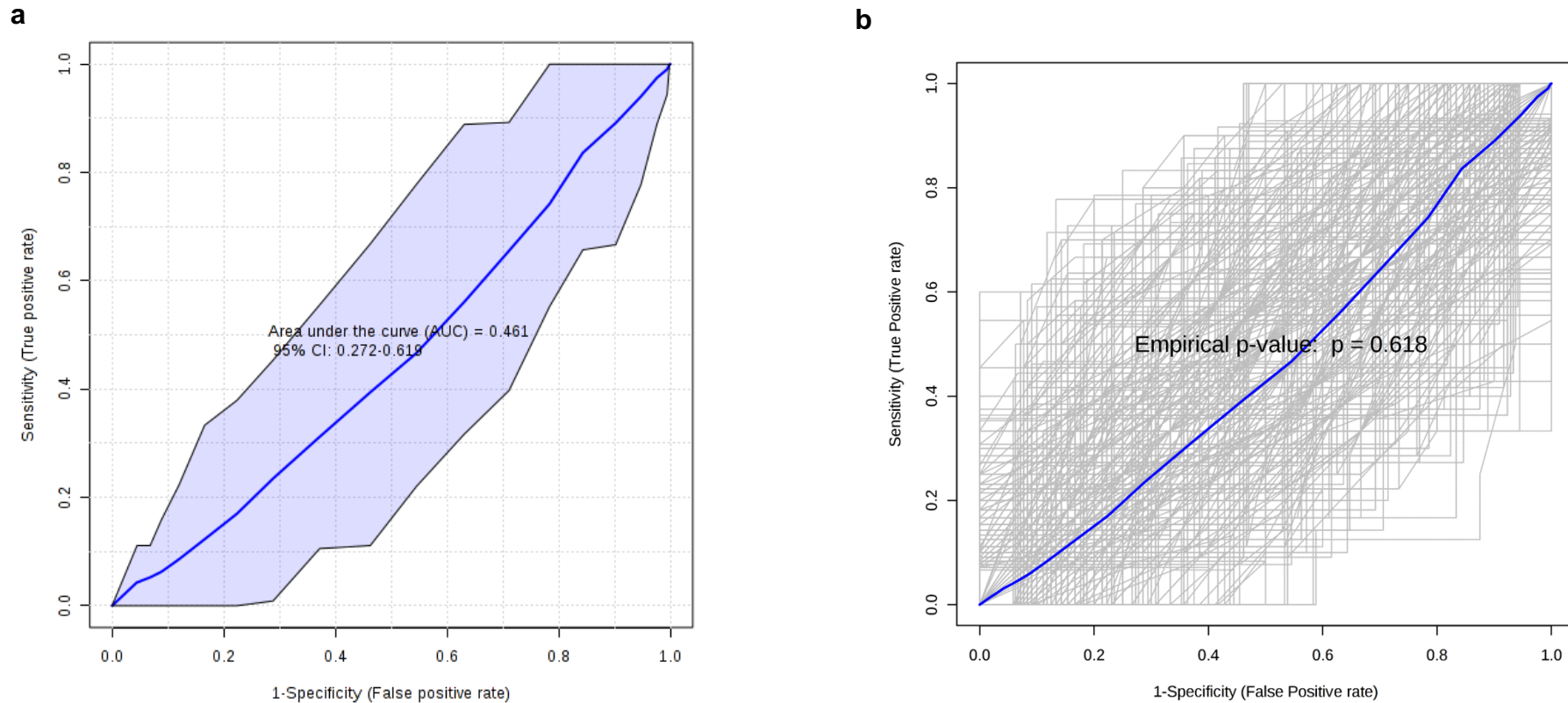


Figure 49. Univariate ROC curve of model prediction performance through multivariate random forest selection algorithm (a), and ROC curve permutation test (b). 100 cross validations were performed to generate the model. The 95% confidence interval is shown in blue. High area under the curve (AUC) values and a curve approaching the upper left corner of the plot indicate strong performance of the model in classifying 10 unlabelled samples. A sampling distribution was built by resampling the data used in the multivariate random forest selection algorithm with 1000 permutation times to assess the AUC performance measure. The high P value displayed on the permutation test (b) indicates the AUC values from each permutation were statistically different, and the model is not robust

5.6 Biomarker identification

The top five significant biomarker mass bins for *E. coli* containing clinical urine samples, identified through random forest analysis, t-tests and volcano plot rankings of FDR corrected *P* values/fold changes, were recorded. Accurate masses for each biomarker were determined, and fragmentation patterns were obtained through tandem mass spectrometry. Precursor and product ions were inputted to both LipidMaps and Metlin databases to obtain lipid identifications for significant biomarkers (Table 6). Database hits with a mass deviation of >10ppm, as well as unlikely or impossible lipid classes originating from bacteria, were discarded. Three of the five biomarkers were lipids of the phosphatidylglycerol (PG) class, with the two remaining being a phosphatidylethanolamine (PE) and phosphatidic acid (PA), likely originating from the cell membrane phospholipid bilayer of *E. coli*. All identified biomarkers were present in mass spectra as deprotonated [M-H]⁻ adducts.

The 747.45 *m/z* (747.51 accurate mass) biomarker, matched as PG (16:0/18:1) through nine product ion matches, also returned a Metlin experimental data match (Figure 50). Although other biomarkers were not identified as experimental matches in Metlin, *in silico* matches, along with product ion matches in LipidMaps and literature searches referencing lipid classes in bacterial membranes are sufficient for identification for the purposes of this investigation.

Table 6. Significant biomarker databases matches for *Escherichia coli*, analysed from clinical urine samples. Identified biomarkers were fragmented using MS/MS and their product ions inputted to LipidMaps and Metlin databases for phospholipid matches. PE = phosphatidylethanolamine, PG = phosphatidylglycerol, PA = phosphatidic acid.

Biomarker mass bin <i>m/z</i>	Accurate measured mass <i>m/z</i>	Monoisotopic mass <i>m/z</i>	Mass deviation ppm	Lipid class	LipidMaps product ion hits	Metlin
688.45	688.4940	688.4923 [M-H]-	2.5	PE (16:0/16:1) / (16:1/16:0)	452.2783 loss of sn2 acyl chain 255.2330 sn1 RCOO- ion 253.2173 sn2 RCOO- ion	<i>in silico</i> match C37H72NO8P MID 60365
733.55	733.4996	733.5025 [M-H]-	4.0	PG (16:0/17:1) / (17:1/16:0)	483.27 loss of sn2 acyl chain 465.26 neutral loss of sn2 RCOOH group 403.23 neutral loss of sn1 RCOOH group and glycerol 391.23 neutral loss of sn2 RCOOH group and glycerol 267.23 sn2 RCOO- ion 255.23 sn1 RCOO- ion 153.00 glycerol-3-phosphate ion with loss of H2O	<i>in silico</i> match C39H75O10P MID 79733

Biomarker mass bin <i>m/z</i>	Accurate measured mass <i>m/z</i>	Monoisotopic mass <i>m/z</i>	Mass deviation ppm	Lipid class	LipidMaps product ion hits	Metlin
747.45	747.5141	747.5181 [M-H]-	5.4	PG (16:0/18:1) / (18:1/16:0)	509.29 loss of sn1 acyl chain 491.28 neutral loss of sn1 RCOOH group 483.27 loss of sn2 acyl chain 465.26 neutral loss of sn2 RCOOH group 417.24 neutral loss of sn1 RCOOH group and glycerol 391.23 neutral loss of sn2 RCOOH group and glycerol 281.25 sn2 RCOO- ion 255.23 sn1 RCOO- ion 227.03 glycerophosphoglycerol ion –H2O	Experimental match C40H77O10P MID 40844
773.45	773.5330	773.5338 [M-H]-	1.0	PG (18:1/18:1)	509.29 loss of sn1 acyl chain 491.28 neutral loss of sn2 RCOOH group 417.24 neutral loss of sn1 RCOOH group 281.25 sn1 RCOO- ion 153.00 glycerol-3-phosphate ion with loss of H2O	<i>in silico</i> match C42H79O10P MID 40847
782.65	783.5960	783.5909	6.5	PA (18:1/24:1) / (24:1/18:1)	519.35 loss of sn1 acyl chain 501.34 neutral loss of sn1 RCOOH group 281.25 sn1 RCOO- ion	No hits

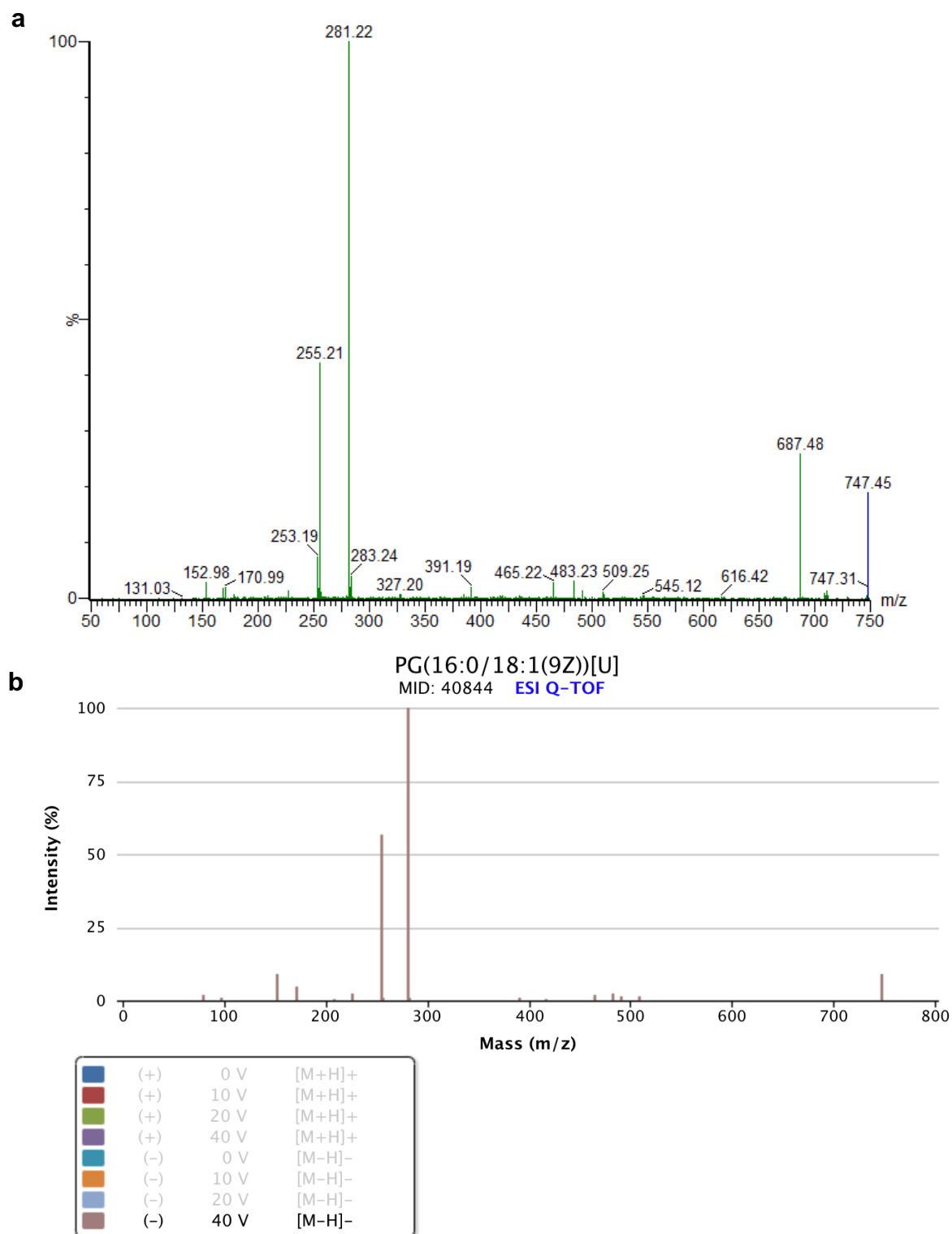


Figure 50. Mass bin 747.45 m/z (PG 16:0/18:1) (Accurate mass 747.52 m/z [M-H]⁻) biomarker product ion spectra from an *Escherichia coli* positive clinical urine sample (a) and Metlin experimental database hit METID40844 (b). Product ion spectra was obtained through quadrupole isolation of the 747.45 m/z peak and a collision energy of 36 eV. PG = phosphatidylglycerol.

5.7 Discussion

5.7.1 Bipolar forceps:

The development of a liquid analysis methodology and successful REIMS analysis of clinical urine samples supports the hypothesis that REIMS applications can be expanded to direct from clinical sample analysis. Bipolar forceps REIMS analysis of spiked urine samples achieved good separation of species in both unsupervised and supervised analysis methods, indicating that the methodology and setup allows detection of species-specific biomarkers above the background spectra of host urine. Species-specific biomarkers identified here were shared with biomarkers identified in the mixed culture analysis experiment (4.1: Species-Specific Biomarkers), however there was little similarity in the top most significant biomarkers. Healthy urine contains a plethora of cells and metabolites, such as epithelial cells, erythrocytes, leukocytes, bacterial cells from the microbiota, lipids, amino acids and other low molecular weight compounds, which although present in low concentrations contribute to the generation of a background spectra with MS analysis (228). This dissimilarity between top ranking biomarkers is therefore likely due to the background spectra produced from urine analysis effecting the relative intensities of biomarkers, and therefore adjusting their significance in speciation. Biomarkers identified in pure cultures were also identified in spiked urine samples, demonstrating that despite the presence of a background urine spectra REIMS is capable of detecting these, and indicating the potential of the technique in clinical sample analysis.

Biomarker intensity was found to decrease in correlation with sample concentration, as shown in Figure 29, signifying a relationship with microbial cell count. The limit of detection of REIMS was not met in this experiment, as biomarkers were detected in the most dilute samples, equating to 10^5 CFU/mL, of *S. aureus* and *P. aeruginosa*, and 10^6 CFU/mL in *E. coli*. For samples spiked with *E. faecalis*, biomarkers were not detected following dilution past 10^9 CFU/mL. The intensity of the top ranking biomarkers for *E. faecalis* were one to three orders of magnitude less at undiluted concentrations than that of the other species, which provides an explanation as to why detection did not occur at relatively high concentrations. Since total numbers of biomarkers discovered

for *E. faecalis* is also lower than other isolates, it can be assumed that the range of metabolites discovered by this method of REIMS is lower for the organism than for others. Objective ii was met in this experiment as a methodology for liquid analysis was evaluated, and detection of biomarkers was successful. Although bipolar forceps REIMS could be utilised in the analysis of clinical urine samples, the method is not scalable to a high-throughput technique due to impracticalities of automating forceps. Therefore, the automated REIMS system was developed in 5.2: LA-REIMS Method Development to attain a commercially viable urine analysis methodology.

5.7.2 LA-REIMS Method Development

Detection of species-specific biomarkers over urine background spectra was shown to be possible in the previous section. Here, a methodology was developed to transfer the technique to clinical sample analysis. Ambient ionisation techniques, such as DESI and Paper Spray MS, have been successfully utilised in the analysis of dried blood spots (DBS), with applications in therapeutic drug monitoring, biomarker discovery and metabolomics (229-231). The methodology of DBS analysis is well established and is extremely simple, involving no sample preparation or separation other than aliquoting around 15 μ L of sample onto cellulose based paper, such as standard filter paper, allowing to dry, and analysing directly by either adding solvent to the paper for Paper Spray MS or spraying charged solvent droplets at the sample in DESI. The method is preferred for many applications due to its cost effectiveness, small sample volume required and analyte stability in dried form allowing storage for long periods of time with no adverse effects (230). In Paper Spray MS, low limits of detection enable analysis of therapeutic drugs in the 0.5 nG/mL range (231). More recently, analysis of dried urine spots (DUS) for therapeutic drug identification has been conducted with ESI-MS and LC-MS/MS, following an extraction and suspension protocol (232, 233).

DUS analysis was investigated with REIMS on cellulose filter paper in order to determine if the technique could be applied to lipid biomarker analysis. Nitrocellulose membranes were also investigated as an alternative to paper, as slower absorption into the membrane allowed for moist membrane analysis rather than the dry analysis conducted on the paper.

Paper analysis showed high spectral richness through high levels of noise from cellulose polymer peaks, equating to low spectral quality. Conversely, nitrocellulose analysis showed lower spectral richness, but low spectral noise and higher quality scores. The high variability observed in quality scores was likely due to the membranes not readily absorbing sample aliquots, meaning the moistness of the analysis zone was variable between samples. Although nitrocellulose membranes were selected as an alternative analysis substrate based on the slower absorption rate, difficulty in fine positioning of the laser resulting in the inability to precisely position the urine sample, membrane and laser beam for consistent analysis. This may be surmountable in the future with improvement of laser control software improvement of the spacial resolution.

Despite the presence of polymer peaks in paper analysis and variable spectral quality in nitrocellulose analysis, some significant biomarkers were identified in both methods. Although top-ranking significant biomarkers identified in the bipolar forceps REIMS analysis of spiked urines experiment were also detected in the two methods investigated here, they were not identified as significant, based on their t-test/fold change volcano plot rank of FDR corrected P value scores and mean decrease accuracy scores from random forest analysis. This suggests that these analysis methods, whilst still having potential for species detection in urine, would yield a different database of significant biomarkers which would not be transferrable with the previous method. Since the background spectra of filter paper is of relatively high intensity, this may act to suppress biomarkers identified in previous experiments. In addition, as droplets of spiked urine added to nitrocellulose membranes did not disperse evenly, analysis was sometimes performed on areas of membrane not fully saturated with sample. This may also contribute to inconsistent biomarker intensity and quantity. Some biomarkers were detected across all sample dilutions, including the most dilute sample equating to $\approx 10^2$ CFU/mL for *E. coli* and *P. aeruginosa*. This is noteworthy as the threshold for a clinical UTI is generally accepted as 10^5 CFU/mL, although a threshold of 10^2 CFU/mL has been proposed for known urinary pathogens, which REIMS is able to meet (223).

While biomarkers were detected at dilute concentrations, speciation was poor across the two methods, likely due to the low biomass of bacterial cells negatively impacting

biomarker peak intensity even in the most concentrated samples. For application to clinical urine samples, improved speciation would likely be seen through concentration of biomass via centrifugation and analysis of urine pellets. Further, a centrifugation and discarding of supernatant step can be easily incorporated into a high throughput and automated workflow, whereas aliquoting urine samples onto filter membranes is tedious and relatively time consuming. This methodology is explored in the Clinical Urine 1 & 2 experiments.

This experiment further shows that REIMS is capable of detecting species-specific biomarkers over the background spectra of urine, in multiple analysis methodologies, and meets objective ii for developing a method for analysis of liquid samples.

5.7.3 Clinical Urine Experiments

Prior to analysis of clinical urine specimens, optimisation of laser power and mode was repeated due to the acquisition of a modified laser probe allowing analysis of pellets from 96 well plates and microcentrifuge tubes, which was previously not possible. At wattages >2 in continuous mode, overheating of biomass caused splattering of pellets, resulting in poor quality spectra and potential contamination of the laser lens as well as neighbouring plate wells. Splattering also permitted laser fire to reach the bottom of plate wells, where contamination of mass spectra by polymer peaks originating from plastic plate material was seen. Splattering was not observed in pulsed mode and a power of 1.5 W was chosen for pellet analysis since a positive balance of spectral richness and quality was attained at this level.

5.7.3.1 Clinical Urine 1

In LA-REIMS analysis of urine sediment pellets from 96 well plates, speciation with both unsupervised and supervised discriminant analysis was poor. Classification of culture negative or positive samples performed better, explained by analysis algorithms identifying biomarkers non-species-specific, but shared by all species examined in this experiment. The structure of the bacterial cell membrane, although lipid-class heterogeneous to an extent between genus and species, contains many conserved

features and molecules. For example, the common glycerophospholipid classes PE, PG and PA are seen throughout bacterial genera, and so classification based on these conserved molecules enables them to perform as ideal biomarkers in a simple culture positive or negative determination for urine samples (14).

Removal of samples other than those with *E. coli* as the causative agent improved classification further and resulted in the identification of 35 biomarkers significant in the classification of *E. coli* containing samples. Samples containing species other than *E. coli* therefore may produce poorer quality spectra with low intensity biomarkers using this REIMS method. Some biomarkers identified as significant were also identified in previous experiments, such as 688.45, 747.55, 691.45 and 719.45 (± 1.0) *m/z*, identified as significant in bipolar forceps REIMS and nitrocellulose/filter paper LA-REIMS. Three of the most significant biomarkers in classifying *E. coli* samples were also of top significance in identifying culture positive samples, suggesting *E. coli* biomarkers dominate the spectra, whilst again suggesting biomarkers for other species are of low significance and intensity.

Some culture positive clinical urine samples were reported to have a low WBC count (<100 cmm) by the ICHNT microbiology laboratory. The presence or absence of WBC is not a steadfast indicator of UTIs, as leukocytes may originate from inflammation sites outside of the urinary tract or persist beyond resolution of a UTI (234). However, a strong positive correlation exists between leukocyturia and bacteriuria, and an increasing WBC count is also associated with a higher bacterial load in urine, as well as serving as a clinical marker considered by biomedical scientists and physicians when determining the presence of a UTI (235). Therefore, in sample screening, a high WBC count and a culture negative status suggests a false negative result, whereas a low WBC count and a culture positive sample status suggests a false positive result, potentially from culture of a member of the urinary tract microbiota or a sample contamination event. Clean catch urine sampling, the recommended collection method in the UK, has up to 26% of specimens contaminated, and up to 47.5% contamination rate in samples obtained from children (236, 237). Follow up culturing of samples and consideration of clinical factors is thus required to reduce false positives.

Determination of ≥ 100 cmm WBC by a urine cytometer or automated microscopy is used in practice in the ICHNT and other clinical laboratories as a marker of infection (235). In cases where a culture positive sample is reported but a < 100 cmm WBC reading is obtained, the laboratory will consider the sample as culture positive until selective or differential culturing is attempted, recommend a new sample be collected or discard the result as inconclusive. Consideration of clinical indicators by a biomedical scientist is required, such as determination of erythrocytes, squamous cells, crystals and casts, and consideration of manual microscopy, patient history, risk factors and symptoms by a physician, in order to definitively class a sample as positive or negative UTI (137). In cases where no bacterial cells are detected through screening, but a ≥ 100 cmm WBC reading is obtained, samples are classed as positive and culturing is performed. This ambiguity, and non-access to clinical decisions made by physicians in sample metadata, means some of the samples analysed in this dataset may be mislabelled. For this reason, a subset of data was constructed containing only *E. coli* positive samples, with WBC counts > 100 cmm, for additional data analysis.

In the WBC corrected data subset, improved separation and clustering of sample class was seen in unsupervised and supervised discriminant analysis, with a stronger model quality score. In random forest analysis, improved classification results were seen based on the top 25 most significant biomarkers, however class specific error for *E. coli* was still high. Volcano plot visualisation of t-test and fold change scores revealed 7x more significant biomarkers significant in *E. coli* classification when compared to the non WBC corrected dataset. This suggests that previously, samples with a low WBC count were mislabelled as UTI positive with *E. coli* as the causative agent, when it is likely an *E. coli* contaminant of very low CFU/mL was present in the sample and therefore distorted the previous model. In a validation model based on the top 15 most significant biomarkers, results were greatly improved, with 73% predictive class accuracy achieved across all cross validations. This implies that although significant based on their FDR corrected *P* value, there were likely some lower ranking biomarkers which performed less well than the top 15 in the previous random forest analysis in terms of classification. It was not possible to determine the biomarker limit of detection in this experiment due to logistical and physical limitations with sample acquisition from the ICHNT

laboratory. Samples were acquired the day after processing by the clinical lab, meaning an aliquot had already been taken from the variable vial volume, with usually around 2-5 mL remaining. ICHNT required 2 mL be returned from each sample to the lab at the end of each work day in the event follow up tests would be required. This meant some samples were unusable for this experiment, and in others 2-3mL was available. Since 2 mL was required for centrifugation to obtain a pellet, the majority of sample volume was utilised, leaving none remaining for CFU enumeration plates. In sample metadata obtained from the clinical laboratory, a crude growth indicator of one to three stars was reported rather than a definitive CFU/mL calculation, due to labour and time constraints and the clinical insignificance of a definitive count when automated screening techniques and follow up culturing is performed.

5.7.3.2 Clinical Urine 2

A second set of clinical urine samples were obtained for analysis from microcentrifuge tubes, rather than 96 well plates, to determine if a superior cell pellet could be recovered from higher centrifuge speeds, leading to improved classification. 67% of samples were *E. coli* positive, with insufficient species replicates obtained for statistical analysis to be carried out on other species.

In the whole dataset, unsupervised and supervised discriminant analysis performed poorly, as in Clinical Urine 1, with PCA showing no class separation and PLS-DA showing overfitting and a poor model quality score. A WBC count corrected data subset was generated as in Clinical Urine 1, however no significant improvement was observed in discriminant analysis and random forest analysis also performed poorly, with no *E. coli* samples classified correctly and only one biomarker identified for *E. coli*. Similarly, only 3 biomarkers significant in classification of *E. coli* containing samples were identified from t-test/fold change volcano plot, compared to the 235 identified in Clinical Urine 1. In the validation model, random forest yielded a different set of biomarkers than the random forest algorithm ran earlier, further demonstrating a weak and unreliable model. Although a 49% predictive class accuracy was observed, this value is likely overoptimistic, as during significant feature selection for the random forest classification model only one biomarker was identified. Instead, features were selected

based on the overall AUC rankings within the validation model, rather than their significance rankings in the previous Clinical Urine 1 experiment. This means that the biomarkers the prediction model is based upon likely perform well in classification for this dataset but are not applicable or significant when transferred to other datasets containing the same sample classes.

Due to poor model robustness, overfitting and failure to identify common significant biomarkers between different statistical tests, the methodology used in this experiment is not optimal. Although a pellet of greater biomass was obtained through this method, the angled rotor of the higher speed centrifuge resulted in pellets not being positioned at the bottom of the microcentrifuge tube, but slightly on the wall. This meant positioning of the laser beam for optimal vaporisation was difficult and so pellets were manually moved to the bottom of microcentrifuge tubes with sterile pipette tips. The resulting irregular surface of the pellet may have affected laser ablation, resulting in inconsistent vaporisation. Since the automated Tecan system is programmed to analyse each plate well or tube for 2 seconds before moving on to the next sample, this may not have been sufficient time to acquire enough analyte, affecting the total ion current. The modified laser probe consists of an extended aspiration rod protruding from the lens, meaning the generated vapour has a further distance to travel, both in air and transport tubing, before reaching the mass spectrometer inlet and the lens is a greater distance from the sample surface (Figure 9). The greater distance travelled by the vapour provides an increased opportunity for ion clusters to adhere to tubing walls, as well as an increased potential for vapour to dissipate before reaching the aspiration rod. The LA-REIMS setup, including the laser module and automated Tecan system, reside entirely within a class 2 biological safety cabinet, where airflow is constant and potentially turbulent due to moving parts. Turbulent or constant airflow may disrupt vapour entry to the aspiration rod, especially as the REIMS interface relies only on the native vacuum of the instrument for suction. Ion transfer from ion sources is directly linked to MS analysis sensitivity (238).

In Clinical Urine 1, the close spacing of wells of the plate and dimensions of the wells mean less movement of the laser module and a tighter fit of the aspiration rod,

potentially reducing vapour disturbance from turbulent air. Additionally, the shallower depth of the plate wells allows the aspiration rod to reach a few millimetres away from the sample pellet, whereas the longer length of the microcentrifuge tube means the aspiration rod tip is at least 10 mm away from pellet surface. It is possible that these factors effect TIC and biomarker discovery in Clinical Urine 2 more than Clinical Urine 1, providing an explanation for poorer performance. The REIMS interface and LA-REIMS system is a research prototype and under continuous improvement through feedback to WRC from the research group. It is likely that in future versions of the system these issues would be addressed, and data analysis would improve. Incorporation of a laminar flow device into other atmospheric pressure interfaces has improved ion transfer at distances up to 6 m by reducing ion losses (239). The introduction of a similar device to LA-REIMS could improve biomarker detection rates, similar to the venturi gas system utilised with the iKnife surgical device, on which microbiology REIMS is based on, where the MS instrument is often situated longer distances from the analysis point (240). Currently, however, data generated from the Clinical Urine 1 experiment yielded much better model robustness and over 234 significant biomarkers for classifying *E. coli*, therefore the previous experiment methodology is preferred to Clinical Urine 2.

5.7.3.3 Biomarker Analysis

Five of the most significant biomarkers for *E. coli* classification in Clinical Urine 1 were investigated. Analysis of biomarkers from other species was not conducted due to poor discriminatory power and low numbers of significant biomarkers for species other than *E. coli* in Clinical Urine 1, and the majority of samples obtained in Clinical Urine 2 containing only *E. coli*, with insufficient replicates of other species. Three of the five investigated biomarkers were lipids of the phosphatidylglycerol (PG) class, with the two remaining being a phosphatidylethanolamine (PE) and phosphatidic acid (PA), likely originating from the phospholipid bilayer of the cell membrane. In *E. coli*, phosphatidylglycerol phosphate synthase is essential for the biosynthesis of PGs, and mutants lacking the gene encoding the enzyme demonstrate an inability to translocate newly synthesised outer membrane proteins across the inner membrane, indicating the role of PG in the process (241, 242). PG is also required for cell growth, and substitution with phosphatidylinositol (PI) does not restore growth ability. Since PG is associated

with the inner membrane of *E. coli*, its role in forming a membrane associated complex is evidently critical for normal functioning of the cell (243). PE is a precursor to lipopolysaccharide (LPS) formation, with pyrophosphoethanolamine residues in particular synthesised following donation of the PE headgroup to a LPS receptor (244). Deficiency of PE within *E. coli* results in downregulated LPS synthesis leading to impaired adhesion, which is an important virulence factor critical to colonisation and biofilm formation (245, 246).

The membrane of *E. coli* predominantly comprises glycerophospholipids of the PE class at around 75% abundance, followed by PGs at 20% (247). It could be expected that significant class defining biomarkers for *E. coli* would therefore be numerous in the PE class and less so in other classes. Although only five biomarkers were considered here, *m/z* values of identified biomarkers identified in this experiment and not investigated through lipid database searches correlate with previous REIMS work on *E. coli*, in which PGs were the dominating lipid class identified by REIMS (220). It is postulated that the REIMS method favours ionisation of PGs over PEs through an unknown mechanism, since lipid class and carbon chain lengths are shared between this work and that conducted previously, with the majority of PG class. In particular, four of the five glycerophospholipids biomarker (688.45, 733.55, 747.45 and 773.45 *m/z*) lipid classes were also reported as present in clinical *E. coli* isolates analysed in the previous work with bipolar forcep REIMS (220).

5.7.4 Conclusions

Identification of glycerophospholipid biomarkers in *E. coli* containing clinical urine samples supports the hypothesis that REIMS can be used as a culture free identification method for clinical samples. Objectives ii, iii and iv have been met, in that a methodology for analysis of liquid samples has been achieved and applied to clinical urine samples, and several of the most significant biomarker classes have been identified. It has been shown through exploration of different REIMS analysis methods that the number of significant biomarkers produced for the same class can be highly variable, with many more biomarkers identified in spiked urine sample bipolar forceps REIMS than in clinical sample microcentrifuge tube LA-REIMS, for example. This is

likely due to the lower microbial load in clinical samples compared to spiked samples, and the possibility of spectral interference and suppression effects occurring with the associated lower intensity biomarkers. Significant biomarkers for *E. coli* were still identified in clinical samples, however, and classification of culture positive and negative samples was achieved, which is significant since 80% of all UTIs contain *E. coli* as the causative agent (96).

Data correction for WBC count values improved classification results, suggesting the presence of false positives and negatives in the uncorrected dataset which were skewing results. This highlights the ambiguity surrounding clinical UTI determination, as the presence or absence of bacterial cells is not indicative of infection and must be considered along with other clinical markers. Since between 65-80% of all urine samples received into clinical laboratories are culture negative and require no follow up testing, screening of samples to reduce workload is required (133, 248). LA-REIMS of 96 well plates could potentially provide a screening alternative to automated microscopy or cytometry methods, and act as a classification system for positive samples without the requirement for culturing.

In future work, conduction of a large-scale purpose-built study with access to greater volumes of urine would likely improve biomarker discovery and classification accuracy, as well as enabling determination of bacterial cell counts and therefore providing a CFU/mL detection limit for the technique. Since adequate species replicates were not achieved for inclusion in analysis other than *E. coli* in Clinical Urine 2, more samples should be collected to ensure sufficient replicates are obtained.

6 AMI-MS

Acoustic Mist Ionisation Mass Spectrometry is a novel technique developed by AstraZeneca in partnership with Waters (191). The technique couples a unique, automated ultrasonic sampling platform to a standard Waters Xevo QToF instrument. The technology was utilised in this experiment to compare the capability of REIMS and AMI-MS in detecting small metabolites and phospholipids of *K. oxytoca*, cultured in broth media supplemented with an array of 192 different carbon sources. Analysis of small metabolites may enable increased classification accuracy of clinical bacteria and provide further clinically relevant metabolic data, such as antimicrobial resistance. Two REIMS methods, filter paper and pellet LA-REIMS, were compared with the AMI-MS data to determine the optimal methodology.

Spectral profile analysis of the two LA-REIMS methods revealed significantly different spectral profiles in the metabolite 50-400 m/z range, but several shared peaks in the phospholipid 600-1000 m/z range (Figure 51). Filter paper LA-REIMS spectra was dominated by high intensity polymer contamination, likely originating from cellulose within paper fibres, which potentially masks smaller intensity biomarker peaks. AMI-MS spectral profiles were typically three orders of magnitude lower intensity than both REIMS methods in the metabolite 50-400 m/z range, and three-four orders of magnitude lower intensity in the phospholipid 600-1000 m/z range (Figure 52). Mean spectral richness was highest in filter paper REIMS for the metabolite region, although this was likely due to polymer contamination peaks. AMI-MS ranked second in spectral richness, but highest in spectral quality, again indicating the presence of polymer contamination in filter paper REIMS (Figure 53). Filter paper REIMS also displayed highest mean spectral richness scores in the phospholipid region, but manual inspection of spectral profiles shows signal domination by polymer contamination peaks. Pellet LA-REIMS is ranked second in spectral richness, with a tight clustering of box plot scores. Pellet LA-REIMS is also ranked first in spectral quality for the phospholipid region (Figure 53). AMI-MS performed relatively poorly in both spectral richness and quality in the phospholipid region, with a large range of quality scores observed between each sample.

Following normalisation of relative signal intensity, spectral overlay comparison of pellet LA-REIMS and AMI-MS analysis of both metabolite and phospholipid mass regions shows sharing of several common peaks between analysis methods, but with very different relative intensities (Figure 54). The vast majority of features are unique to each analysis method, as shown by scatter plots of mass bins verses FDR corrected P values, generated by non-parametric t-tests (Figure 55). With a $P = 0.05$ cut-off value, only eight features were shared between analysis methods in the metabolite region, and ten in the phospholipid region.

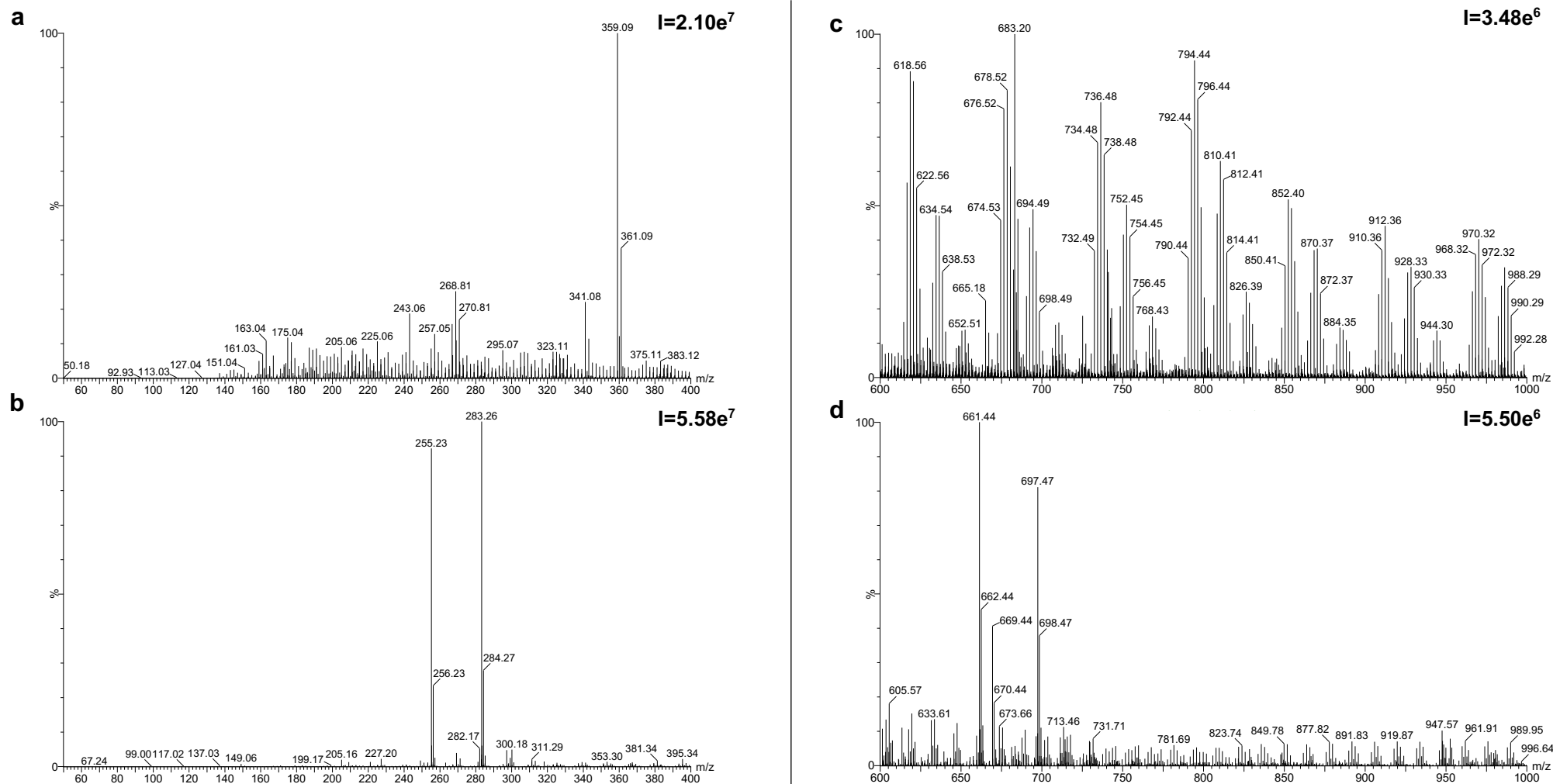


Figure 51. Example spectral profiles of filter paper (a,c) and pellet (b,d) LA-REIMS analysis of a *Klebsiella oxytoca* isolate cultured with a D-Serine carbon source. Filter paper REIMS spectra displays polymer contamination whilst pellet REIMS shows a superior signal:noise ratio. Richer profiles can be seen in the higher (600-1000 m/z) region. Lockmass correction and background subtraction have been applied.

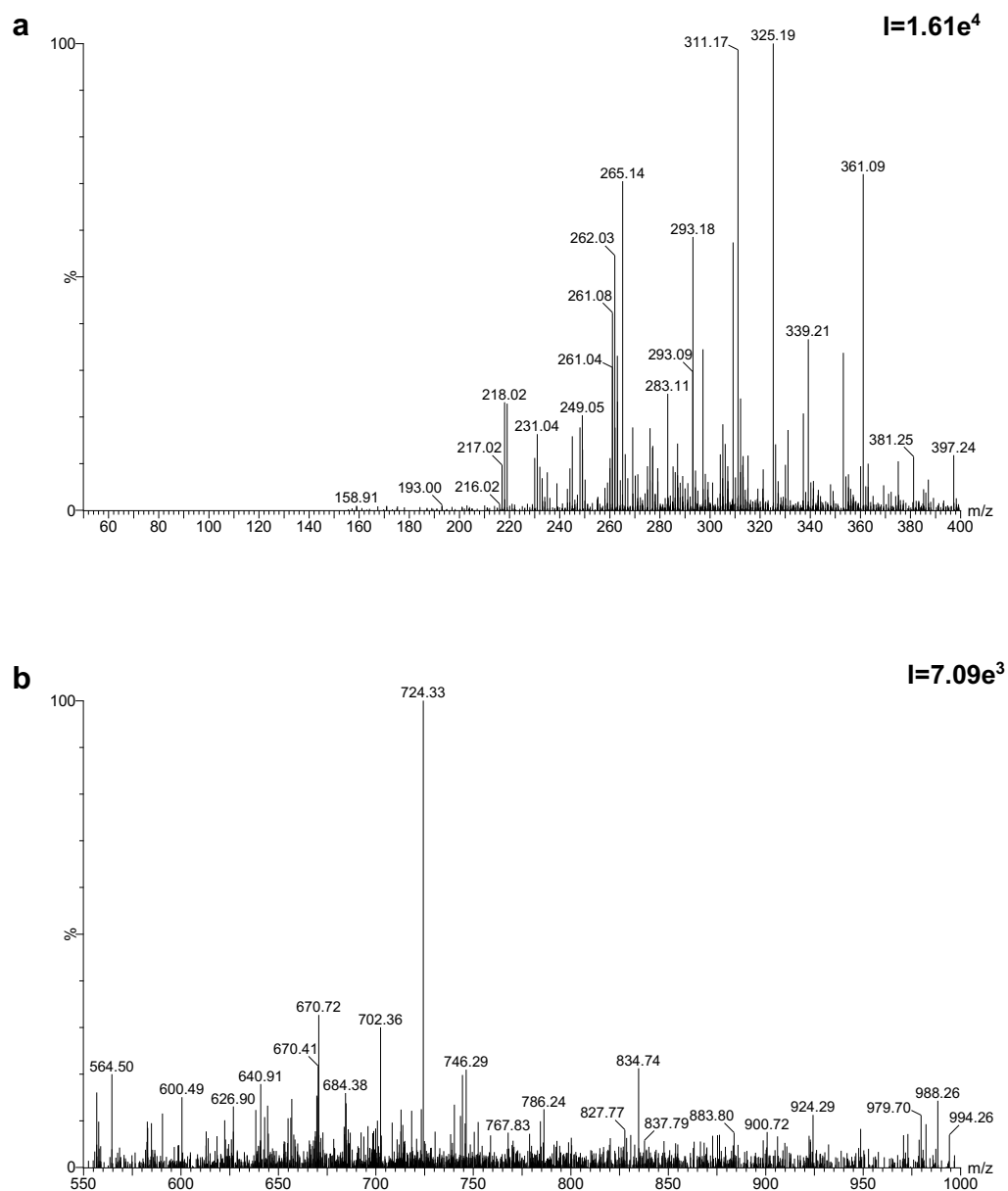


Figure 52. Example AMI-MS spectral profiles of the metabolite 50-400 m/z region (a) and phospholipid 600-1000 m/z region (b) of a *Klebsiella oxytoca* isolate cultured with a D-Serine carbon source. A richer spectral profile is observed in the lower mass region. Lockmass correction and background subtraction have been applied.

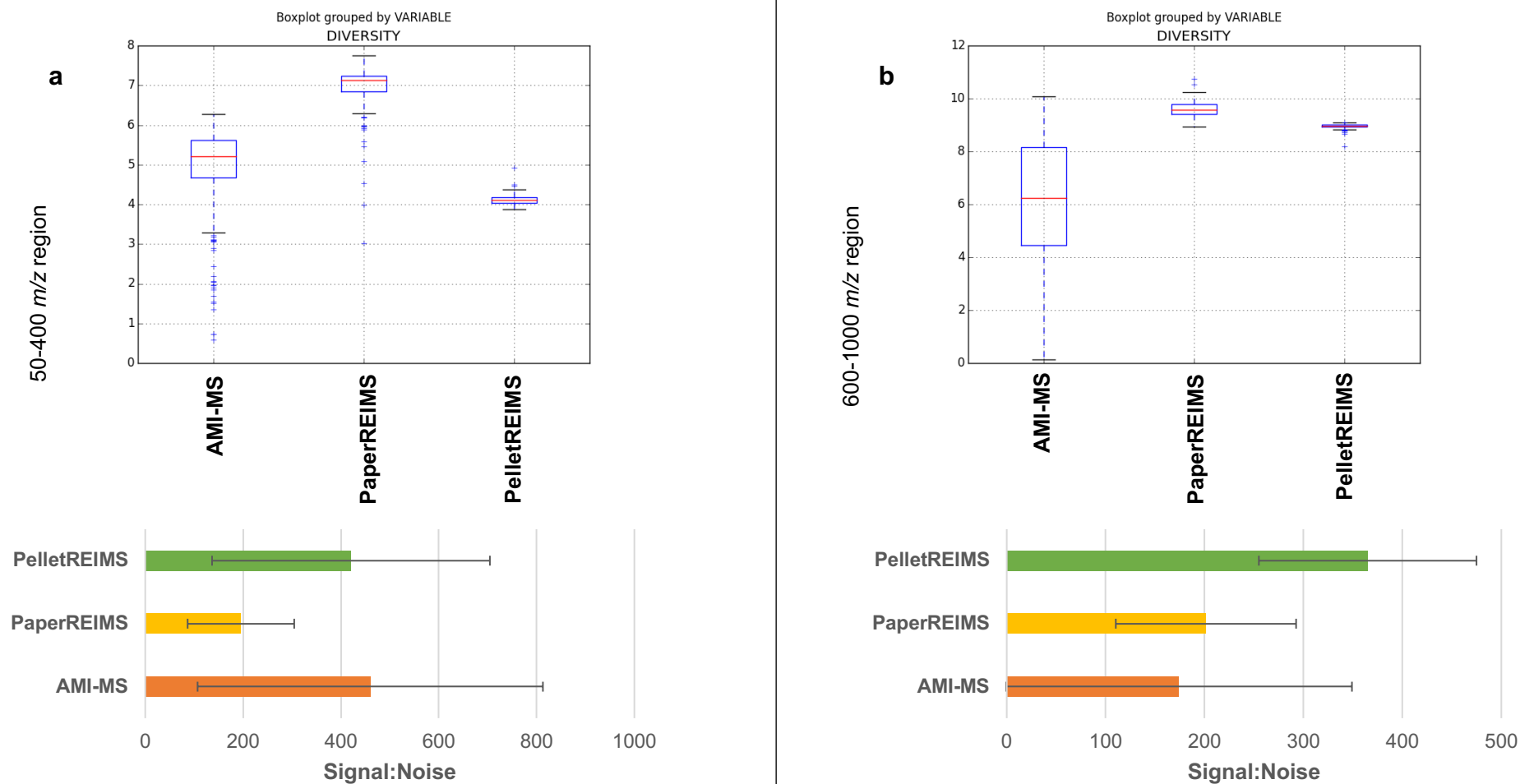


Figure 53. Shannon diversity plot and mean combined spectral quality scores of AMI-MS, filter paper and pellet LA-REIMS analysis of a *Klebsiella oxytoca* isolate of the 50-400 (a) and 600-1000 m/z (b) ranges. Spectral richness displayed through Shannon diversity is greatest in filter paper LA-REIMS, however the low spectral quality score indicates this is due to high background noise at both mass ranges. AMI-MS generates superior spectra in the low mass region analysis, whilst pellet REIMS is optimal for high mass region analysis.

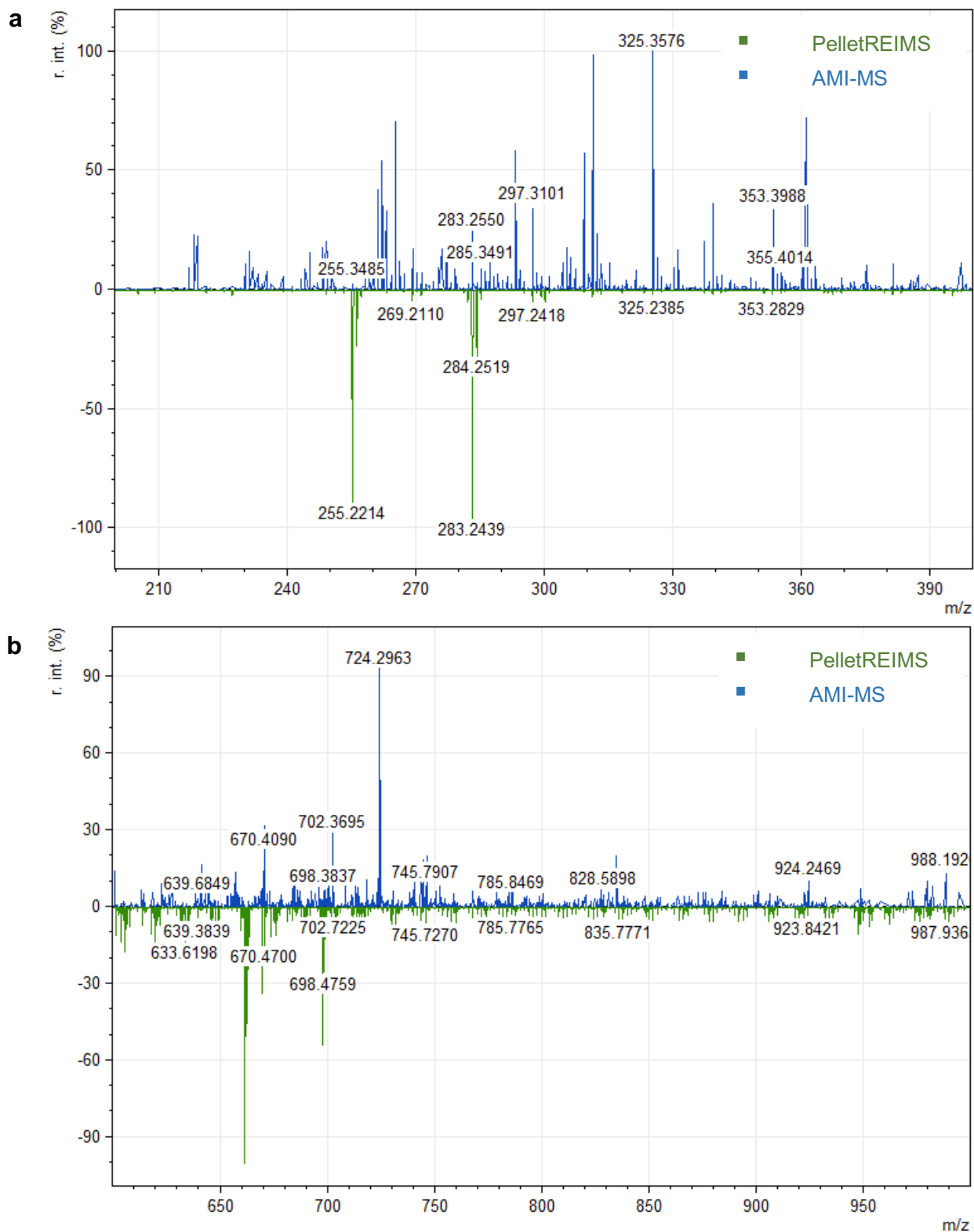


Figure 54. Mirrored spectral comparison of the 50-400 m/z metabolite region (a) and 600-1000 m/z phospholipid region (b) of a single isolate of *Klebsiella oxytoca* cultured with a D-Serine carbon source, analysed with AMI-MS and pellet LA-REIMS. Both methods show a similar spectral richness due to normalised relative intensity. AMI-MS spectra were of lower quality than LA-REIMS in the phospholipid region (b), whilst AMI-MS was of a higher spectral richness in the metabolite region (a). Several peaks were shared in both methods (labelled). Spectra are displayed with normalised relative intensity, with both lockmass correction and background subtraction applied.

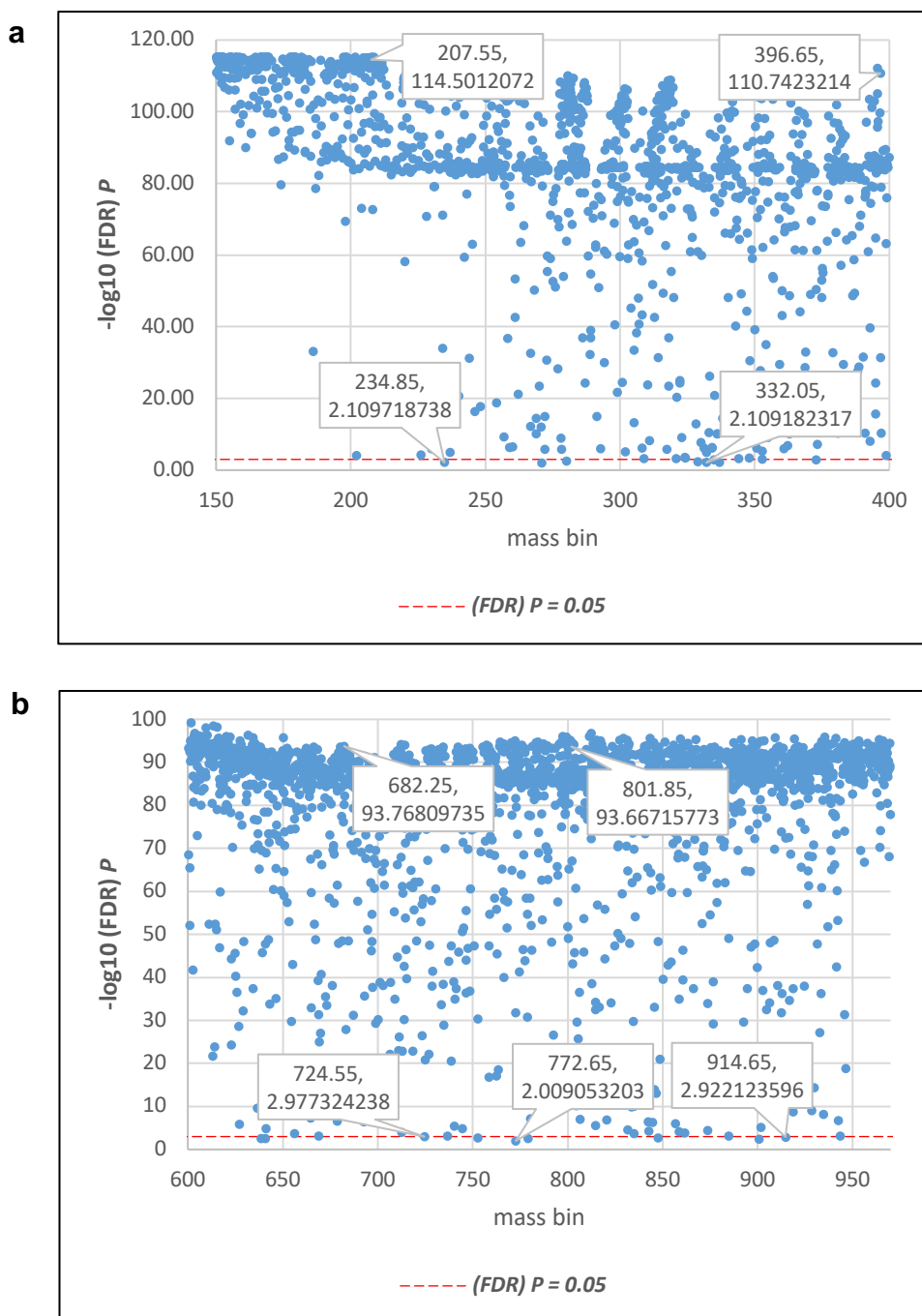


Figure 55. FDR corrected P values ($-\log_{10}$) versus mass bin scatter plot of the 50-400 m/z (a) and 600-1000 m/z (b) range of pellet LA-REIMS and AMI-MS analysed *Klebsiella oxytoca* cultured with a range of 192 carbon sources. FDR corrected P values were generated through a non-parametric t-test with a cut-off of $P = 0.05$. The majority of mass bins are above the P cut-off and therefore significantly different between the two MS analysis methods. Mass bins with FDR corrected P values < 0.05 are displayed below the red dash line, and indicate mass bins not significantly different between the two analysis methods. Pellet LA-REIMS and AMI-MS produce different mass spectra with little similarity.

6.1 Discussion

AMI-MS is in development by Waters and AstraZeneca as an alternative to REIMS for rapid, high-throughput analysis of liquid samples with no significant sample preparation steps. Here, the two techniques were compared to examine spectral profile changes between the two methods, using *K. oxytoca* as a model organism. This experiment forms part of a larger study screening metabolic activities and phospholipid changes within bacterial cell membranes cultured with different carbon sources, with an aim to apply one of the technologies to the investigation of microbial activity within tumours. This larger project is not presented here as it falls outside of the scope of this thesis, however the work completed meets the objective of expanding REIMS to new applications and works toward the culture free analysis of clinical samples.

K. oxytoca was chosen as the investigation isolate of interest as it is a clinically significant pathogen associated with a wide range of both community and nosocomial acquired infections, frequently isolated from mucus, blood, wound swabs and urine, with increasing incidence rates (249). In addition, isolates of *K. oxytoca* were readily obtained from the ICHNT laboratory due to high culture frequency, and the organism is easily grown in aerobic conditions within 24 hours. A rich spectral profile is produced through REIMS analysis which makes it ideal for comparison with other methods. Between the analysis techniques, only a few spectral features were shared and significantly different spectral profiles were produced for both low and high mass regions. This was expected due to the techniques employing a different ionisation method, however highlights that the methods are not directly comparable. In particular, AMI-MS appears to be better suited to analysis of small metabolites in the lower 50-400 m/z region, whilst pellet LA-REIMS is better suited to phospholipid analysis in the higher 600-1000 m/z region, due to their respectively higher spectral richness and quality scores within these ranges.

Metabolomic studies using mass spectrometry consider small intracellular or extracellular molecules of typically <400 m/z range, which together provide an outlook on the physiological state of cells and illuminate active biochemical processes (250). In

clinical microbiology, analysis of the metabolome can provide information on active antimicrobial resistance patterns and virulence factors, which can be important in the treatment of infection. Resistance mechanisms develop via gene mutations, horizontal gene transfer and selective pressure, and are accompanied by significant changes in gene expression and protein synthesis (251). Transcribed protein changes effect the overall metabolism of the cell, with a global impact of increased resistance to a particular antimicrobial drug. It is the changes in metabolism and thus metabolites which metabolome studies can detect, usually through high sensitivity LC-MS projects which provide a wealth of data, but require targeted method development, significant sample preparation and long run times with low throughput.

A study investigating metabolome differences between methicillin resistant (MRSA) and methicillin susceptible (MSSA) *S. aureus* strains revealed a catalogue of 210 metabolites through LC-MS analysis (252). This metabolic fingerprint allowed classification of MRSA and MSSA strains. In particular, UDP-N-acetyl-mannosaminouronic acid and UDP-N-acetyl-fucosamine were two discrimination-significant metabolites found to accumulate within MRSA strains, and are involved in cell capsule synthesis (252). Metabolome studies have also shown unique metabolite patterns associated with resistance in other clinically significant bacteria, including *Campylobacter jejuni* (253). The LC-MS methods employed here, however, are not transferrable to clinical laboratories, due to long run times and the relative simplicity of performing antimicrobial susceptibility testing on live cultures. Whilst never capable of approaching the separation or quantification power of LC-MS, AMI-MS could offer a quick-fire way of matching metabolic fingerprint profiles of bacteria to a reference database for screening or rapid identification of active pathways with clinical relevance. An optimal methodology would be to use REIMS and AMI-MS in a complimentary fashion, since throughput is high for both methods, sample preparation is minimal and low sample volumes are required. As 3 μ L is required for AMI and 40 μ L for LA-REIMS, the majority of original sample will remain for follow up testing or manual plate culture following aliquoting, as volumes in excess of 15 mL are routinely collected for blood, sputum and urine (254, 255). Since phospholipid class profiles are diverse and distinct between bacterial species, LA-REIMS analysis can be considered optimal for species

classification and diagnostic applications, whilst AMI-MS can be considered optimal for obtaining data on active metabolic pathways, for example antibiotic resistance mechanisms.

6.1.1 Conclusions

The experiment discussed here provides a proof of principle for the complementary nature of the two methods on a single bacterial isolate. Coupled together, the two platforms have potential as an all in one analysis technique providing a wide array of clinically significant data. Objective v has been met through this work, as it has been demonstrated that ionisation of *K. oxytoca* metabolites is possible with AMI-MS, and the strengths and weakness of the technique compared to REIMS have begun to be explored. In future work, LC-MS would ideally be utilised alongside REIMS and AMI-MS in the metabolome analysis of a broader range of microorganisms, including resistant and susceptible strains such as MRSA and MSSA. Culture based antimicrobial susceptibility testing would allow validation of results obtained. Tandem MS should also be employed to determine identities of the metabolites detected and this should be compared to previous studies for similarities in order to ascertain robustness of the technique.

7 Bovine Mastitis

Centrifugation of 2mL bovine milk in 24 well plates resulted in a surface lipid layer 3-5 mm thick, and a pellet at the bottom of the plate well. An experiment was conducted to determine whether analysis of the surface lipid layer or pellet was superior in classifying mastitis positive or negative samples, with modified cell line probe REIMS. Analysis of lipid layers was also completed using LA-REIMS, directly from 24 well plates. Parameters for analysis were previously determined in an optimisation experiment (3: Optimisation and Setup).

In cell line probe REIMS analysis, spectral richness boxplots of the 600-1000 m/z lipid region show that the sample pellet gives a richer mass spectrum than the lipid layer, although with a wide score range between samples (Figure 56). Spectral quality analysis revealed lipid layer mass spectra to have a much better signal:noise ratio than pellet spectra, meaning the high richness in pellet spectra is likely due to high background noise (Figure 56). Unsupervised PCA of the 600-1000 m/z lipid region of analysed pellets shows no separation of healthy verses mastitis positive milk samples, however supervised PLS-DA shows good class separation with a robust model Q_2 score (Figure 57). Biomarkers for mastitis positive samples were determined through t-test FDR corrected $P < 0.05$, resulting in 442 significant biomarkers found. The top twenty five most significant biomarkers, ranked in order of FDR corrected P value, are displayed in a heatmap (Figure 58). Analysis of the lipid layer resulted in much better separation of mastitis positive verses negative samples, in both unsupervised and supervised discriminatory analysis (Figure 57). Supervised PLS-DA also demonstrated a more robust model than for pellet analysis, represented by a higher Q_2 score.

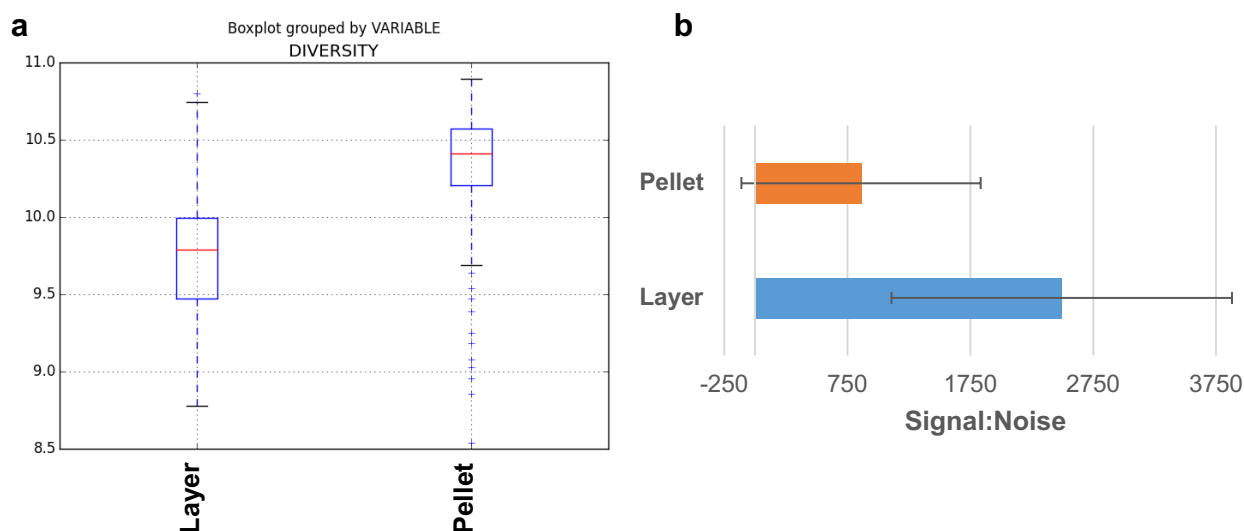


Figure 56. Shannon diversity box plot (a) and mean combined quality scores (b) of bovine mastitis sample pellet and layers analysed with cell line probe REIMS, 600-1000 m/z region. Spectral richness was greater in analysed pellets, however with a large score range. Spectral quality was greater in layer analysis, indicating a high level of background noise in pellet spectra.

Of the 141 mastitis positive samples received, 26% contained *Staphylococcus* spp. as the causative agent, 16% contained *Streptococcus* spp. and 14% *E. coli*. The remaining 41% of samples contained causative agents from 24 other genera, with 6% *Corynebacterium* as the most contributing agent. For the purpose of genera biomarker investigation, only the top three contributing genera were considered, as the remaining 24 were of low sample replicates for sufficient data analysis. These excluded genera were, however, included as “mastitis positive” contributors in the previous models. With lipid layer cell line probe analysed samples grouped into Gram-positive/negative and non-mastitis (“healthy”), supervised analysis shows good clustering of the two Gram classes and good separation of both from non-mastitis samples, with a robust model Q2 score (Figure 57). When samples were classed according to genus, however, clustering of classes appeared good but with poor separation, and a low model Q2 score indicated overfitting (Figure 57). A total of 250 significant biomarkers were found for mastitis positive samples lipid layer analysis, compared with 442 for pellet analysis. The top 25 most significant biomarkers for mastitis lipid layer analysis are displayed in both per sample and averaged heatmaps (Figure 58). Significant biomarkers were also identified for specific genera classes, unlike with pellet analysis (Figure 61).

In LA-REIMS analysis, only the surface lipid layer was analysed, due to superior model results obtained with cell line probe analysis and the difficulty of removing both the lipid layer and liquid sample to reach the pellet with the laser. Good class clustering of mastitis positive and negative (“healthy”) samples was achieved in both unsupervised and supervised analysis, with good separation and a robust model Q₂ score for PLS-DA (Figure 59). A total of 359 biomarkers were identified to be significant in mastitis positive samples, with a top scoring fold change range of 120-649. The top twenty five most significant biomarkers, ranked in order of FDR corrected *P* value, are displayed in a heatmap as previously (Figure 60). A significantly higher fold change in significant biomarkers was observed with LA-REIMS analysis of lipid layers when compared with cell line probe analysis. With Gram-positive/negative and non-mastitis sample class groupings, supervised analysis showed superior class clustering and separation to that of cell line probe REIMS of the lipid layer, with a comparable Q₂ model score. In genera class groupings, again better sample clustering and separation was achieved with a positive Q₂ score, although the low score still suggested overfitting (Figure 59). A slightly lower number of genus specific significant biomarkers were identified in the 600-1000 *m/z* region than with cell line probe analysis (Figure 61). However, when the whole 50-2500 *m/z* acquired region was analysed, several significant biomarkers were identified in the 50-500 *m/z* region which were not identified in cell line probe REIMS. Both unsupervised and supervised discriminatory analysis of this lower mass region alone is capable of separating mastitis positive samples from healthy samples (Figure 59). Gram status separation is, however, very poor, along with genera separation which showed poor clustering and an overfitted PLS-DA model (Supplementary 18).

Direct three component PLS-DA plot comparisons of lipid layer analysis using cell line probe and LA-REIMS further shows that the LA-REIMS method is better at separating mastitis positive from negative samples (Figure 62). Both methods were capable of distinguishing between bacteria genera in mastitis positive samples, but low Q₂ model scores indicate overfitting, especially for cell line probe REIMS. In both methods the majority of significant genus specific biomarkers were identified for *Staphylococcus* spp., likely due to this being the most common causative agent out of all genera present.

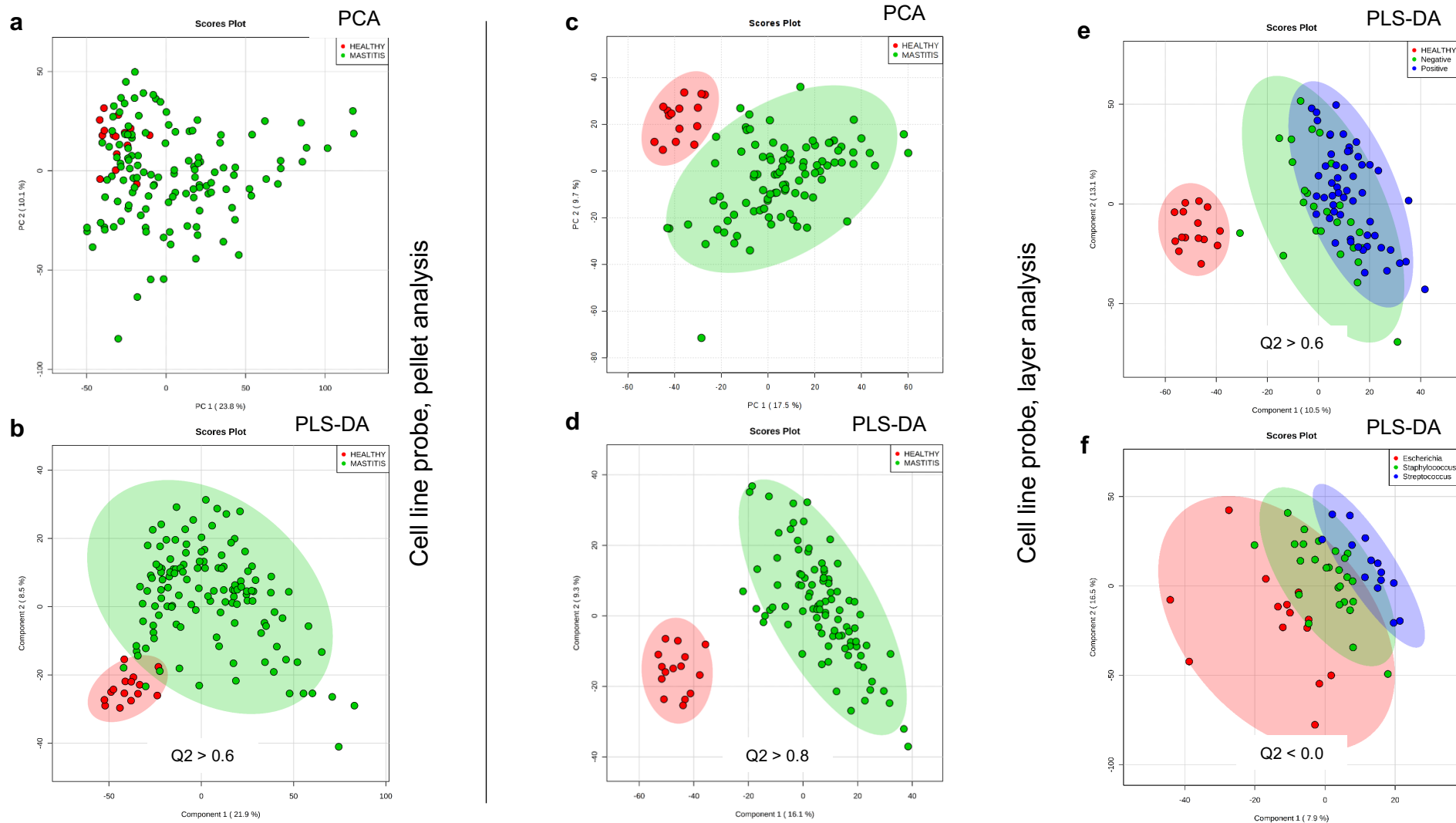


Figure 57. PCA and PLS-DA plots of cell line probe pellet (a,b) and lipid layer (c-f) REIMS analysis of bovine milk 600-1000 m/z range. Supervised PLS-DA shows good separation with a robust model Q_2 score of mastitis positive versus negative samples in pellet analysis (b). In lipid layer analysis, superior discrimination was achieved in both unsupervised (c) and supervised (d) models. Good clustering of Gram-negative and Gram-positive bacterial species containing samples was also achieved with a PLS-DA model (e), however a PLS-DA model investigating discrimination of distinct genera resulted in overfitting and a low Q_2 score (f).

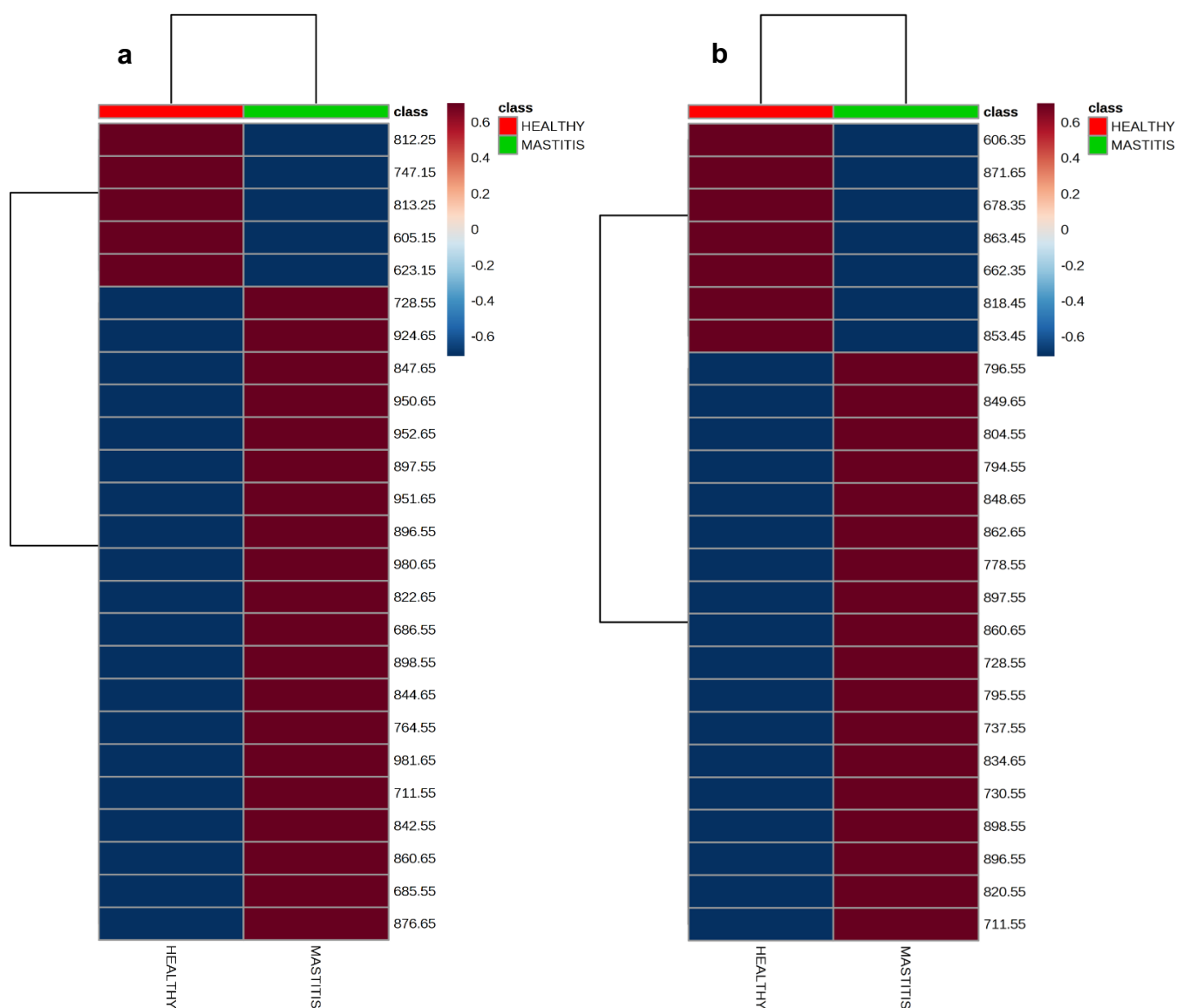


Figure 58. Heatmap clustering of top 25 most significant features for mastitis positive versus negative classification, determined through t-test FDR corrected *P* value rankings, of sample pellets (a) and lipid layer (b) samples analysed through cell line probe REIMS, 600-1000 *m/z* range. Samples are displayed in rows and biomarkers in columns. Significant, distinct biomarkers for both mastitis positive and negative sample classifications are displayed in group averaged heatmaps. Red colours represent a positive fold change of the biomarker in a sample class, whilst blues represent a negative fold change. Full heatmaps are available in Supplementary 14 and Supplementary 15.

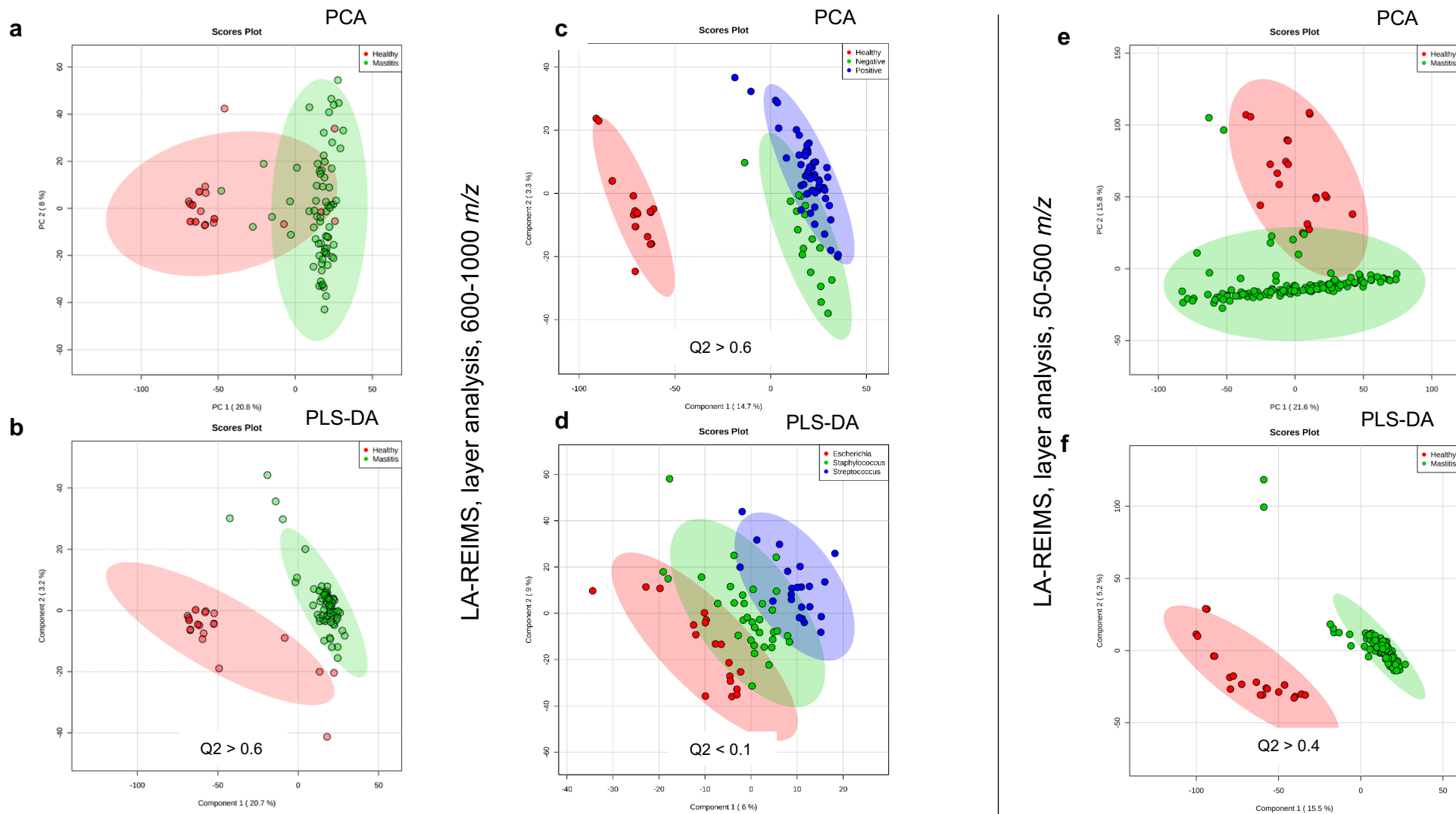


Figure 59. PCA and PLS-DA plots of LA-REIMS lipid layer analysis of bovine milk, 600-1000 (a-d) and 50-500 (e,f) m/z range. Unsupervised PCA (a) and supervised PLS-DA (b) models show good separation and class clustering of mastitis negative and positive samples in analysis of the 600-1000 m/z range. Good clustering and discrimination of Gram-negative and Gram-positive samples was also achieved (c), however a PLS-DA model based on genera classification was weak, with overfitting evident (d). Discrimination of mastitis positive versus negative samples was achieved in the 50-500 m/z range, which was not possible with the cell line probe analysis method (e,f).

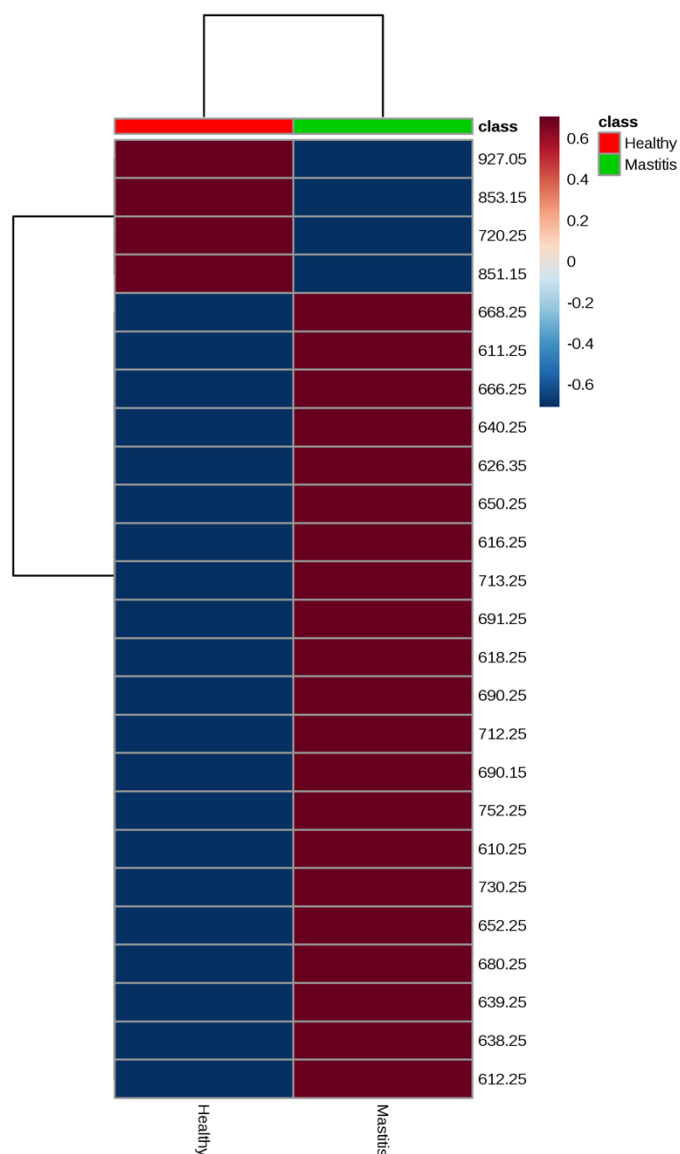


Figure 60. Heatmap clustering of top 25 most significant features for mastitis positive versus negative classification, determined through t-test FDR corrected P value rankings, of lipid layer samples analysed through LA-REIMS, 600-1000 m/z range. Samples are displayed in rows and biomarkers in columns. Distinct biomarkers for both mastitis positive and negative sample classes are displayed in a group averaged heatmap. Red colours represent a positive fold change of the biomarker in a sample class, whilst blues represent a negative fold change. Full heatmap is available in Supplementary 17.

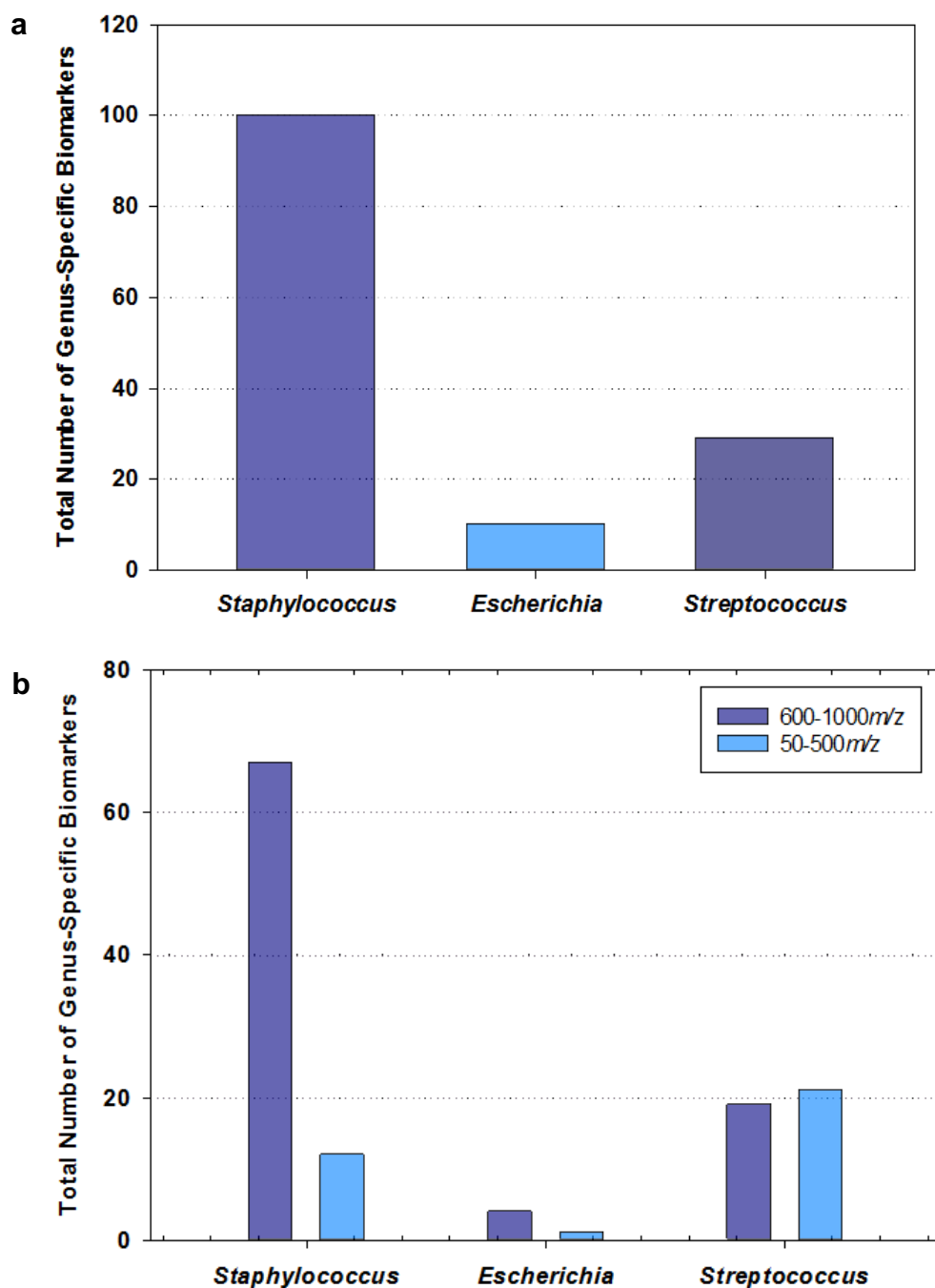


Figure 61. Genera specific biomarkers of mastitis positive bovine milk samples, identified through ANOVA, analysed via lipid layer cell line probe (a) and LA-REIMS (b). Biomarkers were determined to be significant if they were shown to be within the FDR corrected P value cut-off of 0.05. Fewer significant biomarkers were identified in the 600-1000 m/z region with LA-REIMS (b) compared to cell line probe REIMS (a), however biomarkers were identified in the 50-500 m/z region which were not present in the cell line probe analysis method, indicating higher sensitivity for LA-REIMS.

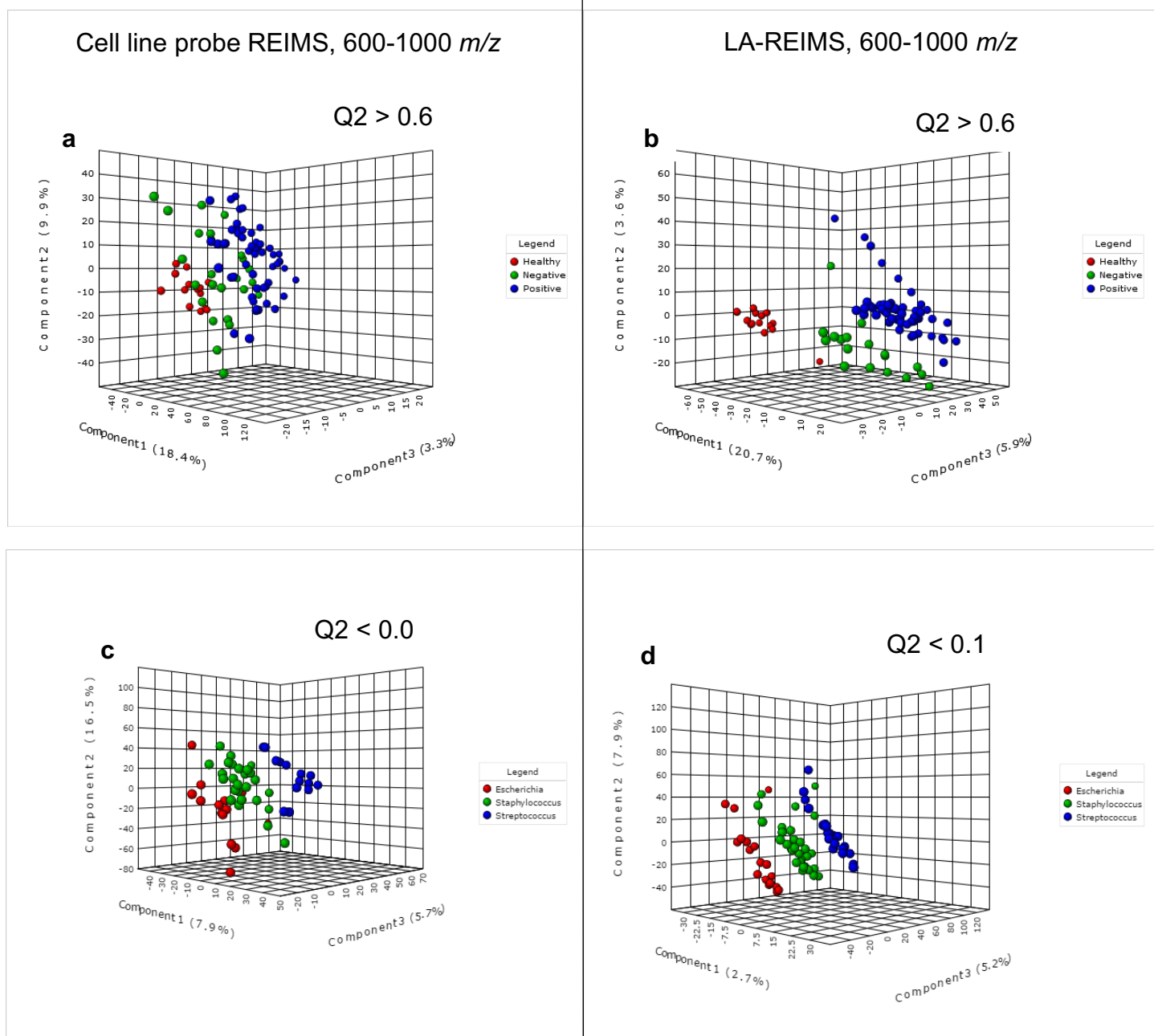


Figure 62. Three component PLS-DA plot comparisons of cell line probe REIMS (a,c) and LA-REIMS (b,d) analysis of bovine mastitis milk samples. Gram-positive, negative and healthy milk sample discrimination was achieved in both cell line probe (a) and LA-REIMS (b), with superior clustering observed in the latter. Genera clustering was observed in both methods (c,d), but with overfitting evident from low Q^2 performance scores in both. Overall, LA-REIMS provides a superior sample analysis platform in throughput, ease of use and sample class discrimination compared to cell line probe REIMS.

7.1 Discussion

The REIMS analysis of non-human veterinary samples is assessed in this experiment, meeting objective vi and supporting the hypothesis that the applications of REIMS are diverse and can be expanded to microbial diagnostics outside of the human clinical laboratory and into the veterinary field. The distribution of species in the received samples is in concordance with the prevalence of causative agents observed within the industry, as the majority of mastitis cases develop from infection by *S. aureus*, *Streptococcus* spp. and *E. coli*, due to their environmental ubiquity and status as common commensals (256).

During centrifugation a large (3-5 mm) lipid layer formed on the sample surface due to the $\approx 4\%$ fat content of milk, along with a sediment layer containing whole cells and nuclei at the bottom of the plate wells. Cell line probe REIMS analysis was conducted on both the lipid and pellet layers in order to determine the optimal analysis method. For LA-REIMS, analysis was conducted on the surface lipid layer only due to the difficulty in firing the laser beam to the bottom of plate wells.

7.1.1 Cell Line Probe REIMS

Separation of mastitis positive and negative samples was achieved with cell line probe analysis in both lipid layer and pellet analysis, with lipid layer analysis yielding superior discrimination. This was likely due to the 600-1000 m/z region of analysis being dominated by lipids and therefore spectral richness and quality of the lipid layer was higher in this region. In mastitis, somatic cell count within milk is increased, mainly due to the influx of leukocytes to the udder in response to infection and inflammation, and epithelial cell shedding into milk (209). The fat composition of milk also changes with mastitis from infection, with both a reduced milk yield and total fat content reduction of $\approx 3.2\%$ observed (257, 258). The fat content of milk is derived from blood plasma lipids and mammary gland synthesis, both of which mastitis impairs through inflammation, tissue damage and altered gland permeability, resulting in decreased total lipid content (258). Lipid composition is altered during mastitis, in that higher amounts of free fatty acids, increased concentrations of short chain fatty acids and lower

concentrations of long chain fatty acids are observed compared to mastitis negative milk (258).

The fat content of milk exists in the form of globules which are dispersed as membrane bound droplets throughout the liquid. Mastitis results in reduced phospholipid content of globules and increased cholesterol, which reduces protection of the globules from lipase activity leading to the observed increase in free fatty acids (259). It is therefore not unexpected that these somatic cell count and compositional changes can be detected with MS methods, and may form the basis of discrimination between the positive and negative mastitis samples. In pellet analysis, separation was likely achieved through analysis of the increased presence of somatic cells in mastitis positive milk, whereas in lipid layer analysis discrimination was likely predominantly based on lipid compositional changes as well as bacterial membrane lipids from lysed cells. This theory is supported by the identification of genera specific biomarkers within lipid layer samples.

A higher quantity of mastitis positive biomarkers were discovered in pellet analysis than in lipid layer analysis, albeit with a reduced discriminatory power, as although compositional changes occur in milk fat content the effect is relatively small in comparison to the markedly increased presence of somatic cells in mastitis positive verses negative milk. Despite this, the superior discriminatory power of lipid layer samples also supports this theory, as REIMS is primarily a lipidomics based technique, and therefore milk fat compositional changes reflected in the lipid layer would be readily detected. In addition, the better model performance in lipid layer analysis over pellet analysis, though with less significant biomarkers identified, may be due to the relatively low fold change of pellet biomarkers compared with lipid layer biomarkers. The highest fold change biomarkers in pellet analysis were 12-17, whereas lipid layer analysis resulted in top ranking biomarkers with a fold change of 17-24.

7.1.2 LA-REIMS

In lipid layer analysis of milk samples with LA-REIMS, more significant class defining biomarkers were seen than with cell line probe analysis for mastitis positive verses

negative classification. This may be due to superior vaporisation and thus ionisation of the sample when compared to the manual hand held probes, which are reliant on user technique and subject to greater TIC variation between spectra acquisitions. In addition, the nature of laser analysis means that no electrode cleaning is required between samples, unlike with a cell line probe, resulting in consistently higher quality vaporisation due to lack of electrode inhibiting charring or build-up after each analysis point. When compared with cell line probe analysis, ≈ 10 fewer genus specific biomarkers were identified in the 600-1000 m/z lipid region for each sample class, although analysis of the entire 50-2500 m/z range revealed class specific biomarkers in the 50-500 m/z region resulting in a greater number of total biomarkers detected. Biomarkers identified in the lipid region were of a much higher fold change (120-649) than comparative analysis with cell line probes (17-24), explaining the greater overall model performance. This is again explained by the greater ionisation efficacy of the LA-REIMS system compared to probe analysis, on which build up, charring and user technique reduce vaporisation potential.

Mastitis positive/negative discrimination was possible through analysis of the 50-500 m/z region alone, which was not achieved in cell line probe analysis, likely due to the reduced ionisation efficacy of the technique and lower fold change biomarkers. However, genus and Gram stain status discrimination was not possible through analysis of the 50-500 m/z region, suggesting the biomarkers responsible for discrimination are sourced from fatty acid compositional alterations in mastitis positive milk verses negative milk, and are not metabolites or free fatty acids liberated from bacterial cells. As discussed previously, the fatty acid composition of milk changes with infection leading to mastitis, in that total free fatty acid content rises due to increased lipolytic activity on fat globules (258, 259). This activity may form the basis of specific fatty acid biomarkers in the 50-500 m/z range which are not present in mastitis negative samples.

7.1.3 Conclusions

LA-REIMS is superior to cell line probe REIMS in the classification of mastitis positive verses negative samples, based on the greater range of discovered biomarkers, higher fold change of significant biomarkers and increased discriminatory power observed in

both PCA and PLS-DA models. Analysis of the lipid layer of milk samples provided greater classification accuracy when compared with pellet analysis as REIMS is primarily a lipidomics based technique, in that ionisation of lipids is more efficient than smaller molecular weight molecules. Overfitting of Gram stain and bacterial genera discriminatory models demonstrates that the methodology utilised here is not optimal for sub-classifications within mastitis positive samples, although the technique shows the capability of as a mastitis positive/negative detection system in this proof of principle study.

Currently, bovine mastitis is diagnosed through a variety of cow side tests, employed routinely or at the onset of symptoms, which primarily focus on measuring total SSC (209). Although such tests are rapid and require little expertise, they do not offer a definitive diagnosis. Instead, an elevated SSC indicates a sample should be sent to a laboratory for follow up testing to be ordered in the form of traditional differential culturing or PCR methods (211). Traditional culturing techniques are labour intensive and typical turnaround time is up to one week, whilst PCR based methods are expensive and require expertise to design targeted primers, multiplex samples and interpret results. REIMS may have the potential to replace such approaches as an automated sampling technique for the detection of mastitis positive milk, as milk lipid or fatty acid compositional changes can be expected to be a more sensitive and specific test for mastitis than SCC based tests (260).

Compositional changes in milk associated with mastitis, such as reduction in long chain fatty acids and the increase of short chain fatty acids, offer an early indicator of disease which REIMS could capitalise on through incorporation of the system to automated milking machines. If the REIMS methodology could be further optimised and improved to include identification of genera specific biomarkers, it may also serve as a stand-alone testing method without the requirement for validation by a central laboratory, saving time, labour and money. It is possible, however, that through the rich native lipid spectral profile of bovine milk and ion suppression effects, REIMS may be limited to detecting changes in milk composition and serving only as a mastitis positive or negative indicator. To investigate this, future work should be focused on investigating

biomarkers identified as significant in mastitis positive discrimination through the use of MS/MS, in order to determine their likely origin. The adaption of the LA-REIMS system with a suitably designed probe enabling analysis of pellets, rather than the lipid layer, would also allow determination of the suitability of REIMS as a SCC test. It may be that REIMS could serve as a dual analysis method of SSC determination and milk compositional changes, by analysing both lipid layer and pellet of centrifuged samples, to provide a more definitive indicator of an active mastitis case.

8 General Discussion

The development of novel REIMS methodologies allowing analysis from a wider range of sample types, for example liquid samples, enables the expansion of the applications of the technology. In this project, detection of species-specific biomarkers in clinical urine samples has been achieved, allowing semi-automated microbial analysis and discrimination of culture negative and positive urine samples without the need for pre-culturing. The versatility of REIMS as a microbial analysis system has been demonstrated through analysis of bovine milk leading to the successful discrimination of mastitis positive and negative samples. In addition, identification of species-specific biomarkers in controlled mixed cultures shows, as a proof of principle, that REIMS may be applied to a wider range of clinical samples containing more than one microbial species.

8.1 REIMS as a Microbial Analysis Technique

Mass spectrometry is well established for bacterial classification due to rapid analysis times compared to traditional methods, high specificity and high sensitivity. MS methods classify based on the fingerprinting of bacterial spectral profiles through a variety of techniques, such as volatile fatty acid methyl esters (FAMES) profiling in gas chromatography (GC)-MS, commercialised in the Sherlock Microbial ID system (MIDI, Delaware, US) in the 1980s, intact glycerophospholipid profiling in fast atom bombardment (FAB)-MS and the more recent protein profiling MALDI-ToF based Bruker Biotyper and VITEK MS (bioMérieux, Marcy-l'Étoile, France) systems (261-263). FAME analysis with GC-MS utilises the specific fatty acid distribution profiles of bacterial species for classification and offers automated analysis, low cost and high throughput at the time of development, however has been largely superseded by newer technologies with higher throughputs, shorter sample preparation protocols and less reliance on stringent culturing conditions.

FAB-MS was the first technique to utilise intact glycerophospholipids for successful bacterial discrimination, and direct ionisation from lysed bacterial cells was achieved, however relatively low signal intensities were recorded for higher molecular weight ions (262). Single bacterial cells are able to modify their fatty acid composition in response

to environmental triggers, effecting their fatty acid profile and highlighting the requirement for stringent, controlled culture conditions for profiles to be used in classification (16). Although lipid composition is altered through fatty acid changes, the phospholipid head group remains, and therefore the lipid class remains the same (15). For this reason, phospholipid profiling of bacteria is more versatile than the fatty acid equivalent, and is preferred due to ease of sample preparation and more consistent results.

Today, MALDI-ToF platforms enjoy almost exclusive use of clinical laboratories due to their rapidity and ease of use, and have revolutionised workflows leading to better outcomes for patients by reducing time until effective antimicrobial drug treatment and enabling shorter hospital stays for in-patients (159, 160). Results obtained from commercial systems offer superior classification rates than conventional biochemical techniques and remove the subjectivity associated with user interpretation (264). A typical clinical microbiology laboratory utilising a commercial MALDI system saves around \$100,000 per year in media costs alone, and is able to offset the purchase price within around three years (159, 160).

Proteins are the most abundant molecule in bacterial cells and are readily ionised, as well as being highly specific to individual species (265). The proteome is transcribed in DNA and is relatively constant despite changing environmental conditions, whilst fatty acids and lipids are more subject to structural changes (265). Classification of bacteria with MALDI systems is based on ribosomal protein fingerprints, meaning that the approach more accurately represents bacterial genotypes, whilst lipidomic approaches can be considered to reflect phenotype (266). Both genotype and phenotype have clinical relevance and are both routinely considered in the identification and diagnosis of infection, however the relationship between the two is not completely understood, with antibiotic resistance identified in genotype studies not always conferring a resistant phenotype (267). It is for this reason that a holistic approach is taken in clinical laboratories, which considers data obtained from multiple analysis approaches. MALDI systems have improved sample turnaround times by 24 hours, but largely still rely on the isolation and culturing of microbial isolates before analysis (268). Protein

extractions are also required, and misclassifications arise between taxonomically close species. In addition, species within poly-microbial samples cannot be discriminated due to the difficulty in deconvoluting mixed spectra and the relatively low instrument resolution of commercial systems (269).

The results presented in Sections 4: Mixed culture analysis and 5: Urine analysis show that REIMS can overcome these limitations in that identification of species-specific biomarkers from poly-microbial mixtures is possible, and classification of culture positive urine samples without pre-culture was achieved. The REIMS ionisation method favours intact glycerophospholipid analysis, as demonstrated in Section 6: AMI-MS, and rich mass spectra are acquired in the 600-1000 m/z range. As such, lipidomic analysis via REIMS offers an alternative approach to the widely used MALDI systems, which may complement and improve current commercialised systems or potentially replace them with further development.

8.2 Application to Specific Sample Types

The obvious application of REIMS from the work presented in this thesis is to clinical urine samples. However, the methodology may be adapted to analysis of blood, sputum and tissue with relative ease. For blood, analysis may be performed on dried blood spots as with other ambient ionisation techniques, and direct analysis of sputum smears on glass slides is feasible (229-231). Tissue analysis requires detection of specific biomarkers above background spectra of the tissue itself and would likely require further optimisation and generation of a biomarker reference catalogue. Successful adaption of REIMS to these sample types would pave the way for an all-encompassing commercial microbial identification system, capable of analysing poly-microbial clinical samples as well as pure isolates.

8.2.1 Competing Technology

MALDI method developments have already enabled diagnostic analysis directly from clinical samples, including blood and urine, without extensive pre-culturing, albeit with sometimes significant sample preparation steps. The previously discussed Sepsityper kit, designed for the Bruker Biotyper, enables detection of bacteria directly from mono-

microbial blood culture broths with high accuracy, although poly-microbial sample analysis is unreliable and usually returns an identification of one causative agent only (270). For urine analysis, several protocols have been developed, including urine short incubation (U-si)-MALDI-ToF, direct analysis following washing and centrifugation, and diafiltration MALDI-ToF (271-273).

In diafiltration, a 10^5 - 10^6 CFU/mL limit of detection is reached by subjecting 15 mL of urine specimens to centrifuge filtration to concentrate the sample, followed by desalting, RBC lysis and washing steps, with the entire process taking around three hours and providing a sensitivity/specificity of 67/100% (273). A direct spotting technique was also evaluated by simply aliquoting urine specimens onto a MALDI plate and covering with matrix, however a 10^8 CFU/mL load was required for identification to occur. The u-si method is performed by aliquoting 200 μ L of urine specimen onto agar medium and incubating for three hours. Following this, a scraping of invisible culture from the inoculation point is taken, directly smeared onto a MALDI plate and covered with matrix prior to analysis. The protocol does not perform well in Gram-positive identifications, and false negative results are relatively high, with a reported sensitivity/specificity of 67/79% (271). Direct protocols are based around centrifugation and washing steps and achieve limits of detection of 10^5 - 10^7 CFU/mL (274). When coupled with automated flow cytometry systems as a screening agent, classification can be improved, with results available in as little as an hour. A sample volume of 10 mL is required however, which is high considering the volume obtained from patients typically does not exceed this and an amount is required to be retained for follow up culturing or other follow on tests. Additionally, multiple studies have found that around 14% of samples with $>10^5$ CFU/mL determined by traditional culture returned no classification result, indicating the potential for false negatives (272, 274).

Importantly, all of the above techniques are sensitive to poly-microbial samples or contaminants, providing either incorrect, incomplete or low confidence classifications in the case of poly-microbial cases, or false positives in the event of contamination. Since a 10^5 CFU/mL limit is the widely accepted clinical indicator of a positive UTI, direct urine analysis with MALDI has thus far proven unreliable.

The MinION portable nanopore sequencing device has also been utilised in clinical urine identification of pathogens without culturing, however identification relied upon high bacterial loads of 10^7 CFU/mL and a sample turnaround time of eight hours, along with a significant sample preparation protocol (275).

The biomarker detection limit of 10^2 CFU/mL identified in 5.7.2: LA-REIMS Method Development for *E. coli* and *P. aeruginosa* indicates that REIMS may be a viable competitor to the existing MALDI methods in clinical urine analysis. This detection limit was, however, achieved with spiked culture negative urine samples and must be evaluated further. As previously mentioned, the detection limit of clinical urine samples was not investigated here due to the limited volume of urine permitted to be taken from the ICHNT microbiology laboratory and the non-reporting of CFU/mL values in their routine workflows.

Swifter identification of causative agents of infection is required to further improve outcomes for patients. Whilst MALDI systems have significantly improved sample turnaround times by 24-48 hours there still exists a need for further improvement. Since there are no extensive sample preparation steps involved in REIMS and analysis of urine specimens can be conducted following a single centrifugation step, samples can be analysed and results obtained as soon as laboratory operators can load samples into the machine, offering a significant further turnaround time improvement.

8.2.2 Non-Human Applications

In addition to applications in the human clinical laboratory, results presented in section 7: Bovine Mastitis demonstrate the system is capable of microbial analysis from non-human samples as well, further expanding the scope of a potential all-encompassing system. Determination of mastitis positive milk and the causative agent of infection is required to be swift in order to reduce economic losses through milk discard for the farmer and maintain animal welfare. Current diagnostics revolve around the use of cow-side tests largely based on SSC enumeration, which provide a rapid, crude indicator of active infection but do not provide insight into a causative agent. A positive cow-side test is therefore followed up with culturing in central laboratories. Some effort has been

placed into developing a MALDI based MS protocol to replace the labour intensive culturing step, where a ribosomal protein extraction method was used to identify causative agents in synthetically spiked milk samples without a pre-culture requirement (210). In this, the previously mentioned Sepsityper kit was utilised to prepare samples by following the standard protocol of multiple centrifugation, lysis and washing steps. Identification of causative agents *S. aureus*, *S. uberis*, *S. agalactiae*, *S. dysgalactiae* and *E. coli* to the species level was successful within the spiked milk, but was dependent on bacterial loads, with 10^6 - 10^8 CFU/mL required for adequate confidence scores (210). To date, the method has not been evaluated on true mastitis milk samples, however the authors suggest a short incubation step of four hours, similar to the U-si-MALDI method, for circumventing potential spectral interference from somatic cells.

The limit of detection was not investigated for mastitis milk samples in this work, as 100% discrimination of mastitis positive and negative samples was achieved in supervised discriminatory analysis with LA-REIMS, with a robust model quality score. Full discrimination of distinct species was not achieved despite the identification of species-specific biomarkers, however the discrimination of mastitis positive versus negative milk samples achieved with LA-REIMS indicates potential of the technology as a rapid classifier of individual cow milk samples. In the current development state, a positive mastitis LA-REIMS result could be followed up with culturing, with isolated bacteria again being subjected to LA-REIMS for a species classification, reducing sample turnaround time by several days. However, since species-specific biomarkers were detected directly from mastitis positive milk samples it is feasible that further development of the platform could lead to reliable classification of the causative agent without pre-culturing, saving up to one week of sample turnaround time (210).

8.3 Strengths and Limitations

The main strengths of REIMS approaches are the lack of sample preparation required and the versatility of the technology in rapid, real-time analysis. Several applications in microbial diagnostics have been investigated here, although REIMS has also been successfully utilised in the classification of filamentous fungi, real-time intraoperative

tissue identification and food product authenticity (276-278). Compared to conventional culturing and biochemical testing methods, REIMS significantly reduces the labour, time and expertise required to obtain results. The LA-REIMS platform offers a viable alternative to existing MALDI based technologies in that it requires no extensive extraction, washing or other sample preparation for analysis, is an ambient ionisation technique allowing direct from sample analysis, is automated, requires no consumables outside of normal MS instrument maintenance and is high throughput. The results presented in this thesis have led to consideration by Waters and Becton Dickinson Diagnostics to integrate the platform into automation systems utilised in clinical laboratories, such as the BD Kiestra. The detection of species-specific biomarkers directly from clinical urine samples, without pre-culture, isolation or extraction and at bacterial loads below that of UTI threshold levels lends appeal to this technique over existing commercially available systems, which are currently unable to do so.

The limitations of the technology include the expense of the Q-ToF instrument, laser platform and Tecan automation robot, as well as their collective size. Ambient ionisation techniques are inherently more effected by environmental conditions including temperature, humidity and air pressure, leading to potential batch effects in data. These effects are largely eliminated by data pre-processing and maintenance of operating conditions however, and have not caused any issues in data analysis throughout the course of this project. Ambient MS techniques in general are not quantitative without the use of an isotopically labelled standard, and so it is not possible with the current setup to determine bacterial load within a sample based on biomarker presence or intensity, making it potentially difficult to distinguish between contaminants and active agents of infection. In direct sample analysis without culturing, however, the bacterial load of any contaminant is likely to be significantly lower than infection causing species, and therefore likely below the limit of detection. Additionally, since biomarkers identified as significant in discrimination of *E. coli* positive clinical urine samples were of different relative intensities and peak values to those identified in culture negative urine samples spiked with an *E. coli* isolate, it can be presumed that the background spectral profile of clinical samples will necessitate building of biomarker databases specific to each sample type. This will be required due to the

presence of some spectral interference and activation of different metabolic pathways in bacterial cells growing in diverse host environments leading to altered lipid profiles.

The technique is also still in developmental phase and more work is required to bring the REIMS platform forward as a commercial competitor to existing technologies. Experimental parameters are currently optimised for each new sample type, however the introduction of the laser platform has greatly reduced the amount of optimisation and method development required, since the issue of electrode geometry, material and voltages has been eliminated, as well as liquid analysis becoming possible. The analysis of liquid urine samples was not satisfactory without a centrifugation step to concentrate the sample, so some sample preparation is still required in such cases.

To become more competitive in the applications discussed in this thesis, the system must become cheaper and more portable to allow implementation as a rapid, point of care UTI diagnosis method, or as a competitor to cow-side mastitis tests. Miniaturisation of the system would also open new applications, for example as a portable, on-site food security instrument, or field based environmental or public health inspection device. A currently viable application is the incorporation of REIMS as a module into existing laboratory automation systems, such as the BD Kiestra (Becton Dickinson, New Jersey, US) which is being adopted by large scale laboratories including the ICHNT microbiology laboratory. Kiestra systems are able to perform automated sample processing, culturing and imaging with add-on modules available. Incorporation of a REIMS system module would expand capability to automated classification.

8.4 Further REIMS Development Work

The intent of this project was to provide proof of principle method developments into the application of REIMS to novel sample types and applications, in addition to protocol optimisation and initial investigatory data analysis. Time constraints did not allow in depth pursuit of biomarker discovery and identification, which would be useful in generation of a repository database for reference in later investigations. The application of REIMS to microbial speciation is still a recent development and analysis of more

sample types is required, along with validation using conventional techniques. This is currently underway as part of the larger MicrobeID group project and individual PhD projects investigating analysis of sputum, blood, stool and swabs.

Further development as a clinical urine diagnostic platform should aim to elucidate the limit of detection, as a priority, in order to better assess the performance of REIMS against existing technologies. To achieve this, a large-scale study with independent ethics permissions to exclusively collect large volumes of clinical urine is required, so cell counts may be determined. This should be coupled with urine culture, MALDI identification and a genomic identification method to obtain a comparative sensitivity and specificity for REIMS. In addition, existing MS methods are unable to classify poly-microbial samples directly and although this was shown to be possible in REIMS as a proof of principle (4: Mixed culture analysis), it was not investigated further. Consideration of poly-microbial urine samples in the above large-scale study would further reveal the performance of REIMS against existing technologies.

For non-human applications including microbial testing for mastitis, a limit of detection should again be sought for REIMS. The advent of miniaturised MS systems, such as the Suitcase ToF, are interesting approaches to the lack of portability of MS techniques (279). Coupling of an LA-REIMS platform to such a device may open the possibility of REIMS being deployed as a cow-side test, although the rapidity and cost of the system would likely find better use in a central laboratory, which can provide classification results on the same day.

8.5 Conclusions

A novel methodology was developed to apply the existing REIMS technique to liquid samples, involving only a centrifugation step. The LA-REIMS method was found to be superior in analysis of multiple liquid sample types when compared to monopolar and bipolar handheld probes. The methodology was optimised and evaluated in direct from sample microbial identification without pre-culturing. When applied to clinical urine and bovine mastitis milk samples, detection of species-specific biomarkers was achieved

in both, as well as robust classification of positive and negative samples. Comparison to existing MALDI systems showed that REIMS has several advantages in limit of detection and shorter sample preparation times. Detection of species-specific biomarkers was also achieved in synthetic mixed cultures, demonstrating the capability of the method in poly-microbial sample analysis. To further evaluate the method, a large-scale study should be conducted on a single application, including a greater depth of biomarker discovery and identification.

In the case of clinical urine analysis, a greater number of samples should be obtained to enable analysis from a wider range of causative agents with sufficient biological replicates, including poly-microbial samples. A greater sample volume should be sought to allow CFU/mL enumeration and comparison of the REIMS limit of detection with existing technologies, such as MALDI. Classification results should also be validated with genomic methods.

In mastitis analysis, comparison of the LA-REIMS technique with existing SSC based cow-side tests should be completed, as well evaluation of the limit of detection with MALDI. A dedicated investigation into the identification of mastitis specific biomarkers should be completed by utilising MS/MS. Further development of the method, as well as potentially coupling the technique with AMI-MS, should be prioritised to enable stronger classification of causative agents of mastitis by analysing both lipid layer and cell pellets following centrifugation. Since both AMI and REIMS utilise the same MS system this would not be restrictive in costs or require additional expertise.

Within this thesis, it has been demonstrated that REIMS can be used for direct from sample microbial analysis with applications in the human clinical microbiology laboratory and veterinary diagnostic laboratory. Further development of the methodologies discussed in this work would likely result in REIMS becoming a viable competing technology to existing MS based systems utilised for microbial identification within the clinical laboratory.

Publications List

- i. Cameron SJ, Bodai Z, Temelkuran B, Perdones-Montero A, Bolt F, **Burke A**, Alexander-Hardiman K, Salzet M, Fournier I, Rebec M, Takats Z. Utilisation of Ambient Laser Desorption Ionisation Mass Spectrometry (ALDI-MS) Improves Lipid-Based Microbial Species Level Identification. *Scientific Reports*. 2019;9(1):3006.
- ii. Bodai Z, Cameron SJ, Bolt F, Simon D, Schaffer R, Karancsi T, Balog J, Rickards T, **Burke A**, Hardiman K, Abda J, Rebec M, Takats Z. Effect of Electrode Geometry on the Classification Performance of Rapid Evaporative Ionization Mass Spectrometric (REIMS) Bacterial Identification. *Journal of the American Society for Mass Spectrometry*. 2018;29(1):26-33.
- iii. Cameron SJ, Bolt F, Perdones-Montero A, Rickards T, Hardiman K, Abdolrasouli A, **Burke A**, Bodai Z, Karancsi T, Simon D, Schaffer R, Rebec M, Balog J, Takats Z. Rapid Evaporative Ionisation Mass Spectrometry (REIMS) Provides Accurate Direct from Culture Species Identification within the Genus Candida. *Scientific Reports*. 2016;6:36788.
- iv. Bolt F, Cameron SJ, Karancsi T, Simon D, Schaffer R, Rickards T, Hardiman K, **Burke A**, Bodai Z, Perdones-Montero A, Rebec M, Balog J, Takats Z. Automated High-Throughput Identification and Characterization of Clinically Important Bacteria and Fungi using Rapid Evaporative Ionization Mass Spectrometry. *Analytical Chemistry*. 2016;88(19):9419-26.

Conference Presentations

- i. **Burke A**, Cameron SJ, Bolt F, Perdones-Montero A, Bodai Z, Hardiman K, Rebec M, Takats Z. Matrix Free Laser Assisted Rapid Evaporative Ionisation Mass Spectrometry (LA-REIMS) in Clinical Urine Sample Microbial Analysis. Mass Spectrometry: Applications to the Clinical Lab, 10th Annual Conference and Exhibits 2018 (MSACL US). Palm Springs, California, US.
Young Investigator Travel Grant Awarded.
- ii. **Burke A**, Cameron SJ, Bolt F, Perdones-Montero A, Bodai Z, Rickards T, Hardiman K, Rebec M, Takats Z. Rapid Evaporative Ionisation Mass Spectrometry (REIMS) as a Novel Approach to Pathogen Detection Directly from Clinical Diagnostic Samples. Mass Spectrometry: Applications to the Clinical Lab, 9th Annual Conference and Exhibits 2017 (MSACL US). Palm Springs, California, US.
Young Investigator Travel Grant Awarded.
- iii. **Burke A**, Cameron SJ, Bolt F, Bodai Z, Rickards T, Hardiman K, Perdones-Montero A, Abdolrasouli A, Rebec M, Karancsi T, Simon D, Takats Z. Rapid Evaporative Ionisation Mass Spectrometry (REIMS) as a Novel Approach to Microbial Community Profiling. Mass Spectrometry: Applications to the Clinical Lab, 3rd Annual Conference and Exhibits 2016 (MSACL EU). Salzburg, Austria.
Young Investigator Travel Grant Awarded.
- iv. **Burke A**, Bolt F, Cameron SJ, Strittmatter N, Rickards T, Hardiman K, Inglese I, Galea D, Abdolrasouli A, Rebec M, Karancsi T, Simon D, Veselkov K, Takats Z. Rapid Evaporative Ionisation Mass Spectrometry (REIMS) as a Novel Approach to Microbial Community Profiling. Mass Spectrometry: Applications to the Clinical Lab, 8th Annual Conference and Exhibits 2016 (MSACL US). Palm Springs, California, US.
Young Investigator Travel Grant Awarded.
- v. Bolt F, Cameron SJ, Abdolrasouli A, Rhodes J, Rickards T, Hardiman K, **Burke A**, Balog J, Karancsi T, Simon D, Schaffer R, Bodai Z, Rebec M, Takats Z. Identification of Fungi using Rapid Evaporative Ionisation Mass Spectrometry.

64th Annual Conference of the American Society of Mass Spectrometry 2016
(ASMS). San Antonio, Texas, US.

Stand in presenter.

9 References

1. King P. Haemophilus influenzae and the lung (Haemophilus and the lung). Clinical and translational medicine. 2012;1(1):10.
2. Vollmer W, Blanot D, De Pedro MA. Peptidoglycan structure and architecture. FEMS microbiology reviews. 2008;32(2):149-67.
3. Daniel RA, Errington J. Control of cell morphogenesis in bacteria: two distinct ways to make a rod-shaped cell. Cell. 2003;113(6):767-76.
4. Kohanski MA, Dwyer DJ, Collins JJ. How antibiotics kill bacteria: from targets to networks. Nature Reviews Microbiology. 2010;8(6):423.
5. Pinho MG, Errington J. Dispersed mode of Staphylococcus aureus cell wall synthesis in the absence of the division machinery. Molecular microbiology. 2003;50(3):871-81.
6. Trachtenberg S. Mollicutes—wall-less bacteria with internal cytoskeletons. Journal of structural biology. 1998;124(2-3):244-56.
7. Gram C. Ueber die isolirte Färbung der Schizomyceten in Schnitt-und Trockenpräparaten. Fortschritte der Medicin. 1884;2:185-9.
8. Cabeen MT, Jacobs-Wagner C. Bacterial cell shape. Nature Reviews Microbiology. 2005;3(8):601.
9. Braun V. Covalent lipoprotein from the outer membrane of Escherichia coli. Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes. 1975;415(3):335-77.
10. Fisher JF, Mobashery S. β -lactam resistance mechanisms: Gram-positive bacteria and Mycobacterium tuberculosis. Cold Spring Harbor perspectives in medicine. 2016:a025221.
11. Marrakchi H, Lanéelle M-A, Daffé M. Mycolic acids: structures, biosynthesis, and beyond. Chemistry & biology. 2014;21(1):67-85.
12. Halebian S, Harris B, Finegold SM, Rolfe RD. Rapid method that aids in distinguishing Gram-positive from Gram-negative anaerobic bacteria. Journal of clinical microbiology. 1981;13(3):444-8.
13. Ghuyssen J-M, Goffin C. Lack of Cell Wall Peptidoglycan versus Penicillin Sensitivity: New Insights into the Chlamydial Anomaly. Antimicrobial agents and chemotherapy. 1999;43(10):2339-44.
14. Sohlenkamp C, Geiger O. Bacterial membrane lipids: diversity in structures and pathways. FEMS microbiology reviews. 2016;40(1):133-59.
15. Zhang Y-M, Rock CO. Membrane lipid homeostasis in bacteria. Nature Reviews Microbiology. 2008;6(3):222.
16. Mansilla MC, Cybulski LE, Albanesi D, de Mendoza D. Control of Membrane Lipid Fluidity by Molecular Thermosensors. Journal of bacteriology. 2004;186(20):6681-8.
17. Annous BA, Becker LA, Bayles DO, Labeda DP, Wilkinson BJ. Critical role of anteiso-C15: 0 fatty acid in the growth of Listeria monocytogenes at low temperatures. Applied and environmental microbiology. 1997;63(10):3887-94.
18. Klein W, Weber MH, Marahiel MA. Cold shock response of Bacillus subtilis: isoleucine-dependent switch in the fatty acid branching pattern for membrane adaptation to low temperatures. Journal of bacteriology. 1999;181(17):5341-9.
19. Mendonca AF, Amoroso TL, Knabel SJ. Destruction of gram-negative food-borne pathogens by high pH involves disruption of the cytoplasmic membrane. Applied and environmental microbiology. 1994;60(11):4009-14.
20. Taormina P, Beuchat L. Survival of Listeria monocytogenes in commercial food-processing equipment cleaning solutions and subsequent sensitivity to sanitizers and heat. Journal of applied microbiology. 2002;92(1):71-80.
21. Giotis ES, McDowell DA, Blair IS, Wilkinson BJ. Role of Branched-Chain Fatty Acids in pH Stress Tolerance in *Listeria monocytogenes*. Applied and environmental microbiology. 2007;73(3):997-1001.
22. Torsvik V, Øvreås L. Microbial diversity and function in soil: from genes to ecosystems. Current opinion in microbiology. 2002;5(3):240-5.
23. Garrity GM. A New Genomics-Driven Taxonomy of *Bacteria* and *Archaea*: Are We There Yet? Journal of Clinical Microbiology. 2016;54(8):1956-63.
24. Valdes AM, Walter J, Segal E, Spector TD. Role of the gut microbiota in nutrition and health. Bmj. 2018;361:k2179.
25. Bull MJ, Plummer NT. Part 1: The Human Gut Microbiome in Health and Disease. Integrative medicine (Encinitas, Calif). 2014;13(6):17-22.

26. Huurre A, Kalliomaki M, Rautava S, Rinne M, Salminen S, Isolauri E. Mode of delivery - effects on gut microbiota and humoral immunity. *Neonatology*. 2008;93(4):236-40.
27. Guarner F, Malagelada JR. Gut flora in health and disease. *Lancet*. 2003;361(9356):512-9.
28. Wu HJ, Wu E. The role of gut microbiota in immune homeostasis and autoimmunity. *Gut microbes*. 2012;3(1):4-14.
29. Smythies LE, Sellers M, Clements RH, Mosteller-Barnum M, Meng G, Benjamin WH, et al. Human intestinal macrophages display profound inflammatory anergy despite avid phagocytic and bacteriocidal activity. *The Journal of clinical investigation*. 2005;115(1):66-75.
30. Ivanov, II, Atarashi K, Manel N, Brodie EL, Shima T, Karaoz U, et al. Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell*. 2009;139(3):485-98.
31. Mazmanian SK, Liu CH, Tzianabos AO, Kasper DL. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell*. 2005;122(1):107-18.
32. Macpherson AJ, Martinic MM, Harris N. The functions of mucosal T cells in containing the indigenous commensal flora of the intestine. *Cellular and molecular life sciences : CMLS*. 2002;59(12):2088-96.
33. Lane ER, Zisman TL, Suskind DL. The microbiota in inflammatory bowel disease: current and therapeutic insights. *Journal of inflammation research*. 2017;10:63-73.
34. Perencevich M, Burakoff R. Use of antibiotics in the treatment of inflammatory bowel disease. *Inflammatory bowel diseases*. 2006;12(7):651-64.
35. Frank DN, Robertson CE, Hamm CM, Kpadeh Z, Zhang T, Chen H, et al. Disease phenotype and genotype are associated with shifts in intestinal-associated microbiota in inflammatory bowel diseases. *Inflammatory bowel diseases*. 2011;17(1):179-84.
36. van Nood E, Vrieze A, Nieuwdorp M, Fuentes S, Zoetendal EG, de Vos WM, et al. Duodenal infusion of donor feces for recurrent *Clostridium difficile*. *The New England journal of medicine*. 2013;368(5):407-15.
37. Moayyedi P, Surette MG, Kim PT, Libertucci J, Wolfe M, Onischi C, et al. Fecal Microbiota Transplantation Induces Remission in Patients With Active Ulcerative Colitis in a Randomized Controlled Trial. *Gastroenterology*. 2015;149(1):102-9.e6.
38. Toruner M, Loftus EV, Harmsen WS, Zinsmeister AR, Orenstein R, Sandborn WJ, et al. Risk Factors for Opportunistic Infections in Patients With Inflammatory Bowel Disease. *Gastroenterology*. 2008;134(4):929-36.
39. Chinen J, Shearer WT. Secondary immunodeficiencies, including HIV infection. *Journal of Allergy and Clinical Immunology*. 2010;125(2):S195-S203.
40. Rosen FS, Cooper MD, Wedgwood RJ. The primary immunodeficiencies. *New England Journal of Medicine*. 1995;333(7):431-40.
41. Barreto ML, Teixeira MG, Carmo EH. Infectious diseases epidemiology. *Journal of epidemiology and community health*. 2006;60(3):192-5.
42. Carapetis JR, Steer AC, Mulholland EK, Weber M. The global burden of group A streptococcal diseases. *The Lancet infectious diseases*. 2005;5(11):685-94.
43. Lamagni T, Efstratiou A, Vuopio-Varkila J, Jasir A, Schalen C. The epidemiology of severe *Streptococcus pyogenes* associated disease in Europe. *Eurosurv* 10 (9), 179-184. 2005.
44. Wessels MR, Bronze MS. Critical role of the group A streptococcal capsule in pharyngeal colonization and infection in mice. *Proceedings of the National Academy of Sciences*. 1994;91(25):12238-42.
45. Stollerman GH, Dale JB. The importance of the group a streptococcus capsule in the pathogenesis of human infections: a historical perspective. *Clinical Infectious Diseases*. 2008;46(7):1038-45.
46. Bisno AL, Brito M, Collins C. Molecular basis of group A streptococcal virulence. *The Lancet infectious diseases*. 2003;3(4):191-200.
47. Paz Maya S, Dualde Beltran D, Lemerrier P, Leiva-Salinas C. Necrotizing fasciitis: an urgent diagnosis. *Skeletal radiology*. 2014;43(5):577-89.
48. Lolis E, Bucala R. Therapeutic approaches to innate immunity: severe sepsis and septic shock. *Nature reviews Drug discovery*. 2003;2(8):635.
49. Herwald H, Cramer H, Mörgelin M, Russell W, Sollenberg U, Norrby-Teglund A, et al. M protein, a classical bacterial virulence determinant, forms complexes with fibrinogen that induce vascular leakage. *Cell*. 2004;116(3):367-79.

50. Brown EJ. The molecular basis of streptococcal toxic shock syndrome. *New England Journal of Medicine*. 2004;350(20):2093-4.
51. Lowy FD. Staphylococcus aureus infections. *New England journal of medicine*. 1998;339(8):520-32.
52. Wertheim HF, Melles DC, Vos MC, van Leeuwen W, van Belkum A, Verbrugh HA, et al. The role of nasal carriage in Staphylococcus aureus infections. *The Lancet infectious diseases*. 2005;5(12):751-62.
53. Solberg CO. Spread of Staphylococcus aureus in hospitals: causes and prevention. *Scandinavian journal of infectious diseases*. 2000;32(6):587-95.
54. Crossley KB, Jefferson KK, Archer GL, Fowler VG. Staphylococci in human disease: John Wiley & Sons; 2009.
55. Peschel A. How do bacteria resist human antimicrobial peptides? *Trends in microbiology*. 2002;10(4):179-86.
56. O'Brien LM, Walsh EJ, Massey RC, Peacock SJ, Foster TJ. Staphylococcus aureus clumping factor B (ClfB) promotes adherence to human type I cytokeratin 10: implications for nasal colonization. *Cellular microbiology*. 2002;4(11):759-70.
57. Fridkin SK, Hageman JC, Morrison M, Sanza LT, Como-Sabetti K, Jernigan JA, et al. Methicillin-resistant Staphylococcus aureus disease in three communities. *New England Journal of Medicine*. 2005;352(14):1436-44.
58. Robinson DA, Kearns AM, Holmes A, Morrison D, Grundmann H, Edwards G, et al. Re-emergence of early pandemic Staphylococcus aureus as a community-acquired methicillin-resistant clone. *The Lancet*. 2005;365(9466):1256-8.
59. Control CfD, Prevention. Vancomycin-resistant Staphylococcus aureus--New York, 2004. *MMWR Morbidity and mortality weekly report*. 2004;53(15):322.
60. Paulsen J, Mehl A, Askim Å, Solligård E, Åsvold BO, Damås JK. Epidemiology and outcome of Staphylococcus aureus bloodstream infection and sepsis in a Norwegian county 1996-2011: an observational study. *Bmc Infect Dis*. 2015;15:116-.
61. Hall-Stoodley L, Costerton JW, Stoodley P. Bacterial biofilms: from the natural environment to infectious diseases. *Nature reviews microbiology*. 2004;2(2):95.
62. Mack D, Fischer W, Krokotsch A, Leopold K, Hartmann R, Egge H, et al. The intercellular adhesin involved in biofilm accumulation of Staphylococcus epidermidis is a linear beta-1, 6-linked glucosaminoglycan: purification and structural analysis. *Journal of bacteriology*. 1996;178(1):175-83.
63. Donlan RM, Costerton JW. Biofilms: Survival Mechanisms of Clinically Relevant Microorganisms. *Clinical Microbiology Reviews*. 2002;15(2):167-93.
64. Rogers KL, Fey PD, Rupp ME. Coagulase-negative staphylococcal infections. *Infect Dis Clin N Am*. 2009;23(1):73-98.
65. Otto M. Staphylococcus epidermidis—the 'accidental' pathogen. *Nature reviews microbiology*. 2009;7(8):555.
66. Chu V, Miro JM, Hoen B, Cabell CH, Pappas PA, Jones P, et al. Coagulase-negative staphylococcal prosthetic valve endocarditis—a contemporary update based on the International Collaboration on Endocarditis: prospective cohort study. *Heart*. 2009;95(7):570-6.
67. Sandre RM, Shafraan SD. Infective endocarditis: review of 135 cases over 9 years. *Clinical infectious diseases*. 1996;22(2):276-86.
68. Fernandez-Guerrero ML, Verdejo C, Azofra J, de Górgolas M. Hospital-acquired infectious endocarditis not associated with cardiac surgery: an emerging problem. *Clinical infectious diseases*. 1995;20(1):16-23.
69. Peleg AY, Hooper DC. Hospital-acquired infections due to gram-negative bacteria. *New England Journal of Medicine*. 2010;362(19):1804-13.
70. Foxman B. Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. *The American journal of medicine*. 2002;113(1):5-13.
71. Kunin CM. Urinary tract infections in females. *Clinical Infectious Diseases*. 1994;18(1):1-10.
72. Engel JD, Schaeffer AJ. Evaluation of and antimicrobial therapy for recurrent urinary tract infections in women. *Urologic Clinics of North America*. 1998;25(4):685-701.
73. Kostakioti M, Hultgren SJ, Hadjifrangiskou M. Molecular blueprint of uropathogenic Escherichia coli virulence provides clues toward the development of anti-virulence therapeutics. *Virulence*. 2012;3(7):592-3.

74. Hooton TM. Uncomplicated urinary tract infection. *New England Journal of Medicine*. 2012;366(11):1028-37.
75. Hannan TJ, Totsika M, Mansfield KJ, Moore KH, Schembri MA, Hultgren SJ. Host-pathogen checkpoints and population bottlenecks in persistent and intracellular uropathogenic *Escherichia coli* bladder infection. *FEMS microbiology reviews*. 2012;36(3):616-48.
76. Levison ME, Kaye D. Treatment of complicated urinary tract infections with an emphasis on drug-resistant gram-negative uropathogens. *Current infectious disease reports*. 2013;15(2):109-15.
77. Lichtenberger P, Hooton TM. Complicated urinary tract infections. *Current infectious disease reports*. 2008;10(6):499-504.
78. Hooton TM, Bradley SF, Cardenas DD, Colgan R, Geerlings SE, Rice JC, et al. Diagnosis, prevention, and treatment of catheter-associated urinary tract infection in adults: 2009 International Clinical Practice Guidelines from the Infectious Diseases Society of America. *Clinical infectious diseases*. 2010;50(5):625-63.
79. Nicolle LE. Catheter-related urinary tract infection. *Drugs & aging*. 2005;22(8):627-39.
80. Flores-Mireles AL, Walker JN, Caparon M, Hultgren SJ. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nature reviews microbiology*. 2015;13(5):269.
81. Nicolle LE. Asymptomatic bacteriuria—important or not? : *Mass Medical Soc*; 2000.
82. Warren JW, Abrutyn E, Hebel JR, Johnson JR, Schaeffer AJ, Stamm WE. Guidelines for antimicrobial treatment of uncomplicated acute bacterial cystitis and acute pyelonephritis in women. *Clinical Infectious Diseases*. 1999;29(4):745-59.
83. Ferry SA, Holm SE, Stenlund H, Lundholm R, Monsen TJ. Clinical and bacteriological outcome of different doses and duration of pivmecillinam compared with placebo therapy of uncomplicated lower urinary tract infection in women: the LUTIW project. *Scandinavian journal of primary health care*. 2007;25(1):49-57.
84. Rubin RH, Shapiro ED, Andriole VT, Davis RJ, Stamm WE. Evaluation of new anti-infective drugs for the treatment of urinary tract infection. *Clinical Infectious Diseases*. 1992;15(Supplement_1):S216-S27.
85. Nicolle LE, Bradley S, Colgan R, Rice JC, Schaeffer A, Hooton TM. Infectious Diseases Society of America guidelines for the diagnosis and treatment of asymptomatic bacteriuria in adults. *Clinical Infectious Diseases*. 2005;643-54.
86. Chenoweth CE, Gould CV, Saint S. Diagnosis, management, and prevention of catheter-associated urinary tract infections. *Infectious Disease Clinics*. 2014;28(1):105-19.
87. Knottnerus BJ, Geerlings SE, van Charante EPM, ter Riet G. Women with symptoms of uncomplicated urinary tract infection are often willing to delay antibiotic treatment: a prospective cohort study. *Bmc Fam Pract*. 2013;14(1):71.
88. Babich T, Zusman O, Elbaz M, Ben-Zvi H, Paul M, Leibovici L, et al. Empirical Antibiotic Treatment Does Not Improve Outcomes in Catheter-Associated Urinary Tract Infection: Prospective Cohort Study. *Clinical Infectious Diseases*. 2017;65(11):1799-805.
89. Bischoff S, Walter T, Gerigk M, Ebert M, Vogelmann R. Empiric antibiotic therapy in urinary tract infection in patients with risk factors for antibiotic resistance in a German emergency department. *Bmc Infect Dis*. 2018;18(1):56.
90. Tenke P, Köves B, Johansen TE. An update on prevention and treatment of catheter-associated urinary tract infections. *Current opinion in infectious diseases*. 2014;27(1):102-7.
91. Horcajada J, Shaw E, Padilla B, Pintado V, Calbo E, Benito N, et al. Healthcare-associated, community-acquired and hospital-acquired bacteraemic urinary tract infections in hospitalized patients: a prospective multicentre cohort study in the era of antimicrobial resistance. *Clin Microbiol Infect*. 2013;19(10):962-8.
92. Darouiche RO, Cadle RM, Zenon GJ, 3rd, Markowski J, Rodriguez M, Musher DM. Progression from asymptomatic to symptomatic urinary tract infection in patients with SCI: a preliminary study. *The Journal of the American Paraplegia Society*. 1993;16(4):219-24.
93. Dethlefsen L, Huse S, Sogin ML, Relman DA. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS biology*. 2008;6(11):e280.
94. Wilcox MH. *Clostridium difficile* infection and pseudomembranous colitis. *Best Practice & Research Clinical Gastroenterology*. 2003;17(3):475-93.
95. De La Cochetière MF, Durand T, Lalande V, Petit JC, Potel G, Beaugerie L. Effect of Antibiotic Therapy on Human Fecal Microbiota and the Relation to the Development of *Clostridium difficile*. *Microbial Ecology*. 2008;56(3):395-402.

96. Stamm WE. Scientific and clinical challenges in the management of urinary tract infections. *The American journal of medicine*. 2002;113(1):1-4.
97. Mulvey MA, Schilling JD, Martinez JJ, Hultgren SJ. Bad bugs and beleaguered bladders: interplay between uropathogenic *Escherichia coli* and innate host defenses. *Proceedings of the National Academy of Sciences*. 2000;97(16):8829-35.
98. Khandelwal P, Abraham SN, Apodaca G. Cell biology and physiology of the uroepithelium. *Am J Physiol-Renal*. 2009;297(6):F1477-F501.
99. Niveditha S, Pramodhini S, Umadevi S, Kumar S, Stephen S. The isolation and the biofilm formation of uropathogens in the patients with catheter associated urinary tract infections (UTIs). *Journal of clinical and diagnostic research: JCDR*. 2012;6(9):1478.
100. Jacobsen SM, Shirtliff ME. *Proteus mirabilis* biofilms and catheter-associated urinary tract infections. *Virulence*. 2011;2(5):460-5.
101. Goldhar J. Nonfimbrial adhesins of *Escherichia coli*. *Advances in experimental medicine and biology*. 1996;408:63-72.
102. Baorto DM, Gao Z, Malaviya R, Dustin ML, van der Merwe A, Lublin DM, et al. Survival of FimH-expressing enterobacteria in macrophages relies on glycolipid traffic. *Nature*. 1997;389(6651):636-9.
103. Malaviya R, Abraham SN. Clinical implications of mast cell-bacteria interaction. *Journal of molecular medicine (Berlin, Germany)*. 1998;76(9):617-23.
104. Mulvey MA, Lopez-Boado YS, Wilson CL, Roth R, Parks WC, Heuser J, et al. Induction and evasion of host defenses by type 1-piliated uropathogenic *Escherichia coli*. *Science*. 1998;282(5393):1494-7.
105. Mctaggart LA, Rigby R, Elliott T. The pathogenesis of urinary tract infections associated with *Escherichia coli*, *Staphylococcus saprophyticus* and *S. epidermidis*. *Journal of medical microbiology*. 1990;32(2):135-41.
106. Fukushi Y, Orikasa S, Kagayama M. An electron microscopic study of the interaction between vesical epithelium and *E. Coli*. *Investigative urology*. 1979;17(1):61-8.
107. Sobel JD. Pathogenesis of urinary tract infection. Role of host defenses. *Infect Dis Clin North Am*. 1997;11(3):531-49.
108. Haraoka M, Hang L, Frendeus B, Godaly G, Burdick M, Strieter R, et al. Neutrophil recruitment and resistance to urinary tract infection. *J Infect Dis*. 1999;180(4):1220-9.
109. Spencer JD, Schwaderer AL, Becknell B, Watson J, Hains DS. The innate immune response during urinary tract infection and pyelonephritis. *Pediatr Nephrol*. 2014;29(7):1139-49.
110. Measley RE, Jr., Levison ME. Host defense mechanisms in the pathogenesis of urinary tract infection. *The Medical clinics of North America*. 1991;75(2):275-86.
111. Hopkins WJ, Gendron-Fitzpatrick A, Balish E, Uehling DT. Time Course and Host Responses to *Escherichia coli* Urinary Tract Infection in Genetically Distinct Mouse Strains. *Infection and immunity*. 1998;66(6):2798-802.
112. Hvidberg H, Struve C, Kroghfelt KA, Christensen N, Rasmussen SN, Frimodt-Møller N. Development of a Long-Term Ascending Urinary Tract Infection Mouse Model for Antibiotic Treatment Studies. *Antimicrobial agents and chemotherapy*. 2000;44(1):156-63.
113. Foxman B, Gillespie B, Koopman J, Zhang L, Palin K, Tallman P, et al. Risk factors for second urinary tract infection among college women. *American journal of epidemiology*. 2000;151(12):1194-205.
114. Foxman B. The epidemiology of urinary tract infection. *Nat Rev Urol*. 2010;7(12):653.
115. Foxman B, Marsh J, Gillespie B, Rubin N, Koopman JS, Spear S. Condom use and first-time urinary tract infection. *Epidemiology*. 1997;637-41.
116. Foxman B, Manning SD, Tallman P, Bauer R, Zhang L, Koopman JS, et al. Uropathogenic *Escherichia coli* are more likely than commensal *E. coli* to be shared between heterosexual sex partners. *American journal of epidemiology*. 2002;156(12):1133-40.
117. Schaberg DR, Weinstein RA, Stamm WE. Epidemics of nosocomial urinary tract infection caused by multiply resistant gram-negative bacilli: Epidemiology and control. *The Journal of infectious diseases*. 1976;133(3):363-6.
118. Freeman JT, McBride SJ, Heffernan H, Bathgate T, Pope C, Ellis-Pegler RB. Community-onset genitourinary tract infection due to CTX-M-15-producing *Escherichia coli* among travelers to the Indian subcontinent in New Zealand. *Clinical infectious diseases*. 2008;47(5):689-92.
119. Foxman B. Urinary tract infection syndromes: occurrence, recurrence, bacteriology, risk factors, and disease burden. *Infect Dis Clin N Am*. 2014;28(1):1-13.

120. Gupta K, Bhadelia N. Management of urinary tract infections from multidrug-resistant organisms. *Infect Dis Clin N Am*. 2014;28(1):49-59.
121. Chen Y-H, Ko W-C, Hsueh P-R. Emerging resistance problems and future perspectives in pharmacotherapy for complicated urinary tract infections. *Expert opinion on pharmacotherapy*. 2013;14(5):587-96.
122. Shaikh S, Fatima J, Shakil S, Rizvi SM, Kamal MA. Antibiotic resistance and extended spectrum beta-lactamases: Types, epidemiology and treatment. *Saudi journal of biological sciences*. 2015;22(1):90-101.
123. Woodford N, Ward M, Kaufmann M, Turton J, Fagan E, James D, et al. Community and hospital spread of *Escherichia coli* producing CTX-M extended-spectrum β -lactamases in the UK. *J Antimicrob Chemoth*. 2004;54(4):735-43.
124. Bush K, Fisher JF. Epidemiological expansion, structural studies, and clinical challenges of new beta-lactamases from gram-negative bacteria. *Annual review of microbiology*. 2011;65:455-78.
125. Nordmann P, Poirel L, Walsh TR, Livermore DM. The emerging NDM carbapenemases. *Trends in microbiology*. 2011;19(12):588-95.
126. Paterson DL. Resistance in gram-negative bacteria: Enterobacteriaceae. *Am J Infect Control*. 2006;34(5):S20-S8.
127. Zhanel GG, Lawson CD, Adam H, Schweizer F, Zelenitsky S, Lagacé-Wiens PR, et al. Ceftazidime-avibactam: a novel cephalosporin/ β -lactamase inhibitor combination. *Drugs*. 2013;73(2):159-77.
128. McNulty C. *Diagnosis of Urinary Tract Infections: Quick Reference Tool for Primary Care*. Public Health England. 2002.
129. Karam MRA, Oloomi M, Mahdavi M, Habibi M, Bouzari S. Vaccination with recombinant FimH fused with flagellin enhances cellular and humoral immunity against urinary tract infection in mice. *Vaccine*. 2013;31(8):1210-6.
130. Goering RV, Mims CA. *Mims' Medical Microbiology*. 4th ed. Philadelphia: Mosby Elsevier; 2008.
131. Joyanes P, Conejo MD, Martinez-Martinez L, Perea EJ. Evaluation of the VITEK 2 system for the identification and susceptibility testing of three species of nonfermenting gram-negative rods frequently isolated from clinical samples. *Journal of clinical microbiology*. 2001;39(9):3247-53.
132. Sterry-Blunt RE, S. Randall K, J. Doughton M, H. Aliyu S, A. Enoch D. Screening urine samples for the absence of urinary tract infection using the sediMAX automated microscopy analyser. *Journal of medical microbiology*. 2015;64(6):605-9.
133. Broeren MA, Bahçeci S, Vader HL, Arents NL. Screening for urinary tract infection with the Sysmex UF-1000i urine flow cytometer. *Journal of clinical microbiology*. 2011.
134. Falbo R, Sala MR, Signorelli S, Venturi N, Signorini S, Brambilla P. Bacteriuria Screening by Automated Whole-Field-Image-Based Microscopy Reduces the Number of Necessary Urine Cultures. *Journal of clinical microbiology*. 2012;50(4):1427-9.
135. Aydin O, Ellidag HY, Eren E, Yilmaz N. High False Positives and False Negatives in Yeast Parameter in an Automated Urine Sediment Analyzer / Brojni Lažno Pozitivni I Lažno Negativni Rezultati Za Parametar Gljivice Na Automatskom Analizatoru Urina. 2015;34(3):332.
136. Oyaert M, Delanghe J. Progress in automated urinalysis. *Annals of laboratory medicine*. 2019;39(1):15-22.
137. Wang J, Zhang Y, Xu D, Shao W, Lu Y. Evaluation of the Sysmex UF-1000i for the diagnosis of urinary tract infection. *American journal of clinical pathology*. 2010;133(4):577-82.
138. Kadkhoda K, Manickam K, DeGagne P, Sokolowski P, Pang P, Kontzie N, et al. UF-1000i™ flow cytometry is an effective screening method for urine specimens. *Diagnostic microbiology and infectious disease*. 2011;69(2):130-6.
139. Pieretti B, Brunati P, Pini B, Colzani C, Congedo P, Rocchi M, et al. Diagnosis of bacteriuria and leukocyturia by automated flow cytometry compared with urine culture. *Journal of clinical microbiology*. 2010;48(11):3990-6.
140. Yang W, Yu X, Liu D, Guan X. High false positive rate of white blood cells in urine samples of pregnant women may be caused by epithelial cells being misclassified by the sysmex UF-1000i urine flow cytometer. *Cytometry Part B: Clinical Cytometry*. 2018.
141. Zaman Z, Roggeman S, Verhaegen J. Unsatisfactory performance of flow cytometer UF-100 and urine strips in predicting outcome of urine cultures. *Journal of clinical microbiology*. 2001;39(11):4169-71.

142. Previtali G, Ravasio R, Seghezzi M, Buoro S, Alessio MG. Performance evaluation of the new fully automated urine particle analyser UF-5000 compared to the reference method of the Fuchs-Rosenthal chamber. *Clinica Chimica Acta*. 2017;472:123-30.
143. De Rosa R, Grosso S, Lorenzi G, Bruschetta G, Camporese A. Evaluation of the new Sysmex UF-5000 fluorescence flow cytometry analyser for ruling out bacterial urinary tract infection and for prediction of Gram negative bacteria in urine cultures. *Clinica Chimica Acta*. 2018;484:171-8.
144. Linko S, Kouri TT, Toivonen E, Ranta PH, Chapoulaud E, Lalla M. Analytical performance of the Iris iQ200 automated urine microscopy analyzer. *Clinica chimica acta*. 2006;372(1-2):54-64.
145. Alves L, Ballester F, Camps J, Joven J. Preliminary evaluation of the Iris IQ™ 200 automated urine analyser. *Clinical Chemistry and Laboratory Medicine (CCLM)*. 2005;43(9):967-70.
146. Stürenburg E, Kramer J, Schön G, Cachovan G, Sobottka I. Detection of significant bacteriuria by use of the iQ200 automated urine microscope. *Journal of clinical microbiology*. 2014;JCM. 00112-14.
147. Kocer D, Sarıgüzel FM, Karakukcu C. Cutoff values for bacteria and leukocytes for urine sediment analyzer FUS200 in culture-positive urinary-tract infections. *Scandinavian journal of clinical and laboratory investigation*. 2014;74(5):414-7.
148. Shayanfar N, Tobler U, von Eckardstein A, Bestmann L. Automated urinalysis: first experiences and a comparison between the Iris iQ200 urine microscopy system, the Sysmex UF-100 flow cytometer and manual microscopic particle counting. *Clinical Chemical Laboratory Medicine*. 2007;45(9):1251-6.
149. Poropatich CO, Mendoza SM, Hitlan JJ, Wilkinson DS. Inconsistent detection of bacteriuria with the Yellow IRIS automated urinalysis workstation. *Laboratory Medicine*. 1988;19(8):499-501.
150. Van der Pol B, Cobas (R) 4800: a fully automated system for the detection of Chlamydia trachomatis and Neisseria gonorrhoeae. *Expert Rev Mol Diagn*. 2013;13(2):131-40.
151. Blocker ME, Krysiak RG, Behets F, Cohen MS, Hobbs MM. Quantification of Chlamydia trachomatis elementary bodies in urine by ligase chain reaction. *Journal of clinical microbiology*. 2002;40(10):3631-4.
152. Crotchfelt KA, Welsh LE, DeBonville D, Rosenstrauss M, Quinn TC. Detection of Neisseria gonorrhoeae and Chlamydia trachomatis in genitourinary specimens from men and women by a coamplification PCR assay. *Journal of clinical microbiology*. 1997;35(6):1536-40.
153. Petti CA. Detection and identification of microorganisms by gene amplification and sequencing. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 2007;44(8):1108-14.
154. Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, Gordon JI. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Science translational medicine*. 2009;1(6):6ra14.
155. Huse SM, Ye YZ, Zhou YJ, Fodor AA. A Core Human Microbiome as Viewed through 16S rRNA Sequence Clusters. *PloS one*. 2012;7(6).
156. Angly FE, Dennis PG, Skarshewski A, Vanwonterghem I, Hugenholtz P, Tyson GW. CopyRighter: a rapid tool for improving the accuracy of microbial community profiles through lineage-specific gene copy number correction. *Microbiome*. 2014;2:11.
157. Clark AE, Kaleta EJ, Arora A, Wolk DM. Matrix-assisted laser desorption ionization-time of flight mass spectrometry: a fundamental shift in the routine practice of clinical microbiology. *Clin Microbiol Rev*. 2013;26(3):547-603.
158. Strittmatter N, Rebec M, Jones EA, Golf O, Abdolrasouli A, Balog J, et al. Characterization and identification of clinically relevant microorganisms using rapid evaporative ionization mass spectrometry. *Analytical chemistry*. 2014;86(13):6555-62.
159. Tran A, Alby K, Kerr A, Jones M, Gilligan PH. Cost Savings Realized by Implementation of Routine Microbiological Identification by Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry. *Journal of clinical microbiology*. 2015;53(8):2473.
160. Ge M-C, Kuo A-J, Liu K-L, Wen Y-H, Chia J-H, Chang P-Y, et al. Routine identification of microorganisms by matrix-assisted laser desorption ionization time-of-flight mass spectrometry: Success rate, economic analysis, and clinical outcome. *Journal of Microbiology, Immunology and Infection*. 2017;50(5):662-8.
161. The Standards Unit MS, PHE. UK Standards for Microbiology Investigations: Matrix-Assisted Laser Desorption/Ionisation - Time of Flight Mass Spectrometry (MALDI-TOF MS) Test Procedure 2015.
162. Wang XH, Zhang G, Fan YY, Yang X, Sui WJ, Lu XX. Direct identification of bacteria causing urinary tract infections by combining matrix-assisted laser desorption ionization-time of flight mass spectrometry with UF-1000i urine flow cytometry. *Journal of microbiological methods*. 2013;92(3):231-5.

163. Ferreira L, Sanchez-Juanes F, Munoz-Bellido JL, Gonzalez-Buitrago JM. Rapid method for direct identification of bacteria in urine and blood culture samples by matrix-assisted laser desorption ionization time-of-flight mass spectrometry: intact cell vs. extraction method. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*. 2011;17(7):1007-12.
164. Kok J, Thomas LC, Olma T, Chen SCA, Iredell JR. Identification of Bacteria in Blood Culture Broths Using Matrix-Assisted Laser Desorption-Ionization Sepsityper (TM) and Time of Flight Mass Spectrometry. *PloS one*. 2011;6(8).
165. Schubert S, Weinert K, Wagner C, Gunzl B, Wieser A, Maier T, et al. Novel, Improved Sample Preparation for Rapid, Direct Identification from Positive Blood Cultures Using Matrix-Assisted Laser Desorption/Ionization Time-of-Flight (MALDI-TOF) Mass Spectrometry. *J Mol Diagn*. 2011;13(6):701-6.
166. Tadros M, Petrich A. Evaluation of MALDI-TOF mass spectrometry and Sepsityper Kit (TM) for the direct identification of organisms from sterile body fluids in a Canadian pediatric hospital. *Can J Infect Dis Med*. 2013;24(4):191-4.
167. Christner M, Rohde H, Wolters M, Sobottka I, Wegscheider K, Aepfelbacher M. Rapid Identification of Bacteria from Positive Blood Culture Bottles by Use of Matrix-Assisted Laser Desorption-Ionization Time of Flight Mass Spectrometry Fingerprinting. *Journal of clinical microbiology*. 2010;48(5):1584-91.
168. Prod'homme G, Bizzini A, Durussel C, Bille J, Greub G. Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry for Direct Bacterial Identification from Positive Blood Culture Pellets. *Journal of clinical microbiology*. 2010;48(4):1481-3.
169. Cooks RG, Ouyang Z, Takats Z, Wiseman JM. Ambient mass spectrometry. *Science*. 2006;311(5767):1566-70.
170. Takats Z, Wiseman JM, Gologan B, Cooks RG. Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. *Science*. 2004;306(5695):471-3.
171. Zhang JI, Talaty N, Costa AB, Xia Y, Tao WA, Bell R, et al. Rapid direct lipid profiling of bacteria using desorption electrospray ionization mass spectrometry. *International Journal of Mass Spectrometry*. 2011;301(1):37-44.
172. Meetani MA, Shin YS, Zhang S, Mayer R, Basile F. Desorption electrospray ionization mass spectrometry of intact bacteria. *Journal of mass spectrometry*. 2007;42(9):1186-93.
173. Song Y, Talaty N, Datsenko K, Wanner BL, Cooks RG. In vivo recognition of *Bacillus subtilis* by desorption electrospray ionization mass spectrometry (DESI-MS). *The Analyst*. 2009;134(5):838-41.
174. Takats Z, Wiseman JM, Cooks RG. Ambient mass spectrometry using desorption electrospray ionization (DESI): instrumentation, mechanisms and applications in forensics, chemistry, and biology. *Journal of Mass Spectrometry*. 2005;40(10):1261-75.
175. Dunham SJ, Ellis JF, Li B, Sweedler JV. Mass spectrometry imaging of complex microbial communities. *Accounts of chemical research*. 2016;50(1):96-104.
176. Watrous J, Hendricks N, Meehan M, Dorrestein PC. Capturing bacterial metabolic exchange using thin film desorption electrospray ionization-imaging mass spectrometry. *Analytical chemistry*. 2010;82(5):1598-600.
177. Roach PJ, Laskin J, Laskin A. Nanospray desorption electrospray ionization: an ambient method for liquid-extraction surface sampling in mass spectrometry. *The Analyst*. 2010;135(9):2233-6.
178. Watrous J, Roach P, Alexandrov T, Heath BS, Yang JY, Kersten RD, et al. Mass spectral molecular networking of living microbial colonies. *Proceedings of the National Academy of Sciences*. 2012;109(26):E1743-E52.
179. Gross JH. Direct analysis in real time—a critical review on DART-MS. *Analytical and bioanalytical chemistry*. 2014;406(1):63-80.
180. Cody RB, Laramée JA, Durst HD. Versatile new ion source for the analysis of materials in open air under ambient conditions. *Analytical chemistry*. 2005;77(8):2297-302.
181. Shelley JT, Wiley JS, Chan GC, Schilling GD, Ray SJ, Hieftje GM. Characterization of direct-current atmospheric-pressure discharges useful for ambient desorption/ionization mass spectrometry. *Journal of the American Society for Mass Spectrometry*. 2009;20(5):837-44.
182. Fox A. Mass Spectrometry for Species or Strain Identification after Culture or without Culture: Past, Present, and Future. *Journal of clinical microbiology*. 2006;44(8):2677-80.
183. Ehrhardt CJ, Chu V, Brown T, Simmons TL, Swan BK, Bannan J, et al. Use of Fatty Acid Methyl Ester Profiles for Discrimination of *Bacillus cereus* T-Strain Spores Grown on Different Media. *Applied and environmental microbiology*. 2010;76(6):1902-12.

184. Pierce CY, Barr JR, Cody RB, Massung RF, Woolfitt AR, Moura H, et al. Ambient generation of fatty acid methyl ester ions from bacterial whole cells by direct analysis in real time (DART) mass spectrometry. *Chemical communications*. 2007(8):807-9.
185. Cody RB, McAlpin CR, Cox CR, Jensen KR, Voorhees KJ. Identification of bacteria by fatty acid profiling with direct analysis in real time mass spectrometry. *Rapid Communications in Mass Spectrometry*. 2015;29(21):2007-12.
186. Lin C-H, Liao W-C, Chen H-K, Kuo T-Y. Paper spray-MS for bioanalysis. *Bioanalysis*. 2014;6(2):199-208.
187. Espy RD, Muliadi AR, Ouyang Z, Cooks RG. Spray mechanism in paper spray ionization. *International Journal of Mass Spectrometry*. 2012;325-327:167-71.
188. Wlekinski M, Li Y, Bag S, Sarkar D, Narayanan R, Pradeep T, et al. Zero volt paper spray ionization and its mechanism. *Analytical chemistry*. 2015;87(13):6786-93.
189. Wei P, Bag S, Pulliam CJ, Snyder DT, Pielak RM, Cooks RG. Analysis of bacteria using zero volt paper spray. *Analytical Methods*. 2016;8(8):1770-3.
190. Liu J, Wang H, Manicke NE, Lin J-M, Cooks RG, Ouyang Z. Development, characterization, and application of paper spray ionization. *Analytical chemistry*. 2010;82(6):2463-71.
191. Sinclair I, Stearns R, Pringle S, Wingfield J, Datwani S, Hall E, et al. Novel Acoustic Loading of a Mass Spectrometer: Toward Next-Generation High-Throughput MS Screening. *J Lab Autom*. 2016;21(1):19-26.
192. Zhang H. Acoustic dispensing-mass spectrometry: the next high throughput bioanalytical platform for early drug discovery. *Bioanalysis*. 2017;9(21):1619-21.
193. Balog J, Sasi-Szabo L, Kinross J, Lewis MR, Muirhead LJ, Veselkov K, et al. Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. *Science translational medicine*. 2013;5(194):194ra93.
194. Strittmatter N, Jones EA, Veselkov KA, Rebec M, Bundy JG, Takats Z. Analysis of intact bacteria using rapid evaporative ionisation mass spectrometry. *Chemical communications*. 2013;49(55):6188-90.
195. Bolt F, Cameron S, Karancsi T, Simon D, Rickards T, Hardiman K, et al. Automated High-Throughput Identification and Characterisation of Clinically Important Bacteria and Fungi using Rapid Evaporative Ionisation Mass Spectrometry (REIMS). *Analytical chemistry*. 2016.
196. Cameron SJS, Bolt F, Perdonés-Montero A, Rickards T, Hardiman K, Abdolrasouli A, et al. Rapid Evaporative Ionisation Mass Spectrometry (REIMS) Provides Accurate Direct from Culture Species Identification within the Genus *Candida*. *Scientific reports*. 2016;6.
197. Tomley FM, Shirley MW. Livestock infectious diseases and zoonoses. The Royal Society; 2009.
198. Olea-Popelka F, Muwonge A, Perera A, Dean AS, Mumford E, Erlacher-Vindel E, et al. Zoonotic tuberculosis in human beings caused by *Mycobacterium bovis*—a call for action. *The Lancet Infectious Diseases*. 2017;17(1):e21-e5.
199. Issa MdA, Soares Filho PM, Fonseca Júnior AA, Hodon MA, Santos LCd, Reis JKPd, et al. Comparative study of *Mycobacterium bovis* primary isolation methods. *Brazilian journal of microbiology*. 2017;48(1):139-44.
200. Bakshi C, Shah D, Verma R, Singh R, Malik M. Rapid differentiation of *Mycobacterium bovis* and *Mycobacterium tuberculosis* based on a 12.7-kb fragment by a single tube multiplex-PCR. *Veterinary microbiology*. 2005;109(3-4):211-6.
201. Monaghan M, Doherty M, Collins J, Kazda J, Quinn P. The tuberculin test. *Veterinary microbiology*. 1994;40(1-2):111-24.
202. Halasa T, Huijps K, Østerås O, Hogeveen H. Economic effects of bovine mastitis and mastitis management: a review. *Veterinary quarterly*. 2007;29(1):18-31.
203. Viguier C, Arora S, Gilmartin N, Welbeck K, O'Kennedy R. Mastitis detection: current trends and future perspectives. *Trends in biotechnology*. 2009;27(8):486-93.
204. Bradley AJ. Bovine mastitis: an evolving disease. *Veterinary Journal*. 2002;164(2):116-28.
205. DeGraves FJ, Fetrow J. Economics of mastitis and mastitis control. *The Veterinary Clinics of North America Food Animal Practice*. 1993;9(3):421-34.
206. Dettileux J, Kehrli Jr M, Stabel J, Freeman A, Kelley D. Study of immunological dysfunction in periparturient Holstein cattle selected for high and average milk production. *Veterinary immunology and immunopathology*. 1995;44(3-4):251-67.
207. Sordillo LM, Streicher KL. Mammary gland immunity and mastitis susceptibility. *Journal of mammary gland biology and neoplasia*. 2002;7(2):135-46.

208. Pankey J, Drechsler P, Wildman E. Mastitis prevalence in primigravid heifers at parturition. *Journal of dairy science*. 1991;74(5):1550-2.
209. Schukken YH, Wilson DJ, Welcome F, Garrison-Tikofsky L, Gonzalez RN. Monitoring udder health and milk quality using somatic cell counts. *Veterinary research*. 2003;34(5):579-96.
210. Barreiro JR, Gonçalves JL, Braga PAC, Dibbern AG, Eberlin MN, dos Santos MV. Non-culture-based identification of mastitis-causing bacteria by MALDI-TOF mass spectrometry. *Journal of dairy science*. 2017;100(4):2928-34.
211. Hiitö H, Riva R, Autio T, Pohjanvirta T, Holopainen J, Pyörälä S, et al. Performance of a real-time PCR assay in routine bovine mastitis diagnostics compared with in-depth conventional culture. *Journal of Dairy Research*. 2015;82(2):200-8.
212. Osborne J. Notes on the use of data transformations. *Practical assessment, research and evaluation*. 2005;9(1):42-50.
213. Balog J, Sasi-Szabo L, Kinross J, Lewis MR, Muirhead LJ, Veselkov K, et al. Intraoperative Tissue Identification Using Rapid Evaporative Ionization Mass Spectrometry. *Science translational medicine*. 2013;5(194).
214. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*. 2001;58(2):109-30.
215. Wold S, Johansson E, Cocchi M. PLS—partial least-squares projections to latent structures, 3D QSAR in drug design. *Theory Methods and Applications* Leiden: ESCOM Science Publishers. 1993.
216. Rokach L. Decision forest: Twenty years of research. *Information Fusion*. 2016;27:111-25.
217. Xia J, Sinelnikov IV, Han B, Wishart DS. MetaboAnalyst 3.0—making metabolomics more meaningful. *Nucleic acids research*. 2015.
218. Fahy E, Sud M, Cotter D, Subramaniam S. LIPID MAPS online tools for lipid research. *Nucleic Acids Res*. 2007;35(Web Server issue):W606-12.
219. Guijas C, Montenegro-Burke JR, Domingo-Almenara X, Palermo A, Warth B, Hermann G, et al. METLIN: A Technology Platform for Identifying Knowns and Unknowns. *Analytical chemistry*. 2018;90(5):3156-64.
220. Strittmatter N. Development of novel mass spectrometric methods for the characterisation and identification of microorganisms [Doctoral]: Imperial College London; 2016.
221. Hilt EE, McKinley K, Pearce MM, Rosenfeld AB, Zilliox MJ, Mueller ER, et al. Urine Is Not Sterile: Use of Enhanced Urine Culture Techniques To Detect Resident Bacterial Flora in the Adult Female Bladder. *Journal of clinical microbiology*. 2014;52(3):871-6.
222. Brubaker L, Wolfe AJ. The new world of the urinary microbiota in women. *American Journal of Obstetrics and Gynecology*. 2015;213(5):644-9.
223. Hooton TM, Roberts PL, Cox ME, Stapleton AE. Voided Midstream Urine Culture and Acute Cystitis in Premenopausal Women. *New England Journal of Medicine*. 2013;369(20):1883-91.
224. Scohy A, Noël A, Boeras A, Brassinne L, Laurent T, Rodriguez-Villalobos H, et al. Evaluation of the Bruker® MBT Sepsityper IVD module for the identification of polymicrobial blood cultures with MALDI-TOF MS. *European Journal of Clinical Microbiology & Infectious Diseases*. 2018;37(11):2145-52.
225. Bouza E, Burillo A, Munoz P, Guinea J, Marín M, Rodríguez-Crèixems M. Mixed bloodstream infections involving bacteria and *Candida* spp. *J Antimicrob Chemoth*. 2013;68(8):1881-8.
226. Royo-Cebrecos C, Gudíol C, Ardanuy C, Pomares H, Calvo M, Carratalà J. A fresh look at polymicrobial bloodstream infection in cancer patients. *PloS one*. 2017;12(10):e0185768.
227. Pavlaki M, Poulakou G, Drimousis P, Adamis G, Apostolidou E, Gatselis NK, et al. Polymicrobial bloodstream infections: epidemiology and impact on mortality. *Journal of global antimicrobial resistance*. 2013;1(4):207-12.
228. Khamis MM, Adamko DJ, El-Aneed A. Mass spectrometric based approaches in urine metabolomics and biomarker discovery. *Mass Spectrom Rev*. 2017;36(2):115-34.
229. Wiseman JM, Kennedy JH. Analysis of dried blood spots using DESI mass spectrometry. *Methods in molecular biology*. 2014;1198:291-7.
230. Déglon J, Thomas A, Mangin P, Staub C. Direct analysis of dried blood spots coupled with mass spectrometry: concepts and biomedical applications. *Analytical and bioanalytical chemistry*. 2012;402(8):2485-98.
231. Manicke NE, Abu-Rabie P, Spooner N, Ouyang Z, Cooks RG. Quantitative Analysis of Therapeutic Drugs in Dried Blood Spot Samples by Paper Spray Mass Spectrometry: An Avenue to Therapeutic Drug Monitoring. *Journal of the American Society for Mass Spectrometry*. 2011;22(9):1501-7.

232. Lee Y, Lai KKY, Sadrzadeh SMH. Simultaneous detection of 19 drugs of abuse on dried urine spot by liquid chromatography–tandem mass spectrometry. *Clinical biochemistry*. 2013;46(12):1118-24.
233. Otero-Fernandez M, Cocho JA, Tabernero MJ, Bermejo AM, Bermejo-Barrera P, Moreda-Pineiro A. Direct tandem mass spectrometry for the simultaneous assay of opioids, cocaine and metabolites in dried urine spots. *Analytica chimica acta*. 2013;784:25-32.
234. Franz M, Hörl WH. Common errors in diagnosis and management of urinary tract infection. I: Pathophysiology and diagnostic techniques. *Nephrol Dial Transpl*. 1999;14(11):2746-53.
235. Pieretti B, Brunati P, Pini B, Colzani C, Congedo P, Rocchi M, et al. Diagnosis of bacteriuria and leukocyturia by automated flow cytometry compared with urine culture. *Journal of clinical microbiology*. 2010;48(11):3990-6.
236. Alam M, Coulter J, Pacheco J, Correia J, Ribeiro M, Coelho M, et al. Comparison of urine contamination rates using three different methods of collection: clean-catch, cotton wool pad and urine bag. *Annals of tropical paediatrics*. 2005;25(1):29-34.
237. Tosif S, Baker A, Oakley E, Donath S, Babl FE. Contamination rates of different urine collection methods for the diagnosis of urinary tract infections in young children: an observational cohort study. *J Paediatr Child H*. 2012;48(8):659-64.
238. Zhou X, Ouyang Z. Ion transfer between ion source and mass spectrometer inlet: electro-hydrodynamic simulation and experimental validation. *Rapid communications in mass spectrometry : RCM*. 2016;30 Suppl 1:29-33.
239. Garimella S, Xu W, Huang G, Harper JD, Cooks RG, Ouyang Z. Gas-flow assisted ion transfer for mass spectrometry. *Journal of Mass Spectrometry*. 2012;47(2):201-7.
240. Golf O, Strittmatter N, Karancsi T, Pringle SD, Speller AV, Mroz A, et al. Rapid evaporative ionization mass spectrometry imaging platform for direct mapping from bulk tissue and bacterial growth media. *Analytical chemistry*. 2015;87(5):2527-34.
241. Kusters R, Breukink E, Gallusser A, Kuhn A, De Kruijff B. A dual role for phosphatidylglycerol in protein translocation across the Escherichia coli inner membrane. *Journal of biological chemistry*. 1994;269(2):1560-3.
242. de Vrije T, de Swart RL, Dowhan W, Tommassen J, de Kruijff B. Phosphatidylglycerol is involved in protein translocation across Escherichia coli inner membranes. *Nature*. 1988;334(6178):173-5.
243. Xia W, Dowhan W. Phosphatidylinositol cannot substitute for phosphatidylglycerol in supporting cell growth of Escherichia coli. *Journal of bacteriology*. 1995;177(10):2926-8.
244. Hasin M, Kennedy EP. Role of phosphatidylethanolamine in the biosynthesis of pyrophosphoethanolamine residues in the lipopolysaccharide of Escherichia coli. *Journal of Biological Chemistry*. 1982;257(21):12475-7.
245. Lu Q, Wang J, Faghihnejad A, Zeng H, Liu Y. Understanding the molecular interactions of lipopolysaccharides during E. coli initial adhesion with a surface forces apparatus. *Soft Matter*. 2011;7(19):9366-79.
246. Yu C, Li M, Sun Y, Wang X, Chen Y. Phosphatidylethanolamine Deficiency Impairs Escherichia coli Adhesion by Downregulating Lipopolysaccharide Synthesis, Which is Reversible by High Galactose/Lactose Cultivation. *Cell Communication & Adhesion*. 2017;23(1):1-10.
247. Oursel D, Loutelier-Bourhis C, Orange N, Chevalier S, Norris V, Lange CM. Lipid composition of membranes of Escherichia coli by liquid chromatography/tandem mass spectrometry using negative electrospray ionization. *Rapid Communications in Mass Spectrometry*. 2007;21(11):1721-8.
248. Jolkkonen S, Paattiniemi E-L, Kärpänäja P, Sarkkinen H. Screening of Urine Samples by Flow Cytometry Reduces the Need for Culture. *Journal of clinical microbiology*. 2010;48(9):3117-21.
249. Singh L, Cariappa MP, Kaur M. Klebsiella oxytoca: An emerging pathogen? *Medical journal, Armed Forces India*. 2016;72(Suppl 1):S59-s61.
250. Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. *Mass Spectrom Rev*. 2007;26(1):51-78.
251. Derewacz DK, Goodwin CR, McNees CR, McLean JA, Bachmann BO. Antimicrobial drug resistance affects broad changes in metabolomic phenotype in addition to secondary metabolism. *Proceedings of the National Academy of Sciences*. 2013;110(6):2336-41.
252. Aros-Calt S, Muller BH, Boudah S, Ducruix Cl, Gervasi G, Junot C, et al. Annotation of the Staphylococcus aureus metabolome using liquid chromatography coupled to high-resolution mass spectrometry and application to the study of methicillin resistance. *J Proteome Res*. 2015;14(11):4863-75.

253. Li H, Xia X, Li X, Naren G, Fu Q, Wang Y, et al. Untargeted metabolomic profiling of amphenicol-resistant *Campylobacter jejuni* by ultra-high-performance liquid chromatography–mass spectrometry. *J Proteome Res.* 2014;14(2):1060-8.
254. Joyce SM. Sputum analysis and culture. *Annals of Emergency Medicine.* 1986;15(3):325-8.
255. Weinstein MP. Current blood culture methods and systems: clinical concepts, technology, and interpretation of results. *Clinical infectious diseases.* 1996;23(1):40-6.
256. Barreiro JR, Ferreira CR, Sanvido GB, Kostrzewa M, Maier T, Wegemann B, et al. Short communication: Identification of subclinical cow mastitis pathogens in milk by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Journal of dairy science.* 2010;93(12):5661-7.
257. Kitchen B. Review of the progress of dairy science: Bovine mastitis: milk compositional changes and related diagnostic test. *J Dairy Res.* 1981;48:167-88.
258. Randolph H, Erwin R. Influence of mastitis on properties of milk. X. Fatty acid composition. *Journal of dairy science.* 1974;57(8):865-8.
259. Erwin R, Randolph H. Influence of mastitis on properties of milk XI. Fat globule membrane. *Journal of dairy science.* 1975;58(1):9-12.
260. Pyörälä S. Indicators of inflammation in the diagnosis of mastitis. *Veterinary research.* 2003;34(5):565-78.
261. Holland R, Wilkes J, Rafii F, Sutherland J, Persons C, Voorhees K, et al. Rapid identification of intact whole bacteria based on spectral patterns using matrix-assisted laser desorption/ionization with time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry.* 1996;10(10):1227-32.
262. Heller D, Cotter RJ, Fenselau C, Uy O. Profiling of bacteria by fast atom bombardment mass spectrometry. *Analytical chemistry.* 1987;59(23):2806-9.
263. Miller LT. Single derivatization method for routine analysis of bacterial whole-cell fatty acid methyl esters, including hydroxy acids. *Journal of clinical microbiology.* 1982;16(3):584-6.
264. Van Veen S, Claas E, Kuijper EJ. High-throughput identification of bacteria and yeast by matrix-assisted laser desorption ionization-time of flight mass spectrometry in conventional medical microbiology laboratories. *Journal of clinical microbiology.* 2010;48(3):900-7.
265. Ryzhov V, Fenselau C. Characterization of the protein subset desorbed by MALDI from whole bacterial cells. *Analytical chemistry.* 2001;73(4):746-50.
266. Suarez S, Ferroni A, Lotz A, Jolley KA, Guerin P, Leto J, et al. Ribosomal proteins as biomarkers for bacterial identification by mass spectrometry in the clinical microbiology laboratory. *Journal of microbiological methods.* 2013;94(3):390-6.
267. Ruppé E, Cherkaoui A, Lazarevic V, Emonet S, Schrenzel J. Establishing genotype-to-phenotype relationships in bacteria causing hospital-acquired pneumonia: a prelude to the application of clinical metagenomics. *Antibiotics.* 2017;6(4):30.
268. Tan KE, Ellis BC, Lee R, Stamper PD, Zhang SX, Carroll KC. Prospective evaluation of a matrix-assisted laser desorption ionization-time of flight mass spectrometry system in a hospital clinical microbiology laboratory for identification of bacteria and yeasts: a bench-by-bench study for assessing the impact on time to identification and cost-effectiveness. *Journal of clinical microbiology.* 2012;50(10):3301-8.
269. Picaud V, Giovannelli J-F, Truntzer C, Charrier J-P, Giremus A, Grangeat P, et al. Linear MALDI-ToF simultaneous spectrum deconvolution and baseline removal. *Bmc Bioinformatics.* 2018;19(1):123.
270. Hariu M, Watanabe Y, Oikawa N, Seki M. Usefulness of matrix-assisted laser desorption ionization time-of-flight mass spectrometry to identify pathogens, including polymicrobial samples, directly from blood culture broths. *Infection and drug resistance.* 2017;10:115-20.
271. Haiko J, Savolainen LE, Hilla R, Pätäri-Sampo A. Identification of urinary tract pathogens after 3-hours urine culture by MALDI-TOF mass spectrometry. *Journal of microbiological methods.* 2016;129:81-4.
272. Zboromyrska Y, Rubio E, Alejo I, Vergara A, Mons A, Campo I, et al. Development of a new protocol for rapid bacterial identification and susceptibility testing directly from urine samples. *Clin Microbiol Infect.* 2016;22(6):561.e1-e6.
273. Burnham C-AD, DeMarco ML. Diafiltration MALDI-TOF Mass Spectrometry Method for Culture-Independent Detection and Identification of Pathogens Directly From Urine Specimens. *American journal of clinical pathology.* 2014;141(2):204-12.

274. Huang B, Zhang L, Zhang W, Liao K, Zhang S, Zhang Z, et al. Direct Detection and Identification of Bacterial Pathogens from Urine with Optimized Specimen Processing and Enhanced Testing Algorithm. *Journal of clinical microbiology*. 2017;55(5):1488-95.
275. Wain J, Schmidt K, Mwaigwisya S, O'Grady J, Livermore DM, Crossman LC, et al. Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. *J Antimicrob Chemoth*. 2016;72(1):104-14.
276. Verplanken K, Stead S, Jandova R, Poucke CV, Claereboudt J, Bussche JV, et al. Rapid evaporative ionization mass spectrometry for high-throughput screening in food analysis: The case of boar taint. *Talanta*. 2017;169:30-6.
277. Balog J, Perenyi D, Guallar-Hoyas C, Egri A, Pringle SD, Stead S, et al. Identification of the Species of Origin for Meat Products by Rapid Evaporative Ionization Mass Spectrometry. *Journal of agricultural and food chemistry*. 2016;64(23):4793-800.
278. Alexander J, Gildea L, Balog J, Speller A, McKenzie J, Muirhead L, et al. A novel methodology for in vivo endoscopic phenotyping of colorectal cancer based on real-time analysis of the mucosal lipidome: a prospective observational study of the iKnife. *Surgical endoscopy*. 2017;31(3):1361-70.
279. Ecelberger SA, Cornish TJ, Collins BF, Lewis DL, Bryden WA. Suitcase TOF: a man-portable time-of-flight mass spectrometer. *Johns Hopkins APL technical digest*. 2004;25(1):14-9.

10 Supplementary

Supplementary Table 1. Xevo G2-XS Q-ToF instrument parameters in MS/MS mode.

Instrument Calibration:

Calibration File:
Parameters
MS1 Static: None
MS1 Scanning:
Mass: 50 Da to 2500 Da.
Resolution: 0.0/0.0
Ion Energy: 0.0
Reference File: ESI_NaFormate_Neg
Acquisition File: 2018_10_17_NEGCAL

MS1 Scan Speed Compensation: None
MS2 Static: None
MS2 Scanning: None
MS2 Scan Speed Compensation: None
Calibration Time: 11:55
Calibration Date: 10/17/18
Coefficients
MS1 Static: None
MS2 Static: None

Instrument Configuration:

Lteff 1800.0
Veff 6336.30
Resolution 22000
Min Points in Peak 2
Acquisition Device WatersADC
ADC Trigger Threshold (V) -1.00
ADC Input Offset (V) -1.59
Average Single Ion Intensity 27

ADC Amplitude Threshold 3
ADC Centroid Threshold -1
ADC Ion Area Threshold 3
ADC Ion Area Offset 15
ADC Pushes Per IMS Increment 1
TargetEnhancement Delay Coefficient 2.1500
TargetEnhancement Delay Offset 0.0000

Experimental Instrument Parameters

Instrument Parameter Filename
C:\Users\Administrator\Desktop\Adam\Bacteria_
heatedcoil.ipr (MODIFIED)
Polarity ES-
Analyser Sensitivity Mode
Capillary (kV) 2.3000
Sampling Cone 80.0000
Source Temperature (°C) 100
Source Offset 50
Desolvation Temperature (°C) 280
Cone Gas Flow (L/Hr) 0.0
Desolvation Gas Flow (L/Hr) 800.0
LM Resolution 5.0
HM Resolution 17.5
Aperture 1 0.0
Pre-filter 2.0
Ion Energy 0.8
Manual Collision Energy TRUE
Collision Energy 36.0
Detector 2725
DetectorCache 0
Sample Infusion Flow Rate (µL/min) 5
Sample Flow State LC
Sample Fill Volume (µL) 250
Sample Reservoir Wash
LockSpray Infusion Flow Rate (µL/min) 10
LockSpray Flow State Infusion
LockSpray Reservoir B
LockSpray Capillary (kV) 2.50
Use Manual LockSpray Collision Energy FALSE
Collision Energy 6.0
Focus2 41.0
Acceleration1 30.0
Acceleration2 110.0
Aperture2 30.0
Transport1 30.0
Transport2 30.0
Steering -0.20
Tube Lens 25
Pusher 1900.0
Pusher Offset -0.25
Puller 1400.0
Pusher Cycle Time (µs) Automatic
Pusher Width (µs) Automatic
Collector 60
Collector Pulse 10.0
Stopper 10

Stopper Pulse 20.0
Entrance 17
Static Offset 120
Puller Offset 0.00
Reflectron Grid (kV) 1.703
Flight Tube (kV) 9.00
Reflectron (kV) 1.602
TWave Entrance 2.00
Static Offset 120.00
TWave Exit 15.00
TWave Trap Height 4.00
TWave Extract Height 15.00
Trap Wave Velocity (m/s) 260
Trap Wave Height (V) 0.2
Cell 2 Gradient 1
Cell2 Offset B 0
Cell2 Offset C 0
Step Wave 1 In Manual Control OFF
Enable Reverse Operation OFF
Step Wave 1 In Velocity (m/s) 300.0
Step Wave 1 In Height 15.0
Step Wave 1 Out Manual Control OFF
Step Wave 1 Out Velocity (m/s) 300.0
Step Wave 1 Out Height 15.0
Step Wave 2 Manual Control OFF
Step Wave 2 Velocity (m/s) 300.0
Step Wave 2 Height 1.0
Use Manual Step Wave DC OFF
Step Wave TransferOffset 25.0
Step Wave DiffAperture1 3.0
Step Wave DiffAperture2 0.0
Use Automatic RF Settings FALSE
StepWave1RFOffset 250.0
StepWave2RFOffset 300.0
Backing 1.66e0
Collision 1.82e-1
TOF 7.44e-7
Collision RF Offset 600
Collision RF Initial 200
Collision RF Final 600
MS Profile Type Auto P
MSProfileMass1 100
MSProfileDwellTime1 20
MSProfileRampTime1 20
MSProfileMass2 300
MSProfileDwellTime2 20
MSProfileRampTime2 40

MSProfileMass3	500	PusherInterval	118.000000
Use Automatic RF Settings	FALSE	PusherOffset	0.250000
AutoStepWave1RFOffset	300	LockMassValidSigma	5
AutoStepWave2RFOffset	350		

Acquisition mass range

Start mass	50.000	Start mass	0.000
End mass	2500.000	End mass	0.000
Calibration mass range			

Function Parameters - Function 1 - TOF MSMS FUNCTION

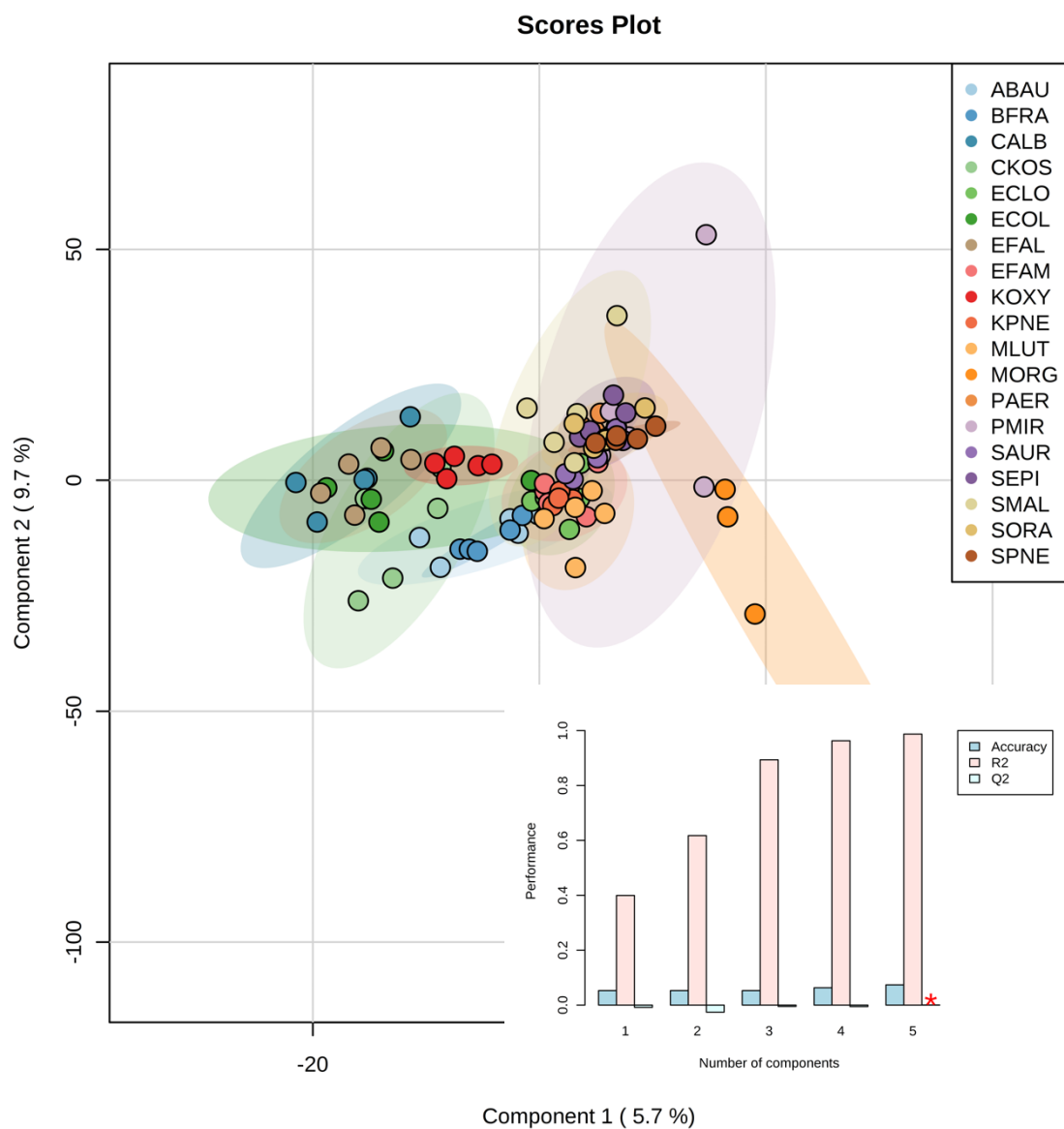
Scan Time (sec)	1.000	ADC Pusher Frequency (µs)	118.0
Interscan Time (sec)	0.014	ADC Pusher Width (µs)	2.75
Set Mass	733.5	Use Tune Page Cone Voltage	YES
Start Mass	50.0	Using Auto Collision Energy (eV)	6.000000
MSMS End Mass	2500.0	Sensitivity	Normal
Start Time (mins)	0.00	Dynamic Range	Normal
End Time (mins)	0.00	Save Collapsed Retention Time Data	No
Data Format	Continuum	Use Rule File Filtering	No
ADC Sample Frequency (GHz)	3.0	Calibration	Dynamic 2

Function 1

Scans in function:	66
Cycle time (sec):	1.014
Scan duration (sec):	1.000
Inter Scan Delay (sec):	0.014
Start and End Time(mins):	Always on
Ionisation mode:	ES-
Data type:	Enhanced Mass
Function type:	TOF Daughter
Mass range:	50 to 2500
Collision Energy:	

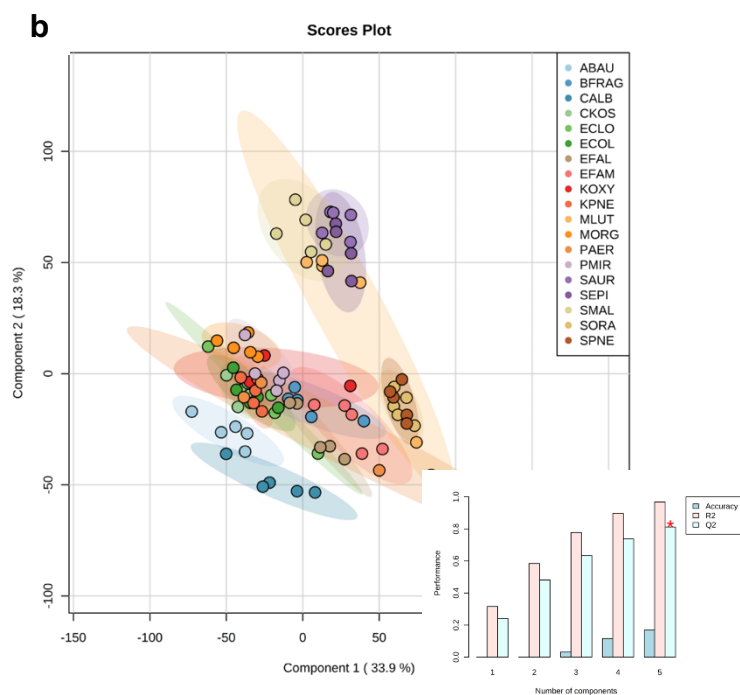
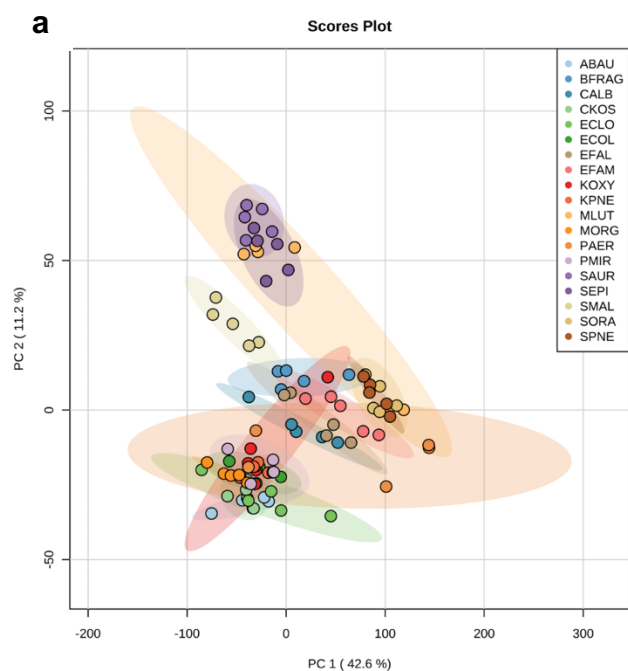
Supplementary Table 2. Modified cell line probes for REIMS were manufactured by the WRC to a range of specifications. Probes with varying electrode material, length and insulations were utilised to find an optimal configuration.

Probe Colour	Band Identifier	Material	Electrode Spacer	Electrode Length	Electrode Insulation
Yellow	None	304 annealed stainless steel (tactile)	None	20mm	No
Yellow	None	304 annealed stainless steel (tactile)	None	22.5mm	No
Yellow	None	304 annealed stainless steel (tactile)	None	25mm	No
Green	Yellow	304 stainless steel	Yes	24mm	Yes - long
Green	Green	Carbon steel, spring steel (full hard)	Yes	24mm	Yes - long
Green	Brown	Stainless steel (1/2 hard)	Yes	24mm	Yes - long
Green	White	Stainless steel, soft, annealed	Yes	24mm	Yes - long
Green	Purple	Low-carbon steel (full hard)	Yes	24mm	Yes - long
Green	Black	Stainless steel (full hard)	Yes	24mm	Yes - long
Green	Grey	Stainless steel (full hard)	Yes	24mm	Yes - long
Blue	Blue	Stainless steel (1/2 hard)	None	23mm	Yes - short
Blue	White	Stainless steel, soft, annealed	None	23mm	Yes - short
Blue	Grey	Stainless steel (full hard)	None	23mm	Yes - short

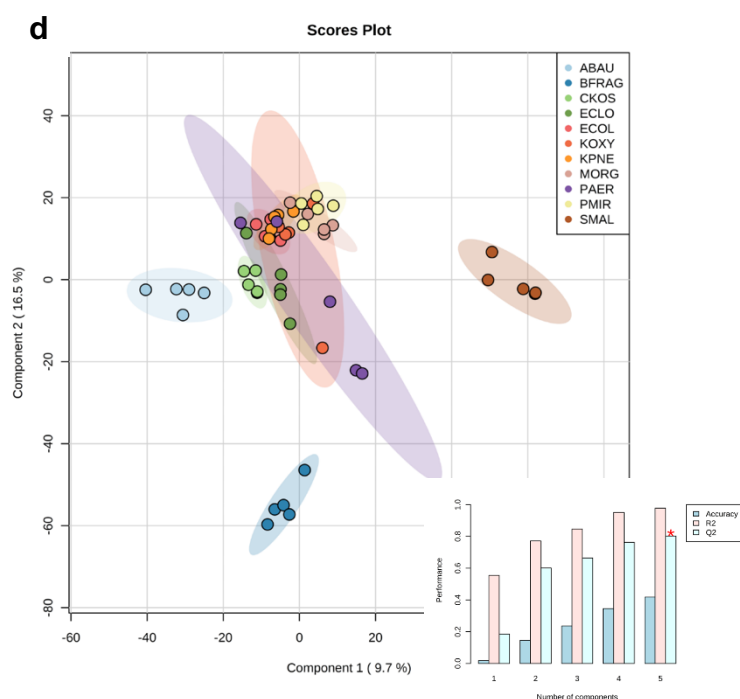
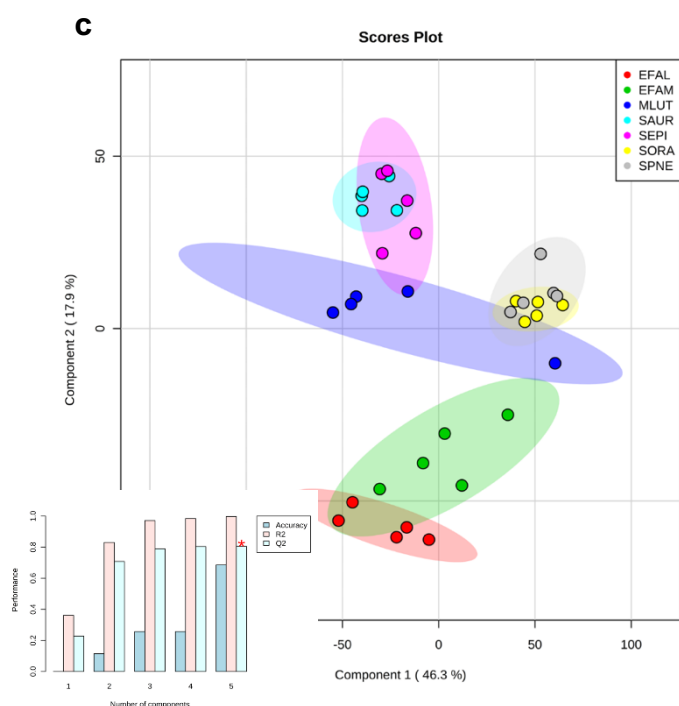


* denotes number of components from which model building results in the highest Q₂ score

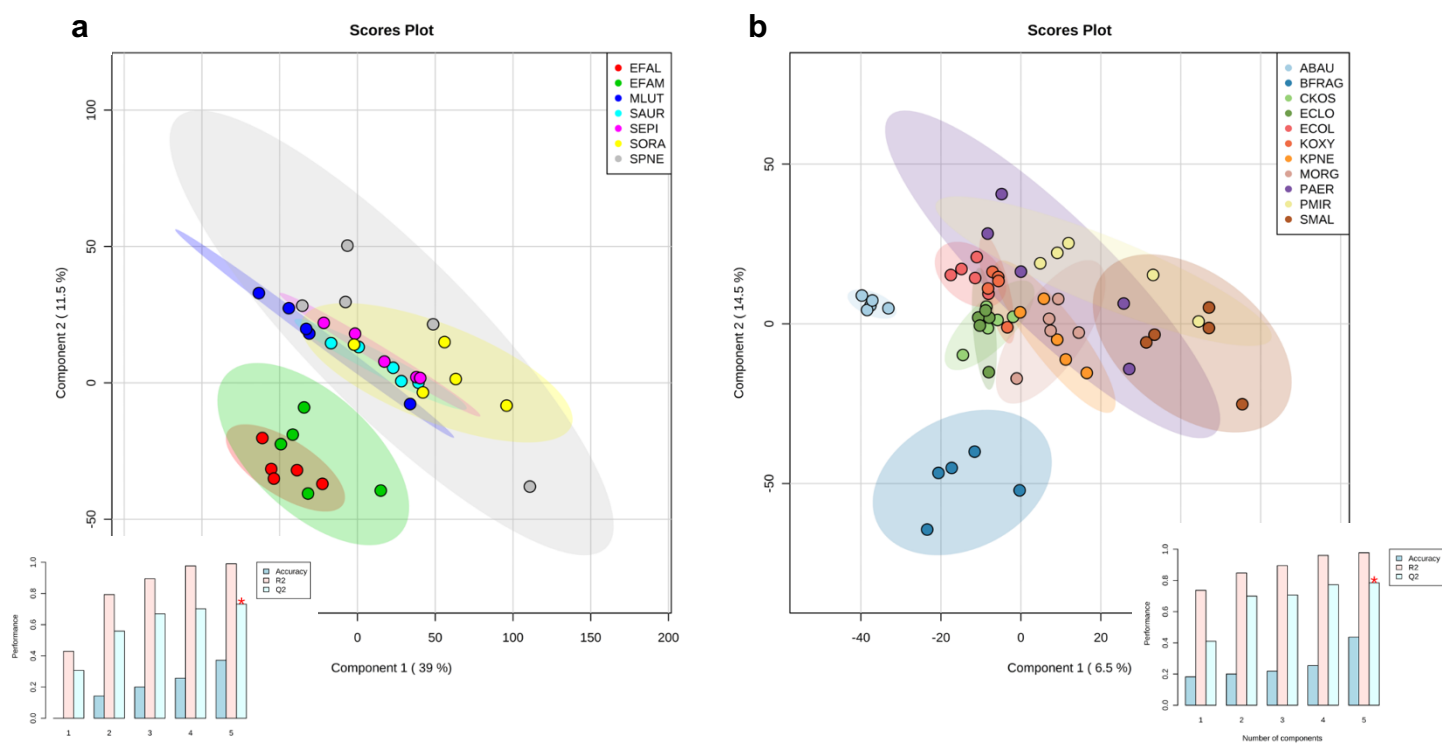
Supplementary 1. Partial least squares discriminant analysis of all bacterial species replicates for cell line probe REIMS (pellet analysis). Weak clustering of species replicates is seen with poor separation.



* denotes number of components from which model building results in the highest Q2 score

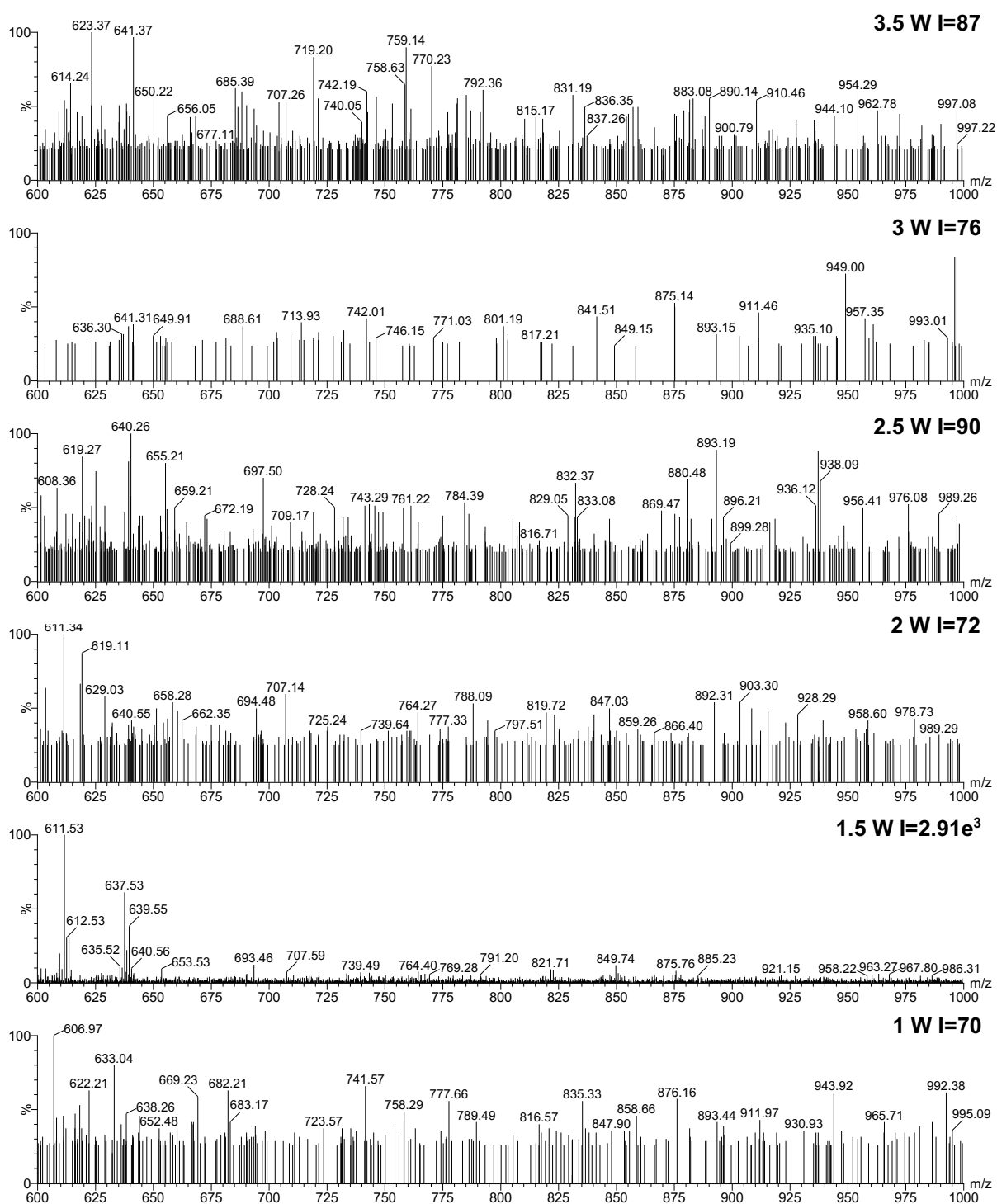


Supplementary 2. Two-dimensional principal component analysis (a), partial least squares discriminant analysis (b) of all bacterial species replicates and partial least squares discriminant analysis of Gram-positive (c) and Gram-negative (d) bacterial species replicates for direct from agar media monopolar handheld probe (iKnife) REIMS analysis. Clustering of identical species replicates can be seen, with some species separation when all isolates are analysed together. Species separation is greatly improved when Gram-negative and Gram-positive isolates are analysed separately.

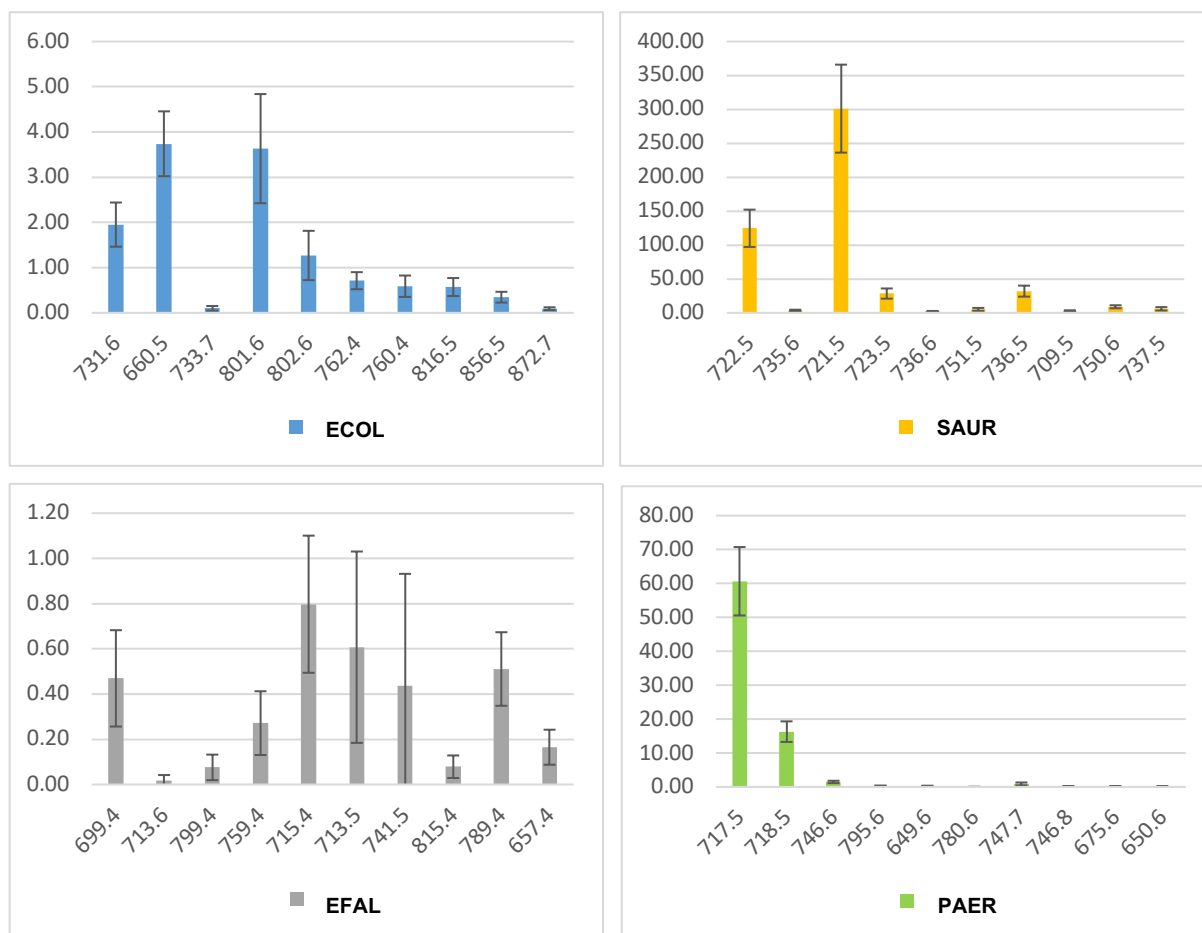


* denotes number of components from which model building results in the highest Q2 score

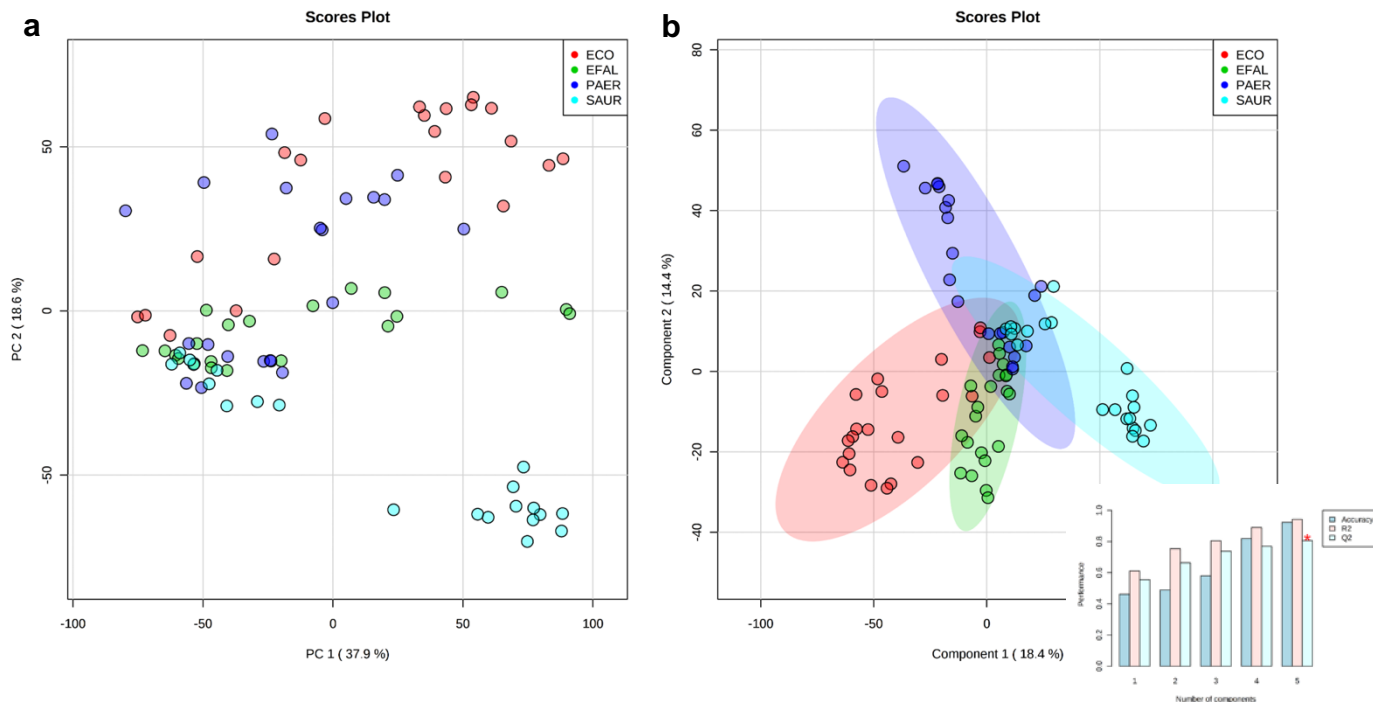
Supplementary 3. Partial least squares discriminant analysis of Gram-positive (a) and Gram-negative (b) bacterial species replicates for direct from agar media bipolar forceps REIMS analysis. Strong clustering of identical species replicates can be seen, with good species separation.



Supplementary 4. Mass spectra of the 600-1000 m/z range produced from LA-REIMS with a lens distance of 2 mm from bovine milk lipid layer surface. Background subtraction and lock mass correction applied. No signal was achieved in the majority of laser powers tested.

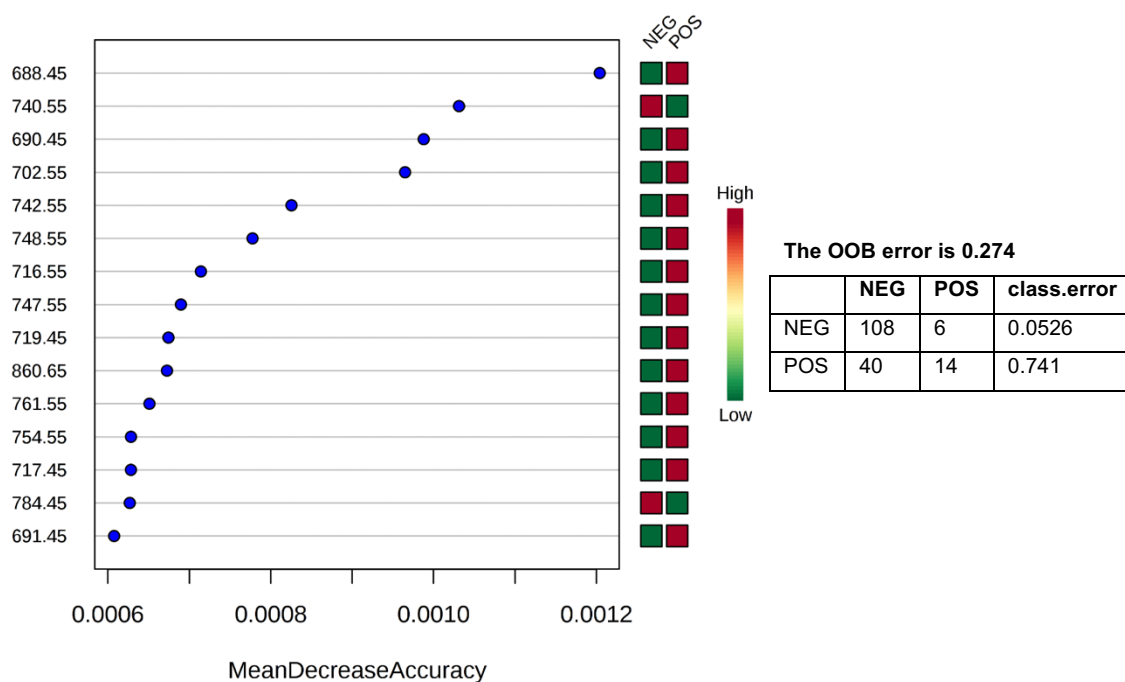


Supplementary 5. Top 10 significant species-specific biomarkers, ranked by FDR corrected P value, detected in artificially mixed bacterial cultures. The mean relative intensity from all concentrations containing the target species is represented by bar height.

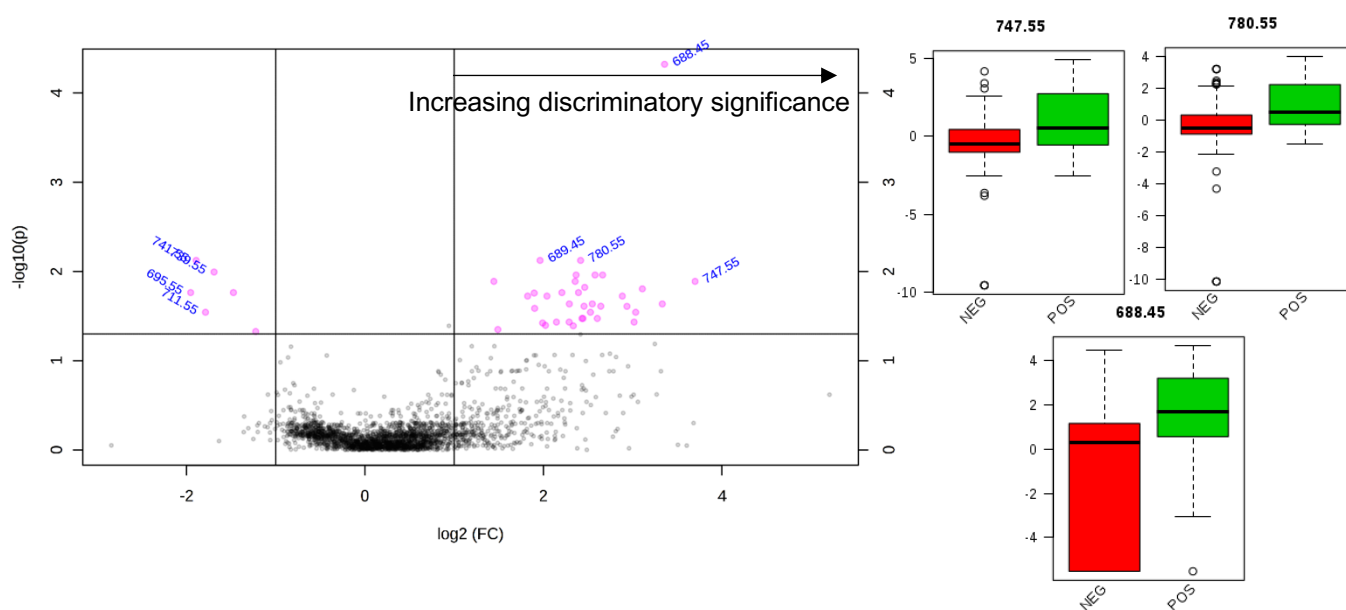


* denotes number of components from which model building results in the highest Q2 score

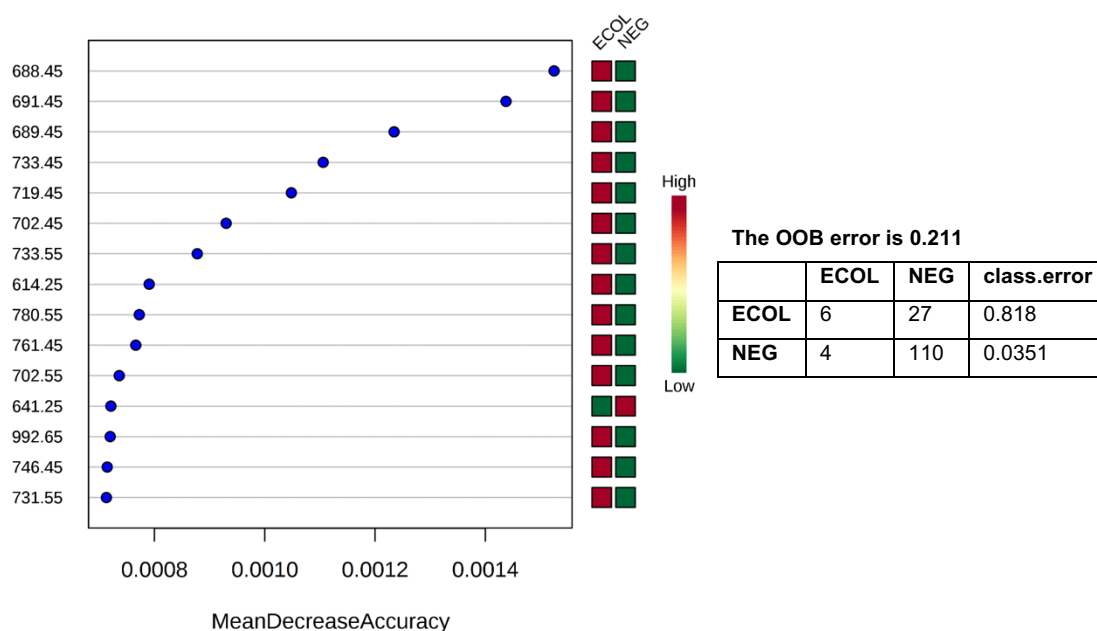
Supplementary 6. PCA (a) and PLS-DA (b) of four bacterial isolates spiked into healthy, culture negative donor urine at dilutions up to 10^7 x, analysed with bipolar REIMS (pellet analysis).



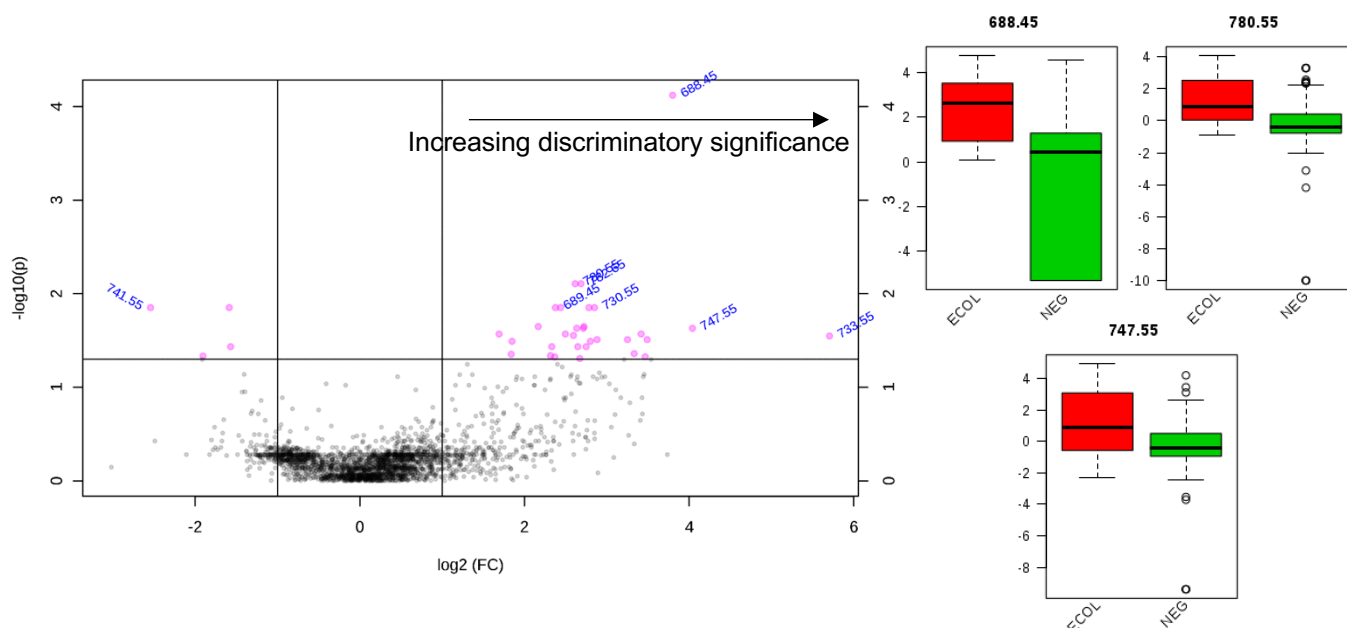
Supplementary 7. Clinical urine pellet LA-REIMS random forest analysis, 600-1000 m/z range. Misclassification of culture positive samples is high, represented by a high class error. Top-ranked biomarkers for culture positive samples are displayed in a feature importance matrix, ranked by mean decrease accuracy.



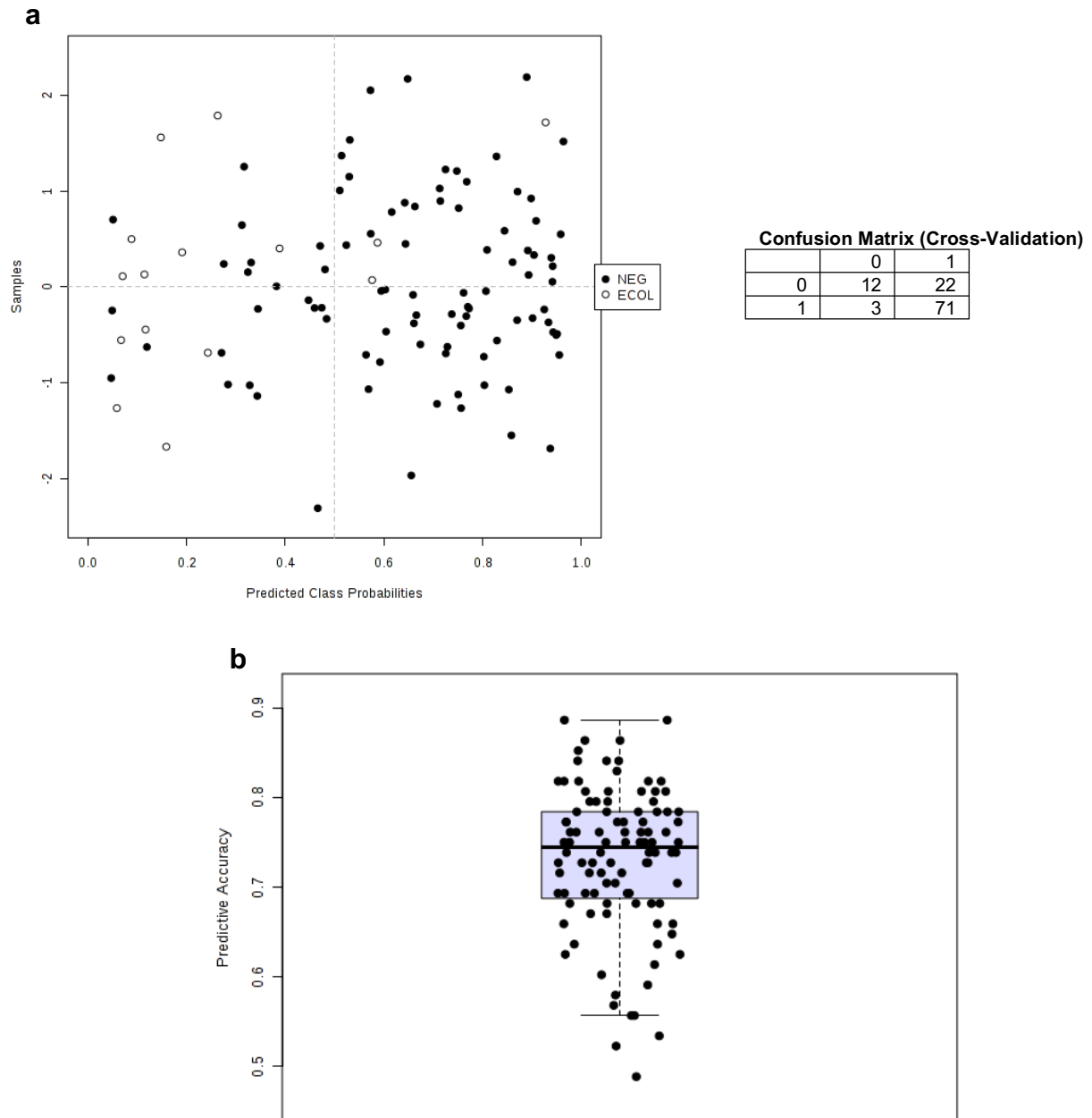
Supplementary 8. Volcano plot of $\log_2$ fold change (x-axis) and $-\log_{10}$ FDR corrected P value scores (y-axis) for culture positive/negative clinical urine sample pellets analysed with LA-REIMS. Features appearing to the top right of the plot have both high fold change and P score in culture positive samples, suggesting they are class specific and significant in classification. Box plots of three top ranking features show a significant increase in significance in culture positive samples verses culture negative.



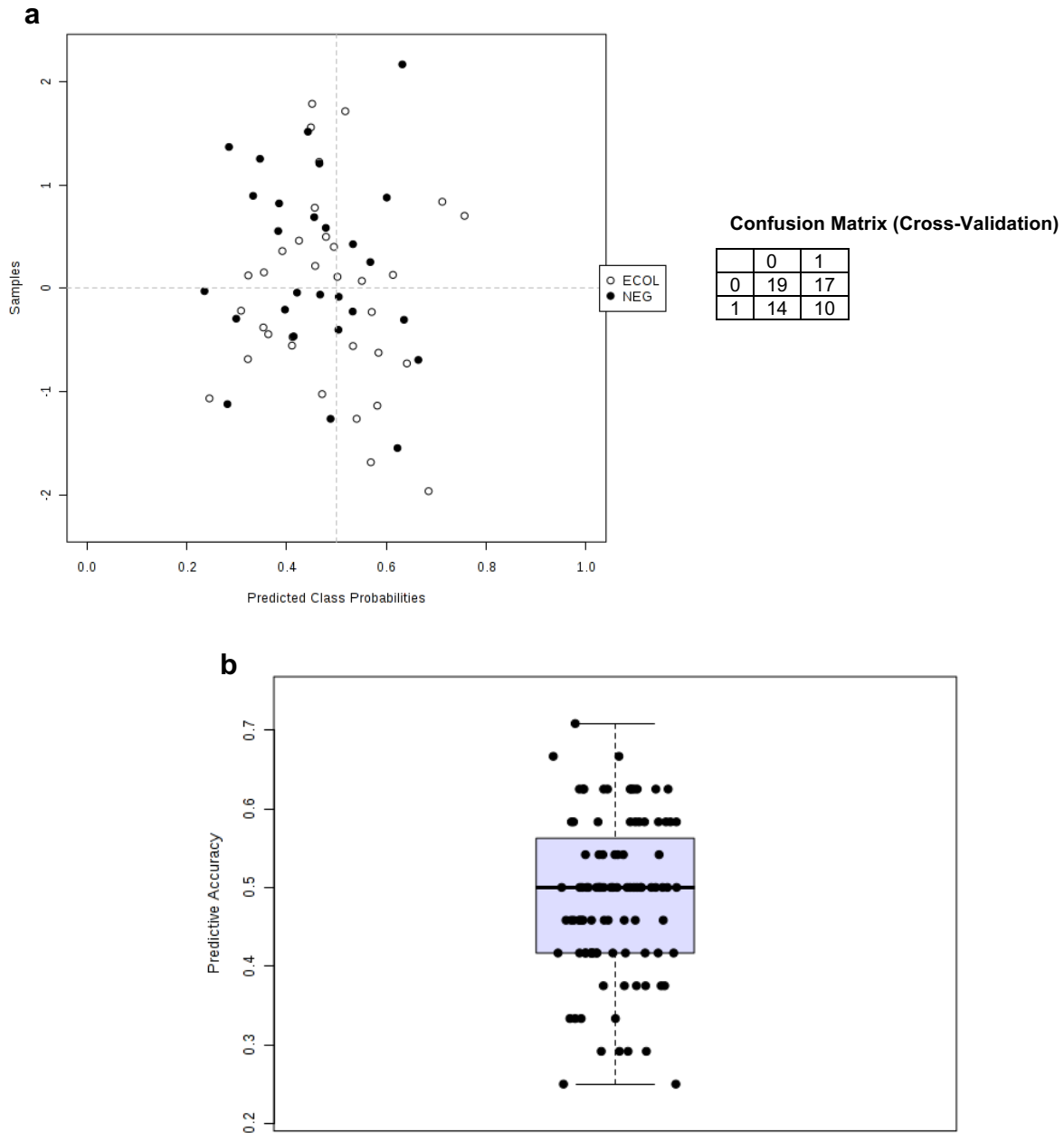
Supplementary 9. Clinical urine pellet LA-REIMS random forest analysis, 600-1000 m/z range. Misclassification of *E. coli* containing samples is high, represented by a high class error. Top-ranked biomarkers for *E. coli* containing urine samples are displayed in a feature importance matrix, ranked by mean decrease accuracy.



Supplementary 10. Volcano plot of $\log_2$ fold change (x-axis) and $-\log_{10}$ FDR corrected P value scores (y-axis) for culture positive *E. coli* containing and culture negative clinical urine sample pellets analysed with LA-REIMS. Features appearing to the top right of the plot have both high fold change and P score in *E. coli* containing samples, suggesting they are class specific and significant in classification. Box plots of three top ranking features show a significant increase in significance *E. coli* samples verses culture negative.



Supplementary 11. Predicted class probability scatter plot (a) and box plot (b) from a random forest validation model created by inserting 10 unlabelled samples to the classification model created previously. Overall model prediction accuracy is 0.731, from 100x cross-validations (Clinical Urine 1, WBC corrected dataset).



Supplementary 12. Predicted class probability scatter plot (a) and box plot (b) from a random forest validation model created by inserting 10 unlabelled samples to the classification model created previously. Overall model prediction accuracy is 0.486, from 100x cross-validations (Clinical Urine 2, WBC corrected dataset).

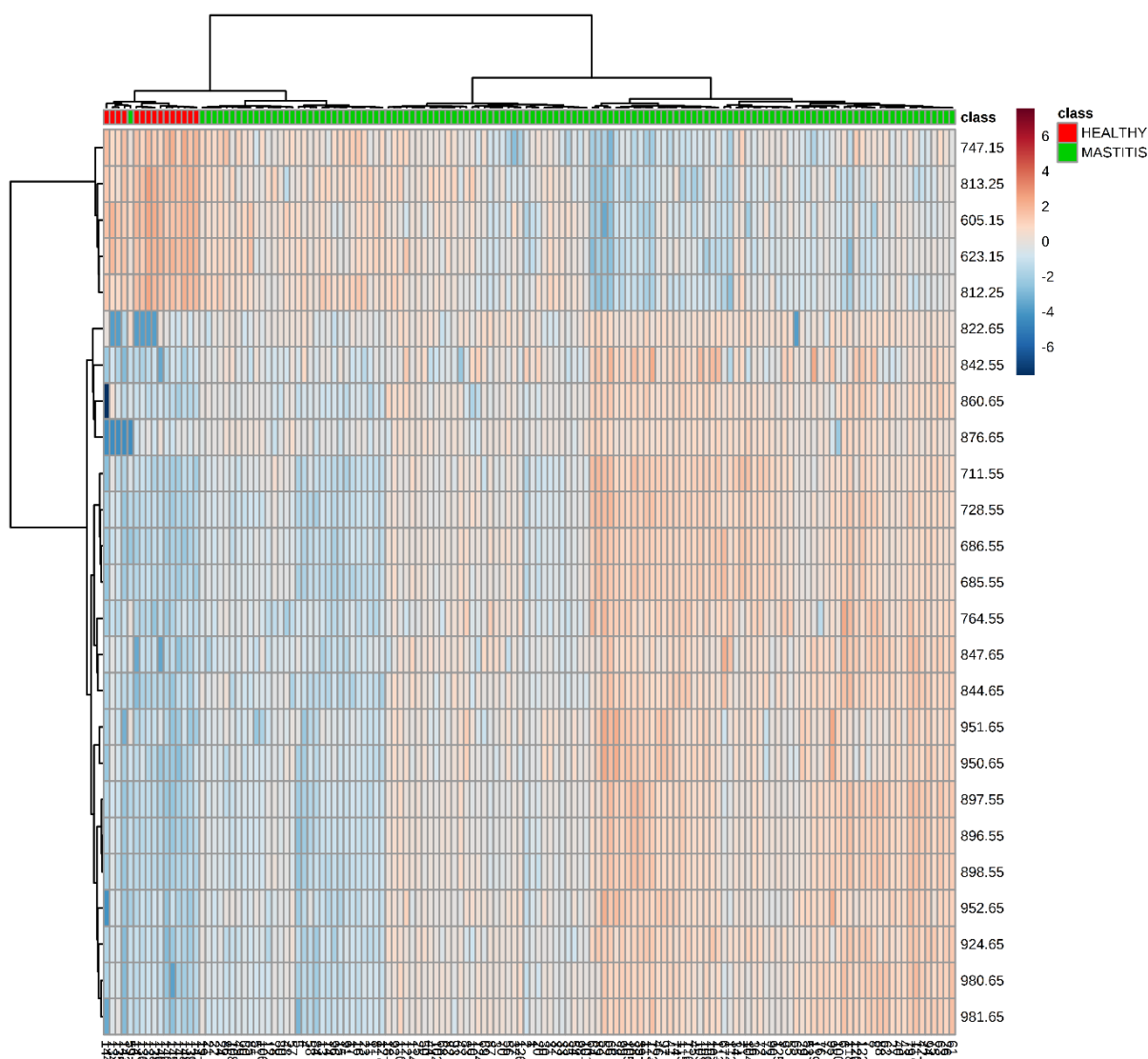
PM1 MicroPlate™ Carbon Sources

A1 Negative Control	A2 L-Arabinose	A3 N-Acetyl-D-Glucosamine	A4 D-Saccharic Acid	A5 Succinic Acid	A6 D-Galactose	A7 L-Aspartic Acid	A8 L-Proline	A9 D-Alanine	A10 D-Trehalose	A11 D-Mannose	A12 Dulcitol
B1 D-Serine	B2 D-Sorbitol	B3 Glycerol	B4 L-Fucose	B5 D-Gluconic Acid	B6 D-Gluconic Acid	B7 D,L-α-Glycerol-Phosphate	B8 D-Xylose	B9 L-Lactic Acid	B10 Formic Acid	B11 D-Mannitol	B12 L-Glutamic Acid
C1 D-Glucose-6-Phosphate	C2 D-Galactonic Acid-γ-Lactone	C3 D,L-Malic Acid	C4 D-Ribose	C5 Tween 20	C6 L-Rhamnose	C7 D-Fructose	C8 Acetic Acid	C9 α-D-Glucose	C10 Maltose	C11 D-Melibiose	C12 Thymidine
D-1 L-Asparagine	D2 D-Aspartic Acid	D3 D-Glucosaminic Acid	D4 1,2-Propanediol	D5 Tween 40	D6 α-Keto-Glutaric Acid	D7 α-Keto-Butyric Acid	D8 α-Methyl-D-Galactoside	D9 α-D-Lactose	D10 Lactulose	D11 Sucrose	D12 Uridine
E1 L-Glutamine	E2 m-Tartaric Acid	E3 D-Glucose-1-Phosphate	E4 D-Fructose-6-Phosphate	E5 Tween 80	E6 β-Hydroxy Glutaric Acid-γ-Lactone	E7 β-Hydroxy Butyric Acid	E8 β-Methyl-D-Glucoside	E9 Adonitol	E10 Maltotriose	E11 2-Deoxy Adenosine	E12 Adenosine
F1 Glycyl-L-Aspartic Acid	F2 Citric Acid	F3 m-Inositol	F4 D-Threonine	F5 Fumaric Acid	F6 Bromo Succinic Acid	F7 Propionic Acid	F8 Mucic Acid	F9 Glycolic Acid	F10 Glyoxylic Acid	F11 D-Cellobiose	F12 Inosine
G1 Glycyl-L-Glutamic Acid	G2 Tricarballic Acid	G3 L-Serine	G4 L-Threonine	G5 L-Alanine	G6 L-Alanyl-Glycine	G7 Acetoacetic Acid	G8 N-Acetyl-β-D-Mannosamine	G9 Mono Methyl Succinate	G10 Methyl Pyruvate	G11 D-Malic Acid	G12 L-Malic Acid
H1 Glycyl-L-Proline	H2 p-Hydroxy Phenyl Acetic Acid	H3 m-Hydroxy Phenyl Acetic Acid	H4 Tyramine	H5 D-Psicose	H6 L-Lyxose	H7 Glucuronamide	H8 Pyruvic Acid	H9 L-Galactonic Acid-γ-Lactone	H10 D-Galacturonic Acid	H11 Phenylethylamine	H12 2-Aminoethanol

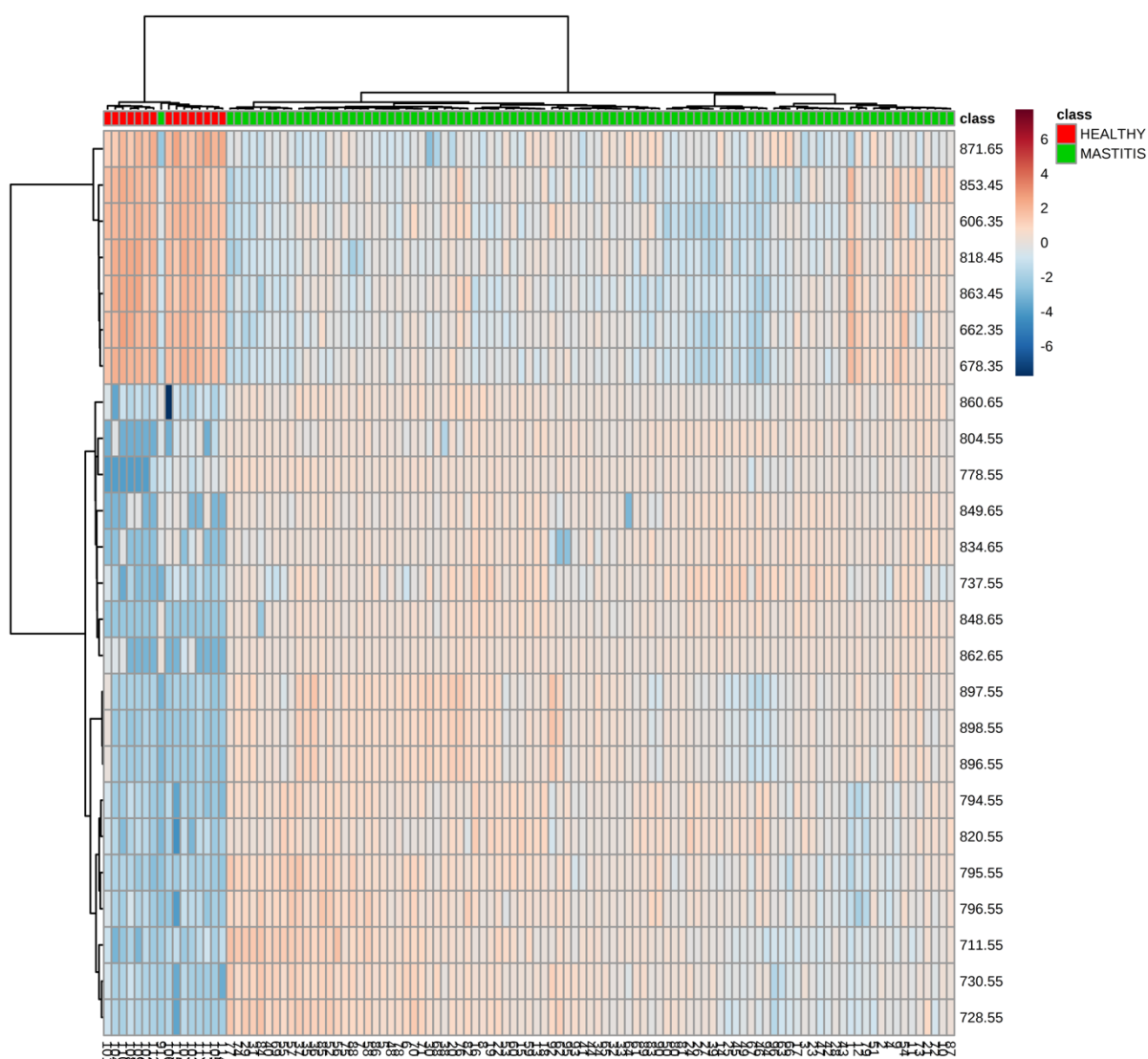
PM2A MicroPlate™ Carbon Sources

A1 Negative Control	A2 Chondroitin Sulfate C	A3 α-Cyclodextrin	A4 β-Cyclodextrin	A5 γ-Cyclodextrin	A6 Dextrin	A7 Gelatin	A8 Glycogen	A9 Inulin	A10 Laminarin	A11 Mannan	A12 Pectin
B1 N-Acetyl-D-Galactosamine	B2 N-Acetyl-Neuraminic Acid	B3 β-D-Allose	B4 Amygdalin	B5 D-Arabinose	B6 D-Arabitol	B7 L-Arabitol	B8 Arbutin	B9 2-Deoxy-D-Ribose	B10 l-Erythritol	B11 D-Fucose	B12 3-O-β-D-Galactopyranosyl-D-Arabinose
C1 Gentiobiose	C2 L-Glucose	C3 Lactitol	C4 D-Melezitose	C5 Maltitol	C6 α-Methyl-D-Glucoside	C7 β-Methyl-D-Galactoside	C8 3-Methyl Glucose	C9 β-Methyl-D-Glucuronic Acid	C10 α-Methyl-D-Mannoside	C11 β-Methyl-D-Xyloside	C12 Palatinose
D1 D-Raffinose	D2 Salicin	D3 Sedoheptulosan	D4 L-Sorbose	D5 Stachyose	D6 D-Tagatose	D7 Turannose	D8 Xylitol	D9 N-Acetyl-D-Glucosaminitol	D10 γ-Amino Butyric Acid	D11 α-Amino Valeric Acid	D12 Butyric Acid
E1 Capric Acid	E2 Caproic Acid	E3 Citraconic Acid	E4 Citramalic Acid	E5 D-Glucosamine	E6 2-Hydroxy Benzoic Acid	E7 4-Hydroxy Benzoic Acid	E8 β-Hydroxy Butyric Acid	E9 γ-Hydroxy Butyric Acid	E10 α-Keto-Valeric Acid	E11 Itaconic Acid	E12 5-Keto-D-Gluconic Acid
F1 D-Lactic Acid Methyl Ester	F2 Malonic Acid	F3 Melibionc Acid	F4 Oxalic Acid	F5 Oxalomalic Acid	F6 Quinic Acid	F7 D-Ribono-1,4-Lactone	F8 Sebacic Acid	F9 Sorbic Acid	F10 Succinamic Acid	F11 D-Tartaric Acid	F12 L-Tartaric Acid
G1 Acetamide	G2 L-Alaninamide	G3 N-Acetyl-L-Glutamic Acid	G4 L-Arginine	G5 Glycine	G6 L-Histidine	G7 L-Homoserine	G8 Hydroxy-L-Proline	G9 L-Isoleucine	G10 L-Leucine	G11 L-Lysine	G12 L-Methionine
H1 L-Ornithine	H2 L-Phenylalanine	H3 L-Pyrroglutamic Acid	H4 L-Valine	H5 D,L-Carnitine	H6 Sec-Butylamine	H7 D,L-Octopamine	H8 Putrescine	H9 Dihydroxy Acetone	H10 2,3-Butanediol	H11 2,3-Butanedione	H12 3-Hydroxy 2-Butanone

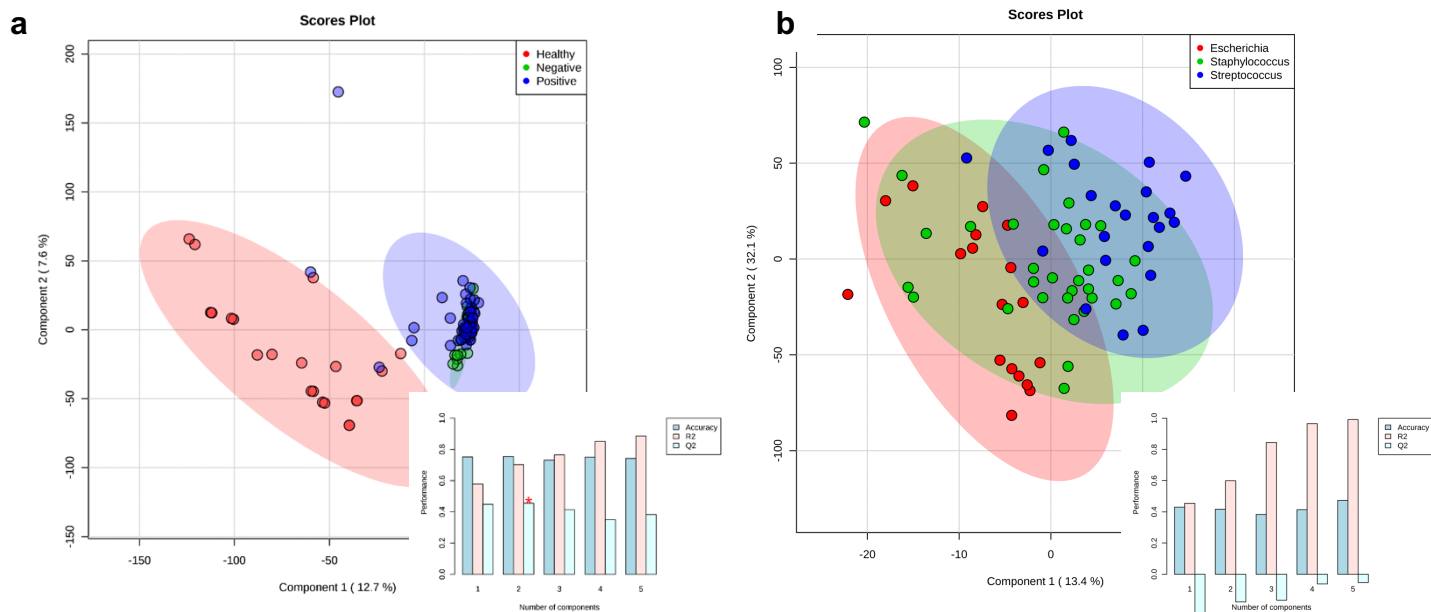
Supplementary 13. Carbon sources used in culture of a single *K. oxytoca* isolate on phenotype microarray plates (BIOLOG, California, US).



Supplementary 14. Heatmap clustering of top 25 most significant features for mastitis positive verses negative classification, determined through t-test FDR corrected *P* value rankings, of sample pellets analysed through cell line probe REIMS, 600-1000 *m/z* range. Samples are displayed in rows and biomarkers in columns. Distinct biomarkers for both mastitis positive and negative sample classes are displayed. Red colours represent a positive fold change of the biomarker in a sample class, whilst blues represent a negative fold change.



Supplementary 15. Heatmap clustering of top 25 most significant features for mastitis positive versus negative classification, determined through t-test FDR corrected P value rankings, of lipid layer samples analysed through cell line probe REIMS, 600-1000 m/z range. Samples are displayed in rows and biomarkers in columns. Distinct biomarkers for both mastitis positive and negative sample classes are displayed. Red colours represent a positive fold change of the biomarker in a sample class, whilst blues represent a negative fold change.



* denotes number of components from which model building results in the highest Q2 score

Supplementary 16. PLS-DA plots, with Gram-positive/negative versus mastitis negative bovine milk samples (a), and culture positive Genus groupings (b), of bovine milk lipid layer LA-REIMS analysis from 24 well plates, 50-500 m/z range. Data groupings by Gram status show good class clustering, but poor separation with a mediocre Q2 score from cross validation (a). Genus data groupings show poor clustering with a negative Q2 score, indicating overfitting of the model (b).

SPRINGER NATURE

Title: Urinary tract infections: epidemiology, mechanisms of infection and treatment options

Author: Ana L. Flores-Mireles, Jennifer N. Walker, Michael Caparon, Scott J. Hultgren

Publication: Nature Reviews Microbiology

Publisher: Springer Nature

Date: Apr 8, 2015

Copyright © 2015, Springer Nature

Logged in as:
Adam Burke
Account #:
3001415763

[LOGOUT](#)

Order Completed

Thank you for your order.

This Agreement between Adam Burke ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

[printable details](#)

License Number	4565880910084
License date	Apr 11, 2019
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature Reviews Microbiology
Licensed Content Title	Urinary tract infections: epidemiology, mechanisms of infection and treatment options
Licensed Content Author	Ana L. Flores-Mireles, Jennifer N. Walker, Michael Caparon, Scott J. Hultgren
Licensed Content Date	Apr 8, 2015
Licensed Content Volume	13
Licensed Content Issue	5
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
High-res required	no
Will you be translating?	no
Circulation/distribution	501 to 1000
Author of this Springer Nature content	no
Title	Expanding the Applications of Rapid Evaporative Ionisation Mass Spectrometry (REIMS): Microbial Detection, Diagnostics and Metabolic Profiling
Institution name	Imperial College London
Expected presentation date	Jul 2019
Portions	Figures 1-2
Requestor Location	Adam Burke Mass Spectrometry, 5L17, 5th Floor Lab Block, Charing Cross Hospital London, English W6 8RD United Kingdom Attn: Adam Burke
Total	0.00 GBP

[ORDER MORE](#)
[CLOSE WINDOW](#)

Copyright © 2019 [Copyright Clearance Center, Inc.](#) All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#).
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Supplementary 18. Republish Authorisation of Urinary Tract Infections Figure 2, Figure 3 (original publication Figures 1-2) (70)



RightsLink®

Home

Account
Info

Help



Title: Novel Acoustic Loading of a Mass Spectrometer

Author: Ian Sinclair, Rick Stearns, Steven Pringle, et al

Publication: SLAS Technology

Publisher: SAGE Publications

Date: 02/01/2016

Copyright © 2016, © SAGE Publications

Logged in as:

Adam Burke

LOGOUT

Gratis Reuse

Permission is granted at no cost for use of content in a Master's Thesis and/or Doctoral Dissertation. If you intend to distribute or sell your Master's Thesis/Doctoral Dissertation to the general public through print or website publication, please return to the previous page and select 'Republish in a Book/Journal' or 'Post on intranet/password-protected website' to complete your request.

BACK

CLOSE WINDOW

Copyright © 2019 Copyright Clearance Center, Inc. All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#).
Comments? We would like to hear from you. E-mail us at customer@copyright.com

Supplementary 19. Republish Authorisation of AMI-MS Figure 5. (original publication Figure 4 (b)) (191)