



**University of
Nottingham**

UK | CHINA | MALAYSIA

Multiplexed Antigen Microarray Development for Comprehensive Immune Health Profiling: A Proof of Concept Study

Chloe Thompson, BSc

20162525

Supervisors: Prof. Patrick Tighe and Dr Tanya Monaghan

Thesis submitted to the University of Nottingham, School of Life
Sciences for the degree of Master of Research in Immunology.

Word Count: 27,825

Acknowledgements

I would like to thank my supervisors, Prof. Patrick Tighe and Dr Tanya Monaghan, for providing support and guidance throughout my Master of Research. I have learnt a lot from their knowledge and enthusiasm for scientific research and their lab expertise. I have thoroughly enjoyed my year of research alongside them both.

I would like to thank Hannah Jackson for her helpful guidance in statistics and data analysis, from which I have learnt a lot. Special thanks to Niamh for being a constant friend and lab partner across the year and all other members of the lab who have helped me along the way.

Lastly, I would like to thank my family and friends for their continuous support and encouragement outside the laboratory.

I sincerely thank you all.

Abstract

This proof-of-concept study introduces a versatile multi-antigen microarray platform for comprehensive serological profiling of adaptive humoral immune responses. Such a platform is required as autoantibodies can be detected in sera years before symptom manifestation and the clinical onset of autoimmune disease. However, traditional ELISAs used to assess one antibody biomarker at a time are relatively time-consuming, costly and reductive. Alternatively, to detect broad-scale antibody reactivities, we developed a custom microarray that can simultaneously quantify IgG and IgA responses against 60 antigens, including autoantigens, pathogen-associated, and viral recall antigens, using sub-microlitre volumes of patient sera. Our platform was applied to 128 individuals: 11 diagnosed with cystic fibrosis, 63 with *C. difficile* infection (CDI), 24 with recurrent CDI (pre- and post-faecal microbiota transplantation; FMT), and 30 healthy controls. Antigens were printed on nitrocellulose slides, probed with sera and scanned for antibody-antigen interactions. Across sixty antigens, 93% of intra-assay variabilities and 90% of inter-assay variabilities fell below the precision limit of <20% coefficient of variation (CV). Distinct heterogeneity was observed in adaptive humoral responses across samples, enabling patient clustering based on antigen-specific IgG and IgA responses. Profiles displayed rare yet strong autoantibody responses across samples, indicating the presence of preclinical or natural autoantibody repertoires. Recurrent CDI sera was tested pre- and post-FMT, and found that FMT did not significantly alter humoral responses after 12 weeks, indicating the need for long-term analysis in larger cohorts. Lastly, differences in antigen-specific IgG and IgA humoral responses distinguished initial and recurrent CDI patients, providing insights into CDI adaptive immunity. In summary, our preliminary assay highlights the potential for efficient and scalable immune health screening. With further research, implementing multiplexed antigen microarrays in healthcare systems could enable annual immune health screening as an early warning system for immune-mediated diseases and multimorbidity, with potential advantages in throughput and cost-efficiency.

Table of Contents

Acknowledgements	2
Abstract.....	3
List of Figures	7
List of Tables	9
Abbreviations.....	11
Introduction	13
1.1 The Immune System	13
1.2 Autoimmune Disease	16
1.2.1 Risk Factors	19
1.2.2 Preclinical Autoimmunity	23
1.2.3 Multimorbidity.....	26
1.2.4 Rising Impact and Prevalence of Autoimmune Disease	26
1.3 The Gut Microbiome and The Immune System	27
1.4 Patient Cohort of Interest: <i>Clostridium difficile</i> Infection.....	29
1.4.1 Immune Responses to CDI.....	30
1.4.2 Treatment: Faecal Microbiota Transplantation (FMT)	33
1.5 Patient Cohort of Interest: Cystic Fibrosis	34
1.6 Serological Profiling and Antibodies as Biomarkers	35
1.6.1 Serological Profiling	35
1.6.2 Antibodies as Biomarkers	36
1.6.3 Established Disease-specific Antibody Biomarkers	36
1.7 Serological Antibody Profiling Techniques	39
1.7.1 ELISA	39
1.7.2 Microarray Technology.....	42
1.8 Current Screening Platform for Autoimmune Disease	43

1.9	Aims and Hypothesis.....	45
2	Materials.....	47
2.1	Chemicals and Reagents	47
2.2	Antigens	49
3	Methods.....	54
3.1	Samples	54
3.2	SDS-PAGE	55
3.3	Carbamylated BSA Production.....	55
3.4	Printing.....	56
3.5	Slide Surfaces	57
3.6	Serum Dilutions.....	58
3.7	Optimising Assay Reagents	59
3.7.1	Secondary Antibody Optimisation.....	59
3.7.2	Streptavidin-cy5 Optimisation.....	60
3.7.3	Blocking Buffer Optimisation.....	61
3.8	Final Optimised Microarray Protocol.....	62
3.9	Microarray Data Analysis	63
3.10	Co-efficient of Variation: Intra- and Inter-assay	64
3.11	Statistical Analysis	64
4	Results.....	66
4.1	Generation of The Multi-antigen Microarray Platform.....	66
4.2	Assay Development and Optimisation.....	69
4.2.1	Slide Surface Assessment	69
4.2.2	Serum Dilution	73
4.2.3	Nitrocellulose Background Fluorescence	76
4.2.4	Optimisation of Assay Reagents.....	77

4.2.4.1	Secondary Antibody and Detection Reagent Dilution	77
4.2.4.2	Blocking Buffers.....	80
4.2.5	Scan Settings.....	83
4.2.6	Optimised Assay Protocol.....	85
4.3	Microarray Platform Validation	88
4.3.1	Intra- and Inter-assay Precision.....	88
4.3.2	Testing Sample Quality and Antibody Reactivity.....	92
4.4	Patient Demographics.....	93
4.5	Application of The Multi-antigen Microarray	94
4.5.1	Hierarchical Clustering of Humoral Antibody Responses.....	95
4.5.2	Humoral Responses of rCDI Patients Before and After FMT....	100
4.5.3	Humoral Responses of CDI and rCDI Patients	110
4.5.4	Overview of Headline Findings.....	118
5	Discussion	120
5.1.1	Optimisation and Validation.....	120
5.1.2	Antibody Response Profiles and Autoantibody Responses.....	121
5.1.3	Hierarchical Clustering Analysis.....	123
5.1.4	Humoral Responses of rCDI Patients Before and After FMT....	126
5.1.5	Humoral Responses of CDI and rCDI Patient Cohorts	129
5.2	Strengths of The Antigen Microarray Platform	133
5.3	Limitations and Future Research	134
5.4	Conclusions	137
6	Bibliography.....	138
7	Appendix.....	147

List of Figures

Figure 1. Summary of the Immune system.	15
Figure 2. Summary of the process from normal to pathological autoimmunity and eventual autoimmune disease manifestation.	18
Figure 3. Environmental Factors and Autoimmune Disease.	22
Figure 4. Pathogenesis of autoimmune disease; transitioning from preclinical to overt autoimmune disease establishment.	25
Figure 5. Immune responses in CDI alongside Fecal Microbiota Transplantation (FMT) for restoring microbial and immune homeostasis.	32
Figure 6. Serological antibody detection methods.	41
Figure 7. The current diagnostic framework for autoimmune diseases.	44
Figure 8. The cassette cover plate format.	58
Figure 9. The layout of the assay reagent optimisation experiments.	61
Figure 10. SDS-PAGE gels for antigen molecular weight validation before microarray selection.	67
Figure 11. Schematic summarising the major steps involved in generating and processing multi-antigen microarrays.	68
Figure 12. Comparison of signal intensities produced from a serial dilution of IVIG probed on six slide surfaces.	71
Figure 13. Summary of the optimisation tests assessing spot intensities and morphology across four slide surfaces.	72
Figure 14. Comparison of signal intensities across a range of serum dilutions on nitrocellulose slides.	75
Figure 15. Scanned image at 635nm of an array probed with control serum diluted 1:160 in blocking buffer (3% BSA-PBS) on a nitrocellulose slide.	76
Figure 16. Comparison of the S/B ratios of two secondary antibodies when probed across a range of dilutions on nitrocellulose slides.	79
Figure 17. Comparison of the S/B ratios of streptavidin-cy5 when probed across a range of dilutions on a nitrocellulose slide.	80
Figure 18. Identification of the optimal blocking buffer.	82

Figure 19. Comparison of the background intensity produced when scanning arrays at 100% and 33% laser power.	84
Figure 20. An overview of the antigen microarray optimisation process.	85
Figure 21. Overview of the optimised assay protocol and its results.	87
Figure 22. Intra- and Inter-assay precision of the multi-antigen microarray platform, including 60 antigens.....	89
Figure 23. Testing sample quality and antibody reactivity of 103 serum samples through IgG responses to two positive-control antigens.....	92
Figure 24. Heatmaps displaying the hierarchical clustering of IgG responses across CF, CDI, rCDI and healthy controls.....	97
Figure 25. Heatmaps displaying the hierarchical clustering of IgA responses across CF, CDI, rCDI and healthy controls.....	99
Figure 26. IgG responses in patients Pre-FMT, 12 weeks Post-FMT and healthy controls across autoantigens (A), positive-control (B), pathogen-associated (C) and viral recall antigens (D).	102
Figure 27. IgA responses in patients Pre-FMT, 12 weeks Post-FMT and healthy controls across autoantigens (A), positive-controls (B), pathogen-associated (C) and viral recall antigens (D).....	107
Figure 28. IgG responses in patients with CDI and rCDI across autoantigens (A), positive-controls (B), pathogen-associated (C) and viral recall antigens (D).	112
Figure 29. IgA responses in patients with CDI and rCDI across autoantigens (A), positive-controls (B), pathogen-associated (C) and viral recall antigens (D).	116

List of Tables

Table 1.1. Functions of the five major classes of human Ig.	14
Table 1.2. Summary of the differences between Pathogenic and Natural Autoantibodies.	17
Table 1.3. Risk factors for autoimmune disease.	19
Table 1.4. Serological antibody biomarkers with established associations for autoimmune diseases.	38
Table 2.1. Blocking buffers.	47
Table 2.2. Slide surface chemistries.	48
Table 2.3. Assay reagents.	48
Table 2.4. Auto-antigens and pathogen-associated antigens. Including antigen concentrations and suppliers.	49
Table 2.5. Viral recall antigens. Including antigen concentrations and suppliers.	52
Table 2.6 Data on the seroprevalence of viral pathogens in the general population.	53
Table 3.1 Subject participant groups, sample numbers and storage conditions.	54
Table 4.1. Summary of the Intra- and Inter-assay precision data for each antigen.	90
Table 4.2. CDI and CF patient demographics.	93
Table 4.3. Recurrent CDI FMT patients and age-matched healthy control demographics.	93
Table 4.4 Summary of comparative analyses in the following application of the multi-antigen microarray results section.	94
Table 4.5. ANOVA type II test results across antigen subset GLM models for IgG responses.	101
Table 4.6. Pairwise comparisons of antigen-specific IgG responses between the patient cohorts.	103
Table 4.7. ANOVA type II test results across antigen subset GLM models for IgA responses.	106

Table 4.8. Pairwise comparisons of antigen-specific IgA responses between the patient cohorts.....	108
Table 4.9. ANOVA type II test results across antigen subset GLM models for CDI vs. rCDI IgG responses.....	110
Table 4.10. Pairwise comparisons of antigen-specific IgG responses between CDI and rCDI patients.....	113
Table 4.11. ANOVA type II test results across antigen subset GLM models for CDI vs. rCDI IgA responses.....	114
Table 4.12. Pairwise comparisons of antigen-specific IgA responses between CDI and rCDI patients.....	117
Table 4.13 Summary of the headline findings for the application of the multi-antigen microarray results section.....	119

Abbreviations

AFU – Arbitrary fluorescent units

AIC – Akaike information criterion

AMA – Anti-mitochondrial antibody

ANA – Anti-nuclear antibody

ANOVA – Analysis of variance

B635nm – Background intensity at 635nm

BPI – Bactericidal permeability-increasing protein

BSA – Bovine serum albumin

BSA-NHS – Bovine serum albumin-N-hydroxy succinimide

cANCA – Antinuclear cytoplasmic antibodies

CDI – *C. difficile* infection

CeD – Celiac disease

CF – Cystic Fibrosis

CMV – Cytomegalovirus

CV – Coefficient of variation (%)

Df – Degrees of freedom

DNA – deoxyribonucleic acid

dsDNA – Double-stranded deoxyribonucleic acid

DTT – Dithiothreitol

EBV – Epstein-Barr virus

ELISA – Enzyme-linked immunosorbent assay

F635nm – Foreground intensity at 635nm

FDA – Food and drug administration

FMT – Faecal microbiota transplantation

GAL – GenePix array list

GI – Gastrointestinal

GLM – Generalised linear model

GPR – GenePix results

HC – Healthy control

HCoV – Human Coronavirus (HKU1, 229E, OC43 – seasonal)

HDV – Hepatitis D

HEV – Hepatitis E

HSV1 – Herpes simplex virus 1

HSV2 – Herpes simplex virus 2

IBD – Inflammatory Bowel disease

Ig – Immunoglobulin

IVIG – Intravenous immunoglobulin

LR Chisq – Likelihood ratio Chi-squared

MS – Multiple sclerosis

OD – Optical density

PBS – Phosphate buffered saline

PBST – Phosphate buffered saline with Tween20

PMT – Photomultiplier tube

RA - Rheumatoid arthritis

rCDI – Recurrent *C. difficile* infection

S/B ratio – Signal-to-background ratio

SDS-Page – Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

SE – Standard error

SjS – Sjogren’s disease

SLE – Systemic lupus erythematosus

T1D – Type I diabetes

TcdA – *C. difficile* toxin A

TcdB – *C. difficile* toxin B

VZV – Varicella zoster virus

Introduction

1.1 The Immune System

The immune system is a complex network of organs, cells, proteins and chemicals that work in synergy to protect the body from foreign antigens. It has "two lines of defence", including innate and adaptive immunity, demonstrated in Figure 1. The first line of defence is the innate immune response, establishing non-specific host-barrier defensive mechanisms immediately or within a few hours of encountering a foreign antigen. Key components include physical barriers like the skin and mucous membranes and cellular defences such as neutrophils, macrophages, dendritic and natural killer cells. The second line of defence is the adaptive immune response, which is antigen-dependent and antigen-specific, with the capacity for immunological memory. These two lines of defence act in tandem, so a failure in either system can induce immunopathological conditions such as autoimmune disease, chronic inflammation, hypersensitivity or immunodeficiency.

The adaptive immune system splits into cell-mediated and humoral (antibody-mediated) immunity. Cell-mediated immunity is orchestrated by T-cells, including helper T cells (CD4+), which stimulate other immune cells and coordinate the immune response, while cytotoxic T cells (CD8+) directly kill infected cells. In contrast, humoral immunity involves the production of antigen-specific antibodies, also known as immunoglobulins (Ig), by

transformed B cells known as plasma cells. Antibodies have various functions, including neutralising pathogens, marking them for destruction, activating complement proteins, and aiding in killing infected cells.

B cells produce five major antibody isotypes: IgM, IgG, IgA, IgE, and IgD, distinguished by the type of heavy chain they contain. The classes of antibodies have different biological functions and target different pathogens in distinct body compartments. The primary functions of each class of human immunoglobulin (Ig) are outlined in Table 1.1.

Table 1.1. Functions of the five major classes of human Ig.

Ig isotype	Functions
IgM	<ul style="list-style-type: none"> • The first Ig to be produced (innate response). • Mainly found in blood (and lymph). • Involved in opsonisation and complement fixation.
IgG	<ul style="list-style-type: none"> • Main Ig during the adaptive immune response. • Principle isotype in blood and extracellular fluid. • Neutralisation, opsonisation and complement fixation.
IgA	<ul style="list-style-type: none"> • Principle isotype in secretions to protect mucosal surfaces. • Second most prevalent Ig in the serum. • Direct neutralisation or prevention of binding to the mucosal surface.
IgE	<ul style="list-style-type: none"> • Lower levels in the blood and extracellular fluid. • Associated with hypersensitivity and allergic reaction.
IgD	<ul style="list-style-type: none"> • Functions primarily unknown – may be involved in homeostasis.

(Janeway CA Jr, Travers P and Walport M, 2001; Marshall *et al.*, 2018)

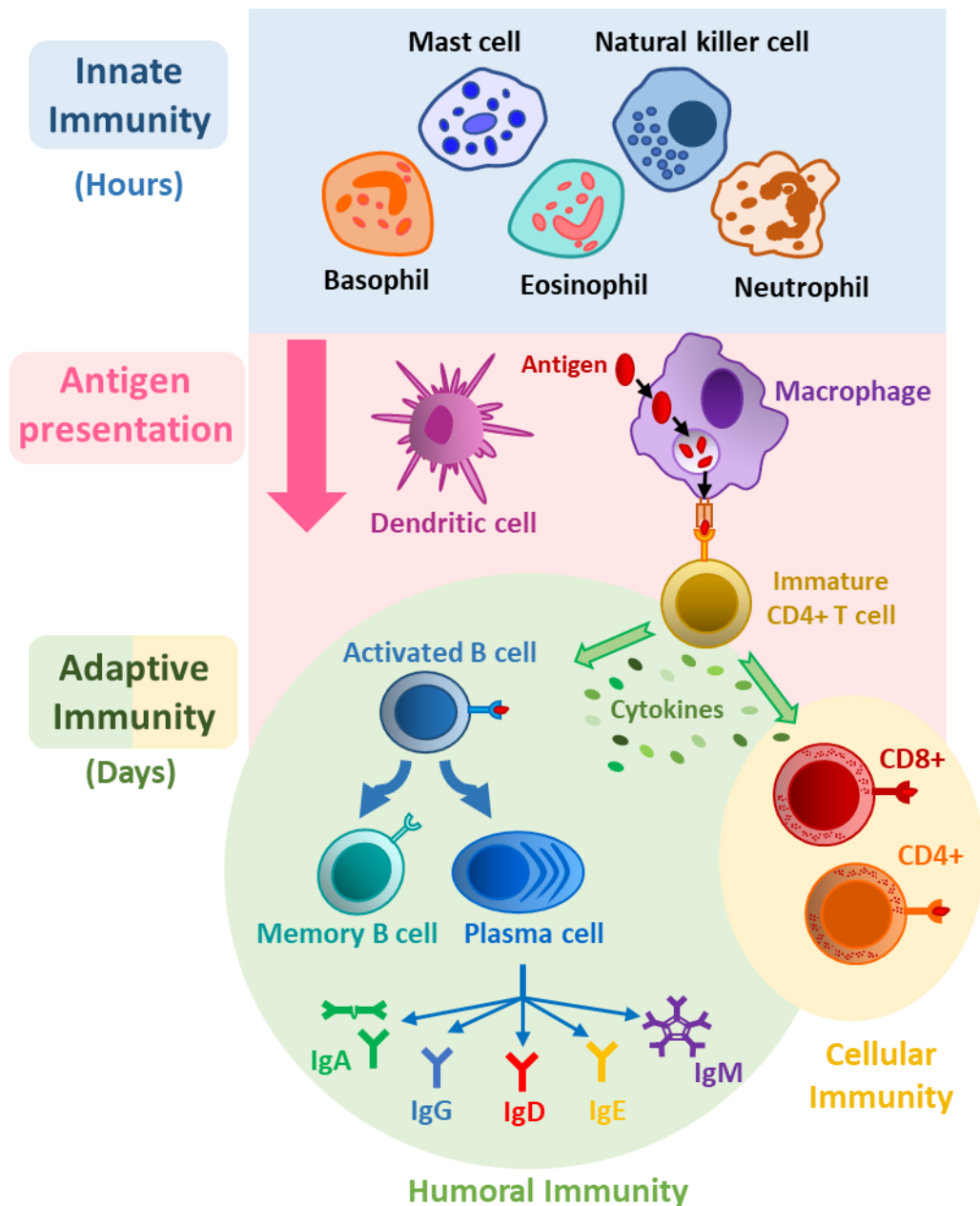


Figure 1. Summary of the Immune system.

The immune system comprises two branches: Innate and Adaptive immunity. Innate immunity provides immediate but non-specific protection, whereas adaptive immunity provides antigen-specific and long-lasting protection through immunological memory. The adaptive branch is split into humoral (antibody-mediated) and cell-mediated immunity (T-cells). Humoral immunity involves the production of antibodies, also known as immunoglobulins (Igs), by B cells. These antibodies circulate in the bloodstream and can neutralise or tag pathogens for destruction. There are five classes of Igs: IgG, IgA (serological and secretory), IgM, IgD and IgE. Memory B cells are also formed during an immune response, allowing the body to mount a faster and stronger defence upon re-exposure to the same pathogen. Cell-mediated immunity is orchestrated by T cells and does not involve antibodies. T-cells include helper T-cells (CD4+) and cytotoxic T-cells (CD8+).

1.2 Autoimmune Disease

Autoimmune diseases, consisting of over 80 known conditions, are a group of chronic disorders characterised by an aberrant immune response directed towards the host. This dysregulation in the immune system is caused by the transition from normal to pathological autoimmunity, demonstrated in Figure 2. A breach in immunological tolerance leads to the immune system erroneously attacking the body's own healthy cells, resulting in auto-reactive T-cells and B-cells, auto-antibodies and inflammation (Marshall *et al.*, 2018).

As illustrated in Figure 2, pathogenic autoantibodies that react with self-antigens and cause tissue damage are distinct from natural autoantibodies produced during non-pathogenic autoimmunity, thought to eliminate cell debris to maintain homeostasis (Nagele *et al.*, 2013). Everyone, healthy or disease-afflicted, will have natural autoantibodies as they are ubiquitous, although present at varying levels (Nagele *et al.*, 2013). Table 1.2 summarises some known differences between pathogenic and natural autoantibodies.

It is hypothesised that risk factors such as genetic predispositions and environmental triggers breach tolerance and induce the transition from harmless autoimmunity to pathological (Wang, Wang and Gershwin, 2015). For instance, in genetically predisposed individuals, chronic activation of the immune system may initiate the expansion of natural auto-polyreactive clones, leading to the development of pathogenic autoantibodies and autoimmune disease (Avrameas, Alexopoulos and Moutsopoulos, 2018).

Table 1.2. Summary of the differences between pathogenic and natural autoantibodies.

	Pathogenic Autoantibodies	Natural Autoantibodies
Type	Mainly, IgG somatically mutated, class-switched autoantibodies	Mainly IgM, occasionally low titre IgG, unmutated polyreactive autoantibodies
Target	Bind to self-antigens with high affinity	Bind to self-antigens with low-medium affinity but high avidity
Effector Functions/ Mechanisms of Action	Triggers an immune response against self-tissues by mechanisms such as IgG Fc direct activation of effector pathways, immune complex formation and release of inflammatory cytokines	First line of defence against infection due to a broad reactivity for a wide variety of microbial components (exogenous antigens such as bacteria, viruses and fungi)
	Disturb homeostatic pathways, including cell clearance, antigen-receptor signalling or cell effector function pathways	Maintain immune tolerance (anti-idiotypic mechanisms) and homeostasis (particularly B-cell homeostasis)
	Contribute to autoimmune pathology	Perform housekeeping functions (cell debris clearance/apoptosis)

(Elkon and Casali, 2008; Lobo, 2016)

Moreover, autoimmune diseases are antigen-specific, with many classified as a result of the organs and tissues in which the aberrant immune response targets. A well-known example is rheumatoid arthritis (RA), where the immune system mistakenly sends auto-antibodies to synovial joints, causing joint damage and deformity (Yap *et al.*, 2018). The ensuing clinical pathology of autoimmune disease arises when autoimmunity causes chronic tissue damage, whether organ-specific or systemic. Currently, the aetiology and specific mechanisms causing autoimmune disease are relatively unknown. However, research believes its causation is influenced by a complex interplay of roles involving genetics, environmental factors, and gut microbiota composition (Vetrano *et al.*, 2022). This proposed multifactorial aetiology would explain the heterogeneous nature of autoimmune disease between conditions and patients with the same condition.

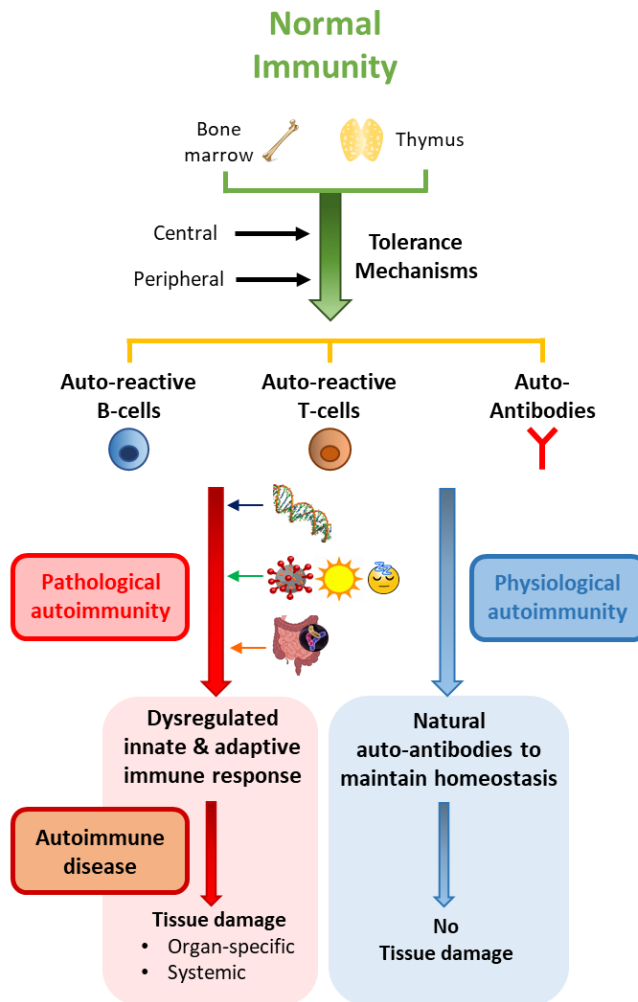


Figure 2. Summary of the process from normal to pathological autoimmunity and eventual autoimmune disease manifestation.

A small proportion of autoreactive T and B cells are present in normal individuals. The presence of risk factors (genetics, the environment, gut microbiota) prompts the transition from harmless to pathological. Modified from (Wang, Wang and Gershwin, 2015).

1.2.1 Risk Factors

Some individuals have an increased risk of developing chronic autoimmune disease, with a myriad of factors contributing to this susceptibility. Table 1.3 details some of the many risk factors, including genetic predispositions, environmental influence, gut microbiota dysbiosis, hormonal fluctuations, gender, age, medications, pre-existing autoimmune conditions, and extensive medical histories.

Table 1.3. Risk factors for autoimmune disease.

Risk Factor	Known Information	Reference
Genetic predisposition	<ul style="list-style-type: none"> • HLA susceptible • Non-HLA loci • Epigenetics • Family history 	(Wang, Wang and Gershwin, 2015)
Environment	<ul style="list-style-type: none"> • Microorganisms • Infections/molecular mimicry • Xenobiotics, e.g. tobacco, UV, drugs • Nutrients (Vitamin D deficiency)/Diet • Geographical location 	(Wang, Wang and Gershwin, 2015)
Gender	<ul style="list-style-type: none"> • Gender bias towards women (rate of 2:1) • Of all patients diagnosed with AID, ~ 80% are women 	(Angum <i>et al.</i> , 2020)
Hormonal fluctuations	<p>AID can be triggered by:</p> <ul style="list-style-type: none"> • Periods of extensive stress • Pregnancy • Menopause 	(Angum <i>et al.</i> , 2020)
Gut microbiota	<ul style="list-style-type: none"> • Gut barrier dysfunction • Gut dysbiosis (imbalance in the composition of the gut microbiota) 	(Martel <i>et al.</i> , 2022)
Age	<ul style="list-style-type: none"> • Age-related autoimmunity (development of AID in the elderly) • The tendency for autoreactivity increases with age 	(Goronzy and Weyand, 2012)
Pre-existing autoimmune disease	<ul style="list-style-type: none"> • AIDs tend to cluster in individuals with a pre-existing AID • Prone to multimorbidity (many diseases co-occurring at once) 	(Bernstein, Wajda and Blanchard, 2005)
Medications	<ul style="list-style-type: none"> • 110 drugs implicated in causing AID from 10 drug classes, including antibiotics, antipsychotics, antidepressants • Drug-induced lupus 	(Chang and Gershwin, 2011)
Extensive medical history	<ul style="list-style-type: none"> • History of chronic inflammation • Obesity • Sleep disorders • Depression (treatment-resistant) 	(Hsiao <i>et al.</i> , 2015; Tsigalou, Vallianou and Dalamaga, 2020; Chan <i>et al.</i> , 2023)

AID; Autoimmune disease, HLA; Human leukocyte antigen

Current research emphasises the complex interplay between genetic predispositions and environmental factors in the development of autoimmune diseases. While genetic predispositions account for approximately 30% of autoimmune diseases, environmental influences are thought to wield a significant 70% contribution to the breakdown in tolerance, encompassing a diverse array of elements ranging from infections and gut dysbiosis to dietary components and toxic agents (Vojdani, Pollard and Campbell, 2014). Figure 3 illustrates a comprehensive array of factors that impact autoimmunity, focusing on environmental influences.

Infectious agents, comprising bacteria, viruses and fungi, are believed to trigger autoimmunity through mechanisms such as molecular mimicry, epitope spreading and bystander activation (Floreani, Leung and Gershwin, 2016). Molecular mimicry, in particular, is a mechanism where foreign antigens share structural or sequence similarities with self-antigens, leading to immune responses that inadvertently target host tissues through cross-reactivity. Notably, molecular mimicry has been implicated in the pathogenesis of many autoimmune diseases, including MS, Graves' disease and diabetes mellitus (Cusick, Libbey and Fujinami, 2012). Moreover, disruption to the gut microbiome, known as dysbiosis, has emerged as another pivotal environmental factor that is associated with impaired intestinal barrier function and dysregulation of the mucosal immune system, ultimately contributing to the breakdown of tolerance mechanisms (Khan and Wang, 2020). Additionally, dietary components like salt (sodium chloride) and

toxic chemicals such as mercury, pesticides, silica and smoking have been identified as potential inducers and exacerbators of autoimmunity.

Ideally, if autoimmunity could be detected early through disease biomarkers, patient prognosis could be improved by removing potential offending triggers such as those highlighted in Figure 3 to halt or reverse autoimmune disease development before irreversible tissue damage.

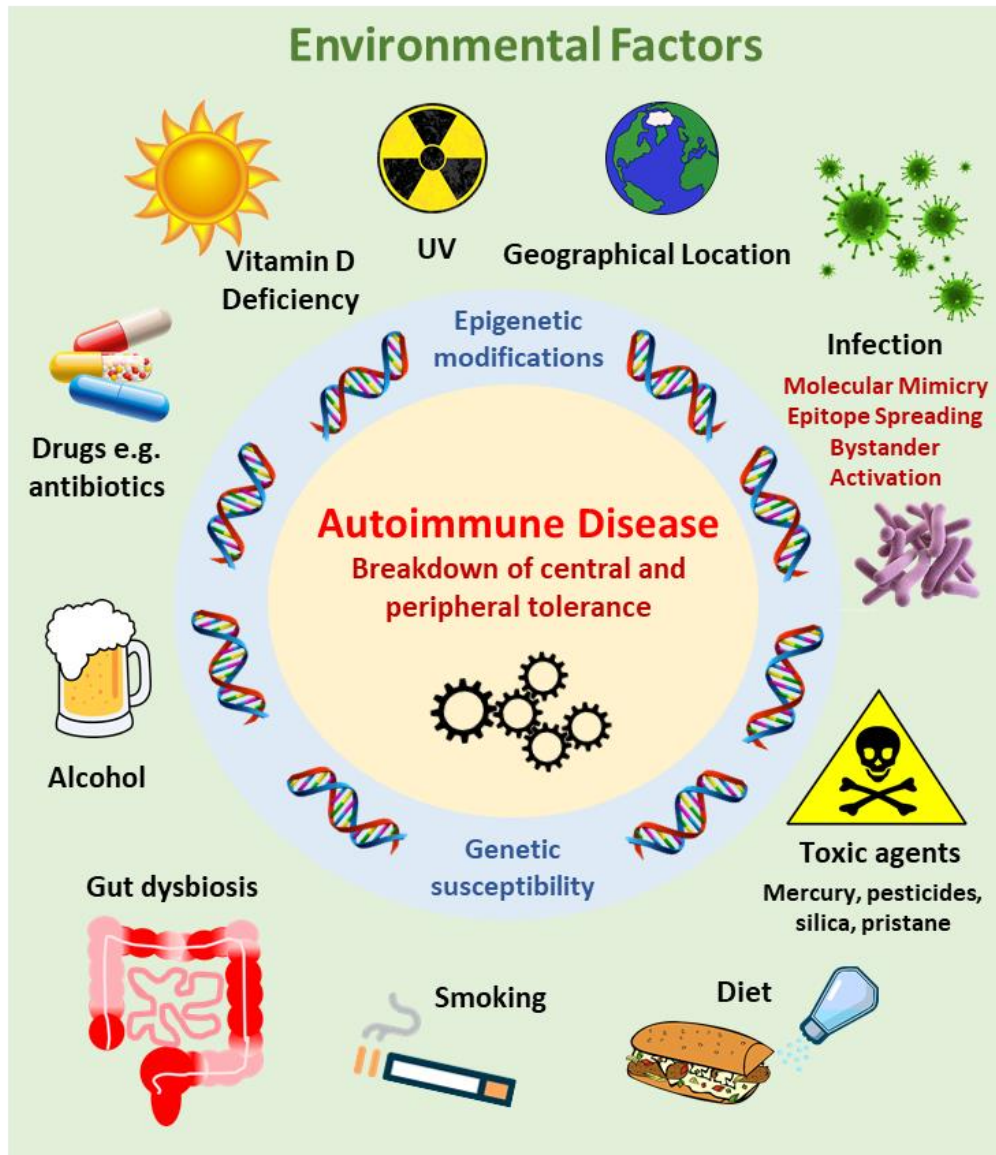


Figure 3. Environmental Factors and Autoimmune Disease.

Environmental Factors known to induce or exacerbate autoimmune disease development are displayed. Infectious triggers such as viral and microbial antigens are hypothesised to induce autoimmunity through mechanisms such as molecular mimicry (foreign antigens resembling self-antigens can lead to immune response cross-reactivity), epitope spreading (diversification of immune response to new antigen epitopes) and bystander activation (activation of immune cells to tissue damage) (Floreani, Leung and Gershwin, 2016). Gut dysbiosis involves dysregulation of the gut microbiome composition and stability, impacting intestinal barrier function and mucosal immunity (Khan and Wang, 2020). Gut dysbiosis may be induced by other environmental factors presented in the figure, such as prolonged antibiotic usage and infection. These environmental factors may contribute to the breakdown of central and peripheral tolerance mechanisms in genetically susceptible individuals, leading to autoimmune diseases.

1.2.2 Preclinical Autoimmunity

The preclinical phase of autoimmune diseases is an initial asymptomatic period, followed by the emergence of non-specific symptoms that precede clinically evident disease manifestations (Frazzei *et al.*, 2022). Figure 4 illustrates the transitional stages in the pathogenesis of autoimmune diseases, from an at-risk state to the establishment of autoimmune disease.

During the preclinical phase, circulating autoantibodies can be detected years before the onset of clinical disease symptoms. For instance, research has shown that autoantibodies can be identified up to 18-20 years prior to the onset of symptoms in Sjögren's syndrome (SjS) (Theander *et al.*, 2015).

Specifically, Theander *et al.* (2015) identified the most commonly detected autoantibodies preceding SjS diagnosis as antinuclear antibodies (ANAs), followed by rheumatoid factor (RF), anti-Ro 60/SSA, anti-Ro 52/SSA, and anti-La/SSB. Recognising early signs of autoimmunity before clinical symptom manifestation can be challenging, especially when dealing with vague and non-specific symptoms. However, the early detection of autoantibody responses targeting specific disease-associated antigens, such as Ro/SSA and La/SSB for SjS, presents a promising screening strategy. Therefore, if autoimmunity could be detected during the preclinical phase, clinical interventions could be initiated to prevent tissue damage and enhance prognosis.

Possibilities for clinical interventions include a range of preventative strategies such as lifestyle changes, dietary modifications and the initiation of

therapy aimed at mitigating or attenuating irreversible tissue damage by reducing attributable risks. Treatment modalities for autoimmune diseases include immunosuppressive medications as well as targeted immunotherapies like recombinant protein therapeutics and B-cell depletion therapy (Jung and Kim, 2022). Employing established treatments during the preclinical phase of autoimmunity may hold significant promise for positively impacting disease prognosis by halting or decelerating disease progression. Additionally, identifying susceptible individuals at early stages of autoimmunity would enable further research into the development of novel therapies focused on prevention and early intervention.

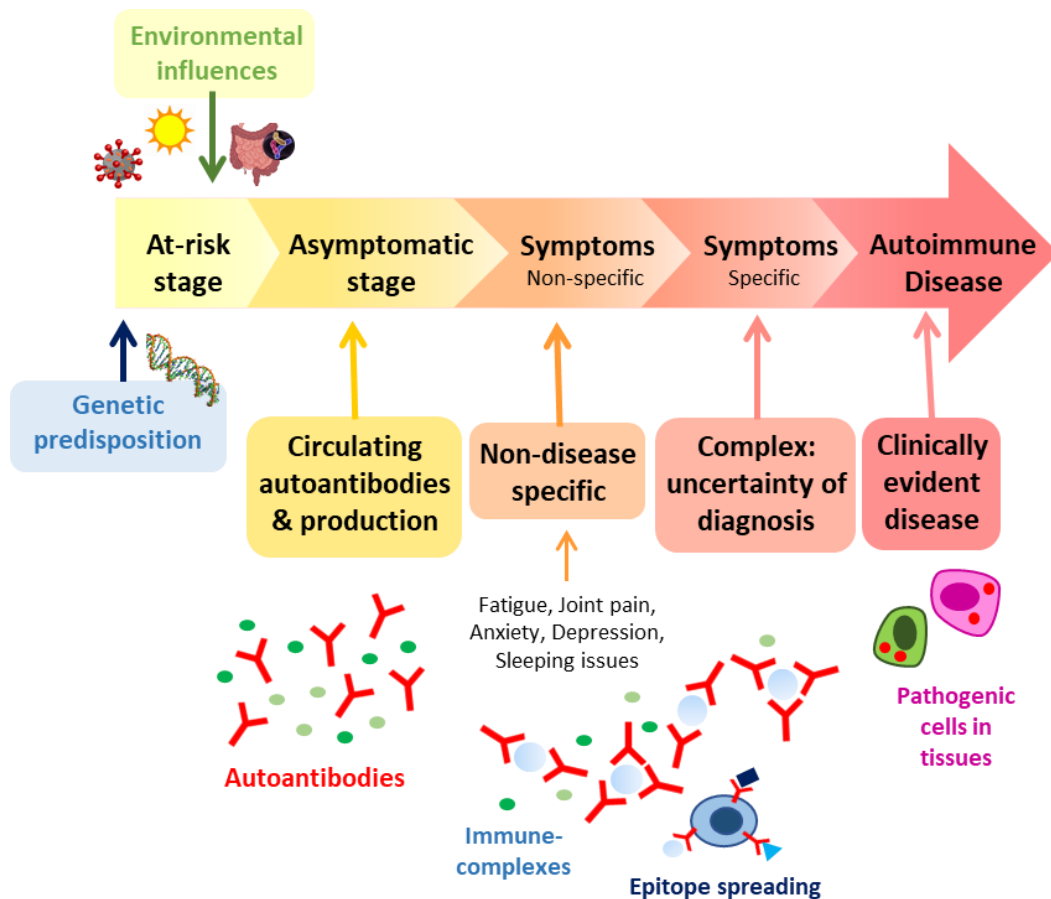


Figure 4. Pathogenesis of autoimmune disease; transitioning from preclinical to overt autoimmune disease establishment.

Exposure to risk factors such as environmental influences in genetically predisposed individuals can induce the transition from normal to pathogenic autoimmunity. Circulating autoantibodies are present in the preclinical asymptomatic and non-disease-specific symptom phases. Although a small proportion of healthy individuals also have detectable autoantibodies. Epitope spreading broadens the autoimmune response towards different epitopes and can lead to the targeting of multiple tissues and organs (Choi and Costenbader, 2022). For some autoimmune diseases such as SLE, the formation and accumulation of immune complexes deposited in tissues can contribute to tissue damage and inflammation (Choi and Costenbader, 2022). These events can lead to the emergence of clinical disease-specific symptoms and overt autoimmune disease establishment. The pathogenesis of autoimmune diseases is still unclear and likely heterogeneous across conditions. Modified from: (Choi and Costenbader, 2022; Frazzei et al., 2022).

1.2.3 Multimorbidity

Multimorbidity is the simultaneous presence of two or more long-term conditions in a patient, which spans physical and mental health conditions and infectious diseases. Autoimmune disease patients tend to have concurrent conditions, including other autoimmune diseases and clinical manifestations, such as gastrointestinal disease. For instance, immune-mediated diseases tend to cluster in individuals with a pre-existing autoimmune condition, hypothesised to stem from shared processes triggered across multiple organs (Bernstein, Wajda and Blanchard, 2005; Buckley *et al.*, 2021). To illustrate, (Robinson *et al.*, 2006) found that 16.5% of patients with RA also had systemic lupus erythematosus (SLE), inflammatory bowel disease (IBD) or SjS. Additionally, gastrointestinal manifestations are common, highly variable and a significant cause of morbidity across autoimmune diseases, particularly systemic disorders, such as SLE, RA, SjS and Bechet's disease (Cojocar *et al.*, 2011). As a result, multimorbidity is important to identify in order for a multidisciplinary approach to treatment; if unidentified, it serves as a risk factor for poor prognosis.

1.2.4 Rising Impact and Prevalence of Autoimmune Disease

Autoimmune diseases have a lasting negative impact on the quality of life for millions of individuals globally due to persistent and chronic tissue damage, necessitating long-term treatment and monitoring. This global burden is expected to escalate further as mounting evidence points to a steady rise in the prevalence of autoimmune diseases, positioning them as a leading public

health concern (Lerner, Jeremias and Matthias, 2016). For example, a study of 22 million individuals in the UK investigating the incidence, prevalence and co-occurrence of autoimmune disease concluded that one in ten individuals are affected, with their prevalence increasing at varying rates depending on the specific autoimmune condition (Conrad et al., 2023). Regarding comorbidities, the study found a higher occurrence of autoimmune diseases like coeliac disease and thyroid diseases in individuals who developed childhood-onset type 1 diabetes (T1D), possibly indicating common genetic or environmental factors (Conrad *et al.*, 2023). To alleviate the increasing strain on healthcare systems, it is becoming more critical to increase early detection, improve diagnostic capabilities for identifying comorbidities, implement effective monitoring strategies, and provide timely treatment for autoimmune diseases.

1.3 The Gut Microbiome and The Immune System

The human gut microbiome constitutes a complex community of microorganisms that reside on the skin and mucosal surfaces of the gastrointestinal (GI) tract. The gut microbiome is the largest symbiotic ecosystem at the centre of human health and disease, comprising bacteria, archaea, eukaryotes and viruses (Sekirov *et al.*, 2010). These microorganisms interact with each other and the host to maintain overall human health by regulating the innate and adaptive immune system and homeostasis (Yang, Chen and Cai, 2023). To sustain optimal health and preserve the delicate balance of the immune system, invading pathogens need to be efficiently

eliminated while self-tolerance is maintained to avert autoimmune responses (Wu and Wu, 2012). To facilitate this equilibrium, the gut microbiota metabolises proteins and complex carbohydrates, synthesises vitamins and produces metabolic products that mediate the cross-talk between the gut epithelium and immune cells (Yoo *et al.*, 2020). As a result, the immune system-gut microbiota interplay ensures a robust mucosal barrier and promotes innate immune responses against invading pathogens.

Concurrently, the gut epithelial cells form a mucosal barrier to segregate the microbiota from immune cells in order to decrease intestinal permeability and limit microbial translocation. However, a compromised relationship between the gut microbiota and the mucosal immune system can lead to the proliferation of pathogenic bacteria, triggering metabolic shifts that disrupt the epithelial barrier, thereby heightening susceptibility to infections (Yoo *et al.*, 2020). An example of a compromised relationship is the negative alteration of gut microbial composition, known as gut dysbiosis, which can cause immune dysregulation by overgrowth of pathogenic bacteria, depletion of beneficial commensal bacteria and/or loss of bacterial diversity (DeGruttola *et al.*, 2016). Consequently, dysbiosis commonly results in disease development such as GI conditions, immune-mediated disorders, metabolic diseases and cancers (Hou *et al.*, 2022). Research is now focusing on understanding the complex interaction loop between the microbiota and the host mucosal and systemic immune response to gain insights into disease development.

1.4 Patient Cohort of Interest: *Clostridium difficile* Infection

Clostridium (Clostridioides) difficile is a gram-positive, spore-forming bacillus, the primary causative agent of in-hospital infectious diarrhoea and a significant cause of morbidity and mortality worldwide. It is transmitted between human hosts through the faecal-oral route, with a high frequency of transmission in healthcare settings (Leffler and Lamont, 2015). Susceptible individuals include those with gut microbiota dysbiosis, mainly due to regular antibiotic use that alters the balance of protective commensal microbes in the gut, enabling pathogen colonisation (Bagdasarian, Rao and Malani, 2015). As *C. difficile* is often resistant to the actions of antibiotics, its spores can germinate, proliferate and produce exotoxins that cause severe inflammation and result in *C. difficile* infection (CDI) (Fischer, 2017). Other contributing risk factors include age, immunosuppression and severe pre-existing conditions, especially those impacting the gastrointestinal tract, e.g. IBD (Leffler and Lamont, 2015). CDI presentation can range from asymptomatic to life-threatening fulminant colitis and death (Bagdasarian, Rao and Malani, 2015). In particular, the presence of exotoxins TcdA and TcdB exert the main pathological effects of disrupting intestinal cells, causing inflammation and infection (Di Bella *et al.*, 2016).

Treatment of acute CDI consists of stopping the antibiotic causing the infection if possible, followed by antibiotics known to kill *C. difficile*, such as oral metronidazole for first-line therapy and vancomycin for severe disease (Monaghan, Boswell and Mahida, 2009). Limitations include a significant

recurrence rate, such as 25% recurrence after vancomycin, and promotion of antibiotic resistance (Fischer, 2017). Fundamentally, an antibiotic-induced condition is treated with another antibiotic, which, while killing CDI, also kills indigenous bacteria of the microbiota that keep CDI at bay, causing greater host susceptibility to recurrent disease (Fischer, 2017).

1.4.1 Immune Responses to CDI

Toxigenic *C. difficile* colonises the colon and, once in its vegetative form, proliferates and produces two exotoxin proteins, TcdA and TcdB, that activate the immune response, as summarised in Figure 5. During CDI, TcdA and TcdB inactivate members of the Rho family of guanosine triphosphatases (Rho GTPases), resulting in colonocyte apoptosis and the weakening of enterocyte tight junctions that compromise intestinal barrier function (Leffler and Lamont, 2015). This toxin-mediated damage allows microbial translocation of intestinal flora across into the mucosa. The intestinal microorganisms entering the mucosa are subsequently exposed to resident immune cells, part of the innate immune response's first line of defence, which triggers the recruitment of a spectrum of immunocytes to the site of infection through cytokine and chemokine signalling (Nibbering *et al.*, 2021). As a result, the robust immune response to the exotoxins and translocated microbiota causes the majority of the damage to the colon and the clinical symptoms of CDI, superseding the damage caused by the toxins alone (Fischer, 2017).

Regarding the adaptive humoral immune response to CDI, T and B cells are triggered to produce anti-TcdA and anti-TcdB pathogen-specific

immunoglobulins (Igs). IgA and IgG are the primary antibodies involved, with IgG neutralising general toxins in the blood and IgA neutralising toxins around the mucosal surface. The fundamental role of antitoxins is to prevent CDI development once infected rather than preventing initial colonisation (Fischer, 2017).

Generally, a better understanding of the complex interactions between the microbiota and the host immune response is required to further elucidate the immune system's role in CDI. Overall, the immune system is a key driving force in the recovery and pathogenesis of CDI.

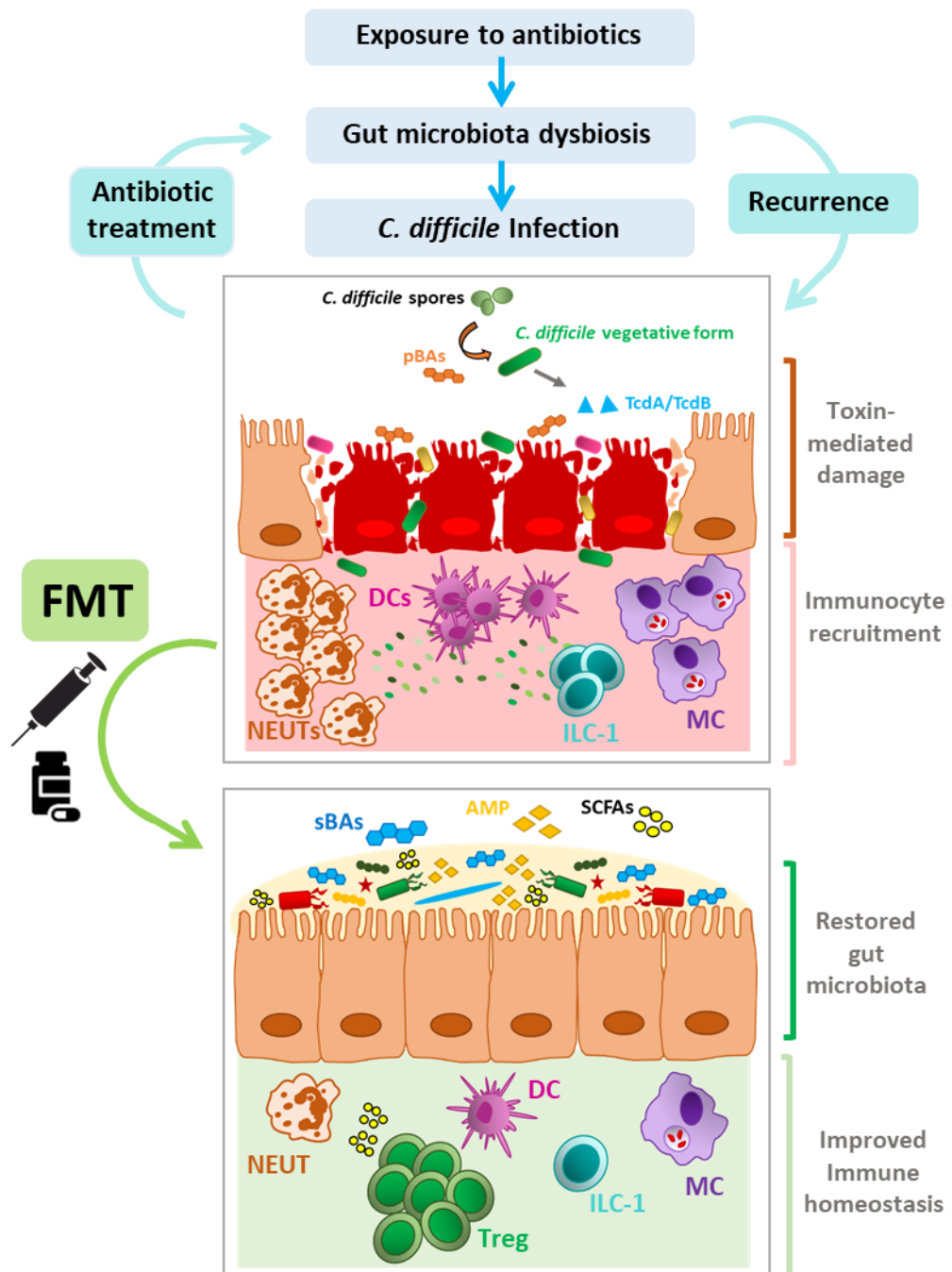


Figure 5. Immune responses in CDI alongside Fecal Microbiota Transplantation (FMT) for restoring microbial and immune homeostasis.

Gut dysbiosis, typically caused by antibiotic exposure, leaves individuals susceptible to *C. difficile* infection (CDI). Primary bile acids (pBAs) induce spore germination into the vegetative form of *C. difficile*, which produces the exotoxins TcdA and TcdB. Resulting toxin-mediated damage involves an impaired intestinal barrier and microbial translocation. Immunocytes are recruited as a first line of defence, including neutrophils (NEUTs), dendritic cells (DCs), macrophages (MC) and innate type I lymphocytes (ILC-1) which release pro-inflammatory cytokines causing tissue damage to the colon. Treating CDI with antibiotics maintains gut dysbiosis and increases the susceptibility for recurrent CDI (rCDI). FMT works by re-establishing the diversity of the gut microbiota, increasing secondary bile acids (sBAs) and short-chain fatty acids (SCFAs). Epithelial cells produce antimicrobial peptides (AMPs) to restore the intestinal barrier. As a result, FMT reduces immune activation by innate immunocytes to improve immune homeostasis in the colon. Treg = Regulatory T-cells. Modified from: (Frisbee and Petri, 2020; Soveral et al., 2022).

1.4.2 Treatment: Faecal Microbiota Transplantation (FMT)

Faecal Microbiota Transplantation (FMT) is an alternative treatment to antibiotics, successfully used to treat rCDI and severe CDI, with a 90% relapse-free cure rate (Frisbee and Petri, 2020). FMT involves administering a mix of faecal material containing functional microbiota from a donor to the intestinal tract of a recipient through a nasogastric or nasoduodenal tube, enema, colonoscopy or capsule (Beyi, Wannemuehler and Plummer, 2022). The treatment aims to restore gut microbiota homeostasis caused by dysbiosis and reinstate colonisation resistance for long-term protection against *C. difficile*, as explained in Figure 5 (Frisbee and Petri, 2020).

FMT is widely used due to its high success rate, with thousands of people undergoing treatment worldwide. Due to this success and the established link between host immunity and the gut microbiota, FMT is now applied more broadly to dysbiosis-associated diseases, including autoimmune diseases, with human clinical trials underway. For instance, FMT may provide a robust treatment option for IBD patients by restoring the disrupted gut microbiota in order to improve local mucosal and systemic immune homeostasis (Yang, Chen and Cai, 2023). Successful treatment would implicate the gut microbiota as a potential modulator in the pathogenesis of autoimmune disease (Belvoncikova, Maronek and Gardlik, 2022). FMT is being investigated to treat a range of other autoimmune diseases, including T1D, RA, MS and SLE.

Currently, investigations into the long-term effects of FMT are ongoing. So far, researchers have closely monitored the effect of FMT on microbial and

metabolic responses in patients with rCDI whilst neglecting the impact on the immune response. However, it is crucial to emphasise that the restoration of gut microbiota by FMT influences various immune pathways critical for CDI recovery; this highlights the necessity of monitoring immune responses before and after FMT (Frisbee and Petri, 2020).

1.5 Patient Cohort of Interest: Cystic Fibrosis

Cystic fibrosis (CF) is an inherited life-shortening condition characterised by mucus build-up in the lungs and digestive tract due to dysfunction of the cystic fibrosis transmembrane conductance regulator (CFTR) protein. Beyond its characteristic manifestations, clinicians have observed that various CF patients have non-CF-related symptoms and aetiology, such as gastrointestinal complaints or arthritis connected to autoimmunity. As a result, it is hypothesised that the CFTR protein may directly alter the immune response, intraluminal dysbiosis and inflammation, establishing an environment conducive to autoimmunity (Chadwick et al., 2023). Examples of autoimmune diseases with a high prevalence in CF patients are coeliac disease, Crohn's disease, vasculitis and CF-associated arthritis. However, symptoms of autoimmune disease can be intersecting, unclear and intermittent with CF-related symptoms, so autoimmune diseases go underdiagnosed and underreported in the CF patient population (Chadwick et al., 2023).

Moreover, CF patients often develop intraluminal dysbiosis and inflammation stemming from alterations to the gut microbiota. These alterations arise in

response to many factors, including CFTR dysfunction, antibiotic medication, dietary patterns, pancreatic malabsorption and environmental influences (Caley *et al.*, 2023). Specifically, CFTR dysregulation manifests in thickened mucus, low intestinal pH, pancreatic enzyme deficiencies and compromised GI motility, affecting over 90% of CF patients (van Dorst, Tam and Ooi, 2022).

Overall, emerging research underscores the pivotal role of immune system dysfunction and gut microbiota dysbiosis in CF pathogenesis, offering an interesting focus for investigations into immune health profiling.

1.6 Serological Profiling and Antibodies as Biomarkers

1.6.1 Serological Profiling

Serological profiling is a non-invasive technique to analyse serum for a spectrum of specific antibodies and antigens. Ultimately, serological profiling explores the host humoral immune response. Each Ig isotype can be evaluated; IgG, IgM and IgA are the common isotypes examined to gain insights into recent, present and prior infections and diseases. Generally, serological profiling assays are utilised for diagnostic purposes such as autoimmune disease diagnosis, testing vaccine responsiveness and verifying an individual's immunity to certain diseases.

1.6.2 Antibodies as Biomarkers

Biomarkers are measurable naturally occurring molecules or genes that, when altered from normal, indicate disease pathogenesis and manifestation.

Biomarkers are often objectively measured in the blood or urine.

Antibodies are common serological biomarkers due to their stability, ubiquity, abundance and existence early in disease development. Self-reactive autoantibodies are particularly valuable biomarkers for autoimmune disease due to the inherent intricacy of symptom classification, diagnosis, monitoring and prognosis (Ahearn *et al.*, 2015). The longer half-lives of autoantibodies in vivo make detection more successful than measuring their associated antigen. For instance, autoantibodies can remain stable in the bloodstream for several weeks, with IgG autoantibodies persisting for three weeks and over years when blood is stored at -80°C (Ma *et al.*, 2022). Hence, researchers regularly utilise autoantibody biomarkers to investigate autoimmunity.

1.6.3 Established Disease-specific Antibody Biomarkers

Serological profiling research has identified many autoantibody biomarkers associated with autoimmune diseases. Table 1.4 provides examples of various established serological antibody biomarkers, their antigenic determinants and clinical associations. Currently, no single biomarker is solely utilised for diagnosing autoimmune diseases due to limited diagnostic accuracy either in terms of low sensitivity or specificity. For example, perinuclear anti-neutrophil cytoplasmic antibodies (pANCA) and anti-*Saccharomyces cerevisiae* antibodies

(ASCA) are recognised biomarkers for IBD with high specificity (low rate of false-positives) but low sensitivity (high rate of false-negatives) (Bourgonje *et al.*, 2022). The low sensitivity of autoantibody biomarkers stems from the inherent heterogeneity between autoimmune disease patients and their immune systems. This heterogeneity is partly due to epitope spreading, a phenomenon involving the diversification of the humoral immune response, wherein the immune system broadens its reaction through new T cells and antibody responses directed to different epitopes of the same antigen (intramolecular spreading) or another antigen (intermolecular spreading) over time (Venkatesha, Durai and Moudgil, 2015). Epitope spreading can contribute to the progression of autoimmune disease and is observed for both organ-specific and systemic autoimmune diseases (Monneaux and Muller, 2002). As a result, not all patients will exhibit the same level of sensitivity to a single antibody biomarker.

Regarding specificity, antinuclear antibodies (ANA) are associated with connective tissue diseases, such as SLE and SjS, with a high sensitivity but low specificity (Grygiel-Górniak, Rogacka and Puszczewicz, 2018). This low specificity is highlighted as low titres of ANA are also detected in up to 35% of the healthy population, with most unlikely to go on and develop an autoimmune disease (Marin *et al.*, 2009). These factors illustrate that ANA alone has limited to no diagnostic utility.

Hence, researchers have concluded that panels to detect a range of antibody biomarkers are a more effective screening approach for autoimmune diseases, especially due to their multifactorial and multimorbid nature.

Table 1.4. Serological antibody biomarkers with established associations for autoimmune diseases.

Antibody biomarkers	Antigenic Determinants	Clinical associations	References
ANAs:		SLE, SjS, neonatal lupus,	(DeMarshall <i>et al.</i> , 2015;
Anti-dsDNA	dsDNA	Inflammatory	Suurmond and Diamond, 2015)
Anti-SSA/Ro	Ribonucleoproteins	myopathy, Diffuse	
Anti-SSB/La	Ribonucleoproteins	scleroderma	
Anti-Sm	Smith		
Anti-Jo1	Histidyl tRNA synthetase		
Anti-Scl-70	DNA topoisomerase I		
Anti-CARP	Carbamylated proteins	RA	(Suurmond and Diamond, 2015)
Anti-TPO	Thyroid peroxidase	Autoimmune thyroiditis, Hashimoto's disease, Grave's disease	(DeMarshall <i>et al.</i> , 2015)
ANCA:		IBD, MPA, EGPA, PSC, RA	(Suurmond and Diamond, 2015)
pANCA	Myeloperoxidase		
cANCA	Proteinase 3		
ASCA	Mannan	IBD	(Bourgonje <i>et al.</i> , 2022)
IAA	Insulin	T1D	(Frommer and Kahaly, 2020)
Anti-Gliadin	Gliadin in wheat	CeD, IBD, gastrointestinal diseases	(Caja <i>et al.</i> , 2011)

ANA, Antinuclear antibody; dsDNA, double-stranded DNA; ANCA, Antineutrophil cytoplasmic antibodies; ASCA, Anti-*Saccharomyces cerevisiae* antibodies; IAA, Insulin autoantibodies; SLE, Systemic lupus erythematosus; SjS, Sjogren's disease; IBD, Inflammatory bowel disease; MPA, Microscopic polyangiitis; EGPA, Eosinophilic granulomatosis with polyangiitis; PSC, Primary sclerosing cholangitis; RA, Rheumatoid arthritis; T1D, Type 1 diabetes; CeD, celiac disease.

1.7 Serological Antibody Profiling Techniques

Serological profiling for antibody biomarkers can be conducted using several methodologies. Since 1970, research has primarily used traditional enzyme-linked immunosorbent assays (ELISA), a classical serological profiling technique used to identify most of the antibody biomarkers outlined in Table 1.4 (Lequin, 2005). Even today, ELISAs remain the mainstay of clinical testing, providing robust and standardised procedures for diagnosing autoimmune diseases. Despite this, rapid technological advances have brought high-throughput multiplexed microarrays to the forefront, enabling unparalleled detection and a new comprehension of antibody repertoires in chronic diseases (Bourgonje *et al.*, 2022).

1.7.1 ELISA

There are many forms of ELISA, all characteristically involving the immobilisation of proteins (antigens or antibodies) onto solid-phase surfaces. These are probed with biological samples to detect specific antibody-to-antigen interactions through an enzymatic signal. Indirect ELISAs are predominantly used to detect serological antibodies, where primary antibodies in the sera bind to coated antigens on plate wells and are detected through an enzyme-labelled secondary antibody, as shown in Figure 6A. The advantages of conducting ELISAs include their high sensitivity and specificity, the ability to quantify with calibration curves, and the ability to use well-established standardised procedures (Bourgonje *et al.*, 2022). However, downsides include its lack of scalability as typically antibody responses are

measured for an antigen at a time and require relatively large amounts of precious sample, antigen and reagents. Additionally, as ELISAs produce an absorbance-based readout, they have a narrow dynamic range dictated by the optical density (OD), requiring samples to be tested at a range of dilutions to pinpoint the linear relationship (Castro and Gourley, 2010).

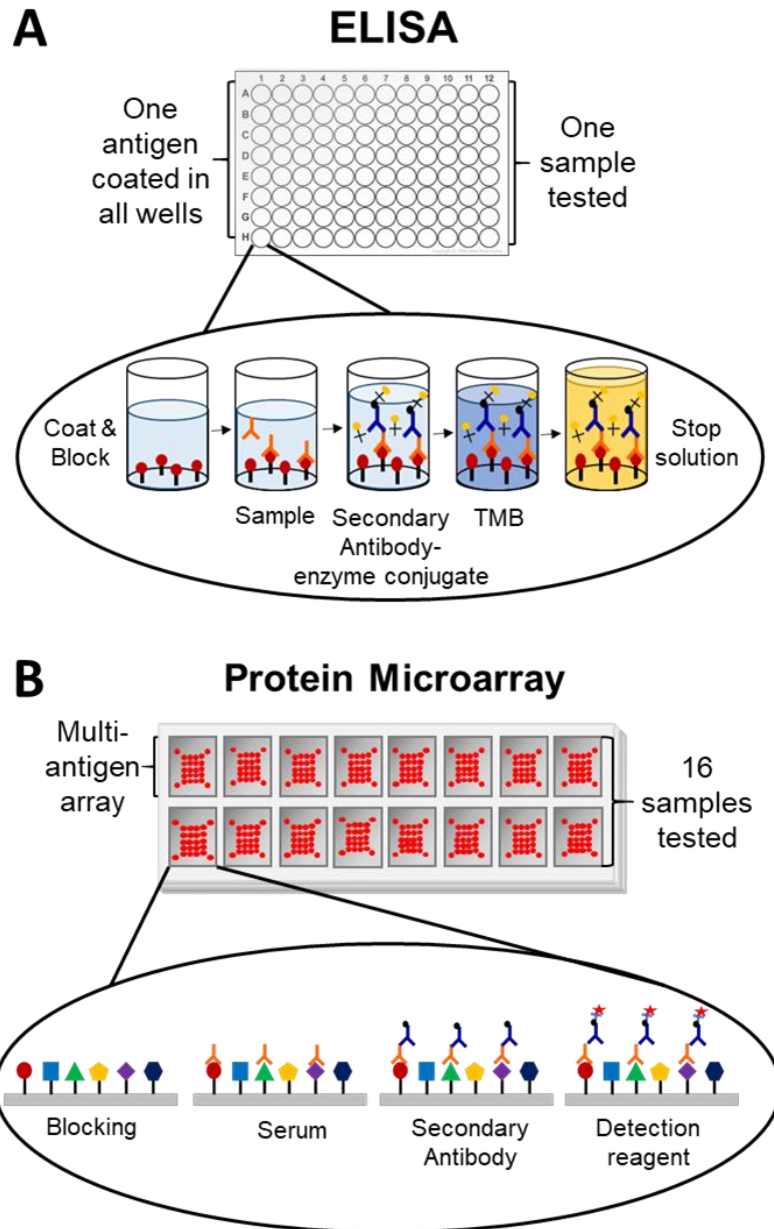


Figure 6. Serological antibody detection methods.

A) Indirect enzyme-linked immunosorbent assays (ELISA) detect primary antibodies specific to a single antigen in a serum sample. B) Protein/antigen microarray platforms enable multiplexed detection of a multitude of antibody-to-antigen interactions for many serum samples at a time.

1.7.2 Microarray Technology

Protein microarrays are miniaturised and multiplexed assay systems that comprise micro amounts of many proteins in a high-density format printed onto chemically-derivatised slide surfaces (Figure 6B). Proteins are either covalently linked to glass surfaces or adsorbed onto slide membranes and probed with samples to detect antibody binding through fluorescent signals. An advantage to protein microarrays over ELISA is the simultaneous detection of numerous analytes from minute samples in a single experiment, providing an efficient screening platform for human sera. Thus, this greater throughput analysis means more data can be obtained per volume of a biological sample (Wilson and Nock, 2003). Additionally, microarray technology overcomes the limited dynamic range of ELISA imposed by OD by measuring antibody-to-antigen interactions through fluorescent intensity. To illustrate, many studies have found that protein microarray data correlates well with ELISA data regarding sensitivity and robustness, if not with a higher sensitivity (Zhu *et al.*, 2015). For instance, an early pioneering study into microarrays conducted by (Robinson *et al.*, 2002) found that antigen microarrays provided lower limits of detection for autoantibodies compared to a conventional ELISA, with arrays producing a 4-8-fold increase in sensitivity for detecting autoantibodies.

In terms of its clinical application, protein microarray serological screening has the potential to detect early disease onset, identify multimorbidity, predict treatment responses, and improve current diagnostic tools. However,

integrating microarrays into clinical settings encounters challenges, including standardisation of protocols, quality control and accessibility to specialised equipment. As such, ELISAs continue to uphold their status as the “gold standard” for clinical testing where reliability, standardisation and accessibility are paramount.

1.8 Current Screening Platform for Autoimmune Disease

Currently, there is no singular test to diagnose autoimmune disease; instead, diagnosis involves a spectrum of techniques, from physically invasive biopsies to less-invasive laboratory testing, including serological profiling (Salamunic, 2010). Figure 7 shows the general diagnostic framework for autoimmune diseases. Assessment typically commences with a physical evaluation, which can be invasive in nature, followed by laboratory tests, including individual immunoassays such as ELISAs conducted to quantify serological autoantibodies. Relatively specific tests include anti-mitochondrial antibodies (AMA) for primary biliary cholangitis and RF for RA. In contrast, non-specific tests investigate ANA, including anti-dsDNA and Anti-Sm autoantibodies associated with SLE (Wang, Wang and Gershwin, 2015).

However, the current platform faces several significant limitations. Firstly, many individual tests are conducted at high costs, and there is inherent subjectivity in the selection of assays based on diagnostic impressions made by a physician. Moreover, these tests are typically administered only once clinically significant symptoms have manifested, hindering their potential for early disease detection. Furthermore, the assays employed in this platform

often focus on measuring a single analyte or biomarker at a time, which limits the capacity to investigate the complexity of autoimmune diseases involving multiple factors or biomarkers. These limitations collectively impede the platform's ability to detect multimorbidity and the early identification of autoimmune diseases.

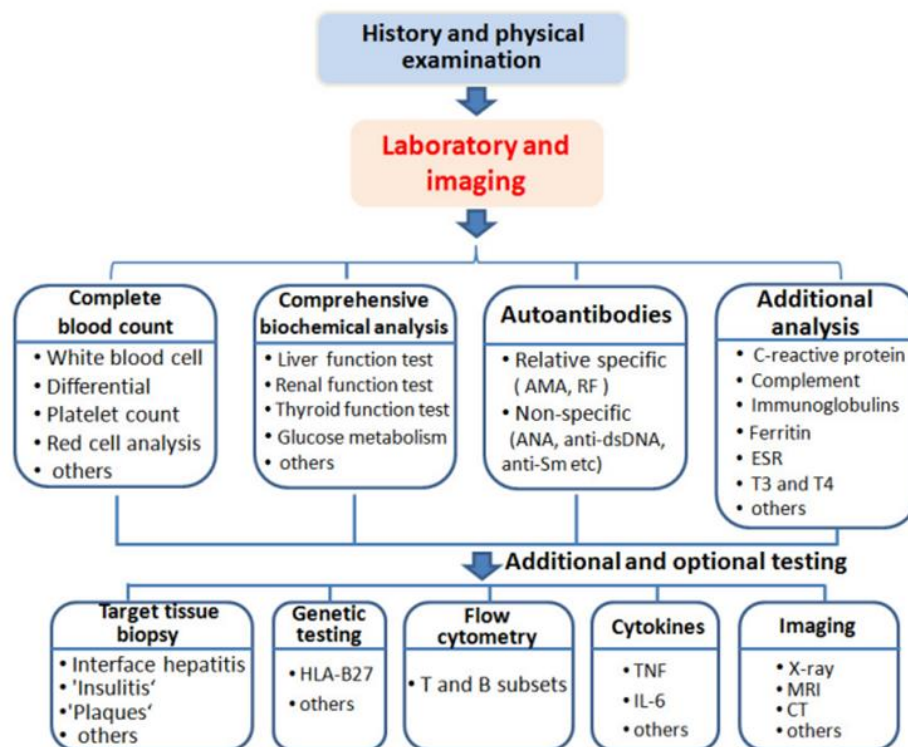


Figure 7. The current diagnostic framework for autoimmune diseases.

AMA, Antimitochondrial antibody; RF, rheumatoid factor; ANA, antinuclear antibody; anti-dsDNA, anti-double-stranded DNA; anti-Sm, anti-Smith antibody. Sourced from (Wang, Wang and Gershwin, 2015).

1.9 Aims and Hypothesis

Autoimmune diseases impact approximately one in ten people in the UK, and their prevalence is rising, placing substantial burdens on healthcare systems and communities (Conrad *et al.*, 2023). Early indicators of pathological autoimmunity can often remain undetected for years, even though autoantibodies have been detected decades before clinical onset. Currently, individuals are screened using costly immunoassays such as ELISAs, which test for single autoantibody biomarkers at a time, lacking scalability. Therefore, there is an urgent need for a comprehensive immune health platform capable of simultaneously screening across a wide array of biomarkers. Such a platform could be used clinically as an early warning system for at-risk patient groups, for detecting multiple health conditions concurrently (multimorbidity) and for precision medicine.

The primary goal of this research was to develop a multiplexed antigen microarray to quantify IgG and IgA antibody responses against a panel of 60 antigens, enabling the analysis of patient humoral immune responses. The antigen panel included autoantigens associated with chronic immune-mediated diseases alongside pathogen-associated and viral recall antigens to investigate infection. The study utilised clinically-defined and accessible cohorts, including individuals diagnosed with cystic fibrosis (CF), *C. difficile* infection (CDI), and those with recurrent *C. difficile* infection (rCDI) before and after FMT. These available cohorts exhibited a heightened susceptibility to

immune-mediated disease due to risk factors, including gut microbiota dysbiosis, a history of inflammation and pre-existing medical conditions.

Our objectives were as follows:

- Develop and validate an antigen microarray platform for quantifying serological IgG and IgA responses across 60 antigens, including autoantigens, positive-control, pathogen-associated and viral recall antigens.
- Screen the sera of CF, CDI, rCDI and healthy controls to generate patient antibody response profiles for comparison.
- Assess the impact of FMT on the humoral responses of patients with rCDI by screening sera pre- and 12 weeks post-FMT.
- Compare the humoral responses of patients with initial CDI and rCDI.

By developing a multiplexed antigen microarray against a panel of 60 antigens, we aim to underscore the clinical potential of such a platform as an early warning system for at-risk patients. Crucially, a platform that could detect sustained levels of disease-associated biomarkers during the preclinical phase of autoimmunity would enable timely interventions geared towards optimising patient outcomes and alleviating strains on our healthcare system. Such a platform offers the opportunity to implement preventative health measures to mitigate known risk factors, such as employing lifestyle and dietary changes. Additionally, pivotal strategies such as regular monitoring of responses to biomarker panels and multimodal treatment approaches could be initiated before irreversible tissue damage occurs.

2 Materials

2.1 Chemicals and Reagents

2.1.1 Printing Buffer

The 2X Printing buffer contained a 2:1 mix of 4M Betaine and 50% (w/v) 2-3 Butanediol.

2.1.2 Blocking Buffers

The selection of Blocking buffers tested during optimisation is displayed in Table 2.1. I-Block™ was chosen for the final assay protocol.

Table 2.1. Blocking buffers.

Blocking Buffers:	Components:	Supplier:
3% v/v of BSA in PBS (3% BSA-PBS)	30% BSA 1X PBS	Sigma, UK, A7284
I-Block™	Casein based 0.2g I-Block in 100ml PBS- Tween-20 (0.1%)	Tropix®, USA, Applied Biosystems, T2015
The Blocking Solution	Casein based Ready to use	Candor Bioscience GmbH, Germany, 110 125
Intercept™ Blocking solution	Ready to use	LI-COR Biosciences, US, 927-85001
SmartBlock™	BSA-free Blocker Ready to use	Candor Bioscience GmbH, Germany, 113 125

2.1.3 Wash Buffer

PBST (PBS-Tween 20): 10X Phosphate-buffered saline (Fisher BioReagents BP399-4) diluted to 1X in distilled water and mixed with 0.05% Tween-20 (Sigma-Aldrich, P1379).

2.1.4 Slide Surfaces

Table 2.2 displays the selection of slide surfaces tested during optimisation.

Aldehyde, Nitrocellulose, Epoxy silane, and BSA-NHS slides were pre-made and provided by Patrick Tighe (supervisor) following established protocols.

Nitrocellulose (made In-house) was chosen for our microarray platform.

Table 2.2. Slide surface chemistries.

Slide Surface:	Supplier:
Aldehyde	Pre-made In-house following (Shi <i>et al.</i> , 2014)
Aldehyde PLUS	Molecular Devices K2080
Nitrocellulose ONCYTE® AVID	Grace Bio-Labs, Inc. 305004
Nitrocellulose	Pre-made In-house – 0.1µm pore nitrocellulose (Millipore, UK)
Epoxy silane	Pre-made In-house following (Nam, Branch and Wheeler, 2006)
BSA-NHS	Pre-made In-house following (MacBeath, 2007)

BSA-NHS, Bovine Serum Albumin-N-hydroxy succinimide

2.1.5 Assay Reagents

The assay reagents used for developing and finalising our microarray protocol are detailed in Table 2.3.

Table 2.3. Assay reagents.

Reagent:	Application:	Supplier:	Diluent used:
Biotinylated anti-human IgG (Goat)	Secondary antibody	Vector Laboratories, BA-3080	The relevant blocking buffer in use
Goat F(ab')₂ anti-human IgA biotin conjugate	Secondary antibody	BIOSOURCE, AHI1109	I-Block™
Anti-human IgG-cy5 conjugate (Goat)	Secondary antibody and detection	Pre-made In-house (By Patrick Tighe – refer to methods 3.5)	3% BSA-PBS
Streptavidin-cy5	Fluorescent dye for detection	Cytiva, PA45001 (Cy®5-Streptavidin)	PBST
Goat serum	Reduce non-specific binding	Sigma, G6767	Secondary antibody diluent

2.2 Antigens

Antigens included on the microarray platform are displayed in Tables 2.4 and 2.5. The 100µg/ml antigen concentration was chosen for optimal functionality and cost-effectiveness. However, certain antigens were only available in small quantities at an affordable price, so a compromise was made to add these antigens at a lower concentration of 50µg/ml. Notably, our focus for the screening assay was not to compare antibody responses between antigens but responses between patients; hence, molecular weight and molarity calculations for each antigen were omitted. Viral recall antigens (Table 2.5) were included on the microarray platform as the majority of the population is expected to have responses to endemic antigens such as EBV and VZV. Table 2.6 displays the approximate seroprevalence of selected viral pathogens in the general population. As such, viral recall antigens help establish if individuals have expected immune responses.

Table 2.4. Auto-antigens and pathogen-associated antigens. Including antigen concentrations and suppliers. Additional information from supplier descriptions included: Native (from a natural source) and recombinant (manufactured artificially).

Antigen	Concentration (µg/ml)	Supplier
Ro (SSA)	100	ATRO2-02 Arotec Native
La (SSB)	100	ALA01-02 Arotec Native
Lactoferrin	100	ATL02-02 Arotec Native
Cenp B Centromere protein B	100	ATC02-02 Arotec Recombinant
Jo-1	100	ATJ01-02 Arotec Native
BPI	100	ATB01-02 Arotec Native

Bactericidal permeability-increasing protein		
RNP/Sm	100	ATR01-02 Arotec Native
Ribosomal P	100	ATR03-02 Arotec Native
cANCA Antineutrophil cytoplasmic antibodies	100	ATP02-02 Arotec Native
pANCA Perinuclear anti-neutrophil cytoplasmic antibodies	100	ATM01-02 Arotec Native
GBM Glomerular basement membrane	100	ATG02-02 Arotec Native
RNP-68k	100	ATR04-02 Arotec Native
Sm Smith	100	ATS01-02 Arotec Native
Scl-70 Scleroderma-70	100	ATS01-02 Arotec Native
Azurocidin	100	ATA01-02 Arotec Native
AGC1	100	H0000176-Q01 Abnova - Recombinant
PL-7	100	15600 Diarect BBI Solutions Recombinant
PL-12	100	15700 Diarect BBI Solutions Recombinant
CK-18 Cytokeratin 18	100	10-663-45633 Genway Recombinant
HMGCR	50	H00003156-P01 Abnova Recombinant
TcdA	50	CDA-TNL-50 Native Antigen Company <i>Clostridium difficile</i> Toxin A
TcdB	50	CDB-TNL-50 Native Antigen Company <i>Clostridium difficile</i> Toxin B
dsDNA double-stranded Deoxyribonucleic acid	100	Diarect 12301 Recombinant
<i>Helicobacter pylori</i> (H. pylori) Cag A	100	Prospec pro-2494-a <i>Helicobacter pylori</i> Cag A Recombinant
rhDecorin	100	R&D 143-DE Recombinant human Decorin
Mannan <i>Saccharomyces cerevisiae</i>	100	Sigma M7504
Laminarin <i>Laminaria digitata</i>	100	Sigma L9634

SSA-60/Ro	100	US Biological S6600-7SA Recombinant
Tetanus toxoid	100	NIBSC 08/218
Mannan <i>Candida albicans</i>	100	NIBSC 77/600
Flagellin <i>S. typhimurium</i>	100	Invitrogen FST 33-01 Flagellin isolated from <i>Salmonella typhimurium</i>
Peptidoglycan <i>S. aureus</i>	100	Sigma 77140-10MG From <i>Staphylococcus aureus</i> cell wall
Elastin from human lung	100	Sigma E7152-2MG
Keratin from Human Epidermis	100	Sigma K0253-1MG
Collagen from human lung	50	Sigma C5983-1MG
TPO Human (Thyroid peroxidase)	100	Prospec enz-285 Recombinant
Thyroglobulin Human thyroid	100	Sigma-Aldrich, 609312
Insulin Human	100	Sigma-Aldrich, 10908
Recombinant Human Transglutaminase 2 (rhTGM2)	100	R&D Systems 4376-TG-050
Gliadin	100	Sigma-Aldrich, G3375
CD20	100	Sino Biological Inc. 11007- H34E
Carbamylated-BSA (BSA – Bovine serum albumin)	100	Made In-house (Refer to methods 3.3)

Table 2.5. Viral recall antigens. Including antigen concentrations and suppliers.

Viral Antigen:	Concentration on array: µg/ml	Supplier:
Parvovirus VP2	100	PV-VP2-100 Native Antigen Recombinant protein VP2 from human Parvovirus B19
Native adenovirus	100	0152-0507 BIO-RAD Purified protein
Influenza A H1N1	100	30R-AI038 Fitzgerald Inactivated antigen Strain A/Taiwan/1/86
CMV Cytomegalovirus	100	The Native Antigen Company Human
Rubeola	100	7604 Meridian Life Science Inc.
HSV1 Herpes simplex virus 1	100	7305 Meridian Life Science Inc.
HSV2 Herpes simplex virus 2	100	7705 Meridian Life Science Inc.
VZV Varicella zoster virus	100	7740 Meridian Life Science Inc.
Mumps	100	8099 Meridian Life Science Inc.
HEV mosaic Hepatitis E	100	hev-235-a Prospec
EBV Epstein–Barr virus	50	EBV-VCA 8202 Meridian Life Science Inc
EBV / VCA	50	Meridian Life Sciences 8202
Rubella	100	6075 Meridian Life Science Inc.
HDV Hepatitis D	100	Prospec hdv-234-a
HCoV-HKU1	50	Sino Biological Inc. 40021-V08H Human coronavirus HKU1 (isolate N1) (HCoV-HKU1) Spike/S1 Protein (S1 Subunit, His Tag)
HCoV-229E	50	Sino Biological Inc. 40601-V08H Human coronavirus (HCoV- 229E) Spike/S1 Protein (S1 Subunit, His Tag)
HCoV-OC43	50	Sino Biological Inc. 40607- V08H1 Human coronavirus (HCoV- OC43) Spike S1 Protein (His Tag)
SARS-CoV-2 Spike S1 (Wu)	50	Sino Biological Inc. 40591-V08H SARS-CoV-2 (2019-nCoV) Spike S1-His Recombinant Protein
SARS-CoV-2 Nucleocapsid (Wu)	50	Sino Biological Inc. 40588 – V08B Nucleocapsid-His Recombinant Protein

Table 2.6 Data on the seroprevalence of viral pathogens in the general population.

Viral Pathogens:	Seroprevalence data in the general population:	Reference:
EBV	94.7%	(Mentzer <i>et al.</i> , 2022)
VZV	92.5%	(Mentzer <i>et al.</i> , 2022)
Parvovirus	40-60% of adults 85% of the elderly	(Marano G <i>et al.</i> , 2015)
HSV1	69.8%	(Mentzer <i>et al.</i> , 2022)
HCoV	>90% of adults for at least 1 HCoV species	(Ogimi <i>et al.</i> , 2020)
HCMV	58.2%	(Mentzer <i>et al.</i> , 2022)

3 Methods

3.1 Samples

Serum samples were obtained from 63 patients with CDI, 24 patients with rCDI before and after FMT, 11 patients with CF and 30 age and gender-matched healthy controls (Table 3.1), all acquired from Dr Tanya Monaghan (supervisor). Patient clinical characteristics are outlined in Tables 4.2 and 4.3 on page 93. Power calculations were not considered as this was a proof of concept study with a known limited sample set.

The CF and CDI sera were historical samples stored since 2009 in a -80°C freezer, which had not been subjected to freeze-thaw. Serum samples were prepared from venous blood samples (prior to this study), and all sample aliquots were stored at -80°C and processed with standard procedures. For the rCDI (FMT) cohort, the ethics were approved by the Health Research Ethics Board of Alberta (HREBA) Biomedical panel, Pro00130973. The CDI and CF cohorts were approved for use by the Nottingham Research Ethics Committee 1, REC Q1020310.

Table 3.1 Subject participant groups, sample numbers and storage conditions.

Subject groups:	Sample number (n):	Sample storage conditions:
CDI	63	-80°C
rCDI	24	-80°C
CF	11	-80°C
HC (no known health conditions)	30	-80°C

Clostridium difficile infection, CDI; recurrent *Clostridium difficile* infection, rCDI; Cystic fibrosis, CF; Healthy control, HC.

3.2 SDS-PAGE

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions was performed to validate antigen molecular weights before their addition onto the microarray platform. Novex™ 12% Tris-Glycine plus 1.0 mm gels (6-200 kDa) were used with Midi Gel adaptors in a 20-well electrophoresis chamber (Criterion™ Cell). Samples were prepared to 0.5µg in 25% v/v of 4X SDS Page sample buffer (LI-COR, C90517-05) with DTT (20mM) and made up to 20µL with distilled water. The prepared antigens were heated at 90°C for 2 minutes and run alongside a 7-175kDa Broad range protein standard (Biolabs, P7708S). The total loading volume for each well was 15µL. The gel was run in 1X Tris-Glycine Running buffer for approximately 2 hours at 125V/25mA (starting voltage = 116V). Gels were stained with 1% w/v of Coomassie Brilliant Blue in 50% v/v of methanol, 10% v/v acetic acid and made up with distilled water to 50ml with rocking overnight. The gels were de-stained with 50% v/v methanol, 10% v/v acetic acid and 40% v/v distilled water to 50mL with rocking and draining/re-addition every 30 minutes until the gels appeared clear. The gels were photographed under visible light using a NuGenius (Syngene™) gel imaging system.

3.3 Carbamylated BSA Production

Chemically carbamylated-bovine serum albumin (BSA) was produced to provide a target for anti-carbamylated protein antibodies. First, Urea (Sigma, U5128) was made up to 8M in distilled water and mixed with 3.3% v/v of 30%

BSA (Sigma-Aldrich, A7284) and 2% v/v of 10X PBS to a total volume of 100 μ L. The solution was vortexed and placed in a 50°C oven for 96 hours. The resulting carbamylated-BSA was concentrated to 1.47mg/ml in 1X PBS using a Microcon-30kDa Centrifugal Filter Unit with Ultracel-30 membrane (Merck Millipore Ltd., MRCFOR030) to remove the urea. The carbamylated-BSA solution was mixed with 75% v/v of 1X PBS to a total volume of 400 μ L, which was transferred onto the Ultracel-30 membrane. The solution was centrifuged at 10,000g for 1 minute 20 seconds and repeated thrice. After every centrifugation, 1X PBS equal to the volume of liquid collected was added to remove and replace the urea. The filter was inverted onto a new tube and centrifuged at 10,000g for 10 seconds to recover the final concentrated carbamylated-BSA.

3.4 Printing

The complete range of antigens selected from and printed are displayed with suppliers and final prepared concentrations in Tables 2.4 and 2.5. Most antigens were diluted to 100 μ g/ml in 75% v/v of 2X printing buffer (2:1 4M Betaine, 50% w/v 2-3 Butanediol), supplemented with distilled water to make a 50 μ L total volume. Antigens at a lower stock concentration (below 200 μ g/ml) were prepared to 50 μ g/ml in 75% v/v of 2X print buffer and made up with distilled water to 50 μ L total volume. A proportion of the prepared antigen solutions were transferred to a 384-well V-bottom polypropylene microarray plate (Thermo-Fisher™) and printed in duplicate onto slides using a BioRobotics Microgrid II with solid pins. The target humidity in the array

chamber was set to 65% during printing. The distance between the spots was set to 0.4mm in a 13x13 format, the pins were set to 2 strikes per spot, the target height printed was 0.5mm, and the dwell time was limited to 0.25s. If left printing overnight, slides remained in the chamber with the set humidity until removal the next day. Slides were kept in dark slide boxes at room temperature in a vacuum desiccator and processed within a week.

3.5 Slide Surfaces

To ascertain a slide surface chemistry which would produce optimal signal intensities and spot morphologies, intravenous immunoglobulin (IVIG; Viagam® Liquid) was serially diluted 2-fold from 100µg/ml across 16 dilutions in 2X print buffer. IVIG is manufactured from pools of plasma from thousands of healthy donors, providing a suitable baseline for testing signal intensity outputs from a range of dilutions. The 16-step IVIG serial dilution was printed in 24 repeated blocks onto six slide surfaces (Table 2.2). Using the cassette cover plate format displayed in Figure 8, the six slides were washed 4X with PBST (480µL per wash) and then incubated with 3% BSA-PBS (240µL) for 1 hour. Once rewashed 4X, the slide arrays were probed with a goat anti-human IgG-cy5 conjugate (0.975 mg/ml, DOL 3.54; made In-house) diluted 1:1000 in PBST (240µL) for 1 hour. Following the final 4X washes, the slide arrays were rinsed briefly in distilled water and spun dry using a microarray slide centrifuge (Labnet™). The slides were scanned with the GenePixPro 4000B microarray scanner at 635nm (Axon GenePix®).

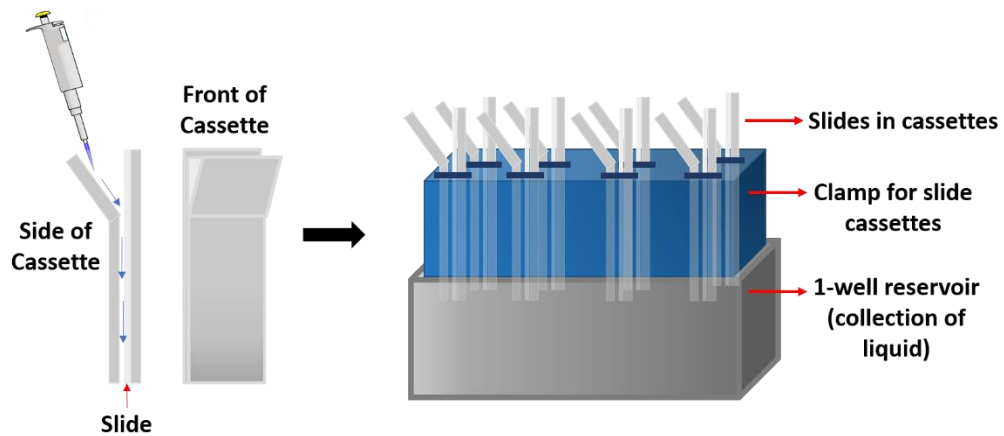


Figure 8. The cassette cover plate format.

The small gap between the slide and the cassette allowed reagents to run evenly over the surface substrate and prevented the slides from drying out. A clamp held the cassettes in place, and the liquid from each step was collected in the 1-well reservoir. The OT-2 robot protocol was programmed to dispense liquid in the top opening of the cassettes with a delay to prevent dripping. (Thermo Scientific™, Shandon™, Plastic Coverplates).

3.6 Serum Dilutions

A negative control experiment was conducted where normal control serum was used to probe antigen arrays at a range of dilutions to assess the specificity of signal intensity, the degree of non-specific binding and background intensity. An array of antigens (Tables 2.4 and 2.5) were printed in duplicate onto three nitrocellulose-coated slides using the above printing conditions. The assay protocol was conducted using an Opentrons OT-2 liquid handling robot in the cassette format displayed in Figure 8. The slides were washed 5X with PBST (480 μ L) and incubated with 3% BSA-PBS (240 μ L) for 1 hour. After another set of 5X washes, the arrays were incubated with normal-control sera (negative control) diluted 1:40, 1:80 and 1:160 in 3% BSA-PBS (240 μ L) for 1 hour. The slides were washed again and then incubated with an anti-human IgG-cy5 conjugate diluted 1:1000 in PBST (240 μ L) for 1 hour. After

a final round of 5X washes, the slides were rinsed briefly in distilled water, spun dry using a microarray slide centrifuge (Labnet™) and then scanned at 635nm with the GenePixPro 4000B microarray (Axon GenePix®).

3.7 Optimising Assay Reagents

The printing alignment was altered so that ten subarrays were printed per nitrocellulose slide (16-subarray max format), with Tetanus toxoid, *Candida albicans* and CMV antigens printed in duplicate. The pattern grid, spacing, strikes, target height, dwell time and humidity remained the same as the printing protocol stated above. Each slide was placed in a 16-multi-well slide cover (ProPlate®, Grace Bio-Labs USA) to be probed, exhibited in Figure 9.

This 16-subarray format was used for all assay reagent optimisation experiments. Regarding volumes, all subarray wells were washed 4X with 200µL of PBST per wash and 150µL of serum and assay reagents per well for each incubation step (consistent for all subsequent experiments).

3.7.1 Secondary Antibody Optimisation

The printed slide comprising ten subarrays was washed 4X with PBST and then blocked with 3% BSA-PBS for 1 hour. Following another set of 4X washes, the five subarrays in column 2 were incubated with serum diluted to 1:160 in 3% BSA-PBS for 1 hour. The other five subarrays in column 1 were incubated with 3% BSA-PBS and no serum (negatives), as displayed in Figure 9. After another round of 4X washes, all slide chambers were incubated for 1 hour with a range of biotinylated anti-human IgG dilutions in 3% BSA-PBS at

1:1000, 1:5000, 1:10,000, 1:20,000 and 1:40,000 in the format displayed in Figure 9. Once rewashed, the subarrays were probed with streptavidin-cy5 diluted 1:10,000 in PBST and incubated for 1 hour. After another 4X washes, the slide was rinsed with distilled water, spun dry and scanned in a 4000B microarray scanner (Axon GenePix®).

This method was repeated with identical arrays tested with an IgG-cy5 conjugate at the same dilutions as the biotinylated anti-human IgG, as displayed in Figure 9. As the antibody was conjugated to the far-red fluorescent dye cy5, the step with the addition of streptavidin-cy5 was removed and replaced with just the addition of PBST. The rest of the steps remained the same.

3.7.2 Streptavidin-cy5 Optimisation

The ten subarrays were washed 4X with PBST and then blocked with 3% BSA-PBS for 1 hour. After washing again, the five subarrays in column 2 were incubated with serum diluted 1:160 in 3% BSA-PBS for 1 hour. The five subarrays in column 1 were incubated with only 3% BSA-PBS and no serum, which were used as negatives, displayed in Figure 9. Upon completion of another four washes, the arrays were incubated with biotinylated anti-human IgG diluted 1:5000 in 3% BSA-PBS. After rewashing 4X, the subarrays were incubated for 1 hour with a range of streptavidin-cy5 dilutions in PBST, including 1:2500, 1:5000, 1:10,000, 1:20,000 and 1:40,000 in the format displayed in Figure 9. After another 4X washes, the slide was rinsed, spun dry and scanned in a 4000B microarray scanner (Axon GenePix®).

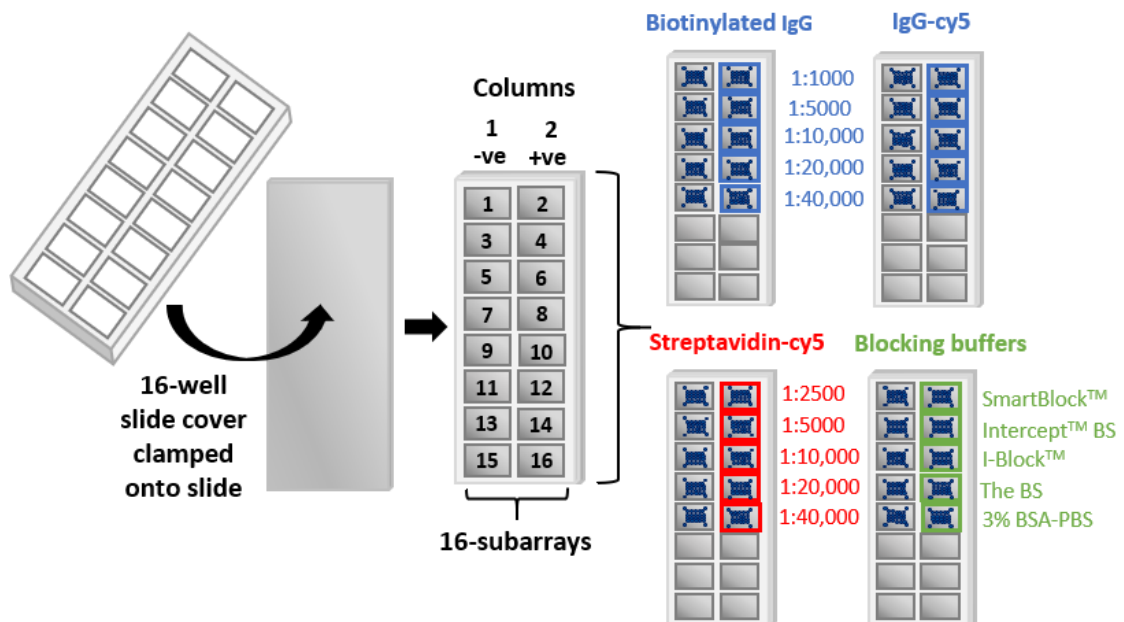


Figure 9. The layout of the assay reagent optimisation experiments.

The reagents tested included biotinylated anti-human IgG, anti-human IgG-cy5 conjugate and streptavidin-cy5, all tested at five dilutions. Five blocking buffers were also tested. Ten identical arrays containing Tetanus toxoid, *Candida albicans* and CMV duplicates were printed onto four slides orientated to fit inside 16-multi-well slide covers (ProPlate®, Grace Bio-Labs USA). Column 1 blocks were negative chambers (-ve), meaning only blocking buffer was added at the serum incubation step. Column 2 blocks were positive chambers (+ve) with the addition of serum diluted 1:160 in blocking buffer. Each row of the slide chambers was treated with separate conditions highlighted in the Figure. A volume of 150µL was added per well for each serum and assay reagent incubation. The BS; The Blocking Solution.

3.7.3 Blocking Buffer Optimisation

The printed slide was washed 4X with PBST and then blocked for 1 hour with shaking in 3% BSA-PBS, The Blocking solution, I-Block™, Intercept™ Blocking buffer and SmartBlock™ in the format displayed in Figure 9. After another four washes, the five subarrays in column 2 were incubated with serum diluted 1:160 in the corresponding blocking buffer for 1 hour with shaking. The five subarrays in column 1 were incubated with only the corresponding blocking buffer and no serum (negative controls). The slide was rewashed 4X

and then incubated for 1 hour with biotinylated anti-human IgG diluted to 1:5000 in the relevant blocking buffer with shaking. After another four washes, the slide was probed with streptavidin-cy5 diluted 1:10,000 in PBST for 20 minutes with shaking. Lastly, after another 4X washes, the slide was rinsed in distilled water, spun dry and scanned in a 4000B microarray scanner (Axon GenePix®).

These optimisation experiments were done alongside a fellow MRes student, Niamh Thomas.

3.8 Final Optimised Microarray Protocol

The printing protocol slide layout was altered to accommodate all antigens when printing 16 arrays per slide. All other printing conditions remained the same as previously stated. The humidity in the chamber remained set at 65%, and slides were stored the same as previously stated.

Printed arrays were placed in 16-multi-well slide covers, washed 4X with PBST (200µL per well) with shaking and blocked with 150µL of I-Block™ per array for 1 hour with shaking. Following another round of four washes, 15 subarrays were incubated with 150µL (per array) of serum samples diluted to 1:160 in I-Block™ for 1 hour with shaking. The remaining subarray was incubated with only 150µL of I-Block™ and no serum; this was used as a negative control. Slides were washed as above and then probed with 150µL (per array) of biotinylated anti-human IgG (Vector Laboratories, BA-3080) diluted 1:5000 in I-Block™ with 3% v/v of goat serum (Sigma, G6767) for 1

hour with shaking. Slides were rewashed, followed by the addition of 150 μ L (per array) of streptavidin-cy5 (Cytiva, PA45001) diluted 1:10,000 in PBST and incubated for 20 minutes with shaking. After a final four washes, the arrays were rinsed in distilled water and spun dry before scanning with the GenePixPro 4000B microarray scanner at 635nm (635 PMT Gain = 500, Laser Power = 33%, Axon GenePix[®]). All incubation steps were conducted at room temperature, and all reagents were allowed to equilibrate to room temperature before use.

When testing IgA responses, the IgG antibody was replaced with an incubation of 150 μ l (per array) of biotinylated goat anti-human IgA (BioSource, AHI1109) diluted to 1:2500 in I-Block[™] with 3% v/v of goat serum for 1 hour with shaking. All other steps remained the same.

3.9 Microarray Data Analysis

The GenePix Pro analysis software version 6.1.0.4 (Axon GenePix) was used to align a template GenePix array list (GAL) file to the scanned images to generate GenePix Result (GPR) files. The median foreground intensity of each feature (F635nm) was normalised by the subtraction of the median local background (B635nm) to determine the corrected signal intensity (F635nm – B635nm); this was conducted in J-Express Pro 2012 (Norwegian Microarray Consortium). For the optimisation of assay reagent experiments, the signal-to-background ratio (S/B ratio) was calculated by dividing the median fluorescent signal (F635nm) by the median background signal (B635nm) to

establish signal strength relative to the background. All optimisation and validation graphs were made using GraphPad Prism 9.4.0.

3.10 Co-efficient of Variation: Intra- and Inter-assay

Intra- and inter-assays were conducted to evaluate the repeatability and precision of the experimental microarray platform for detecting primary antibodies in sera. Intra-assays involved repeating array probing with the same sera on the same slide on the same day, while inter-assays were performed with a 3-day gap between replicated experiments to assess slide-to-slide consistency. A total of 60 antigens were tested, and the inter- and intra-assay coefficient of variation % (CV) was calculated using the standard deviation and mean values of the replicates. This experiment was designed to ensure the accuracy and consistency of the microarray platform and can provide valuable information on the precision of the assay, which is critical for reliable results.

3.11 Statistical Analysis

Heatmaps and hierarchical clustering (Euclidean distance) of patient antibody response profiles were conducted in Instant Clue (Nolte *et al.*, 2018). All ensuing statistical analysis was conducted in R version 4.3.1 utilising the R Studio IDE Version 2023.6.1.524 (Posit team, 2023). First, to assess the impact of patient status (Pre-FMT, Post-FMT and healthy control) on antigen-specific antibody responses, a Generalised Linear Model (GLM) with Gaussian distribution was fitted to the data. Initially, a fully saturated model was fitted,

including fixed effects of antigens and patient status and an interaction term between these two variables to assess the impact on the mean of the response variable (antigen-specific antibody responses). To determine an optimal model, step-wise deletion of fixed effects was conducted, and the resulting Akaike Information Criterion (AIC) was assessed by ranking all combinations. The final and best-fitting model was the fully saturated model previously described. Subsets of antigens (autoantigens, positive-control, pathogen-associated and viral recall antigens) were modelled separately. To analyse the significance of the data, ANOVA type II tests were run on the GLM models using the Anova function from the R package "car" (Fox and Weisberg, 2019). Pairwise comparisons using Tukey's p-value adjustment were conducted using the R package "emmeans" (Lenth, 2023). The R package "ggplot2" (Wickham, 2016) was used to plot and visualise the data. Subsequently, the same process was repeated but fitted onto another dataset to assess the impact of patient condition (CDI vs. rCDI) on antigen-specific antibody responses.

4 Results

4.1 Generation of The Multi-antigen Microarray Platform

SDS-PAGE gels were used to validate the presence and size of antigens before their addition onto a 384-well microarray plate for printing. Two SDS-PAGE gels were conducted to compare the antigen molecular weights to those the manufacturer stated, as shown in Figure 10. Overall, 29 antigens previously stored in a -20°C freezer for over six months were selected for validation due to the potential for degradation or modifications (oxidation, denaturation, aggregation). Antigens stored for a short period in a -80°C freezer or recently purchased were excluded from the validation process due to a more stable storage environment and recent validation by the producer. A 7-175 kDa broad-range protein standard was used to ensure that a wide range of molecular weights could be estimated accurately. Of the 29 antigens, 24 showed bands at their expected molecular weights (Figure 10, A, B) and were added to the 384-well plate. Notably, the histone antigen showed no band present (Figure 10A), although this may have been due to migration off the bottom of the gel.

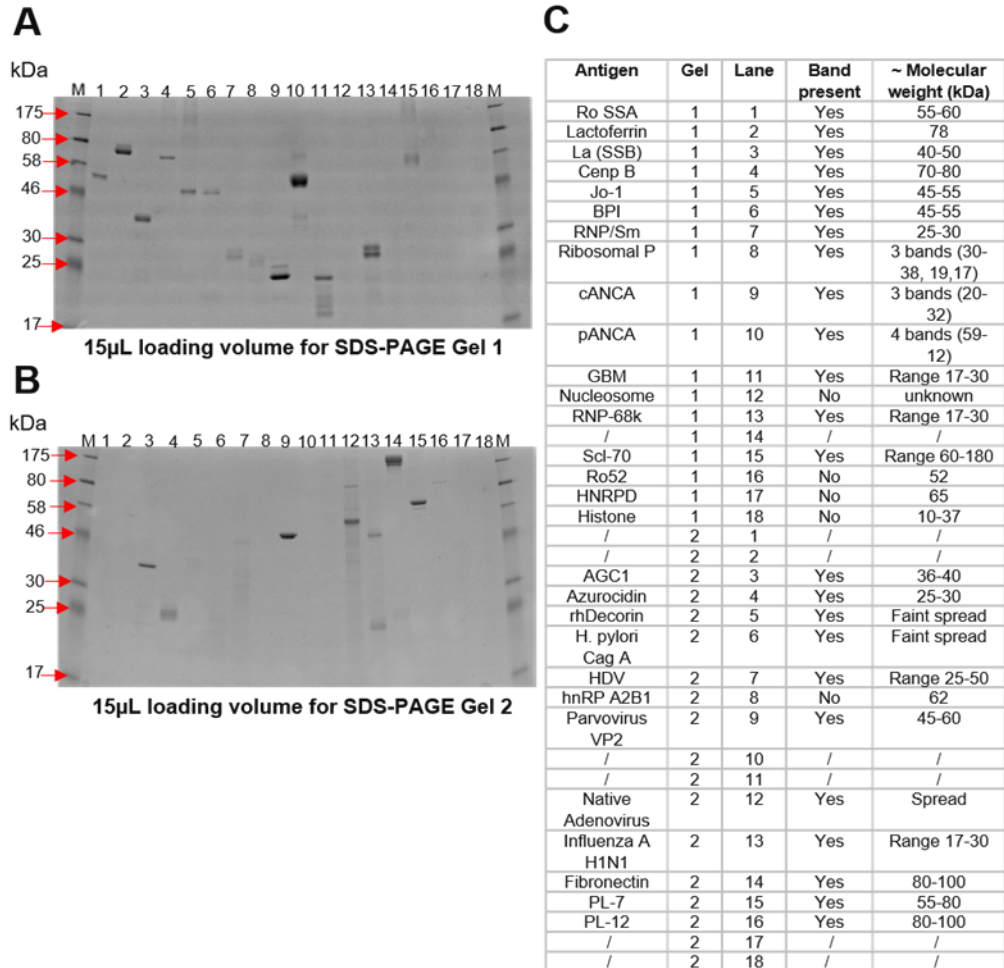


Figure 10. SDS-PAGE gels for antigen molecular weight validation before microarray selection.

Panels A and B show 12% Tris-Glycine gels (NovexTM, 6-200 kDa) that were used with Midi Gel adaptors in a 20-well electrophoresis chamber (CriterionTM Cell) to validate antigen molecular weights, under reducing conditions. Bands that were not present when expected were excluded from the microarray selection. Lanes labelled M indicate the 7-175 kDa broad range protein standard (Biolabs, P7708S), for molecular weight estimation. The numbered lanes of Gel 1 and 2 correspond to antigens displayed in panel C. Gels were stained with Coomassie Brilliant Blue and imaged using the Nu Genius (Syngene) gel documentation system. Notably, due to the size range over which the gel was optimal the histone antigen may have migrated off the end of the gel.

After SDS PAGE validation, the 60 selected antigens were prepared to their specified final concentrations in the print buffer and arrayed onto microarray slides. Figure 11 displays the major steps involved in generating and processing the multi-antigen microarray protocol. The array probing and scanning steps required optimisation to determine a final assay protocol before testing patient samples.

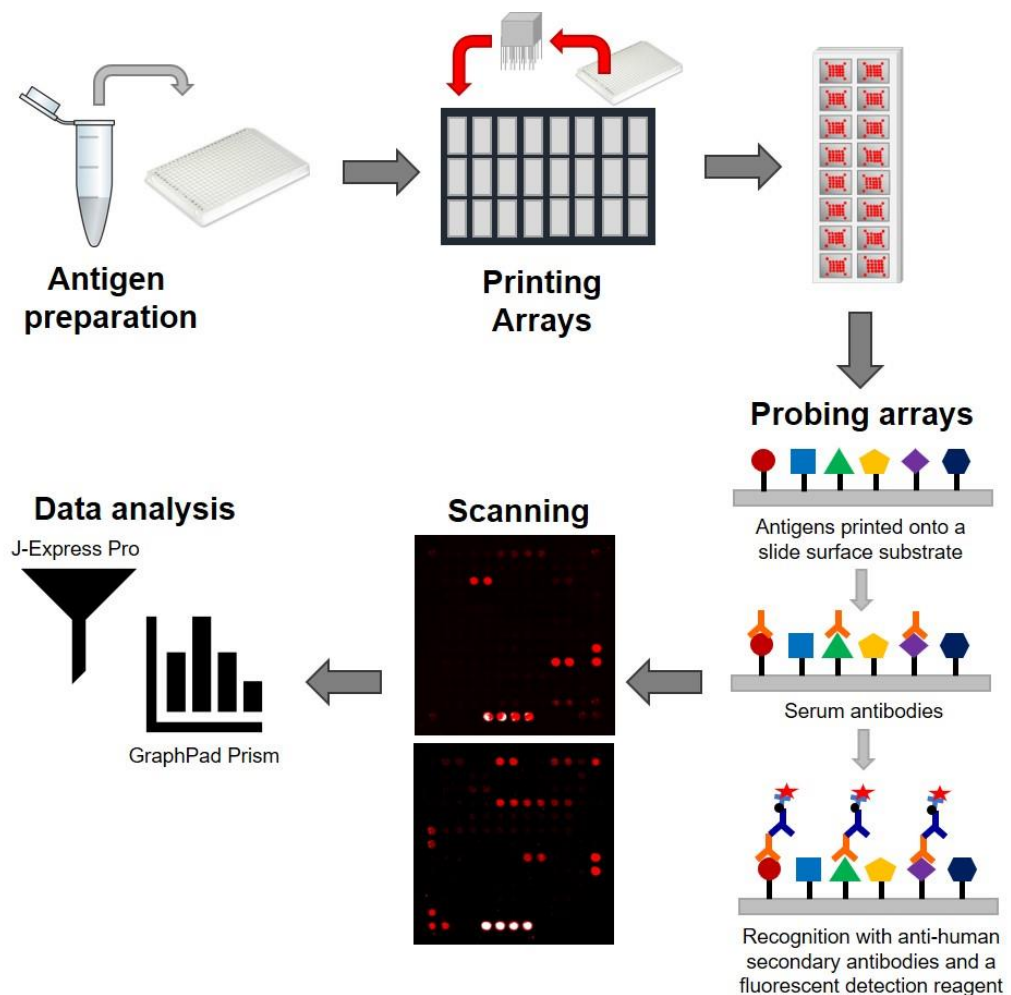


Figure 11. Schematic summarising the major steps involved in generating and processing multi-antigen microarrays.

Initially validated antigens and controls are prepared in a print buffer to their chosen dilutions and transferred to a 384-well microarray plate for printing in duplicate on microarray slides. Slides are subsequently blocked, incubated with patient sera, and probed with a secondary antibody and fluorescent detection reagent. Arrays are washed at the end of each incubation, spun dry before scanning, and processed with a microarray image analysis program.

4.2 Assay Development and Optimisation

4.2.1 Slide Surface Assessment

Choosing an optimal slide surface and serum dilution is crucial for developing a microarray platform to obtain high signals and uniform spot morphology. To evaluate various slide surfaces, a comprehensive comparison was conducted using a two-fold serial dilution of 16 intravenous immunoglobulin (IVIG) concentrations, ranging from 100 μ g/ml to 0.003 μ g/ml. These IVIG concentrations were printed in 24 repeated blocks on six different slides, including aldehyde PLUS, aldehyde made in-house, nitrocellulose ONCYTE[®] AVID, nitrocellulose made in-house, epoxy silane made in-house, and BSA-NHS made in-house. Each slide was blocked, probed with an anti-human IgG-cy5 conjugate and then scanned at 635nm to compare signal intensities.

Figure 12 consists of bar graphs displaying the signal intensities produced by each slide surface across the 16 IVIG dilutions. The commercial (Figure 12C) and lab-made in-house (Figure 12D) nitrocellulose slides exhibited the highest signal intensities at levels equal to or above 20,000 AFUs and signals detectable across all IVIG concentrations. Notably, the commercial nitrocellulose slide displayed significantly greater signal intensities than the lab-made in-house slide, with a difference of approximately 40,000 AFUs observed at an IVIG concentration of 100 μ g/ml. The aldehyde (Figure 12A, B), epoxy silane (Figure 12E), and BSA-NHS (Figure 12F) slides failed to achieve signal intensities above 4,000 AFUs and were unable to detect signal

intensities below 0.781 $\mu\text{g}/\text{ml}$ IVIG. This indicates a limited capacity to quantify low signals compared to the dynamic range of the nitrocellulose slides.

However, it is important to note that the commercial nitrocellulose slide did not exhibit a consistent decrease in signal intensity as the IVIG concentration decreased. Unexpected sharp declines in signal intensity were observed at IVIG concentrations of 12.500 $\mu\text{g}/\text{ml}$ and 0.012 $\mu\text{g}/\text{ml}$. A similar drop in signal intensity was also observed for the epoxy silane slide (Figure 12E) at the IVIG concentration of 6.260 $\mu\text{g}/\text{ml}$. These drops in signal intensity were attributed to droplet marks on the background of the commercial nitrocellulose and lab-made in-house epoxy silane slides, which hindered accurate signal quantification. To address this issue in future experiments, special care was taken to securely clamp cassettes over the slides during the assay procedure to ensure a uniform flow of liquid over the slide surface, preventing droplets.

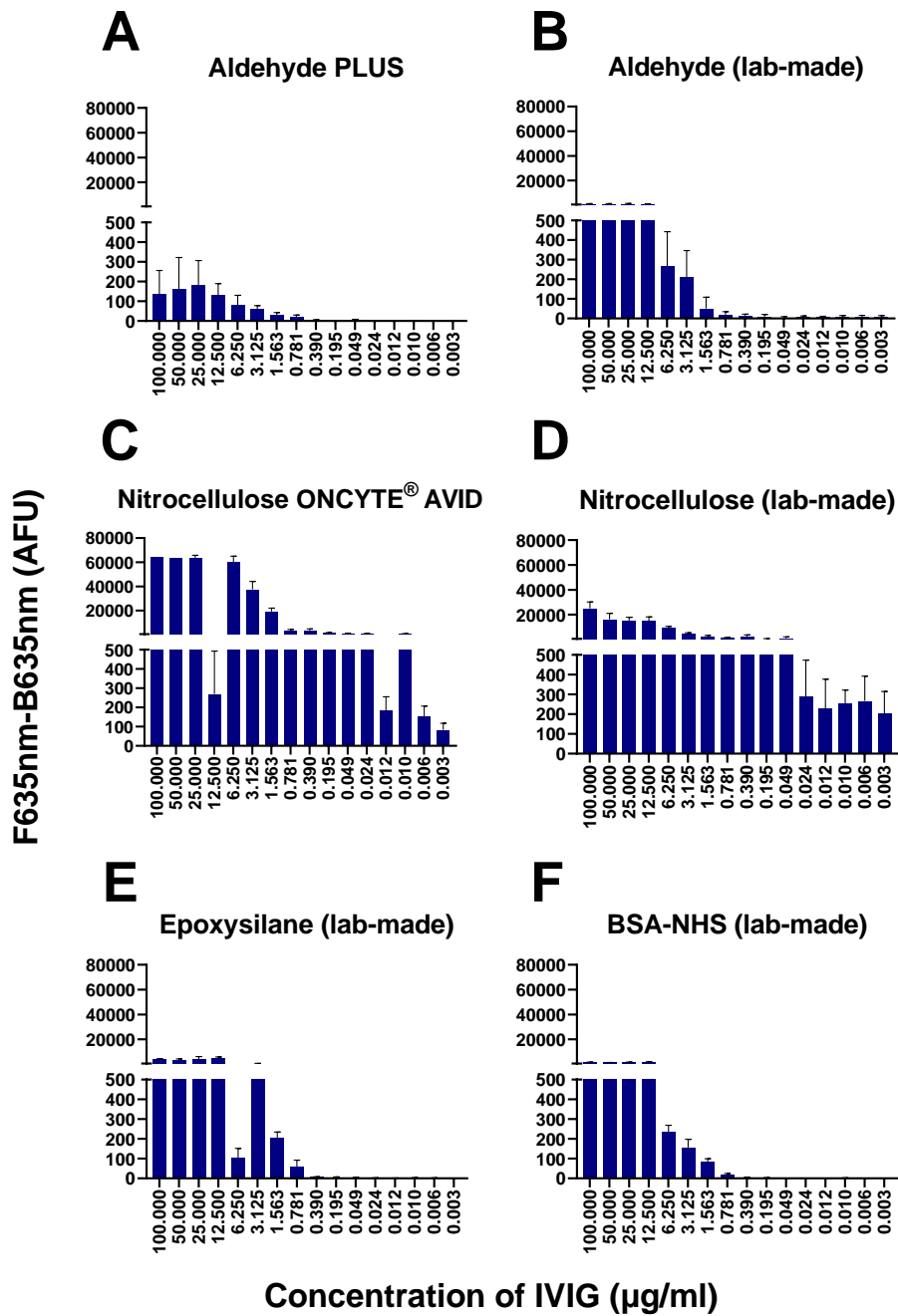


Figure 12. Comparison of signal intensities produced from a serial dilution of IVIG probed on six slide surfaces.

Identical arrays were printed on each slide, comprising 24 repeated blocks of a 16-step IVIG serial dilution (100µg/ml – 0.003µg/ml). All slides were assayed with the same protocol involving incubation with an anti-human IgG-cy5 conjugate after blocking with 3% BSA-PBS. Nitrocellulose slides (C) and (D) exhibited the highest signal intensities and best dynamic range, with signals detectable at 0.003µg/ml IVIG. Means were plotted with error bars representing the standard deviation. N = 24. BSA-NHS; Bovine Serum Albumin-N-hydroxy succinimide, F635nm-B635nm; Foreground at 635nm – Background at 635nm, AFU; Arbitrary fluorescent units.

Alongside signal intensities, the spot morphology produced on different slide surfaces is vital for consistency and accuracy between protein signal detection. Figure 13 compares spot morphologies between aldehyde, nitrocellulose, epoxy silane and BSA-NHS slide surfaces. Nitrocellulose exhibited superior spot morphology regarding circularity, intensity, and uniformity and was consequently selected as the optimal slide surface for our multi-antigen microarray platform.





Slide surface	Spot Morphology	Spot Intensity	Spot variability
Aldehyde		Low	Low
Nitrocellulose		High	Low
Epoxy silane		Moderate	High
BSA-NHS		Moderate	High

Figure 13. Summary of the optimisation tests assessing spot intensities and morphology across four slide surfaces.

Nitrocellulose produced the highest signal intensities with exceptional spot morphology and low spot variability. Aldehyde slides produced the lowest signal intensities but had consistent spot morphology. Epoxy silane slides produced moderate signal intensities but with small, inconsistent protein spots. Lastly, BSA-NHS slides produced moderate signal intensities but inconsistent spot morphology. All antigens were printed in the same print buffer. Spot characteristics were partly made by subjective visual assessments of the scanned images above, particularly regarding spot shape/variability, where consistency and uniformity of the four spots were assessed by eye across the slide surfaces. Spot intensities were determined from the Foreground at 635nm – Background at 635nm results generated from the scanned images, which was the primary decision-making factor when choosing the optimal slide surface.

4.2.2 Serum Dilution

Nitrocellulose slides were identified as the ideal surface for printing our multi-antigen microarrays. The next step was to identify the optimal serum dilution for our protocol. As a result, three identical arrays were printed onto nitrocellulose slides and probed with healthy control serum diluted to 1:40, 1:80 and 1:160 in blocking buffer. The slides were first blocked (3% BSA-PBS) before incubation with the relevant serum dilutions, followed by the addition of biotinylated anti-human IgG (secondary antibody) and streptavidin-cy5 (detection reagent).

Figure 14C shows that the positive-control proteins, Tetanus toxoid and *Candida albicans*, consistently displayed high signal intensities exceeding 30,000 AFUs across all serum dilutions. Specifically, the 1:160 serum dilution yielded the highest signal intensity for *Candida albicans*, reaching 45,000 AFUs (Figure 14C). Concerning viral recall antigens, the 1:40 serum dilution exhibited marginally higher signal intensities than the 1:160 and 1:80 dilutions (Figure 14A). Notably, all serum dilutions consistently demonstrated high signal intensities for CMV and the seasonal coronaviruses (HCoV-229E, HCoV-HKU1, HCoV-OC34), indicating strong IgG antibody responses, which is anticipated for healthy control sera (Figure 14A).

Signal intensities were less consistent across serum dilutions for the autoantigens, with higher signal intensities than expected for healthy control serum (Figure 14B). The 1:40 serum dilution resulted in high signal intensities above 10,000 AFUs and large error bars for several autoantigens due to

increased non-specific binding compared to the 1:80 and 1:160 dilutions (Figure 14B). Although only marginal differences were found, the 1:160 serum dilution was selected due to its slightly lower degree of non-specific binding and the advantage of using smaller volumes of sera for array probing. Consequently, the multi-antigen microarray platform utilised lab-made in-house nitrocellulose slides in combination with serum diluted 1:160 in blocking buffer for all subsequent experiments.

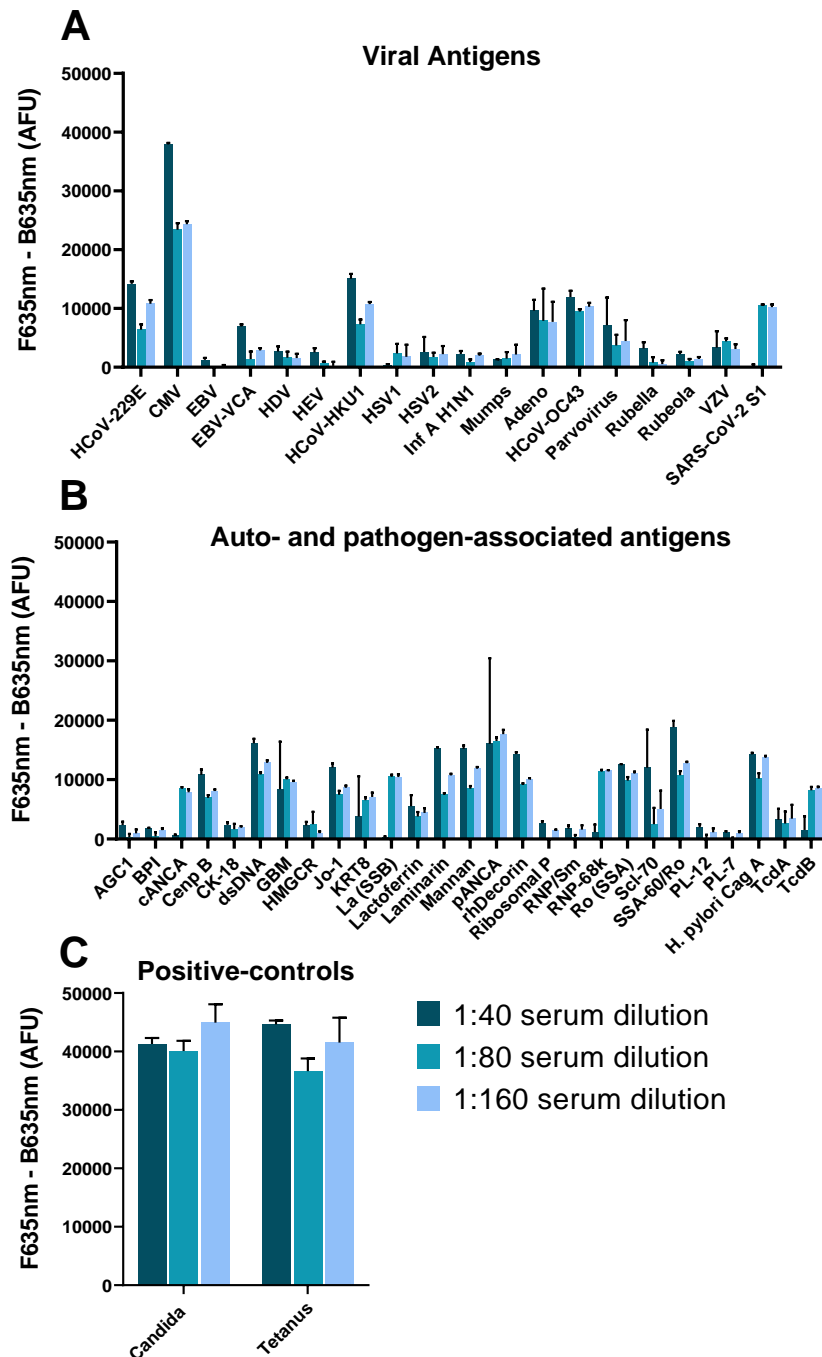


Figure 14. Comparison of signal intensities across a range of serum dilutions on nitrocellulose slides.

Identical printed arrays were tested with serum dilutions (1:40, 1:80, and 1:160) of the same healthy-control sample diluted with 3% BSA-PBS. (A) *Candida albicans* produced the highest signal intensity, observed at 45,000 AFUs for the 1:160 serum dilution. (B) CMV and the seasonal coronaviruses (HCoV-229E, HCoV-HKU1, HCoV-OC34) produced signal intensities over 10,000 AFUs for the 1:40 and 1:160 serum dilutions. (C) Numerous auto-antigens produced signal intensities above 10,000 AFUs, showing a high degree of non-specific binding, particularly at the 1:40 serum dilution, indicating that the 1:160 dilution is marginally favourable. N = 4. Mean values were plotted with error bars portraying standard deviation. F635nm – B635nm; Foreground 635nm – Background 635nm, AFU; Arbitrary fluorescent units.

4.2.3 Nitrocellulose Background Fluorescence

Even though signal intensities were optimal on nitrocellulose slides, we had to consider the high level of non-specific binding and background fluorescence produced when probing our arrays. Figure 15 displays an image scanned at 635nm of the nitrocellulose slide probed with the 1:160 control serum dilution. The scanned image visually demonstrates the high background fluorescence of the probed array. The high background and non-specific binding were attributed to increased protein binding from the control serum to the slide surface or reagents used in the assay, indicating that the assay protocol required optimising.

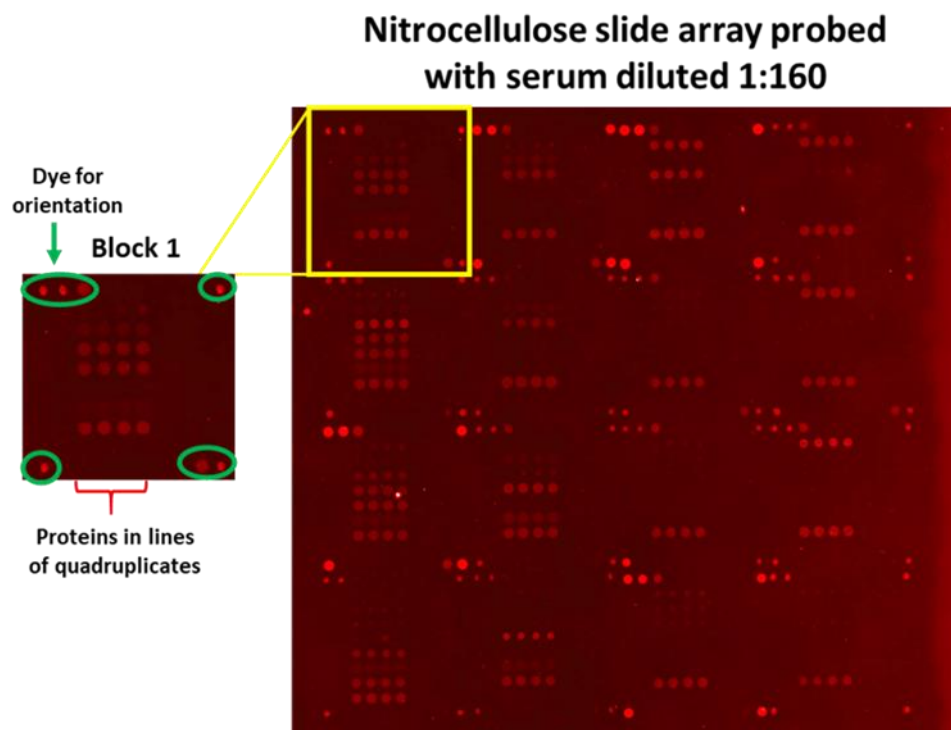


Figure 15. Scanned image at 635nm of an array probed with control serum diluted 1:160 in blocking buffer (3% BSA-PBS) on a nitrocellulose slide.

This Figure demonstrates the high background fluorescence and non-specific binding attributed to the control serum. Despite this issue, the signal intensities (F635nm – B635nm) and spot morphology (visual assessment) appear favourable. The serum used in this experiment was a healthy-control sample (normal individual with no identified health conditions).

4.2.4 Optimisation of Assay Reagents

Optimisation tests were conducted to test modifications to several parameters, including biotinylated anti-human IgG dilution, streptavidin-cy5 dilution, blocking buffers, and scan settings.

4.2.4.1 Secondary Antibody and Detection Reagent Dilution

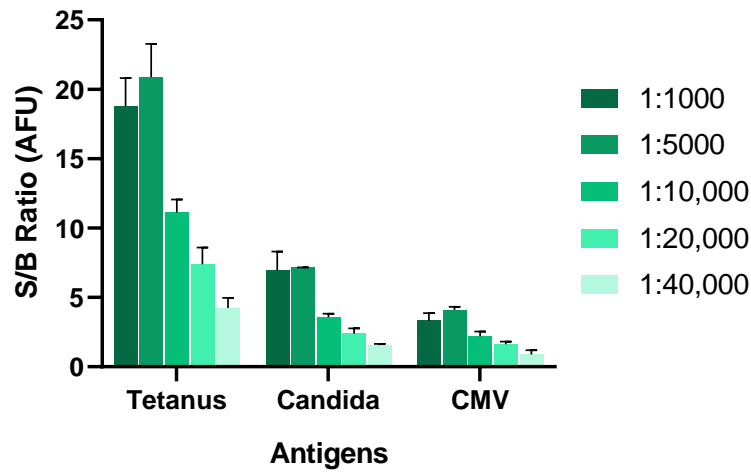
To determine the optimal concentration of secondary antibody and detection reagent, a range of dilutions were tested on arrays of Tetanus toxoid, *Candida albicans* and CMV printed in duplicate. Slides, each with ten printed sub-arrays, were used with 16 multi-well slide covers to test a five-step dilution series, each with a negative control (no added serum). Biotinylated anti-human IgG and an anti-human IgG-cy5 conjugate were tested at dilutions of 1:1000, 1:5000, 1:10,000, 1:20,000 and 1:40,000 in 3% BSA-PBS. Streptavidin-cy5 was tested at dilutions of 1:2500, 1:5000, 1:10,000, 1:20,000 and 1:40,000 in PBST. The aim was to determine which secondary antibody produced the best signal intensities alongside a low background intensity (signal-to-background ratio; S/B ratio) and at which dilution, as well as the best dilution of streptavidin-cy5.

Analysis of the S/B ratios for each secondary antibody dilution revealed that the biotinylated anti-human IgG exhibited consistently higher ratios for most dilutions (Figure 16A). In contrast, the anti-human IgG-cy5 conjugate showed a substantial drop in signal intensity after the 1:1000 dilution, which appeared saturated (Figure 16B). Based on these observations, the biotinylated anti-human IgG was selected as the optimal secondary antibody.

Regarding the best dilution factor for the biotinylated anti-human IgG, the 1:5000 dilution demonstrated the highest S/B ratio across Tetanus, *Candida albicans* and CMV (Figure 16A). The scanned images provide visual evidence that the anti-human IgG 1:5000 dilution (highlighted in green in Figure 16A) had the best balance between signal intensity and background intensity (S/B ratio).

The biotinylated anti-human IgG dilution at 1:5000 was then applied to the assay protocol, testing the optimal streptavidin-cy5 dilution. Figure 17 demonstrates that the highest S/B ratios were for Tetanus and *Candida* at the 1:10,000 streptavidin-cy5 dilution. Regarding CMV, the 1:2500 and 1:5000 dilutions had higher S/B ratios than the 1:10,000 dilution; however, they both showed higher background intensities and spot saturation in the scanned images (Figure 17). The 1:20,000 and 1:40,000 dilutions had a very low background fluorescence, as observed in the scanned images, but with low foreground signal intensities, producing lower S/B ratios (Figure 17). Thus, the streptavidin-cy5 1:10,000 dilution in PBST was chosen for our microarray assay protocol.

A Optimisation of the biotinylated anti-human IgG dilution



B Optimisation of the anti-human IgG-cy5 dilution

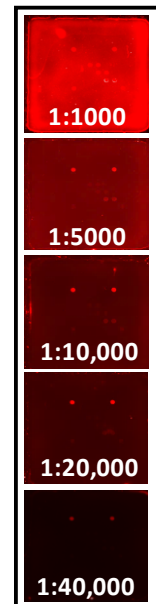
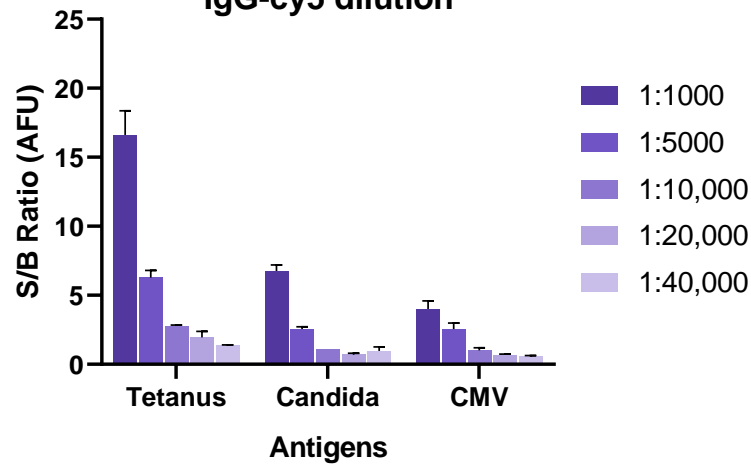


Figure 16. Comparison of the S/B ratios of two secondary antibodies when probed across a range of dilutions on nitrocellulose slides.

(A) Dilution optimisation for goat biotinylated anti-human IgG in 3% BSA-PBS with the subsequent addition of streptavidin-cy5 diluted 1:10,000 in PBST. (B) Dilution optimisation for goat anti-human IgG-cy5 conjugate in 3% BSA-PBS with the subsequent addition of PBST. All other assay steps remained the same. The images scanned at 635nm of each subarray are shown to the right of the relevant graphs. This Figure shows that the biotinylated anti-human IgG at the 1:5000 dilution had the highest S/B ratios and was chosen as the optimal secondary antibody dilution; the green box highlights the 1:5000 dilution image. N = 2. Means were plotted with error bars representing the standard deviation. S/B Ratio; signal-to-background ratio, AFU; Arbitrary fluorescent units.

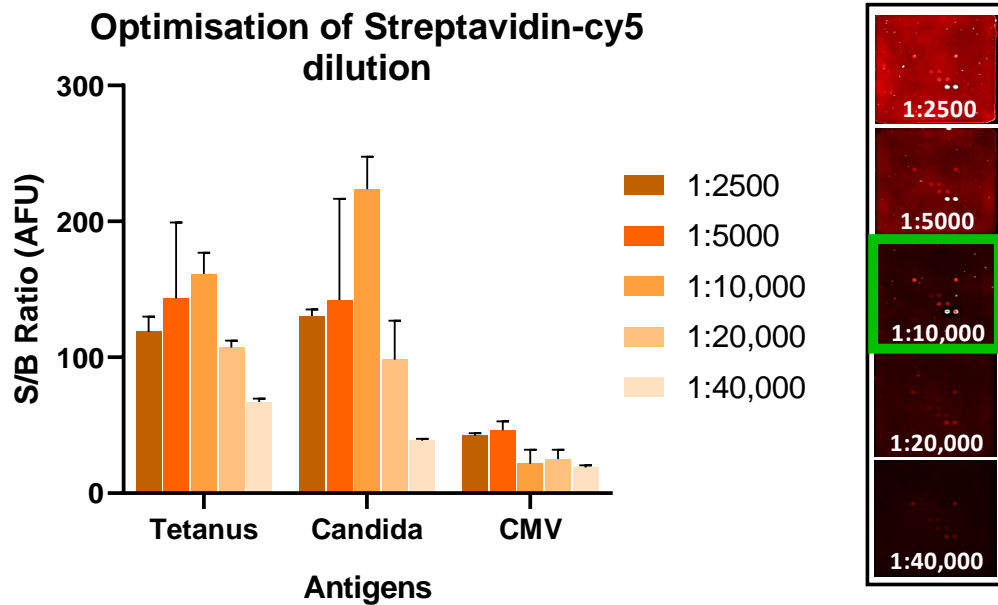


Figure 17. Comparison of the S/B ratios of streptavidin-cy5 when probed across a range of dilutions on a nitrocellulose slide.

Dilution optimisation for streptavidin-cy5 in PBST subsequent to biotinylated anti-human IgG diluted 1:5000 in 3% BSA-PBS. All other assay steps remained the same. This Figure shows that streptavidin-cy5 at the 1:10,000 dilution showed the highest S/B ratios for Tetanus and Candida and the best visual balance between signal and background (image highlighted by the green box). N = 2. Means were plotted with error bars representing the standard deviation. S/B Ratio; signal-to-background ratio, AFU; Arbitrary fluorescent units.

4.2.4.2 Blocking Buffers

Next, a range of blocking buffers was tested to achieve an ideal low background intensity. The assay implemented the optimised secondary antibody dilution (1:5000) and detection reagent dilution (1:10,000). The difference between the protocols was the blocking buffer used in the initial blocking step, as the serum diluent and the secondary antibody diluent. The blocking buffers tested were 3% BSA-PBS, The Blocking Solution, I-Block™, Intercept™ Blocking Solution and SmartBlock™. The blocking buffers were

tested in the same 16-multi-well chamber format as the secondary antibody and detection reagent optimisation tests.

Figure 18A shows that The Blocking solution had the lowest median background intensity at less than 2000 AFUs, followed by I-Block™ and 3% BSA-PBS around 2000 AFU, then Intercept™ Blocking Solution at approximately 3000 AFUs and lastly, SmartBlock™ with a much higher background intensity at approximately 8000 AFUs. The results in Figure 18A correspond with the background intensities visually depicted in Figure 18C, displaying images scanned at 635nm of the sub-arrays probed with each blocking buffer. Figure 18B identifies I-Block™ as having the highest median S/B ratio, followed closely by 3% BSA-PBS. The Blocking Solution, Intercept™ Blocking Solution and SmartBlock™ had the lowest S/B ratios under 1 AFU (Figure 18B). The data presented in Figure 18B corresponds with the scanned images in Figure 18C, which shows Tetanus had the highest signal intensity across all arrays, reaching saturation for SmartBlock™. Overall, we decided that either I-Block™ or 3% BSA-PBS would be a suitable choice of blocking buffer for further experiments.

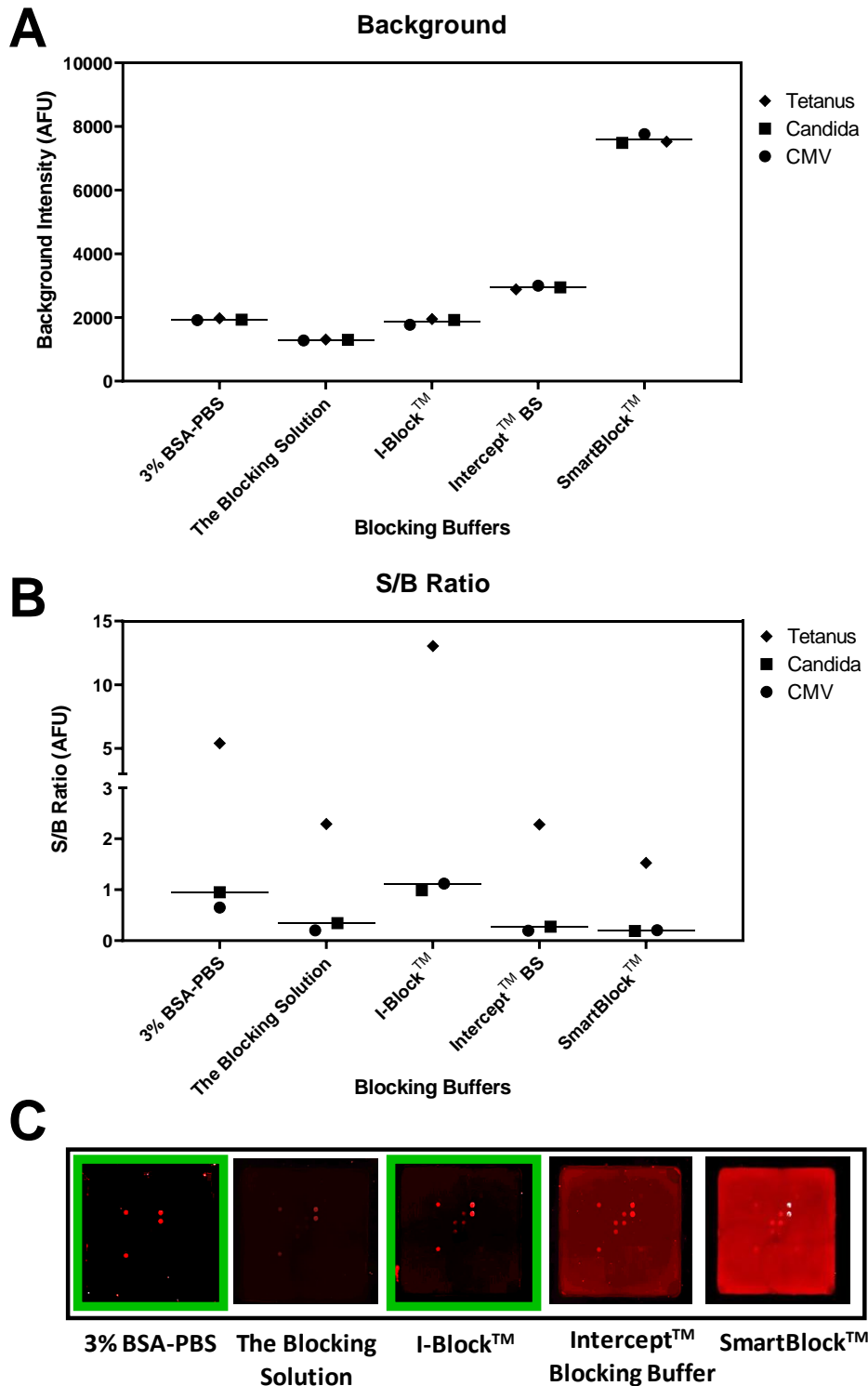


Figure 18. Identification of the optimal blocking buffer.

The sub-arrays were printed with rows of Tetanus, Candida, and CMV antigens, each in duplicate. (A) displays the background intensity at 635nm of the arrays when tested with each blocking buffer. (B) displays the S/B ratios when tested with each blocking buffer. (C) displays images scanned at 635nm of each subarray after probing with the relevant blocking buffers. This Figure demonstrates that I-Block™ and 3% BSA-PBS had the highest S/B ratios and were chosen as the optimal blocking buffers (highlighted with green boxes in part (C)). The same healthy-control serum was used to probe all arrays alongside the same assay protocol. $N = 2$. Means are plotted, and lines represent the grand median. S/B Ratio; signal-to-background ratio, AFU; Arbitrary fluorescent units, Intercept™ BS; Intercept™ Blocking Solution.

4.2.5 Scan Settings

In order to improve signal quantification, reduce background intensity and prevent spot saturation, we decided to test altering the laser power percentage when scanning arrays. We applied this approach to the probed slide used to test for the optimal blocking buffer. The probed slide was scanned at 100% and 33% laser power, with a consistent PMT Gain of 500. Figure 19A demonstrates that decreasing the laser power percentage from 100% to 33% successfully decreased the overall background intensity of all subarrays by nearly half its original intensity. This finding was further supported by the scanned images in Figure 19B, displaying the SmartBlock™ and I-Block™ probed subarrays at both laser power percentages. We observed a visible reduction in background intensity, particularly in the case of SmartBlock™, where the reduction in laser power percentage reduced spot saturation. As a result, we decided to implement the 33% laser power and PMT Gain of 500 settings for all subsequent slide scans. This change was to improve signal quantification and accuracy.

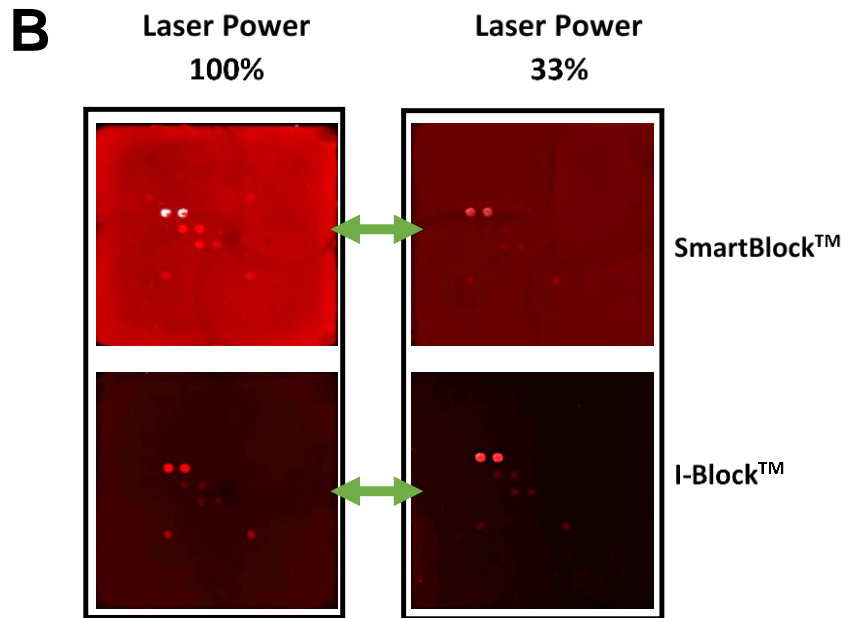
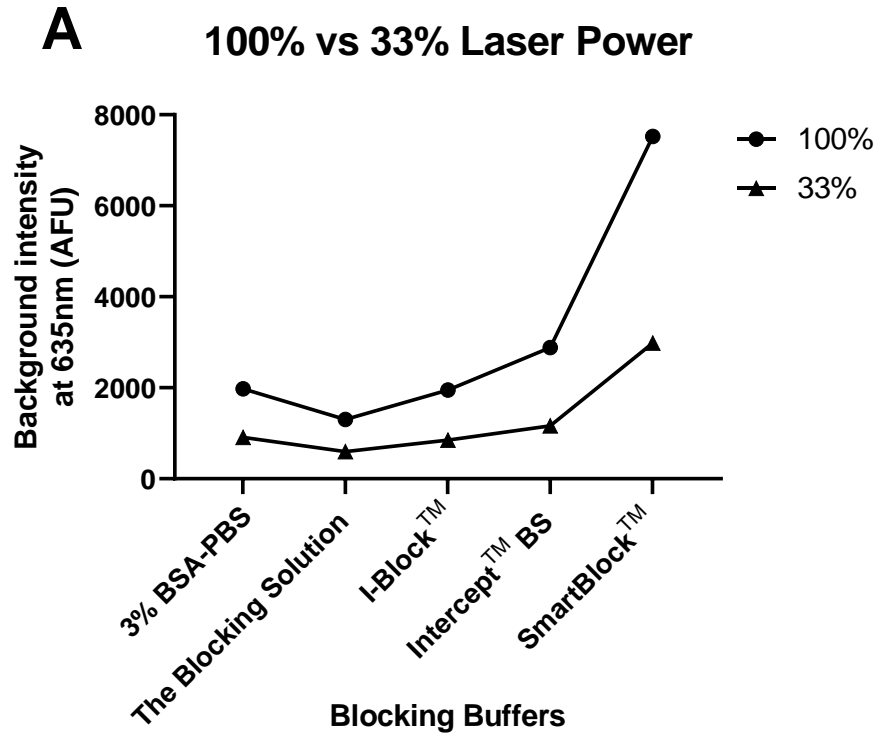


Figure 19. Comparison of the background intensity produced when scanning arrays at 100% and 33% laser power.

The slide containing ten subarrays for testing blocking buffers was scanned at 100% and 33% laser power with the PMT Gain set at 500. Each array consisted of rows of Tetanus, *Candida albicans*, and CMV printed in duplicate. (A) demonstrates that reducing the laser power to 33% reduces background intensity by approximately half the original intensity. (B) displays images scanned at 635nm of the SmartBlock™ and I-Block™ subarrays scanned at 100% and 33% laser power, visually depicting the reduction in background intensity. The means were plotted in (A), n=2. S/B Ratio; signal-to-background ratio, AFU; Arbitrary fluorescent units, Intercept™ BS; Intercept™ Blocking Solution.

4.2.6 Optimised Assay Protocol

An overview of the antigen microarray optimisation process is presented in Figure 20, illustrating the step-by-step development detailed above.

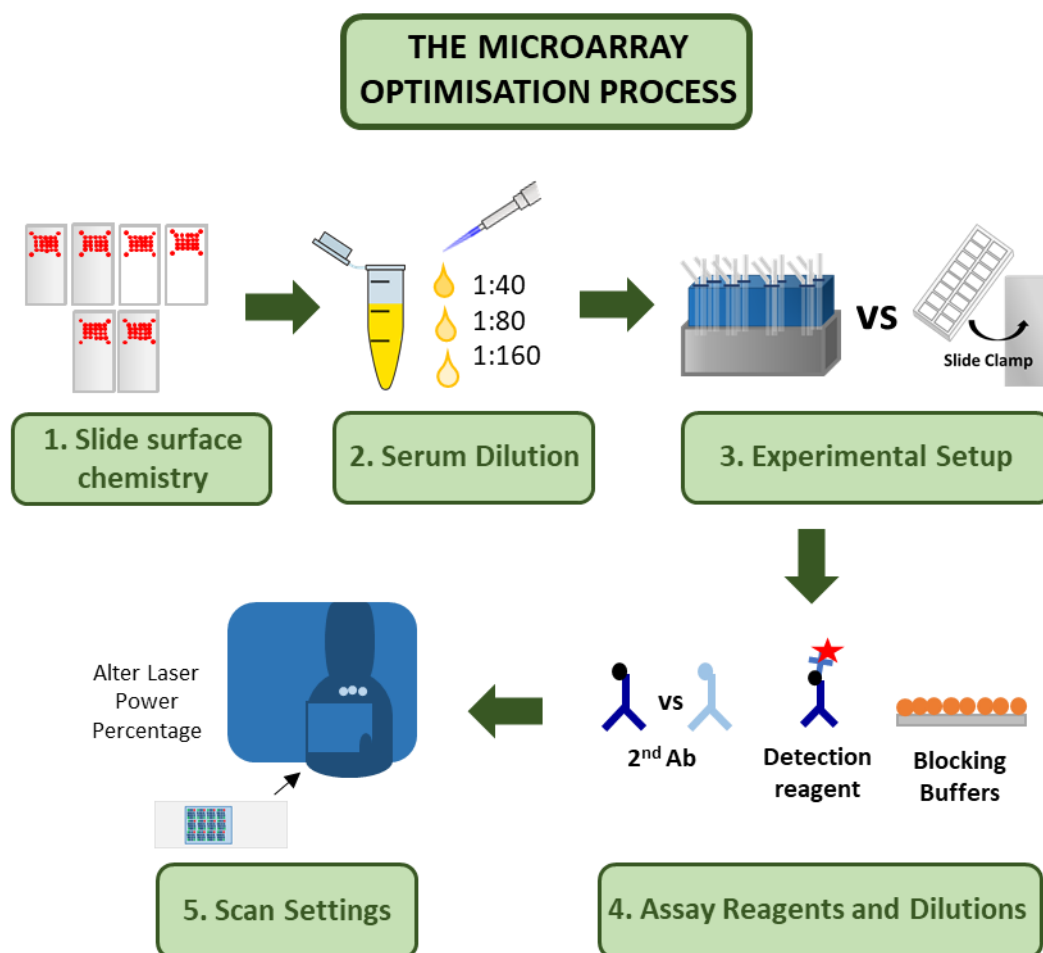


Figure 20. An overview of the antigen microarray optimisation process.

Schematic illustrating the key steps involved in optimising the antigen microarray platform, as detailed in the preceding results. Initial evaluation tested a range of slide surface chemistries to establish high signal intensities and uniform spot morphologies. Subsequently, serum dilutions were tested on nitrocellulose slides to evaluate non-specific binding and background intensity. Next the experimental set up was evaluated and transitioned from a slide cassette format to a 16-chamber multi-well clamp set up to improve background clarity and increase sample throughput. Further optimisation involved assessing assay reagents and their optimal dilutions including secondary antibodies (biotinylated anti-human IgG and anti-human IgG-cy5), detection reagent (Streptavidin-cy5) and a range of five blocking buffers to achieve optimal signal-to-background ratios. Finally, scan settings were fine-tuned on the GenePixPro 4000B microarray scanner at 635nm by testing different laser power percentages when scanning to ensure optimal signal quantification.

The optimised assay protocol for our microarray platform is demonstrated in Figure 21. Optimal dilutions of sera and assay reagents are summarised in Figure 21A, consisting of sera diluted to 1:160 in I-Block™, biotinylated anti-human IgG at a 1:5000 dilution in I-Block™, and streptavidin-cy5 at a 1:10,000 dilution in PBST. Figure 21B displays images scanned at 635nm of arrays probed with the optimised assay protocol, producing high and specific signal intensities alongside a low background intensity. The scanned images demonstrate the success of the optimised protocol in detecting the specific binding of CDI and CF sera antibodies to the respective antigens on the array while maintaining low levels of background and non-specific binding. The arrays were scanned using the optimised settings of 33% laser power and a PMT gain of 500.

Notably, cassette cover plates were used in the surface chemistry experiments, which produced some inconsistencies with the quantification of signal intensities. These inconsistencies were caused by droplet-marked backgrounds produced by flawed clamping of the slide cassette. The 16-multi-well chamber format resulted in consistently clear backgrounds with no clamping issues and thus was selected for the optimised protocol.

Additional steps were implemented into the optimised protocol to further improve assay performance, including the addition of 3% v/v goat serum to the biotinylated anti-human IgG diluent to reduce the cross-reactivity of the secondary antibody. Gentle shaking was applied during each incubation step

to improve protein binding efficacy and during wash steps to enhance the removal of unbound molecules to prevent non-specific binding.

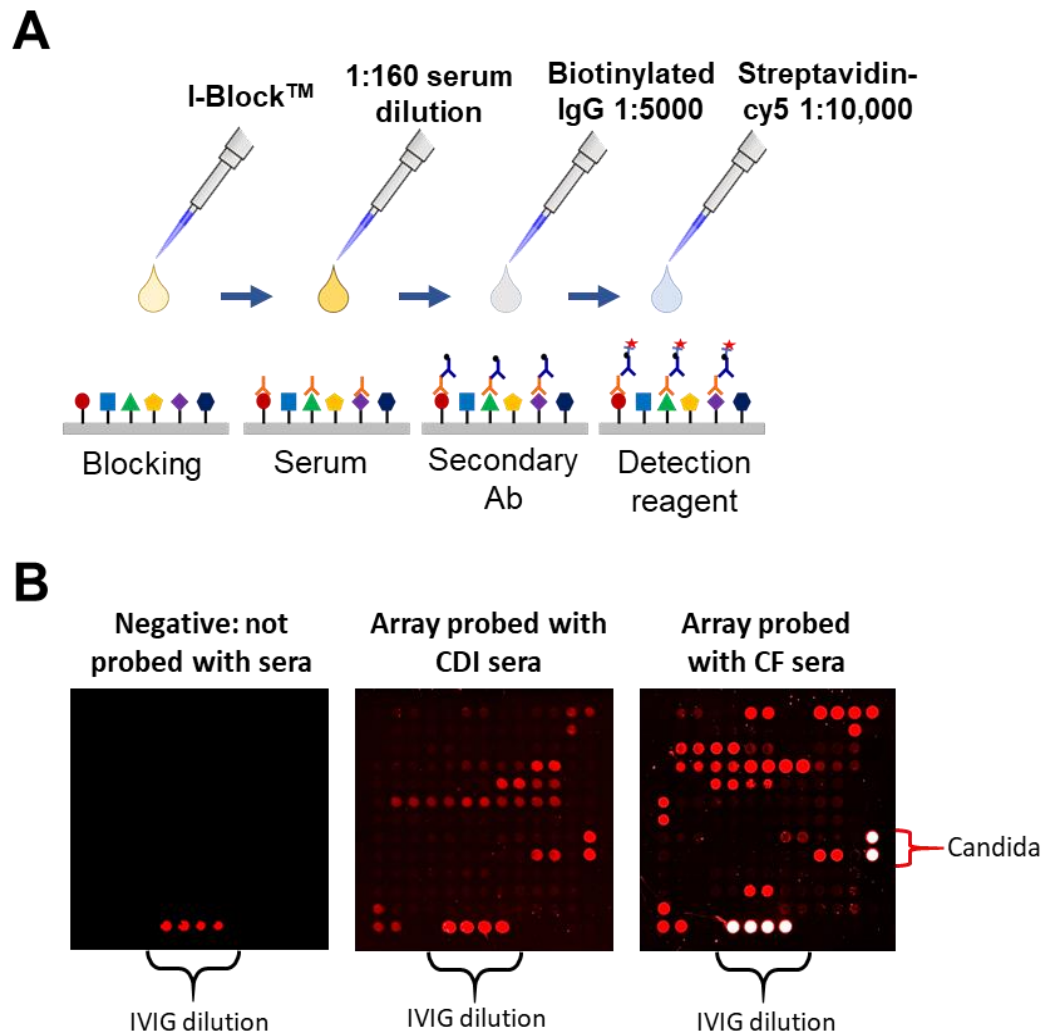


Figure 21. Overview of the optimised assay protocol and its results.

(A) Schematic summarising the optimised assay protocol. (B) Images scanned at 635nm of the optimised microarrays when probed with CDI (*Clostridium difficile* infection) and CF (Cystic fibrosis) serum samples and without sera (negative). The scans show high signal intensities, low background intensity, and low non-specific binding. Intravenous immunoglobulin (IVIG) was printed on the arrays at 100µg/ml and 50µg/ml for array orientation and as a standard for quality control. *Candida albicans* (positive-control antigen) can be seen in both images probed with CDI and CF sera but not in the negative, as expected. The negative control was probed with only I-Block™ and exhibited no autofluorescence of proteins in the absence of sera.

4.3 Microarray Platform Validation

4.3.1 Intra- and Inter-assay Precision

To ensure reliable results, it was important to validate the precision of the microarray platform before testing all patient samples. Hence, seven CF samples were assayed onto two subarrays per slide at two independent time points to determine each antigen's intra- and inter-assay coefficient of variation (CV). Slides were printed with 16 identical subarrays, with all 60 antigens printed in duplicate. Regarding intra-assay variation, 56 out of 60 antigens fell within acceptable limits of precision ($CV \leq 20\%$), as shown in Figure 22. The antigens that exceeded the 20% limit of precision were ribosomal P, rhDecorin, PL-12 and flagellin from *S. typhimurium*. Regarding inter-assay precision measuring slide-to-slide consistency, 54 out of 60 antigens fell within acceptable limits of precision ($CV \leq 20\%$), demonstrated in Figure 22. The six antigens failing to meet the inter-assay level of precision were ribosomal P, HEV, HCoV-OC34, mumps, PL-12 and flagellin from *S. typhimurium*. The intra-assay grand median resided under 10% CV precision and ranged from 5% to 32% across the 60 antigens. The inter-assay grand median resided under 15% CV precision and ranged from 8% to 38%. As expected, the median inter-assay CV, measuring slide-to-slide consistency, was higher than the median intra-assay CV, measuring within-slide consistency. Table 4.1 outlines the intra- and inter-assay CV% data for each antigen of the multi-antigen microarray platform. The antigens surpassing the 20% precision threshold (depicted in red in Figure 22) were still included in

the final array. However, during analysis, these antigens were flagged as potentially less reliable due to their elevated CV%. All 60 antigens presented in Figure 22 demonstrated specific reactivity after background signal subtraction. This study did not establish a threshold between natural and pathological antibody responses, so we cannot comment on the significance of general reactivity across antigens. Ultimately, “natural antibody” levels for each antigen must be determined in order to comment on the general reactivity observed.

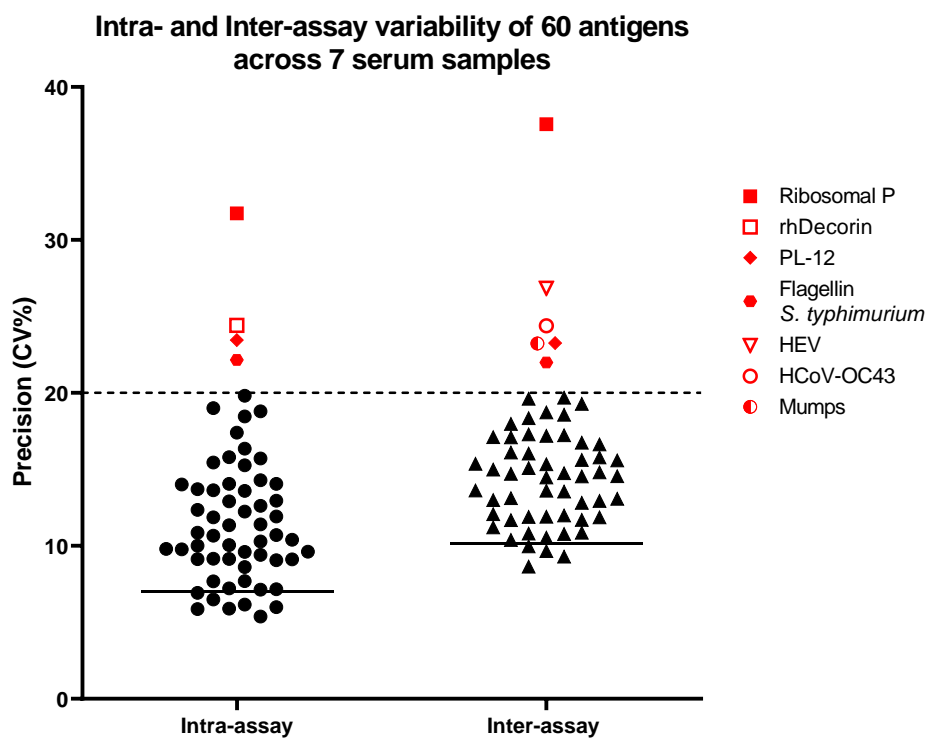


Figure 22. Intra- and Inter-assay precision of the multi-antigen microarray platform, including 60 antigens.

The dashed line represents an ideal precision limit for the intra-assay and inter-assay ($CV \leq 20\%$). The sera of seven CF patients was used to calculate the microarray intra- and inter-assay CVs. Each serum sample was assayed in duplicate onto two subarrays of a slide and replicated at two independent time points. All antigens were printed in duplicate onto each array. The lines for each group represent the grand median, both under the 20% CV limit of precision. Only antigens exceeding the 20% precision limit are labelled, with symbols displayed in red. All 60 antigens presented demonstrated specific reactivity after background signal subtraction. A threshold between natural and pathological antibody responses was not established in this study and so we can't comment on the significance of general reactivity across antigens. CV, Coefficient of variation; CF, Cystic fibrosis.

Table 4.1. Summary of the Intra- and Inter-assay precision data for each antigen.

(* antigens which exceeded detection limits).

Antigens	Intra-assay (CV%)	Inter-assay (CV%)
Tetanus toxoid	11.34	14.80
<i>Candida albicans</i>	13.70	10.86
HCoV-229E	10.39	17.23
CMV	19.00	19.28
EBV	13.63	17.10
EBV-VCA	12.23	13.55
HDV	9.62	9.97
HEV	10.66	26.84*
HCoV-HKU1	19.80	15.09
HSV1	9.12	15.62
HSV2	7.13	18.35
Influenza A H1N1	12.35	11.69
Mumps	15.71	23.23*
Native Adenovirus	10.85	9.65
HCoV-OC34	10.05	24.38*
Parvovirus VP2	7.21	11.86
Rubella	8.60	13.60
Rubeola	6.48	12.94
VZV	9.79	16.10
SARS-CoV-2 Spike S1	9.60	17.28
SARS-CoV-2 Nucleocapsid	5.87	12.07
AGC1	7.16	13.11
Azurocidin Human Neutrophil	15.80	10.38
BPI	5.42	14.62
cANCA	7.70	13.08
Carbamylated-BSA	10.00	15.78
Cenp B	12.94	10.53
CK-18	12.90	15.36
dsDNA	13.60	10.77
Flagellin <i>S. typhimurium</i>	22.16*	21.98*
GBM	10.70	16.03
HMGCR	12.61	19.61
La (SSB)	9.14	8.65

Lactoferrin	9.11	19.69
Laminarin	9.06	9.29
Mannan	11.85	16.74
pANCA	6.15	17.18
Peptidoglycan <i>S. aureus</i>	18.45	16.63
PL-12	23.46*	23.26*
PL-7	5.38	14.71
rhDecorin	24.42*	11.98
Ribosomal P	31.73*	37.56*
RNP/Sm	14.04	10.80
RNP-68k	15.44	12.99
Ro (SSA)	14.01	14.47
Scl-70	14.24	17.25
SSA-60/Ro	9.77	11.68
Jo-1	11.39	11.21
Elastin Human Lung	11.91	11.90
Keratin Human Epidermis	9.15	15.33
<i>H. pylori</i> Cag A	7.69	13.61
TcdA	18.79	14.54
TcdB	15.26	18.57
CD20	16.36	14.54
Gliadin	9.41	12.81
Cardiac troponin	14.00	15.58
Insulin	17.39	11.89
Thyroid peroxidase	10.28	14.99
rhTGM2	5.99	17.09
Thyroglobulin	6.93	14.74

4.3.2 Testing Sample Quality and Antibody Reactivity

Positive-control antigens *Candida albicans* and Tetanus toxoid were included on each array for quality control. These positive-control antigens are from common human pathogens to which most individuals would have strong protective IgG antibody responses. Failure of serum samples to exhibit reactivity to these well-established positive-controls raises questions regarding the suitability of the samples for evaluating humoral antibody responses. Thus, the inclusion of control antigens indirectly assesses serum quality and antibody reactivity for array processing. Figure 23 is a scatterplot of the responses from 103 serum samples for each control antigen. Among the 103 samples tested, 78% exhibited signal intensities exceeding 10,000 AFU for *Candida albicans*, while 65% displayed signal intensities exceeding 10,000 AFUs for Tetanus toxoid, demonstrating robust IgG reactivity from the majority of serum samples to the positive-control antigens.

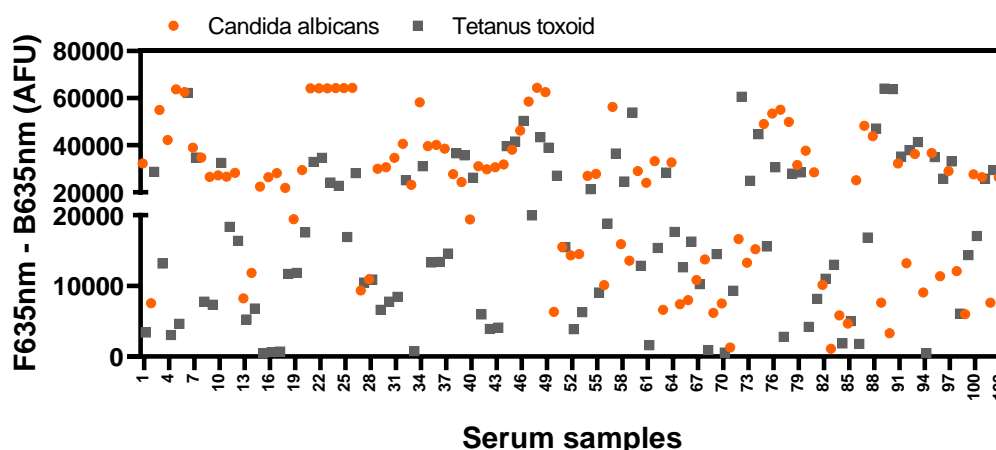


Figure 23. Testing sample quality and antibody reactivity of 103 serum samples through IgG responses to two positive-control antigens.

The positive-control antigens were *Candida albicans* and Tetanus toxoid. IgG responses to the controls are displayed with many strong IgG responses and expected variation between individuals. A negative control of just printing buffer was also incorporated on each array as a further quality control measure. F635nm – B635nm; Foreground 635nm – Background 635nm, AFU; Arbitrary fluorescent units.

4.4 Patient Demographics

This study utilised 152 serum samples from individuals with cystic fibrosis (CF), *C. difficile* infection (CDI), recurrent *C. difficile* infection (rCDI) before and after FMT and healthy control (HC) sera. The patient demographics of the four subject cohorts are detailed in Tables 4.2 and 4.3. Patients with CF and CDI had not undergone any transplants, unlike the rCDI FMT patient cohort, who all had successful responses to the FMT therapy. The medical histories of patients with rCDI were available, and pre-existing autoimmune conditions have been detailed in Table 4.3. Fourteen patients with CDI and two patients with rCDI (FMT cohort) had a history of immunosuppressant use (Tables 4.2 and 4.3); the impact of immunosuppressants on our results is highlighted in the discussion.

Table 4.2. CDI and CF patient demographics.

Subject characteristics	<i>Clostridium difficile</i> infection (CDI) (n = 63)	Cystic Fibrosis (n = 11)
Age – mean years (range)	69 (22-92)	29 (20-47)
Male/Female	26/ 37	4/7
Immunosuppressant	14	N/A
CDI Severity	43 Mild/ 20 Severe	N/A

Table 4.3. Recurrent CDI FMT patients and age-matched healthy control demographics.

The 24 rCDI patients had sets of Pre-FMT and 12 weeks Post-FMT serum samples each, including medical history of autoimmune disease.

Subject characteristics	Healthy Controls (n = 30)	Recurrent <i>Clostridium difficile</i> patients undergoing FMT (n = 24)
Age - mean years (range)	61 (23-85)	54 (20-86)
Male/Female	15/15	5/19
Immunosuppressant	0	2
Ulcerative colitis	0	3
Crohn's disease	0	1
Multiple sclerosis	0	1
Psoriasis	0	1
Hypothyroidism	0	2
Microscopic colitis	0	1

4.5 Application of The Multi-antigen Microarray

The preceding results focused on generating, optimising, and validating the microarray platform. Now, the focus will shift towards applying the antigen microarray for screening clinically defined patient cohorts, including patients diagnosed with CF, CDI, rCDI pre- and post-FMT, and Healthy controls (HC). The microarray was developed to screen humoral immune responses to 60 antigens: autoantigens, positive-control antigens, pathogen-associated antigens and viral recall antigens. Antigens exceeding the 20% CV threshold determined from precision testing (Figure 22) were still included in the final array but flagged as less reliable during analysis. Table 4.4 provides a concise overview of the main comparisons that will be expanded upon within this section, elucidating the multifaceted applications of the multi-antigen microarray.

Table 4.4 Summary of comparative analyses in the following application of the multi-antigen microarray results section.

Comparisons	To Investigate
4.5.1 Hierarchical Clustering Analysis	Commonalities and differences in humoral immune responses across and within subject cohorts (CF, CDI, rCDI, HC)
4.5.2 Humoral Responses of rCDI Patients Before and After FMT	If FMT impacts humoral immune responses in rCDI patients by assessing responses before and 12 weeks after treatment
4.5.3 Humoral Responses of CDI and rCDI Patients	Variations in humoral immune responses between initial CDI and recurrent CDI patient cohorts

4.5.1 Hierarchical Clustering of Humoral Antibody Responses

The multi-antigen microarray platform was probed with sera across our four individual cohorts. Given the high dimensionality of the dataset, we initially used heatmaps to group antibody responses by subject cohort and identify antigens exhibiting positive antibody responses (Appendix 1 and 2 display full 60-antigen heatmaps). Hierarchical clustering analysis was performed on the subsetted positive response antigens to elucidate commonalities and differences in humoral immune response profiles across patient cohorts, specifically evaluating IgG and IgA responses. We opted for hierarchical clustering of antibody responses as subtle distinctions in humoral immune responses may be overlooked when clustering based on patient groups, which requires larger cohorts to ensure reliable clustering.

Figure 24 presents heatmaps displaying the serological IgG responses of CF, CDI, rCDI and HC cohorts across three antigen categories: autoantigens (A), pathogen-associated (B), and viral recall antigens (C). In Figure 24A, B and C, we observed that CDI patients exhibited a distinct clustering pattern, roughly dividing into four groups based on levels of antigen-specific IgG responses. This distinction is particularly evident for the viral recall antigens (Figure 24C), where the largest block of grouped CDI patients displays low IgG responses across the panel of eight viral antigens.

IgG autoantibody responses were observed across patient cohorts (Figure 24A). We identified that the autoantigen, dsDNA, elicited IgG responses across CF, rCDI and HC groups (Figure 24A). Robust IgG responses to gliadin

were detected frequently across CF and rCDI patients (Pre- and Post-FMT). Figure 24A showed that individual samples had occasional and robust IgG responses to carbamylated-BSA, BPI, cANCA, Scl-70, Ro-SSA, thyroid peroxidase, laminarin and mannan. Notably, rare responses like these would be interesting to follow up on but require large cohorts to determine significance. In Figure 24B, a clear clustering pattern emerged, with most CF and rCDI (Pre-FMT and Post-FMT) patients grouping together for pathogen-associated antigens. This clustering was primarily attributed to their robust IgG responses to TcdA and TcdB for rCDI patients. The small proportion of CF and rCDI patients that did not conform to the cluster had lower IgG responses to TcdA and TcdB (Figure 24B).



Figure 24. Heatmaps displaying the hierarchical clustering of IgG responses across CF, CDI, rCDI and healthy controls.

(A) Autoantigens, (B) pathogen-associated antigens, (C) viral recall antigens. The antigens are displayed on the vertical axis of the heatmaps and the patient groups are colour-coded on the horizontal axis. Dark to light blue indicate high and low IgG responses, respectively. Pre-FMT and Post-FMT refer to rCDI patients. The signal intensity values are log2 transformed from arbitrary fluorescent units (AFUs).

Figure 25 displays a set of heatmaps illustrating the serological IgA responses across the same cohorts and antigen subsets. IgA autoantibody responses were mainly observed in CF and CDI patients compared to rCDI and HCs (Figure 25A). In Figure 25A, we observed that the majority of CF patients clustered together, primarily due to their notably robust IgA responses in the autoantigen subset to gliadin, carbamylated-BSA, human insulin, and dsDNA. Notably, small bands of CDI patients had IgA responses to gliadin (Figure 25A). On the other hand, rCDI patients (both Pre- and Post-FMT) displayed relatively low IgA responses across the autoantigen subset, a pattern shared with HCs (Figure 25A). Figure 25B did not show distinct clustering patterns among patient groups for pathogen-associated antigens due to the low IgA responses across this subset, except for the expected robust responses to the positive-control *Candida albicans*. In Figure 25C, we noted that HCs tended to cluster together, primarily attributed to their comparatively low IgA levels across the viral recall antigens.

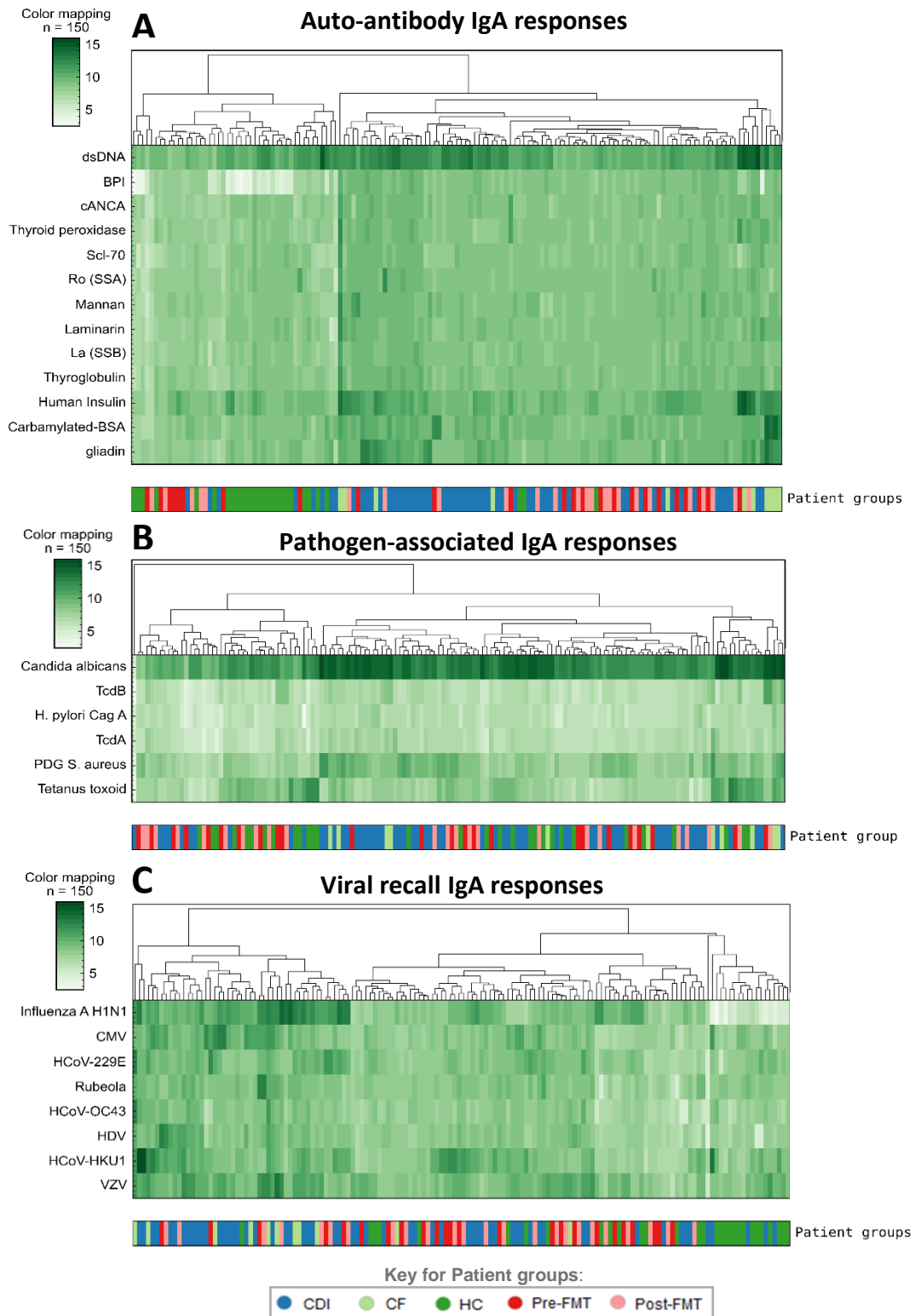


Figure 25. Heatmaps displaying the hierarchical clustering of IgA responses across CF, CDI, rCDI and healthy controls.

(A) Autoantigens, (B) pathogen-associated antigens, (C) viral recall antigens. The antigens are displayed on the vertical axis of the heatmaps and the patient groups are colour-coded on the horizontal axis. Dark to light green indicate high and low IgA responses, respectively. Pre-FMT and Post-FMT refer to rCDI patients. The signal intensity values are log₂ transformed from arbitrary fluorescent units (AFUs).

4.5.2 Humoral Responses of rCDI Patients Before and After FMT

Next, we assessed the impact of FMT on the humoral responses of rCDI patients. Generalised Linear Models (GLMs) were employed to investigate antigen-specific differences in IgG and IgA antibody responses across three patient status cohorts: Pre-FMT, 12 weeks Post-FMT and healthy controls. Our analysis focused exclusively on antigens eliciting positive antibody responses, as identified through the prior heatmap analysis. Antigens were categorised into four subsets: autoantigens, positive-control, pathogen-associated, and viral recall antigens.

4.5.2.1 IgG

The GLMs were assessed for significance with Analysis of Variance (ANOVA) type II tests (see Table 4.5). The results demonstrated that IgG reactivity in all antigen subsets differed in response to antigen type (Table 4.5). Moreover, there were no significant differences in antigen-specific IgG responses between patient status groups for the autoantigen and positive-control antigen subsets (Table 4.5; Figure 26A, B). However, antigen-specific IgG responses did differ across patient status groups in the pathogen-associated and viral recall antigen subsets (Table 4.5; Figure 26C, D).

Table 4.5. ANOVA type II test results across antigen subset GLM models for IgG responses.

Status: Pre-FMT vs. Post-FMT vs. Healthy control. Likelihood Ratio Chi-Square (LR Chisq). $p < 0.05$ *, $p < 0.01$ **, $p < 0.001$ ***.

Autoantigens IgG	X ²	df	p-value
Antigen	38.67	5	2.767e-07 ***
Status	2.60	2	0.2730
Antigen : Status	6.19	10	0.7989
Positive-Control antigens IgG	X ²	df	p-value
Antigen	5.58	1	0.01817 *
Status	0.89	2	0.64075
Antigen : Status	2.24	2	0.32581
Pathogen-associated antigens IgG	X ²	df	p-value
Antigen	78.98	3	< 2.2e-16 ***
Status	34.64	2	3.008e-08 ***
Antigen : Status	37.90	6	1.174e-06 ***
Viral recall antigens IgG	X ²	df	p-value
Antigen	104.845	7	< 2.2e-16 ***
Status	11.408	2	0.003332 **
Antigen : Status	48.551	14	1.065e-05 ***

Pairwise comparisons demonstrated that Pre-FMT and Post-FMT patient groups exhibited significantly elevated IgG responses to the pathogen-associated antigen, TcdA, compared to healthy controls (Table 4.6; Figure 26C; $p < 0.0001$). Notably, the Pre-FMT patient group displayed significantly higher IgG responses to the viral recall antigen, CMV, than healthy controls (Table 4.6; Figure 26D; $p < 0.0001$). No statistically significant differences were found in antigen-specific IgG responses between the pre-FMT and 12 weeks post-FMT rCDI groups across antigen subsets (Table 4.6; $p > 0.05$).

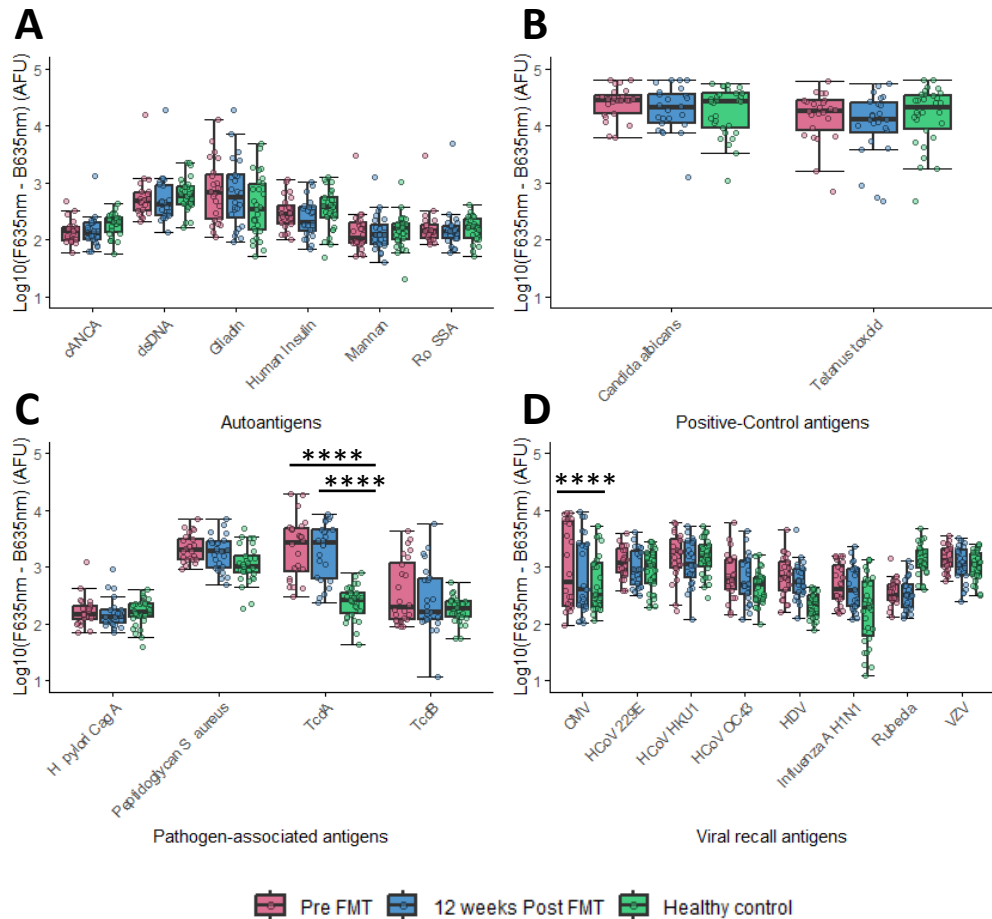


Figure 26. IgG responses in patients Pre-FMT, 12 weeks Post-FMT and healthy controls across autoantigens (A), positive-control (B), pathogen-associated (C) and viral recall antigens (D).

Differences between groups were assessed using Generalised Linear Models (GLM) with fixed effects of patient status (Pre-FMT, Post-FMT, Healthy control) and antigen choice. The models were evaluated for significance through ANOVA type II tests and Tukey's *p*-value adjustment for multiple pairwise comparisons. For Autoantigens (A) and Positive-Control Antigens (B), the analysis found no significant differences in antigen-specific IgG responses between patient status groups. However, for Pathogen-Associated Antigens (C) and Viral Recall Antigens (D), antigen-specific IgG responses differed between patient status groups. Specifically: (C) Both Pre-FMT and Post-FMT patient groups demonstrated significantly higher IgG responses to TcdA compared to healthy controls ($p < 0.0001$, ****). (D) The Pre-FMT patient group exhibited significantly elevated IgG responses to CMV compared to healthy controls ($p < 0.0001$, ****). Sample sizes: Pre-FMT ($n = 24$), 12 weeks Post-FMT ($n = 24$), Healthy control ($n = 30$). Note that the Y-axis scale is log₁₀() due to outliers (strong or weak antibody responses). Error bars display the standard error of the mean. F635nm – B635nm; Foreground 635nm – Background 635nm, AFU; Arbitrary fluorescent units.

Table 4.6. Pairwise comparisons of antigen-specific IgG responses between the patient cohorts.

Status: Pre-FMT, Post-FMT and healthy controls across the autoantigen, positive-control, pathogen-associated and viral recall antigen subsets.

Autoantigens	Comparison	Estimate	SE	df	p-value
cANCA	Pre-FMT vs Post-FMT	-32.02	463	450	1.00
cANCA	Pre-FMT vs. Healthy control	-47.10	440	450	1.00
cANCA	Post-FMT vs. Healthy control	-15.08	440	450	1.00
dsDNA	Pre-FMT vs Post-FMT	-130.71	463	450	1.00
dsDNA	Pre-FMT vs. Healthy control	387.77	440	450	1.00
dsDNA	Post-FMT vs. Healthy control	518.48	440	450	1.00
Gliadin	Pre-FMT vs Post-FMT	-296.44	463	450	1.00
Gliadin	Pre-FMT vs. Healthy control	807.79	440	450	0.93
Gliadin	Post-FMT vs. Healthy control	1104.23	440	450	0.52
Human Insulin	Pre-FMT vs. Post-FMT	59.25	463	450	1.00
Human Insulin	Pre-FMT vs. Healthy control	-76.14	440	450	1.00
Human Insulin	Post-FMT vs. Healthy control	-135.39	440	450	1.00
Mannan	Pre-FMT vs. Post-FMT	66.19	463	450	1.00
Mannan	Pre-FMT vs. Healthy control	68.28	440	450	1.00
Mannan	Post-FMT vs. Healthy control	2.09	440	450	1.00
Ro SSA	Pre-FMT vs. Post-FMT	-60.48	463	450	1.00
Ro SSA	Pre-FMT vs. Healthy control	102.23	440	450	1.00
Ro SSA	Post-FMT vs. Healthy control	162.71	440	450	1.00
Positive-Control Antigens	Comparison	Estimate	SE	df	p-value
<i>Candida albicans</i>	Pre-FMT vs. Post-FMT	3585	4857	150	1.00
<i>Candida albicans</i>	Pre-FMT vs. Healthy control	5321	4608	150	0.86
<i>Candida albicans</i>	Post-FMT vs. Healthy control	1736	4608	150	1.00
Tetanus toxoid	Pre-FMT vs. Post-FMT	2720	4857	150	0.99
Tetanus toxoid	Pre-FMT vs. Healthy control	-3376	4608	150	0.98
Tetanus toxoid	Post-FMT vs. Healthy control	-6096	4608	150	0.77

Pathogen-associated Antigens	Comparison	Estimate	SE	df	p-value
<i>H. pylori</i> Cag A	Pre-FMT vs. Post-FMT	15.40	529	300	1.00
<i>H. pylori</i> Cag A	Pre-FMT vs. Healthy control	28.60	502	300	1.00
<i>H. pylori</i> Cag A	Post-FMT vs. Healthy control	13.20	502	300	1.00
Peptidoglycan <i>S. aureus</i>	Pre-FMT vs. Post-FMT	288.60	529	300	1.00
Peptidoglycan <i>S. aureus</i>	Pre-FMT vs. Healthy control	1091.37	502	300	0.57
Peptidoglycan <i>S. aureus</i>	Post-FMT vs. Healthy control	802.76	502	300	0.91
TcdA	Pre-FMT vs. Post-FMT	1017.90	529	300	0.74
TcdA	Pre-FMT vs. Healthy control	3862.47	502	300	< 0.0001
TcdA	Post-FMT vs. Healthy control	2844.57	502	300	< 0.0001
TcdB	Pre-FMT vs. Post-FMT	171.21	529	300	1.00
TcdB	Pre-FMT vs. Healthy control	637.94	502	300	0.98
TcdB	Post-FMT vs. Healthy control	466.73	502	300	1.00
Viral recall Antigens	Comparison	Estimate	SE	df	p-value
CMV	Pre-FMT vs. Post-FMT	940.60	356	600	0.56
CMV	Pre-FMT vs. Healthy control	1857.34	337	600	<0.0001
CMV	Post-FMT vs. Healthy control	916.74	337	600	0.50
HCoV-229E	Pre-FMT vs. Post-FMT	192.15	356	600	1.00
HCoV-229E	Pre-FMT vs. Healthy control	266.70	337	600	1.00
HCoV-229E	Post-FMT vs. Healthy control	74.56	337	600	1.00
HCoV-HKU1	Pre-FMT vs. Post-FMT	488.52	356	600	1.00
HCoV-HKU1	Pre-FMT vs. Healthy control	269.36	337	600	1.00
HCoV-HKU1	Post-FMT vs. Healthy control	-219.16	337	600	1.00
HCoV-OC34	Pre-FMT vs. Post-FMT	182.46	356	600	1.00
HCoV-OC34	Pre-FMT vs. Healthy control	560.12	337	600	0.99
HCoV-OC34	Post-FMT vs. Healthy control	377.66	337	600	1.00
HDV	Pre-FMT vs. Post-FMT	190.71	356	600	1.00
HDV	Pre-FMT vs. Healthy control	734.23	337	600	0.88
HDV	Post-FMT vs. Healthy control	543.52	337	600	0.99
Influenza A H1N1	Pre-FMT vs. Post-FMT	25.65	356	600	1.00

Influenza A H1N1	Pre-FMT vs. Healthy control	295.45	337	600	1.00
Influenza A H1N1	Post-FMT vs. Healthy control	269.80	337	600	1.00
Rubeola	Pre-FMT vs. Post-FMT	-1.54	356	600	1.00
Rubeola	Pre-FMT vs. Healthy control	-1171.30	337	600	0.08
Rubeola	Post-FMT vs. Healthy control	-1169.76	337	600	0.08
VZV	Pre-FMT vs. Post-FMT	231.50	356	600	1.00
VZV	Pre-FMT vs. Healthy control	365.31	337	600	1.00
VZV	Post-FMT vs. Healthy control	133.81	337	600	1.00

4.5.2.2 IgA

Next, we reiterated the GLM analysis to investigate the influence of patient status on antigen-specific IgA responses. The results demonstrated that IgA reactivity differed in response to antigen type (Table 4.7). For autoantigen, positive-control, and pathogen-associated antigen subsets, there were no significant differences in antigen-specific IgA responses between patient status groups (Table 4.7; Figure 27A, B, C). However, antigen-specific IgA responses did differ across patient status groups in the viral recall antigen subset (Table 4.7; Figure 27D).

Table 4.7. ANOVA type II test results across antigen subset GLM models for IgA responses.

Status: Pre-FMT vs. Post-FMT vs. Healthy control. Likelihood Ratio Chi-Square (LR Chisq). $p < 0.05$ *, $p < 0.01$ **, $p < 0.001$ ***.

Autoantigens IgA	LR Chisq	df	p-value
Antigen	55.554	5	1.004e-10 ***
Status	2.392	2	0.3025
Antigen : Status	2.624	10	0.9890
Positive-Control antigens IgA	LR Chisq	df	p-value
Antigen	50.900	1	9.72e-13 ***
Status	4.709	2	0.09492
Antigen : Status	3.102	2	0.21207
Pathogen-associated antigens IgA	LR Chisq	df	p-value
Antigen	98.028	3	< 2.2e-16 ***
Status	9.322	2	0.009455 **
Antigen : Status	6.059	6	0.416608
Viral recall antigens IgA	LR Chisq	df	p-value
Antigen	40.926	7	8.366e-07 ***
Status	24.814	2	4.090e-06 ***
Antigen : Status	18.430	14	0.1879

Subsequent pairwise comparisons revealed that Post-FMT patients exhibited significantly higher IgA responses to the viral recall antigen, Influenza A H1N1, than healthy controls (Table 4.8; Figure 27D; $p < 0.001$, ***). Our results found no significant differences in antigen-specific IgA responses between the Pre-FMT and Post-FMT rCDI groups across antigen subsets (Table 4.8; $p > 0.05$).

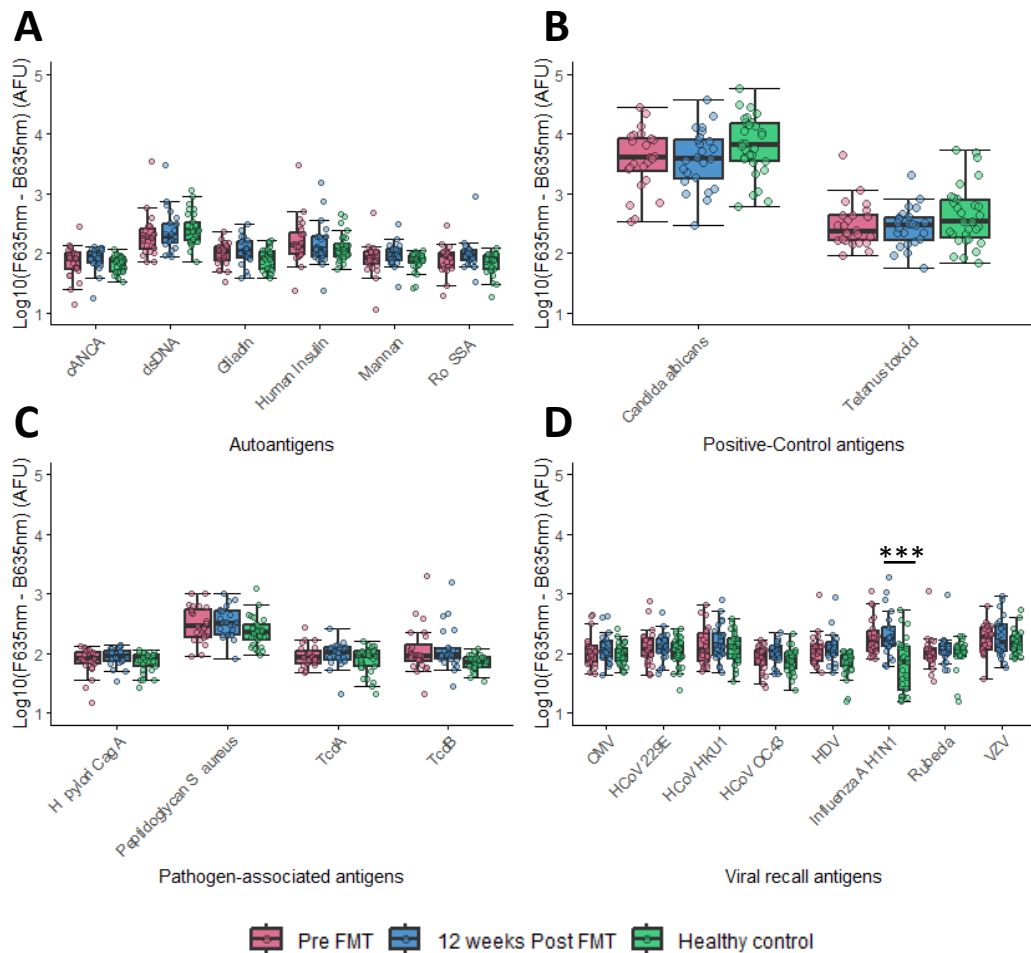


Figure 27. IgA responses in patients Pre-FMT, 12 weeks Post-FMT and healthy controls across autoantigens (A), positive-controls (B), pathogen-associated (C) and viral recall antigens (D).

Differences between groups were assessed using Generalised Linear Models (GLM) with fixed effects of patient status (Pre-FMT, Post-FMT and healthy control) and antigen choice. The models were evaluated for significance through ANOVA type II tests and Tukey's *p*-value adjustment for multiple pairwise comparisons. For Autoantigens (A), Positive-Control Antigens (B) and Pathogen-associated antigens (C), the analysis found no significant differences in antigen-specific IgA responses between patient status groups. However, for Viral Recall Antigens (D), antigen-specific IgG responses did differ between patient status groups. Specifically, the Post-FMT patient group exhibited significantly higher IgA responses to Influenza A H1N1 than healthy controls ($p < 0.001$, ***). Sample sizes: Pre-FMT ($n = 24$), 12 weeks Post-FMT ($n = 24$), Healthy control ($n = 30$). Note that the Y-axis scale is to the $\log_{10}()$ due to outliers (very strong or weak antibody responses). Error bars display the standard error of the mean. F635nm – B635nm; Foreground 635nm – Background 635nm, AFU; Arbitrary fluorescent units.

Table 4.8. Pairwise comparisons of antigen-specific IgA responses between the patient cohorts.

Status: Pre-FMT, Post-FMT and healthy controls across the autoantigen, positive-control, pathogen-associated and viral recall antigen subsets.

Autoantigens	Comparison	Estimate	SE	df	p-value
cANCA	Pre-FMT vs. Post-FMT	-2.31	78.3	450	1.00
cANCA	Pre-FMT vs. Healthy control	13.37	74.2	450	1.00
cANCA	Post-FMT vs. Healthy control	15.68	74.2	450	1.00
dsDNA	Pre-FMT vs. Post-FMT	-25.06	78.3	450	1.00
dsDNA	Pre-FMT vs. Healthy control	18.10	74.2	450	1.00
dsDNA	Post-FMT vs. Healthy control	43.16	74.2	450	1.00
Gliadin	Pre-FMT vs. Post-FMT	-15.65	78.3	450	1.00
Gliadin	Pre-FMT vs. Healthy control	20.95	74.2	450	1.00
Gliadin	Post-FMT vs. Healthy control	36.60	74.2	450	1.00
Human Insulin	Pre-FMT vs. Post-FMT	72.23	78.3	450	1.00
Human Insulin	Pre-FMT vs. Healthy control	143.32	74.2	450	0.90
Human Insulin	Post-FMT vs. Healthy control	71.10	74.2	450	1.00
Mannan	Pre-FMT vs. Post-FMT	-5.96	78.3	450	1.00
Mannan	Pre-FMT vs. Healthy control	20.20	74.2	450	1.00
Mannan	Post-FMT vs. Healthy control	26.15	74.2	450	1.00
Ro SSA	Pre-FMT vs. Post-FMT	-34.69	78.3	450	1.00
Ro SSA	Pre-FMT vs. Healthy control	17.19	74.2	450	1.00
Ro SSA	Post-FMT vs. Healthy control	51.88	74.2	450	1.00
Positive-Control Antigens	Comparison	Estimate	SE	df	p-value
<i>Candida albicans</i>	Pre-FMT vs. Post-FMT	-351	1944	150	1.00
<i>Candida albicans</i>	Pre-FMT vs. Healthy control	-4522	1844	150	0.15
<i>Candida albicans</i>	Post-FMT vs. Healthy control	-4171	1844	150	0.22
Tetanus toxoid	Pre-FMT vs. Post-FMT	116	1944	150	1.00
Tetanus toxoid	Pre-FMT vs. Healthy control	-403	1844	150	1.00
Tetanus toxoid	Post-FMT vs. Healthy control	-518	1844	150	1.00
Pathogen-associated Antigens	Comparison	Estimate	SE	df	p-value
<i>H. pylori</i> Cag A	Pre-FMT vs. Post-FMT	-7.729	52.9	300	1.00
<i>H. pylori</i> Cag A	Pre-FMT vs. Healthy control	7.017	50.1	300	1.00
<i>H. pylori</i> Cag A	Post-FMT vs. Healthy control	14.746	50.1	300	1.00

Peptidoglycan S. aureus	Pre-FMT vs. Post-FMT	-0.125	52.9	300	1.00
Peptidoglycan S. aureus	Pre-FMT vs. Healthy control	107.700	50.1	300	0.59
Peptidoglycan S. aureus	Post-FMT vs. Healthy control	107.825	50.1	300	0.59
TcdA	Pre-FMT vs. Post-FMT	-4.646	52.9	300	1.00
TcdA	Pre-FMT vs. Healthy control	13.588	50.1	300	1.00
TcdA	Post-FMT vs. Healthy control	18.233	50.1	300	1.00
TcdB	Pre-FMT vs. Post-FMT	12.500	52.9	300	1.00
TcdB	Pre-FMT vs. Healthy control	131.933	50.1	300	0.27
TcdB	Post-FMT vs. Healthy control	119.433	50.1	300	0.42
Viral recall Antigens	Comparison	Estimate	SE	df	p-value
CMV	Pre-FMT vs. Post-FMT	-4.9167	43.6	600	1.00
CMV	Pre-FMT vs. Healthy control	27.6208	41.4	600	1.00
CMV	Post-FMT vs. Healthy control	32.5375	41.4	600	1.00
HCoV-229E	Pre-FMT vs. Post-FMT	15.2708	43.6	600	1.00
HCoV-229E	Pre-FMT vs. Healthy control	44.0000	41.4	600	1.00
HCoV-229E	Post-FMT vs. Healthy control	28.7292	41.4	600	1.00
HCoV-HKU1	Pre-FMT vs. Post-FMT	-23.6458	43.6	600	1.00
HCoV-HKU1	Pre-FMT vs. Healthy control	32.1917	41.4	600	1.00
HCoV-HKU1	Post-FMT vs. Healthy control	55.8375	41.4	600	1.00
HCoV-OC34	Pre-FMT vs. Post-FMT	-19.1875	43.6	600	1.00
HCoV-OC34	Pre-FMT vs. Healthy control	7.8042	41.4	600	1.00
HCoV-OC34	Post-FMT vs. Healthy control	26.9917	41.4	600	1.00
HDV	Pre-FMT vs. Post-FMT	-4.7917	43.6	600	1.00
HDV	Pre-FMT vs. Healthy control	72.1250	41.4	600	0.99
HDV	Post-FMT vs. Healthy control	76.9167	41.4	600	0.98
Influenza A H1N1	Pre-FMT vs. Post-FMT	-68.6458	43.6	600	1.00
Influenza A H1N1	Pre-FMT vs. Healthy control	147.8083	41.4	600	0.07
Influenza A H1N1	Post-FMT vs. Healthy control	216.4542	41.4	600	0.0001
Rubeola	Pre-FMT vs. Post-FMT	-3.7708	43.6	600	1.00
Rubeola	Pre-FMT vs. Healthy control	33.1458	41.4	600	1.00
Rubeola	Post-FMT vs. Healthy control	36.9167	41.4	600	1.00
VZV	Pre-FMT vs. Post-FMT	-52.3333	43.6	600	1.00
VZV	Pre-FMT vs. Healthy control	32.1292	41.4	600	1.00
VZV	Post-FMT vs. Healthy control	84.4625	41.4	600	1.00

4.5.3 Humoral Responses of CDI and rCDI Patients

Next, we assessed the differences in humoral immune responses between patients with initial CDI and rCDI. Pre-FMT sera from patients with rCDI was utilised for this comparison. The investigation focused exclusively on antigens eliciting positive antibody responses from the four antigen subsets.

4.5.3.1 IgG

The results demonstrated that IgG reactivity differed in response to antigen type across all antigen subsets (Table 4.9). Furthermore, significant differences were found in antigen-specific IgG responses between patient conditions (CDI and rCDI) across all antigen subsets (Table 4.9; Figure 28A, B, C and D).

Table 4.9. ANOVA type II test results across antigen subset GLM models for CDI vs. rCDI IgG responses.

*Likelihood Ratio Chi-Square (LR Chisq). $p < 0.05$ *, $p < 0.01$ **, $p < 0.001$ ***.*

Autoantigens IgG	LR Chisq	df	p-value
Antigen	73.161	10	1.084e-11 ***
Status	8.172	1	0.004254 **
Antigen : Status	53.769	10	5.357e-08 ***
Positive-Control antigens IgG	LR Chisq	df	p-value
Antigen	30.382	1	3.548e-08 ***
Status	69.183	1	< 2.2e-16 ***
Antigen : Status	0.042	1	0.8382
Pathogen-associated antigens IgG	LR Chisq	df	p-value
Antigen	16.108	3	0.001078 **
Status	43.396	1	4.472e-11 ***
Antigen : Status	54.259	3	9.883e-12 ***
Viral recall antigens IgG	LR Chisq	df	p-value
Antigen	31.863	7	4.305e-05 ***
Status	54.597	1	1.480e-13 ***
Antigen : Status	66.802	7	6.528e-12 ***

In Figure 28A, pairwise comparisons found that rCDI patients exhibited significantly greater IgG responses to the autoantigen, gliadin, compared to patients with CDI (Table 4.10; $p < 0.0001$). Figure 28B shows rCDI patients with significantly higher IgG responses to the positive-control antigens, *Candida albicans* and tetanus toxoid, than CDI patients (Table 4.10; $p < 0.0001$). In Figure 28C, rCDI patients exhibited significantly elevated IgG responses to the pathogen-associated antigens, TcdA and peptidoglycan *S. aureus*, compared to CDI patients (Table 4.10; TcdA $p < 0.0001$, PDG *S. aureus* $p < 0.01$). Finally, Figure 28D displays rCDI patients with significantly higher IgG responses to the viral recall antigens, CMV and HCoV-HKU1, than CDI patients (Table 4.10; CMV $p < 0.0001$, HCoV-HKU1 $p < 0.001$). Across all significant pairwise comparisons, rCDI patients exhibited greater antigen-specific IgG responses than CDI patients.

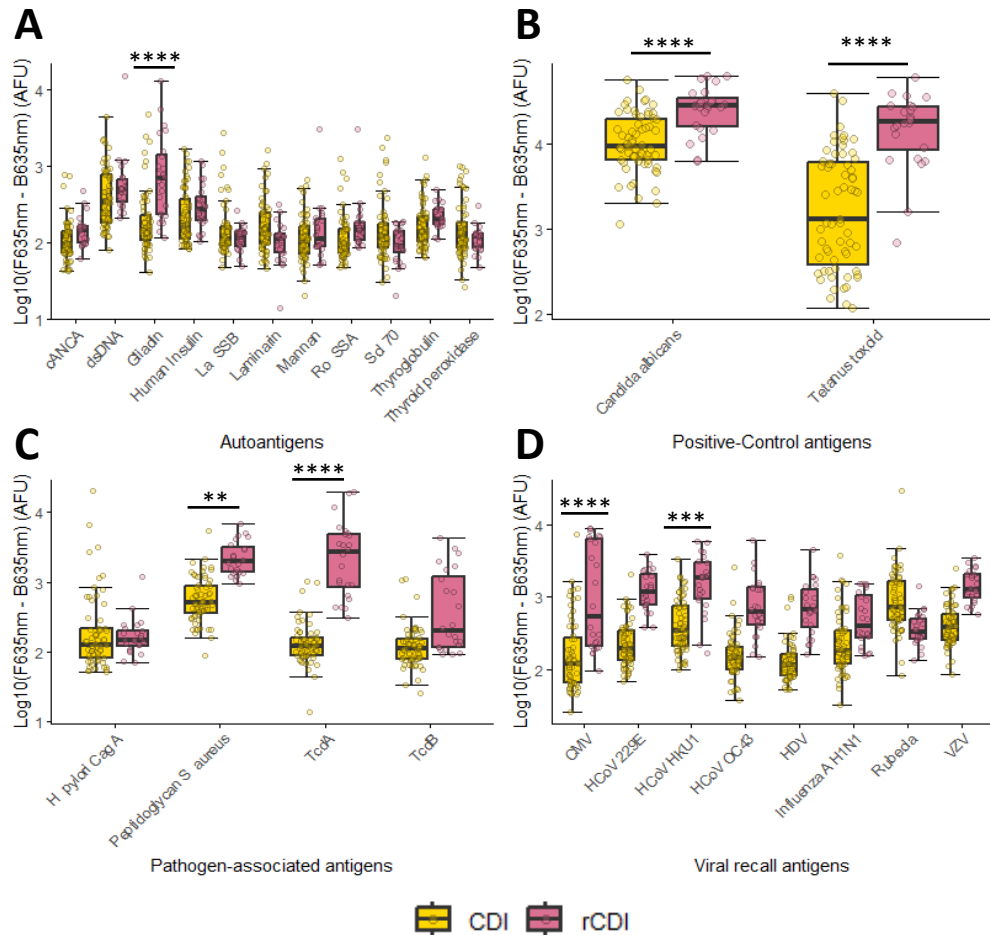


Figure 28. IgG responses in patients with CDI and rCDI across autoantigens (A), positive-controls (B), pathogen-associated (C) and viral recall antigens (D).

Differences between groups were assessed using Generalised Linear Models (GLM) with fixed effects of patient condition (CDI vs. rCDI) and antigen choice. The models were evaluated for significance through ANOVA type II tests and Tukey's *p*-value adjustment for multiple pairwise comparisons. The analysis found significant differences in antigen-specific IgG responses between patient conditions (CDI and rCDI) across all antigen subsets. Pairwise comparisons revealed that (A) rCDI patients had significantly higher IgG responses to gliadin than CDI patients. (B) rCDI patients had significantly greater IgG responses to both *Candida albicans* and tetanus toxoid than CDI patients. (C) rCDI patients had significantly elevated IgG responses to TcdA and peptidoglycan *S. aureus* than CDI patients. (D) rCDI patients had significantly greater IgG responses to both CMV and HCoV-HKU1 than CDI patients. $p < 0.01$ **, $p < 0.001$ ***, $p < 0.0001$ ****. Sample sizes: CDI (n=63), rCDI (n=24). Note that the Y-axis scale is log₁₀() due to outliers (strong or weak antibody responses). Error bars display the standard error of the mean. F635nm – B635nm; Foreground 635nm – Background 635nm, AFU; Arbitrary fluorescent units.

Table 4.10. Pairwise comparisons of antigen-specific IgG responses between CDI and rCDI patients.

Across autoantigen, positive-control, pathogen-associated and viral recall antigen subsets. $p < 0.05$ *, $p < 0.01$ **, $p < 0.001$ ***, $p < 0.0001$ ****.

Autoantigens	Comparison	Estimate	SE	df	p-value
cANCA	CDI vs rCDI	-23.130	177	935	1.00
dsDNA	CDI vs rCDI	-557.649	177	935	0.18
Gliadin	CDI vs rCDI	-1239.109	177	935	<0.0001
Human Insulin	CDI vs rCDI	-47.613	177	935	1.00
Thyroid peroxidase	CDI vs rCDI	91.025	177	935	1.00
La SSB	CDI vs rCDI	107.340	177	935	1.00
Laminarin	CDI vs rCDI	117.149	177	935	1.00
Mannan	CDI vs rCDI	-114.712	177	935	1.00
Ro SSA	CDI vs rCDI	-124.376	177	935	1.00
Scl-70	CDI vs rCDI	138.843	177	935	1.00
Thyroglobulin	CDI vs rCDI	-21.839	177	935	1.00
Positive-Control Antigens	Comparison	Estimate	SE	df	p-value
<i>Candida albicans</i>	CDI vs rCDI	-16276	2701	170	<0.0001
Tetanus toxoid	CDI vs rCDI	-15495	2701	170	<0.0001
Pathogen-associated Antigens	Comparison	Estimate	SE	df	p-value
<i>H. pylori</i> Cag A	CDI vs rCDI	512.3	449	340	0.9472
Peptidoglycan <i>S. aureus</i>	CDI vs rCDI	-1767.6	449	340	0.0025
TcdA	CDI vs rCDI	-3978.9	449	340	<0.0001
TcdB	CDI vs rCDI	-680.5	449	340	0.7986
Viral recall Antigen	Comparison	Estimate	SE	df	p-value
CMV	CDI vs rCDI	-2554.29	344	680	<0.0001
HCoV-229E	CDI vs rCDI	-1163.80	344	680	0.0597
HCoV-HKU1	CDI vs rCDI	-1556.61	344	680	0.0008
HCoV-OC34	CDI vs rCDI	-894.73	344	680	0.4003
HDV	CDI vs rCDI	-806.58	344	680	0.5910
Influenza A H1N1	CDI vs rCDI	-282.88	344	680	1.0000
Rubeola	CDI vs rCDI	1160.80	344	680	0.0613
VZV	CDI vs rCDI	-1094.22	344	680	0.1074

4.5.3.2 IgA

Next, the GLM analysis was reiterated to evaluate the influence of patient condition (CDI vs. rCDI) on antigen-specific IgA responses. The results demonstrated that IgA reactivity differed in response to antigen type across all antigen subsets (Table 4.11). Significant differences were found in antigen-specific IgA responses between patient conditions (CDI and rCDI) in the autoantigen, positive-control and pathogen-associated antigen subsets (Table 4.11; Figure 29A, B, C). However, antigen-specific IgA responses did not significantly differ between patient conditions in the viral recall antigen subset (Table 4.11; Figure 29D).

Table 4.11. ANOVA type II test results across antigen subset GLM models for CDI vs. rCDI IgA responses.

*Likelihood Ratio Chi-Square (LR Chisq). $p < 0.05$ *, $p < 0.01$ **, $p < 0.001$ ***.*

Autoantigens IgA	LR Chisq	df	p-value
Antigen	239.295	10	< 2.2e-16 ***
Status	4.851	1	0.02762 *
Antigen : Status	36.349	10	7.33e-05 ***
Positive-Control antigens IgA	LR Chisq	df	p-value
Antigen	71.597	1	< 2e-16 ***
Status	5.385	1	0.02031 *
Antigen : Status	5.485	1	0.01918 *
Pathogen-associated IgA	LR Chisq	df	p-value
Antigen	132.874	3	< 2e-16 ***
Status	2.035	1	0.15372
Antigen : Status	10.320	3	0.01603 *
Viral recall antigens IgA	LR Chisq	df	p-value
Antigen	38.389	7	2.556e-06 ***
Status	4.631	1	0.0314 *
Antigen : Status	6.238	7	0.5123

In Figure 29A, pairwise comparisons established that CDI patients exhibited significantly higher IgA responses to the autoantigen, dsDNA, than rCDI patients (Table 4.12; $p < 0.0001$). Figure 29B shows that CDI patients demonstrated significantly higher IgA responses to the positive-control antigen, *Candida albicans*, than rCDI patients (Table 4.12; $p < 0.01$). Lastly, in Figure 29C, CDI patients displayed significantly elevated IgA responses to the pathogen-associated antigen, peptidoglycan *S. aureus*, compared to rCDI patients (Table 4.12; $p < 0.05$). Across all significant pairwise comparisons, CDI patients displayed higher antigen-specific IgA responses compared to rCDI patients.

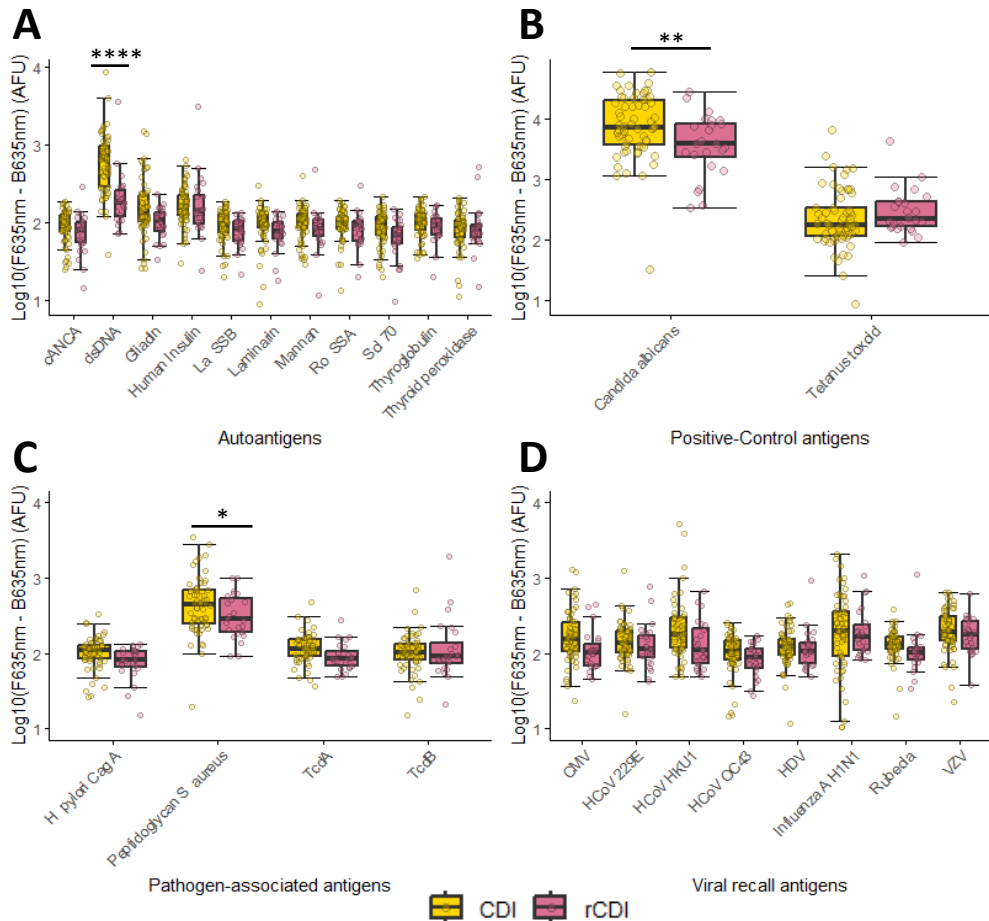


Figure 29. IgA responses in patients with CDI and rCDI across autoantigens (A), positive-controls (B), pathogen-associated (C) and viral recall antigens (D).

Differences between groups were assessed using Generalised Linear Models (GLM) with fixed effects of patient condition (CDI vs. rCDI) and patient choice. The models were evaluated for significance through ANOVA type II tests and Tukey's *p*-value adjustment for multiple pairwise comparisons. The analysis found significant differences in antigen-specific IgA responses between patient conditions (CDI and rCDI) across all antigen subsets except for the viral recall antigens. Pairwise comparisons revealed that: (A) CDI patients had significantly higher IgA responses to dsDNA compared to rCDI patients. (B) CDI patients had significantly higher IgA responses to *Candida albicans* than rCDI patients. (C) CDI patients showed significantly elevated IgA responses to peptidoglycan *S. aureus* compared to rCDI patients. $p < 0.05^*$, $p < 0.01^{**}$, $p < 0.001^{***}$, $p < 0.0001^{****}$. Sample sizes: CDI ($n=63$), rCDI ($n=24$). Note that the Y-axis scale is $\log_{10}()$ due to outliers (strong or weak antibody responses). Error bars display the standard error of the mean. F635nm – B635nm; Foreground 635nm – Background 635nm, AFU; Arbitrary fluorescent units.

Table 4.12. Pairwise comparisons of antigen-specific IgA responses between CDI and rCDI patients.

Across autoantigen, positive-control, pathogen-associated and viral recall antigen subsets. $p < 0.05$ *, $p < 0.01$ **, $p < 0.001$ ***, $p < 0.0001$ ****.

Autoantigens	Comparison	Estimate	SE	df	p-value
cANCA	CDI vs rCDI	14.204	83	924	1.00
dsDNA	CDI vs rCDI	509.403	83	924	< 0.0001
Glialin	CDI vs rCDI	110.242	83	924	1.00
Human Insulin	CDI vs rCDI	-100.956	83	924	1.00
Thyroid peroxidase	CDI vs rCDI	-16.875	83	924	1.00
La SSB	CDI vs rCDI	14.655	83	924	1.00
Laminarin	CDI vs rCDI	23.692	83	924	1.00
Mannan	CDI vs rCDI	7.470	83	924	1.00
Ro SSA	CDI vs rCDI	10.847	83	924	1.00
Scl-70	CDI vs rCDI	20.150	83	924	1.00
Thyroglobulin	CDI vs rCDI	13.525	83	924	1.00
Positive-Control Antigen	Comparison	Estimate	SE	df	p-value
<i>Candida albicans</i>	CDI vs rCDI	6602.0	2002	168	0.0065
Tetanus toxoid	CDI vs rCDI	-30.4	2002	168	1.0000
Pathogen-associated Antigen	Comparison	Estimate	SE	df	p-value
<i>H. pylori</i> Cag A	CDI vs rCDI	24.3	72	336	1.00
Peptidoglycan <i>S. aureus</i>	CDI vs rCDI	235.5	72	336	0.026
TcdA	CDI vs rCDI	30.4	72	336	1.00
TcdB	CDI vs rCDI	-84.6	72	336	0.94
Viral recall Antigen	Comparison	Estimate	SE	df	p-value
CMV	CDI vs rCDI	99.804	73.7	672	0.99
HCoV-229E	CDI vs rCDI	9.441	73.7	672	1.00
HCoV-HKU1	CDI vs rCDI	197.652	73.7	672	0.34
HCoV-OC34	CDI vs rCDI	25.447	73.7	672	1.00
HDV	CDI vs rCDI	-8.864	73.7	672	1.00
Influenza A H1N1	CDI vs rCDI	87.026	73.7	672	1.00
Rubeola	CDI vs rCDI	-2.028	73.7	672	1.00
VZV	CDI vs rCDI	40.077	73.7	672	1.00

4.5.4 Overview of Headline Findings

Overall, this section presents the results of screening clinically defined patient cohorts, including those with CF, CDI, and recurrent CDI (rCDI) pre- and post-FMT, alongside healthy controls. We aimed to elucidate humoral immune responses to a diverse range of antigens across cohorts, with a particular interest in those at heightened susceptibility to immune-mediated disease. The antigen panel included autoantigens associated with chronic immune-mediated diseases alongside pathogen-associated and viral recall antigens to investigate infection. We explored the commonalities and differences in humoral immune responses across and within-subject cohorts through hierarchical clustering analysis. Additionally, we delved into the impact of FMT on humoral responses in rCDI patients and examined variations in immune responses between initial CDI and recurrent CDI patient cohorts. Table 4.13 outlines the headline findings from each comparative analysis, providing a roadmap for the subsequent discussions.

Table 4.13 Summary of the headline findings for the application of the multi-antigen microarray results section.

Comparisons	Headline Findings
4.5.1 Hierarchical Clustering Analysis	Identified positive IgG autoantibody responses across subject cohorts.
	CF patients clustered based on IgG responses to autoantigens.
	CF and rCDI cohorts clustered based on IgG responses to pathogen-associated antigens.
	IgA autoantibody responses were mainly observed in CF and CDI cohorts.
4.5.2 Humoral Responses of rCDI Patients Before and After FMT	There were no significant differences in antigen-specific IgG and IgA responses between the pre-FMT and 12-week post-FMT profiles of rCDI patients across all antigen subsets.
	Pre- and post-FMT rCDI groups had significantly higher IgG responses to TcdA than the HC group.
	Pre-FMT rCDI patients had significantly elevated IgG anti-CMV responses compared to HCs.
4.5.3 Humoral Responses of CDI and rCDI Patients	Across all significant pairwise comparisons, rCDI patients exhibited greater antigen-specific IgG responses than CDI patients.
	Across all significant pairwise comparisons, CDI patients displayed higher antigen-specific IgA responses compared to rCDI patients.

5 Discussion

In this proof-of-concept study, we have developed a preliminary immune health screening platform, enabling the simultaneous assessment of multiple antigen-specific antibody responses. The goal was to develop, optimise and implement an antigen microarray to efficiently profile individuals at heightened risk of chronic immune-mediated disease and healthy controls, offering comprehensive insights into their humoral antibody responses. We show that our multiplexed platform can precisely quantify IgG and IgA autoantibody, pathogen-associated, and viral recall antibody responses across clinically defined patient cohorts. Moreover, our results demonstrate the broad humoral response profiles detected through multiplexed screening. Overall, we showcase the practicality of antigen microarrays in exploring immune health, including identifying potential biomarkers of disease, assessing responses to FMT therapy, and elucidating differences in adaptive humoral immunity between CDI and rCDI patients.

5.1.1 Optimisation and Validation

Our initial experiments focused on developing, optimising and validating an antigen microarray that would allow the simultaneous detection of numerous biomarkers. We optimised multiple parameters such as different slide surfaces, serum dilutions, assay reagent dilutions, blocking buffers and scan settings. These optimisation experiments ensured high signal intensities, low background fluorescence, and low non-specific binding. Regarding validation,

in adherence to FDA guidelines for precision, we established a precision limit of equal to or less than 20% coefficient of variation (CV) (U.S. Department of Health and Human Services, 2018). Our findings underscore the reproducibility of the microarray platform for antibody detection, as 93% of the antigens met this precision limit for intra-assay CV and 90% for inter-assay CV. However, further research is required to test the level of accuracy across our 60 antigens regarding sensitivity and specificity. In addition, future work could correlate microarray and ELISA results to determine the level of agreement in antibody detection.

5.1.2 Antibody Response Profiles and Autoantibody Responses

Once the platform was optimised, we tested clinically-defined patient sera to generate antibody response profiles. We identified positive IgG and IgA autoantibody responses in patients diagnosed with CF, CDI, and rCDI, as well as in healthy controls. Positive responses demonstrate that autoantibodies are present within the systemic circulation, which may reflect preclinical disease or natural autoantibody activity. Many previous studies have identified circulating autoantibodies in pre-symptomatic patients up to 20 years before the onset of the autoimmune disease. For instance, Theander et al., 2015, identified anti-nuclear antibodies, rheumatoid factor, and antibodies against Ro 60/SSA 19-20 years before SjS diagnosis. Johansson et al., 2016, found anti-VCP and anti-CCP antibodies directed against citrullinated antigens, pre-dating symptom onset in RA patients. These findings underscore the potential of circulating autoantibodies as measurable

preclinical disease biomarkers, offering the possibility of earlier identification of disease and intervention. However, IgM and low-titre IgG natural autoantibody responses are also detected in healthy individuals (Elkon and Casali, 2008b). Hence, autoantibodies are not always disease-associated and may reflect natural autoantibody repertoires involved in immune homeostasis and cell debris clearance (Nagele *et al.*, 2013). Therefore, future research could prioritise establishing a discernible threshold to differentiate between physiological levels of natural autoantibodies and pathological levels of autoantibody responses indicative of disease development.

Interestingly, we identified rare yet strong specific IgG autoantibody responses to carbamylated-BSA, BPI, cANCA, Scl-70, Ro-SSA, thyroid peroxidase, laminarin, and mannan. Antibody responses to these markers are associated with chronic immune-mediated diseases such as RA, vasculitis, IBD, scleroderma, SjS, SLE and autoimmune thyroiditis. In particular, we found IgG antibody responses to bactericidal permeability-increasing protein (BPI) and carbamylated-BSA in the CF cohort. Our results are congruent with previous research by Skopelja *et al.*, 2016, who also found autoantibodies directed against BPI and carbamylated proteins in CF patients. Cystic fibrosis is an inherited condition that causes mucus build-up in the lungs and digestive system, which is associated with chronic infection by *Pseudomonas aeruginosa* (Lyczak, Cannon and Pier, 2002). It is hypothesised that the CF transmembrane regulator protein may directly alter the immune response, intraluminal dysbiosis and inflammation, establishing an environment that leads to autoimmunity (Chadwick *et al.*, 2023). Skopelja *et al.*, 2016, indicated

a role for autoimmunity in CF disease severity as BPI was associated with diminished lung function and *Pseudomonas aeruginosa* infection. In contrast, anti-carbamylated responses are biomarkers for RA (Shi *et al.*, 2014). The alignment between our findings and prior research underscores the significance of autoantibody responses for investigating autoimmunity and the utility of multiplexed antigen microarrays for future exploration.

Detecting robust autoantibody responses, such as these rare cases, has potential applications in identifying early or co-morbid pathological autoimmunity. However, it is important to recognise that our limited sample size and one-time sampling restrict the extent of conclusive interpretation. A shortage of samples with established autoimmune and co-morbid diseases constrained our ability to draw statistically significant conclusions regarding potential pre-clinical or multimorbid biomarkers. However, our findings remain preliminary and should serve as a catalyst for future research endeavours.

5.1.3 Hierarchical Clustering Analysis

To gain further insight into our large dataset, we conducted hierarchical clustering to visualise commonalities and distinctions between humoral immune response profiles. We found patients clustered based on their unique responses to specific antigens and the magnitude of antibody responses.

For example, we found CF and rCDI patients clustered in response to our panel of pathogen-associated antigens due to similar patterns in IgG antibody responses. CF and rCDI cohorts showed higher IgG responses to TcdA and

TcdB compared to CDI patients and healthy controls. TcdA and TcdB are exotoxins produced by *C. difficile*, involved in CDI pathogenesis, causing toxin-mediated damage to the intestinal barrier and microbial translocation (Leffler and Lamont, 2015). Our results may reflect repeat exposure to *C. difficile* in CF and rCDI patients due to extended hospital stretches, inducing greater antibody responses. However, our cohort of eleven CF patients was too small to make any statistical comparisons and conclusions. Prior research by Monaghan et al., 2017, found elevated IgA anti-TcdA and anti-TcdB levels in CF patient sera but no elevated IgG anti-TcdA responses compared to CDI and healthy controls, indicating robust anti-toxin responses in CF patients may confer protection against CDI.

Interestingly, rCDI patients displayed greater IgG responses to TcdA and TcdB than those with initial CDI. Within the rCDI cohort, we observed clustering based on elevated antibody responses to the *C. difficile* toxins. However, we did not observe distinct clustering patterns between rCDI and initial CDI patient groups. This observation may be attributed to the chronicity and severity of infection in recurrent cases, characterised by prolonged antigen exposure and recurring immune stimulation. However, this challenges the idea that higher antibody responses to *C. difficile* toxins may confer protection against CDI. Potential explanations for this observation include underlying immune dysfunction in rCDI patients, indicating that factors beyond levels of antibody responses may contribute to the susceptibility to recurrent CDI.

In addition, we observed CF patients grouping due to robust IgA antibody responses to gliadin, carbamylated-BSA and dsDNA. Similarly, multiple previous studies have identified elevated antibody responses to gliadin, a component of wheat, and double-stranded DNA (dsDNA), a self-antigen, in the sera of individuals with CF compared to controls. For instance, Troncone et al., 1994, found elevated IgA antibody titres to the alimentary antigen gliadin. Generally, high titres of anti-gliadin are detected in individuals with celiac disease and gastrointestinal diseases (Ben-Ami Shor, Papageorgiou and Shoenfeld, 2017). Interestingly, CF patients often have non-CF-related symptoms and aetiology, such as gastrointestinal complaints and a high frequency of celiac disease (Chadwick *et al.*, 2023). Yadav et al., 2020, found elevated levels of IgA anti-dsDNA in CF patients compared to age-matched control subjects, hypothesising that the chronic presence of DNA in CF airways may induce autoantibody production against host double-stranded DNA. Hence, our results are congruent with previous research identifying autoantibody responses in CF patients, which may indicate pathological autoimmunity. The collective evidence from our study and previous research suggests that autoimmune processes may contribute to the clinical heterogeneity observed in CF patients, including non-CF-related symptoms and complications. Nevertheless, the limited CF sample size meant definitive statistical conclusions could not be drawn.

Overall, our clustering analysis yielded intriguing findings, but our interpretation of results is limited without larger patient cohorts and repeated sampling. Despite this, our preliminary findings highlight the

importance of comprehensive multiplexed assays for investigating relationships and trends between adaptive humoral responses across multiple immunoglobulin isotypes in disease and healthy control cohorts. In addition, hierarchical clustering of humoral antibody responses has relevant clinical applications for identifying patients with similar disease characteristics, risk of autoimmune disease development, underlying pathological mechanisms, or predicting treatment outcomes.

5.1.4 Humoral Responses of rCDI Patients Before and After FMT

Next, we aimed to gain deeper insights into the humoral immune responses of rCDI patients who had successfully undergone FMT. FMT is a treatment for severe or rCDI that reconstitutes the gut microbiota and modulates several immune pathways critical for recovery (Frisbee and Petri, 2020). Our findings revealed no differences in serological antigen-specific IgG and IgA responses to autoantigens, pathogen-associated and viral recall antigens between the pre-FMT and 12-week post-FMT profiles of rCDI patients. These results may suggest that either the mechanistic action of FMT did not impact the adaptive humoral responses of our rCDI patient cohort or, more likely, the relatively short 12-week observation period was insufficient to elucidate any changes in antibody responses.

Based on prior research, we would have expected an increase in IgG and IgA responses to the *C. difficile* exotoxins, TcdA and TcdB, after FMT. A prior study by Cook et al., 2021, examined the antibody responses of 22 patients with rCDI before and eight weeks after FMT using ELISAs and found increased

proportions of serological anti-toxin antibodies and TcdB-specific Th17 cells, indicating FMT affects both humoral and cellular immunity. As a result, they hypothesised that an enhancement in protective adaptive immunity may have contributed to FMT treatment success. However, it is important to acknowledge that the variance in our results may be attributed to differences between assays, different toxigenic virulence strains between patient cohorts and the nature of memory B cell responses across individuals. Additionally, our studies used binding titres to quantify antibody responses to specific antigens serologically, so our results are less relevant for conferring protective immunity. Lastly, both our studies comprise relatively small sample sizes, limiting the applicability and generalisability of our research findings to the broader patient group. Expanding our investigations to encompass larger cohorts would enhance statistical power and increase the generalisability of our findings.

Notably, our investigation found no differences in autoantibody responses 12 weeks after FMT; this suggested that FMT did not alter the autoimmune profile of patients with rCDI in the short term. At present, it is hard to relate our findings to prior work, as research into the impact of FMT has focused on microbial and metabolic effects on rCDI. However, research into adaptive immunity post-FMT is essential, as the gut microbiota plays a significant role in shaping the immune system, and alterations in the microbiota can influence host immune responses (Yang, Chen and Cai, 2023). This is exemplified as FMT is now being explored as a potential treatment for certain autoimmune conditions, such as IBD. For example, FMT appears effective for

remission induction in ulcerative colitis, demonstrating that restoring the gut microbiota beneficially impacts immune system function (Costello *et al.*, 2017). Although we found no difference in humoral response profiles after FMT, our findings should inform the design of FMT clinical trials. Particularly to emphasise the importance of serologically profiling the humoral responses of patients in the long term after FMT at intervals of at least six months and across large cohorts.

Nonetheless, our results did identify differences in IgG and IgA responses to pathogen-associated and viral recall antigens, distinguishing between the rCDI and HC cohorts. Specifically, pre- and post-FMT groups had higher IgG responses to TcdA than the HC group. The higher anti-toxin antibody responses in rCDI patients may have arisen from repeat exposure to *C. difficile*, eliciting humoral responses on multiple occasions compared to healthy individuals.

Moreover, our analysis identified pre-FMT sera with elevated IgG anti-CMV responses compared to HCs, indicating a heightened immune response to CMV in rCDI patients. This heightened response may be attributed to co-infection or reactivation of latent CMV due to gut microbiota dysbiosis and immune dysregulation. Prior research has established a history of CMV infection as a significant risk factor for *C. difficile* infection, offering insight into our findings (Hung *et al.*, 2015). Interestingly, the post-FMT sera did not significantly differ in anti-CMV IgG antibody responses compared to pre-FMT and HC sera. This observation raises the possibility that FMT may impact the

virome. However, further investigation through more extensive studies and longer-term post-FMT sample analysis would be warranted.

Lastly, post-FMT sera demonstrated elevated IgA anti-influenza A responses compared to HCs. These variations between viral antibody responses are likely attributable to gut dysbiosis and impaired intestinal barriers in rCDI patients from which viruses can translocate into the bloodstream, triggering elevated humoral immune responses.

5.1.5 Humoral Responses of CDI and rCDI Patient Cohorts

In our final phase of analysis, we compared antigen-specific IgG and IgA antibody responses between patients with initial CDI and rCDI to unveil distinctions in their adaptive humoral responses. This investigation revealed significant differences between the two patient groups in antigen-specific IgG and IgA humoral responses. Across the antigen-specific differences, we identified that rCDI patients exhibited higher IgG responses, whereas patients with initial CDI displayed higher IgA responses. IgG is the predominant immunoglobulin isotype in sera during the adaptive humoral response. IgA is the second most prevalent serological immunoglobulin and is pivotal in the host defence of mucosal surfaces. Previous research into the adaptive immunity of rCDI has indicated a reduction in mucosal IgA-producing B cells (Johal, 2004). Therefore, our rCDI sera may show lower antigen-specific IgA responses than initial CDI sera due to compromised local mucosal IgA impacting systemic IgA responses. The literature indicates that mucosal responses may impact the systemic IgA pool by releasing gut antigen-specific

IgA into the bloodstream from the marginal zones of the spleen (Leong and Ding, 2014). However, it remains uncertain to what extent serological IgA reflects local mucosal IgA responses.

Notably, a larger proportion of CDI patients were under immunosuppressive treatments. Immunosuppression typically results in attenuated adaptive humoral responses, including reduced antibody production, which may explain why we observed lower IgG responses in CDI patients to positive control antigens such as *Candida albicans* and tetanus toxoid, as well as viral recall antigens like seasonal human coronavirus HKU1 (HCoV-HKU1) and CMV, compared to rCDI patients.

Interestingly, we found that rCDI patients had elevated IgG antibody responses to gliadin and TcdA compared to CDI patients. Gliadin, a wheat component, triggers antibody production in individuals with gastrointestinal autoimmune diseases like celiac disease (Ben-Ami Shor, Papageorgiou and Shoenfeld, 2017). The gut mucosa is the first line of defence against microbial and alimentary antigens such as gliadin. Conditions that cause pro-inflammatory states, such as *C. difficile* infection, cause damage to the intestinal barrier and dysregulate the immune response and are exacerbated when infections are severe and recurrent (Leffler and Lamont, 2015).

Therefore, ongoing exposure to translocated gliadin in rCDI patients may explain the elevated IgG responses to the antigen compared to CDI patients.

Regarding the exotoxin TcdA, our results demonstrate elevated IgG anti-TcdA responses in rCDI patients compared to CDI patients. This may be due to

repeat exposure to *C. difficile* in rCDI compared to initial CDI patients. As our assay quantifies antibody binding to specific antigens, we cannot comment specifically on neutralising antibody responses that confer protective immunity to CDI. However, many prior studies show contrasting information on the importance of anti-TcdA and anti-TcdB responses in CDI and rCDI disease pathogenesis (Rees and Steiner, 2018).

Patients with initial CDI exhibited elevated IgA responses to the antigens, *Candida albicans*, peptidoglycan from *Staphylococcus aureus* (*S. aureus*), and dsDNA compared to rCDI patients. *Candida albicans* and peptidoglycan from *S. aureus* are antigens associated with opportunistic pathogens that can exist asymptotically within the human microbiome. Gut microbiota dysbiosis alters the balance of commensal microbes in the gut, enabling colonisation by pathogens such as *C. difficile* (Pickard *et al.*, 2017). As a result, CDI impairs the intestinal barrier, increasing permeability and microbial translocation (Leffler and Lamont, 2015). Our results may suggest CDI patients mount more robust IgA-mediated responses to translocated pathogens compared to rCDI patients.

Anti-dsDNA autoantibodies, which target double-stranded DNA, are typically associated with SLE when present in high titres (Lee, 2022). The understanding of IgA anti-dsDNA remains limited, although it is probable that these autoantibodies originate from ongoing self-exposure to host DNA in mucosal areas and are associated with colonic damage or pathological autoimmunity (Lee, 2022). Our results may demonstrate that CDI patients

mount greater IgA anti-dsDNA responses than rCDI patients due to superior IgA-mediated immune responses.

However, it is important to acknowledge the limitations of our study. While we could obtain medical histories for the rCDI patient group, details about existing conditions were unavailable for patients with initial CDI. Therefore, other co-morbidities and treatments may have influenced the differences we identified in humoral responses between initial and recurrent CDI cohorts. Additionally, unequal sample sizes between the CDI and rCDI cohorts, driven by limited sera availability, may have affected our statistical power to detect genuine differences. Future research should prioritise larger, evenly distributed sample cohorts and include control groups with similar or matched co-morbidities to enhance our understanding of CDI and rCDI-associated humoral responses.

Despite these limitations, our study contributes new insights into the antigen repertoire detected by patients with CDI and rCDI and highlights distinctions in their adaptive humoral responses. Further exploration could provide valuable insights into how the adaptive immune system responds to initial versus recurrent CDI to detect potential biomarkers for CDI recurrence and assess the risk of pathological autoimmunity.

5.2 Strengths of The Antigen Microarray Platform

Our developed and optimised antigen microarray offers several advantages over conventional ELISAs. These advantages stem from the multiplexed nature of the assay, enabling the simultaneous quantification of numerous antibody responses while using submicrolitre volumes of sera to conserve valuable samples. Firstly, the assay demonstrates cost efficiency due to the minimal reagent volumes required, including nanolitre volumes of antigens for array printing. Secondly, the microarray assay is efficient and has a high throughput, as demonstrated by the ability to print and probe sixteen arrays per slide, facilitating paralleled probing of many samples. Lastly, the versatility of the microarray platform allows for easy customisation and expansion of the antigen panel to accommodate various research endeavours. In contrast, ELISA assays often necessitate multiple plates with larger sample and reagent volumes, leading to increased costs per test, all while being limited to screening for a single biomarker at a time. As a result, antigen microarrays hold significant promise for antibody response profiling, including autoantibody detection, and with further research and development, they could potentially replace ELISAs for clinical serological screening.

5.3 Limitations and Future Research

In this study, we have developed a preliminary immune health screening platform that, with further research, could be applied clinically as an early warning system for immune-mediated disease and multimorbidity. However, several limitations were encountered, which should inform future research endeavours.

Firstly, our sample sizes were relatively small, and sample numbers were unevenly distributed across patient cohorts. These constraints restricted the generalisability and statistical power of our findings. Future research endeavours should prioritise profiling more extensive and evenly distributed clinically-defined serum samples encompassing various disease states. A larger study would also enable the quantification of humoral response profiles across different demographics, including gender, age and ethnicity.

Secondly, an expansion of the study should involve a larger pool of healthy controls and patients with a broader spectrum of diagnosed autoimmune diseases. This approach would facilitate the exploration of pre-clinical biomarkers of pathological autoimmunity. The extended research could help establish antigen-specific thresholds associated with early or multi-morbid disease states, contributing to a comprehensive understanding of the clinical relevance of the autoantibody responses.

Thirdly, the comparability of humoral immune response profiles was limited in patients undergoing immunosuppressant treatment due to the potential

attenuation of immune responses, consisting of a proportion of the CDI and rCDI cohorts in this study. Immunosuppressants can alter the immune profile, increase susceptibility to infections, and introduce variability in treatment regimens, making interpreting differences in antibody responses between different patient groups challenging. Hence, future research may consider grouping immunosuppressed individuals into a distinct patient cohort and incorporating larger cohorts of patients with parallel medical conditions who are not undergoing immunosuppressant therapy.

Fourthly, availability and cost considerations influenced the selection of antigens on our microarray. For example, our panel lacked a range of *C. difficile* and IBD-related antigens, which would have been interesting to investigate in the disease cohorts. However, our assay is easily expandable, so there is room for including numerous additional pathogen-associated antigens and autoantigens associated with a broader array of autoimmune diseases. Initially, we recommend incorporating a selection of autoantigens linked to the most prevalent immune-mediated conditions, including but not limited to RA, T1D, MS, IBD, celiac disease, and psoriasis. Future research should begin by comprehensively exploring a diverse array of established immunological markers from the literature and then refine the selection to identify the most valuable ones for constructing an optimised marker panel.

Lastly, while our antigen microarray assay was conducted manually, there is potential for automation using liquid handling robots. This automation would

enhance scalability and throughput, reducing human error and allowing for more efficient data collection and analysis.

Additional avenues of research include expanding our array to add capture antibodies alongside antigens for detecting soluble proteins in sera. As a result, proteins associated with pathological autoimmunity, such as inflammatory cytokines, could be detected to provide further insight into an individual's adaptive immunity. Such an approach would effectively merge antigen and antibody microarrays. However, successful implementation would necessitate the identification of a slide surface chemistry capable of accommodating both antigens and antibodies.

5.4 Conclusions

In conclusion, our proof-of-concept study highlights the numerous advantages of our antigen microarray compared to traditional immunoassays, owing to its high-throughput and multiplexed detection capabilities. In addition, we have successfully demonstrated the efficacy of our assay for quantifying serological IgG and IgA autoantibody, pathogen-associated, and viral recall antibody responses across diverse patient cohorts, including those at an elevated risk of chronic immune-mediated diseases, as well as in healthy individuals. We found that FMT did not alter the humoral immune response profiles of rCDI patients after 12 weeks, although further research should be conducted in the long term as a part of FMT clinical trials. Additionally, we demonstrated antigen-specific differences in humoral responses between initial CDI and recurrent CDI to understand differences in adaptive immunity. As a result, our study underscores the breadth and magnitude of adaptive humoral responses that can be detected with our multiplexed assay, showcasing its versatility in measuring the heterogeneity between immune responses. Regarding future clinical applications, our immune health screening platform holds promise for annual profiling to identify early indicators of immune-mediated diseases, facilitate multimorbidity detection, and enable patient stratification for improved prognosis and treatment outcomes.

Our study lays the foundation for a promising immune health screening platform, serving as a catalyst for future research.

6 Bibliography

- Ahearn, J. *et al.* (2015) 'Cardiovascular disease biomarkers across autoimmune diseases', *Clinical Immunology*, 161(1), pp. 59–63. Available at: <https://doi.org/10.1016/j.clim.2015.05.024>.
- Angum, F. *et al.* (2020) 'The Prevalence of Autoimmune Disorders in Women: A Narrative Review', *Cureus* [Preprint]. Available at: <https://doi.org/10.7759/cureus.8094>.
- Avrameas, S., Alexopoulos, H. and Moutsopoulos, H.M. (2018) 'Natural Autoantibodies: An Undersign Hero of the Immune System and Autoimmune Disorders—A Point of View', *Frontiers in Immunology*, 9. Available at: <https://doi.org/10.3389/fimmu.2018.01320>.
- Bagdasarian, N., Rao, K. and Malani, P.N. (2015) 'Diagnosis and Treatment of Clostridium difficile in Adults', *JAMA*, 313(4), p. 398. Available at: <https://doi.org/10.1001/jama.2014.17103>.
- Di Bella, S. *et al.* (2016) 'Clostridium difficile Toxins A and B: Insights into Pathogenic Properties and Extraintestinal Effects', *Toxins*, 8(5), p. 134. Available at: <https://doi.org/10.3390/toxins8050134>.
- Belvoncikova, P., Maronek, M. and Gardlik, R. (2022) 'Gut Dysbiosis and Fecal Microbiota Transplantation in Autoimmune Diseases', *International Journal of Molecular Sciences*, 23(18), p. 10729. Available at: <https://doi.org/10.3390/ijms231810729>.
- Ben-Ami Shor, D., Papageorgiou, N.P. and Shoenfeld, Y. (2017) 'Autoantibodies in Gastrointestinal Autoimmune Diseases', in Manuel Ramos-Casals *et al.* (eds) *Handbook of Systemic Autoimmune Diseases*, pp. 49–66. Available at: <https://doi.org/10.1016/B978-0-444-63707-9.00003-9>.
- Bernstein, C.N., Wajda, A. and Blanchard, J.F. (2005) 'The Clustering of Other Chronic Inflammatory Diseases in Inflammatory Bowel Disease: A Population-Based Study', *Gastroenterology*, 129(3), pp. 827–836. Available at: <https://doi.org/10.1053/j.gastro.2005.06.021>.
- Beyi, A.F., Wannemuehler, M. and Plummer, P.J. (2022) 'Impacts of Gut Microbiota on the Immune System and Fecal Microbiota Transplantation as a Re-Emerging Therapy for Autoimmune Diseases', *Antibiotics*, 11(8), p. 1093. Available at: <https://doi.org/10.3390/antibiotics11081093>.
- Bourgonje, A.R. *et al.* (2022) 'Antibody signatures in inflammatory bowel disease: current developments and future applications', *Trends in Molecular*

- Medicine*, 28(8), pp. 693–705. Available at:
<https://doi.org/10.1016/j.molmed.2022.05.004>.
- Buckley, C.D. *et al.* (2021) 'Immune-mediated inflammation across disease boundaries: breaking down research silos', *Nature Immunology*, 22(11), pp. 1344–1348. Available at: <https://doi.org/10.1038/s41590-021-01044-7>.
- Caja, S. *et al.* (2011) 'Antibodies in celiac disease: implications beyond diagnostics', *Cellular & Molecular Immunology*, 8(2), pp. 103–109. Available at: <https://doi.org/10.1038/cmi.2010.65>.
- Caley, L.R. *et al.* (2023) 'Cystic Fibrosis-Related Gut Dysbiosis: A Systematic Review', *Digestive Diseases and Sciences*, 68(5), pp. 1797–1814. Available at: <https://doi.org/10.1007/s10620-022-07812-1>.
- Castro, C. and Gourley, M. (2010) 'Diagnostic testing and interpretation of tests for autoimmunity', *Journal of Allergy and Clinical Immunology*, 125(2), pp. S238–S247. Available at: <https://doi.org/10.1016/j.jaci.2009.09.041>.
- Chadwick, C. *et al.* (2023) 'Autoimmunity in people with cystic fibrosis', *Journal of Cystic Fibrosis* [Preprint]. Available at:
<https://doi.org/10.1016/j.jcf.2023.03.007>.
- Chang, C. and Gershwin, M.E. (2011) 'Drug-Induced Lupus Erythematosus', *Drug Safety*, 34(5), pp. 357–374. Available at:
<https://doi.org/10.2165/11588500-000000000-00000>.
- Chan, V.K.Y. *et al.* (2023) 'Treatment-resistant depression and risk of autoimmune diseases: evidence from a population-based cohort and nested case-control study', *Translational Psychiatry*, 13(1), p. 76. Available at:
<https://doi.org/10.1038/s41398-023-02383-9>.
- Choi, M.Y. and Costenbader, K.H. (2022) 'Understanding the Concept of Pre-Clinical Autoimmunity: Prediction and Prevention of Systemic Lupus Erythematosus: Identifying Risk Factors and Developing Strategies Against Disease Development', *Frontiers in Immunology*, 13. Available at:
<https://doi.org/10.3389/fimmu.2022.890522>.
- Cojocar, M. *et al.* (2011) 'Gastrointestinal manifestations in systemic autoimmune diseases.', *Maedica*, 6(1), pp. 45–51.
- Conrad, N. *et al.* (2023) 'Incidence, prevalence, and co-occurrence of autoimmune disorders over time and by age, sex, and socioeconomic status: a population-based cohort study of 22 million individuals in the UK', *The Lancet*, 401(10391), pp. 1878–1890. Available at: [https://doi.org/10.1016/S0140-6736\(23\)00457-9](https://doi.org/10.1016/S0140-6736(23)00457-9).

- Cook, L. *et al.* (2021) 'Fecal Microbiota Transplantation for Recurrent Clostridioides difficile Infection Enhances Adaptive Immunity to C difficile Toxin B', *Gastroenterology*, 160(6), pp. 2155-2158.e4. Available at: <https://doi.org/10.1053/j.gastro.2021.01.009>.
- Costello, S.P. *et al.* (2017) 'Systematic review with meta-analysis: faecal microbiota transplantation for the induction of remission for active ulcerative colitis', *Alimentary Pharmacology & Therapeutics*, 46(3), pp. 213-224. Available at: <https://doi.org/10.1111/apt.14173>.
- Cusick, M.F., Libbey, J.E. and Fujinami, R.S. (2012) 'Molecular Mimicry as a Mechanism of Autoimmune Disease', *Clinical Reviews in Allergy & Immunology*, 42(1), pp. 102-111. Available at: <https://doi.org/10.1007/s12016-011-8294-7>.
- DeGruttola, A.K. *et al.* (2016) 'Current Understanding of Dysbiosis in Disease in Human and Animal Models', *Inflammatory Bowel Diseases*, 22(5), pp. 1137-1150. Available at: <https://doi.org/10.1097/MIB.0000000000000750>.
- DeMarshall, C. *et al.* (2015) 'Utility of Autoantibodies as Biomarkers for Diagnosis and Staging of Neurodegenerative Diseases', in, pp. 1-51. Available at: <https://doi.org/10.1016/bs.irn.2015.05.005>.
- Van Dorst, J.M., Tam, R.Y. and Ooi, C.Y. (2022) 'What Do We Know about the Microbiome in Cystic Fibrosis? Is There a Role for Probiotics and Prebiotics?', *Nutrients*, 14(3), p. 480. Available at: <https://doi.org/10.3390/nu14030480>.
- Elkon, K. and Casali, P. (2008) 'Nature and functions of autoantibodies', *Nature Clinical Practice Rheumatology*, 4(9), pp. 491-498. Available at: <https://doi.org/10.1038/ncprheum0895>.
- Fischer, M. (2017) 'Clostridium difficile Infection and the Role of Adaptive Immunity in the Microbiome.', *Gastroenterology & hepatology*, 13(5), pp. 301-303.
- Floreani, A., Leung, P.S.C. and Gershwin, M.E. (2016) 'Environmental Basis of Autoimmunity', *Clinical Reviews in Allergy & Immunology*, 50(3), pp. 287-300. Available at: <https://doi.org/10.1007/s12016-015-8493-8>.
- Fox, J. and Weisberg, S. (2019) *An R Companion to Applied Regression*. Third edition. Thousand Oaks, CA: Sage. Available at: <https://socialsciences.mcmaster.ca/jfox/Books/Companion/> (Accessed: 4 September 2023).
- Frazzei, G. *et al.* (2022) 'Preclinical Autoimmune Disease: a Comparison of Rheumatoid Arthritis, Systemic Lupus Erythematosus, Multiple Sclerosis and

Type 1 Diabetes', *Frontiers in Immunology*, 13. Available at:
<https://doi.org/10.3389/fimmu.2022.899372>.

Frisbee, A.L. and Petri, W.A. (2020) 'Considering the Immune System during Fecal Microbiota Transplantation for *Clostridioides difficile* Infection', *Trends in Molecular Medicine*, 26(5), pp. 496–507. Available at:
<https://doi.org/10.1016/j.molmed.2020.01.009>.

Frommer, L. and Kahaly, G.J. (2020) 'Type 1 diabetes and associated autoimmune diseases', *World Journal of Diabetes*, 11(11), pp. 527–539. Available at: <https://doi.org/10.4239/wjd.v11.i11.527>.

Goronzy, J.J. and Weyand, C.M. (2012) 'Immune aging and autoimmunity', *Cellular and Molecular Life Sciences*, 69(10), pp. 1615–1623. Available at:
<https://doi.org/10.1007/s00018-012-0970-0>.

Grygiel-Górniak, B., Rogacka, N. and Puszczewicz, M. (2018) 'Antinuclear antibodies in healthy people and non-rheumatic diseases – diagnostic and clinical implications', *Rheumatology*, 56(4), pp. 243–248. Available at:
<https://doi.org/10.5114/reum.2018.77976>.

Hou, K. *et al.* (2022) 'Microbiota in health and diseases', *Signal Transduction and Targeted Therapy*, 7(1), p. 135. Available at:
<https://doi.org/10.1038/s41392-022-00974-4>.

Hsiao, Y.-H. *et al.* (2015) 'Sleep Disorders and Increased Risk of Autoimmune Diseases in Individuals without Sleep Apnea', *Sleep*, 38(4), pp. 581–586. Available at: <https://doi.org/10.5665/sleep.4574>.

Hung, Y.P. *et al.* (2015) 'Clinical impact of *Clostridium difficile* colonization', *Journal of Microbiology, Immunology and Infection*, 48(3), pp. 241–248. Available at: <https://doi.org/10.1016/j.jmii.2014.04.011>.

Janeway CA Jr, Travers P and Walport M (2001) *Immunobiology: The Immune System in Health and Disease*. 5th edition. New York: Garland Science.

Johal, S.S. (2004) 'Colonic IgA producing cells and macrophages are reduced in recurrent and non-recurrent *Clostridium difficile* associated diarrhoea', *Journal of Clinical Pathology*, 57(9), pp. 973–979. Available at:
<https://doi.org/10.1136/jcp.2003.015875>.

Johansson, L. *et al.* (2016) 'Antibodies directed against endogenous and exogenous citrullinated antigens pre-date the onset of rheumatoid arthritis', *Arthritis Research & Therapy*, 18(1), p. 127. Available at:
<https://doi.org/10.1186/s13075-016-1031-0>.

Jung, S.M. and Kim, W.-U. (2022) 'Targeted Immunotherapy for Autoimmune Disease', *Immune Network*, 22(1). Available at: <https://doi.org/10.4110/in.2022.22.e9>.

Khan, M.F. and Wang, H. (2020) 'Environmental Exposures and Autoimmune Diseases: Contribution of Gut Microbiome', *Frontiers in Immunology*, 10. Available at: <https://doi.org/10.3389/fimmu.2019.03094>.

Lee, A.Y.S. (2022) 'IgA anti-dsDNA antibodies: A neglected serological parameter in systemic lupus erythematosus', *Lupus*, 31(2), pp. 137–142. Available at: <https://doi.org/10.1177/09612033221074184>.

Leffler, D.A. and Lamont, J.T. (2015) 'Clostridium difficile Infection', *New England Journal of Medicine*, 372(16), pp. 1539–1548. Available at: <https://doi.org/10.1056/NEJMra1403772>.

Lenth, R. (2023) 'emmeans: Estimated Marginal Means, aka Least-Squares Means'. Available at: <https://CRAN.R-project.org/package=emmeans> (Accessed: 4 September 2023).

Leong, K.W. and Ding, J.L. (2014) 'The Unexplored Roles of Human Serum IgA', *DNA and Cell Biology*, 33(12), pp. 823–829. Available at: <https://doi.org/10.1089/dna.2014.2639>.

Lequin, R.M. (2005) 'Enzyme Immunoassay (EIA)/Enzyme-Linked Immunosorbent Assay (ELISA)', *Clinical Chemistry*, 51(12), pp. 2415–2418. Available at: <https://doi.org/10.1373/clinchem.2005.051532>.

Lobo, P.I. (2016) 'Role of Natural Autoantibodies and Natural IgM Anti-Leucocyte Autoantibodies in Health and Disease', *Frontiers in Immunology*, 7. Available at: <https://doi.org/10.3389/fimmu.2016.00198>.

Lyczak, J.B., Cannon, C.L. and Pier, G.B. (2002) 'Lung Infections Associated with Cystic Fibrosis', *Clinical Microbiology Reviews*, 15(2), pp. 194–222. Available at: <https://doi.org/10.1128/CMR.15.2.194-222.2002>.

MacBeath, G. (2007) 'Protein Arrays: Preparation of Microscope Slides', *Cold Spring Harbor Protocols*, 2007(3), p.pdb.prot4629. Available at: <https://doi.org/10.1101/pdb.prot4629>.

Ma, H. *et al.* (2022) 'Autoantibodies - enemies, and/or potential allies?', *Frontiers in Immunology*, 13. Available at: <https://doi.org/10.3389/fimmu.2022.953726>.

Marano G *et al.* (2015) 'Human Parvovirus B19 and blood product safety: a tale of twenty years of improvements', *PMC*, 13(2), pp. 184–196.

Marin, G.G. *et al.* (2009) 'Prevalence of Antinuclear Antibodies in 3 Groups of Healthy Individuals', *JCR: Journal of Clinical Rheumatology*, 15(7), pp. 325–329. Available at: <https://doi.org/10.1097/RHU.0b013e3181bb971b>.

Marshall, J.S. *et al.* (2018) 'An introduction to immunology and immunopathology', *Allergy, Asthma & Clinical Immunology*, 14(S2), p. 49. Available at: <https://doi.org/10.1186/s13223-018-0278-1>.

Martel, J. *et al.* (2022) 'Gut barrier disruption and chronic disease', *Trends in Endocrinology & Metabolism*, 33(4), pp. 247–265. Available at: <https://doi.org/10.1016/j.tem.2022.01.002>.

Mentzer, A.J. *et al.* (2022) 'Identification of host–pathogen-disease relationships using a scalable multiplex serology platform in UK Biobank', *Nature Communications*, 13(1), p. 1818. Available at: <https://doi.org/10.1038/s41467-022-29307-3>.

Monaghan, T., Boswell, T. and Mahida, Y.R. (2009) 'Recent advances in Clostridium difficile -associated disease', *Postgraduate Medical Journal*, 85(1001), pp. 152–162. Available at: <https://doi.org/10.1136/gut.2007.128157>.

Monaghan, T.M. *et al.* (2017) 'High prevalence of subclass-specific binding and neutralizing antibodies against Clostridium difficile toxins in adult cystic fibrosis sera: possible mode of immunoprotection against symptomatic C. difficile infection', *Clinical and Experimental Gastroenterology*, Volume 10, pp. 169–175. Available at: <https://doi.org/10.2147/CEG.S133939>.

Monneaux, F. and Muller, S. (2002) 'Epitope spreading in systemic lupus erythematosus: Identification of triggering peptide sequences', *Arthritis & Rheumatism*, 46(6), pp. 1430–1438. Available at: <https://doi.org/10.1002/art.10263>.

Nagele, E.P. *et al.* (2013) 'Natural IgG Autoantibodies Are Abundant and Ubiquitous in Human Sera, and Their Number Is Influenced By Age, Gender, and Disease', *PLoS ONE*, 8(4), p. e60726. Available at: <https://doi.org/10.1371/journal.pone.0060726>.

Nam, Y., Branch, D.W. and Wheeler, B.C. (2006) 'Epoxy-silane linking of biomolecules is simple and effective for patterning neuronal cultures', *Biosensors and Bioelectronics*, 22(5), pp. 589–597. Available at: <https://doi.org/10.1016/j.bios.2006.01.027>.

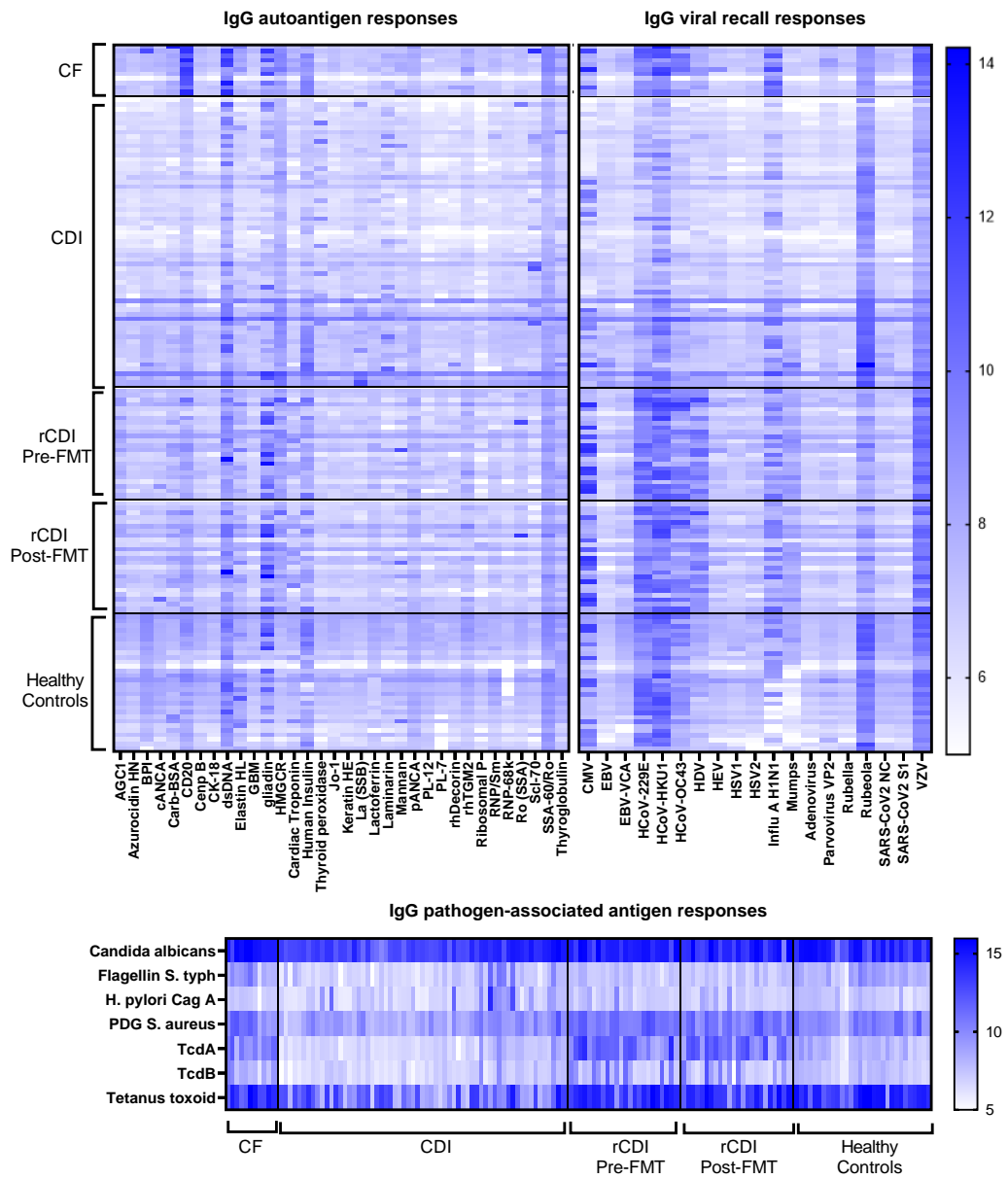
Nibbering, B. *et al.* (2021) 'Host Immune Responses to Clostridioides difficile: Toxins and Beyond', *Frontiers in Microbiology*, 12. Available at: <https://doi.org/10.3389/fmicb.2021.804949>.

- Nolte, H. *et al.* (2018) 'Instant Clue: A Software Suite for Interactive Data Visualization and Analysis', *Scientific Reports*, 8(1), p. 12648. Available at: <https://doi.org/10.1038/s41598-018-31154-6>.
- Ogimi, C. *et al.* (2020) 'What's New With the Old Coronaviruses?', *Journal of the Pediatric Infectious Diseases Society*, 9(2), pp. 210–217. Available at: <https://doi.org/10.1093/jpids/piaa037>.
- Pickard, J.M. *et al.* (2017) 'Gut microbiota: Role in pathogen colonization, immune responses, and inflammatory disease', *Immunological Reviews*, 279(1), pp. 70–89. Available at: <https://doi.org/10.1111/imr.12567>.
- Posit team (2023) 'RStudio: Integrated Development Environment for R.' Boston, MA: Posit Software, PBC. Available at: <http://www.posit.co/> (Accessed: 4 September 2023).
- Rees, W.D. and Steiner, T.S. (2018) 'Adaptive immune response to *Clostridium difficile* infection: A perspective for prevention and therapy', *European Journal of Immunology*, 48(3), pp. 398–406. Available at: <https://doi.org/10.1002/eji.201747295>.
- Robinson, D. *et al.* (2006) 'Co-occurrence and comorbidities in patients with immune-mediated inflammatory disorders: an exploration using US healthcare claims data, 2001-2002.', *Current medical research and opinion*, 22(5), pp. 989–1000. Available at: <https://doi.org/10.1185/030079906X104641>.
- Robinson, W.H. *et al.* (2002) 'Autoantigen microarrays for multiplex characterization of autoantibody responses', *Nature Medicine*, 8(3), pp. 295–301. Available at: <https://doi.org/10.1038/nm0302-295>.
- Salamunic, I. (2010) 'Laboratory diagnosis of autoimmune diseases - new technologies, old dilemmas', *Biochimica Medica*, pp. 45–56. Available at: <https://doi.org/10.11613/BM.2010.006>.
- Sekirov, I. *et al.* (2010) 'Gut Microbiota in Health and Disease', *Physiological Reviews*, 90(3), pp. 859–904. Available at: <https://doi.org/10.1152/physrev.00045.2009>.
- Shi, J. *et al.* (2014) 'Carbamylation and antibodies against carbamylated proteins in autoimmunity and other pathologies', *Autoimmunity Reviews*, 13(3), pp. 225–230. Available at: <https://doi.org/10.1016/j.autrev.2013.10.008>.

- Skopelja, S. *et al.* (2016) 'The role for neutrophil extracellular traps in cystic fibrosis autoimmunity', *JCI Insight*, 1(17). Available at: <https://doi.org/10.1172/jci.insight.88912>.
- Severall, L.F. *et al.* (2022) 'Immunological mechanisms of fecal microbiota transplantation in recurrent *Clostridioides difficile* infection', *World Journal of Gastroenterology*, 28(33), pp. 4762–4772. Available at: <https://doi.org/10.3748/wjg.v28.i33.4762>.
- Suurmond, J. and Diamond, B. (2015) 'Autoantibodies in systemic autoimmune diseases: specificity and pathogenicity', *Journal of Clinical Investigation*, 125(6), pp. 2194–2202. Available at: <https://doi.org/10.1172/JCI78084>.
- Theander, E. *et al.* (2015) 'Prediction of Sjögren's Syndrome Years Before Diagnosis and Identification of Patients With Early Onset and Severe Disease Course by Autoantibody Profiling', *Arthritis & Rheumatology*, 67(9), pp. 2427–2436. Available at: <https://doi.org/10.1002/art.39214>.
- Troncone, R. *et al.* (1994) 'Increased serum antibody levels to dietary antigens in cystic fibrosis', *Acta Paediatrica*, 83(4), pp. 440–441. Available at: <https://doi.org/10.1111/j.1651-2227.1994.tb18139.x>.
- Tsigalou, C., Vallianou, N. and Dalamaga, M. (2020) 'Autoantibody Production in Obesity: Is There Evidence for a Link Between Obesity and Autoimmunity?', *Current Obesity Reports*, 9(3), pp. 245–254. Available at: <https://doi.org/10.1007/s13679-020-00397-8>.
- U.S. Department of Health and Human Services (2018) *Bioanalytical Method Validation Guidance for Industry*. Available at: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/bioanalytical-method-validation-guidance-industry> (Accessed: 23 September 2023).
- Venkatesha, S.H., Durai, M. and Moudgil, K.D. (2015) 'Epitope Spreading in Autoimmune Diseases', in *Infection and Autoimmunity*. Elsevier, pp. 45–68. Available at: <https://doi.org/10.1016/B978-0-444-63269-2.00003-9>.
- Vetrano, S. *et al.* (2022) 'ImmUniverse Consortium: Multi-omics integrative approach in personalized medicine for immune-mediated inflammatory diseases', *Frontiers in Immunology*, 13. Available at: <https://doi.org/10.3389/fimmu.2022.1002629>.
- Vojdani, A., Pollard, K.M. and Campbell, A.W. (2014) 'Environmental Triggers and Autoimmunity', *Autoimmune Diseases*, 2014, pp. 1–2. Available at: <https://doi.org/10.1155/2014/798029>.

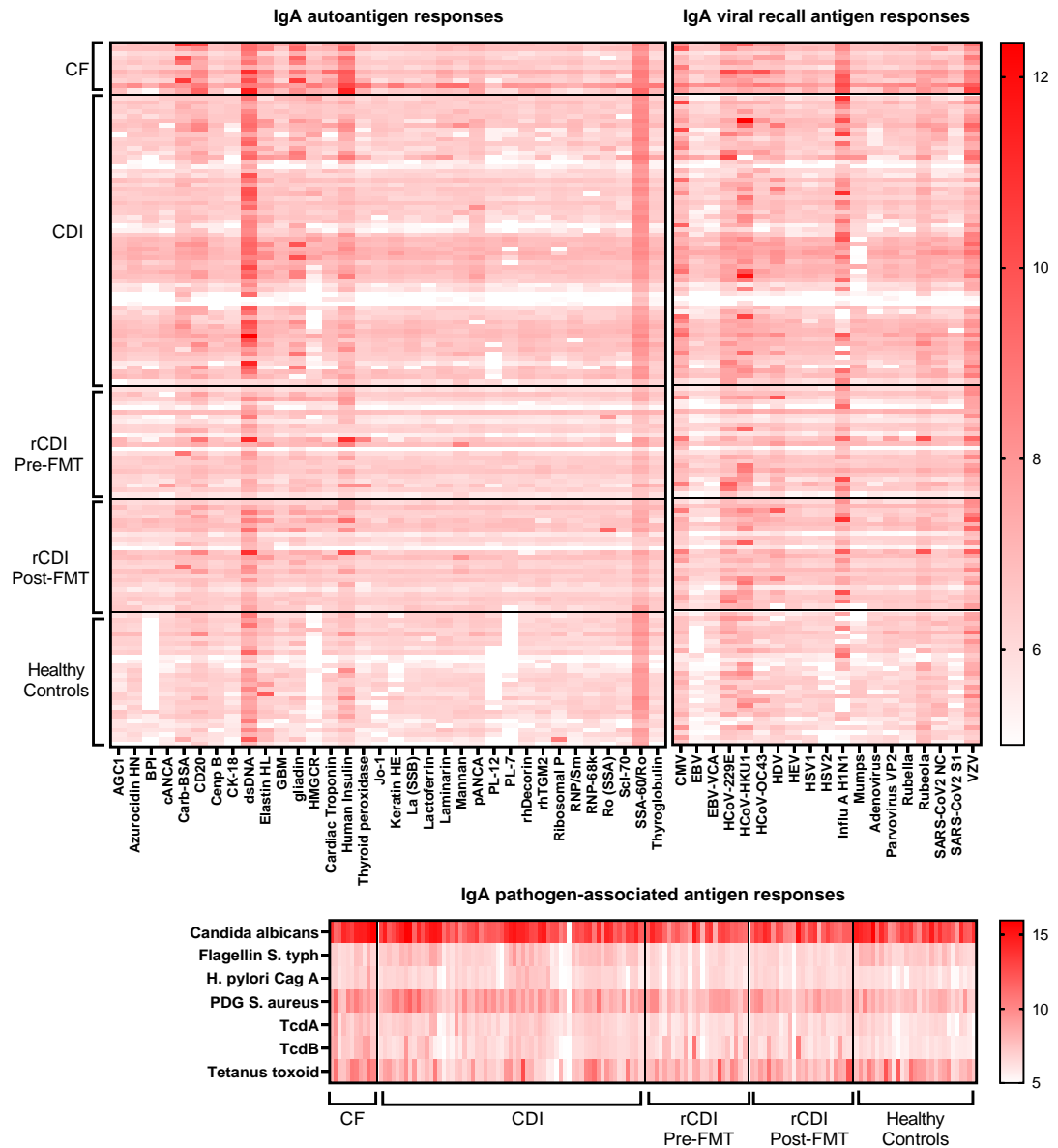
- Wang, L., Wang, F.-S. and Gershwin, M.E. (2015) 'Human autoimmune diseases: a comprehensive update', *Journal of Internal Medicine*, 278(4), pp. 369–395. Available at: <https://doi.org/10.1111/joim.12395>.
- Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York. Available at: <https://ggplot2.tidyverse.org> (Accessed: 4 September 2023).
- Wilson, D.S. and Nock, S. (2003) 'Recent Developments in Protein Microarray Technology', *Angewandte Chemie International Edition*, 42(5), pp. 494–500. Available at: <https://doi.org/10.1002/anie.200390150>.
- Yadav, R. *et al.* (2020) 'IgA autoantibodies directed against self DNA are elevated in cystic fibrosis and associated with more severe lung dysfunction', *Autoimmunity*, 53(8), pp. 476–484. Available at: <https://doi.org/10.1080/08916934.2020.1839890>.
- Yang, R., Chen, Z. and Cai, J. (2023) 'Fecal microbiota transplantation: Emerging applications in autoimmune diseases', *Journal of Autoimmunity*, p. 103038. Available at: <https://doi.org/10.1016/j.jaut.2023.103038>.
- Yap, H.-Y. *et al.* (2018) 'Pathogenic Role of Immune Cells in Rheumatoid Arthritis: Implications in Clinical Treatment and Biomarker Development', *Cells*, 7(10), p. 161. Available at: <https://doi.org/10.3390/cells7100161>.
- Yoo, J. *et al.* (2020) 'Gut Microbiota and Immune System Interactions', *Microorganisms*, 8(10), p. 1587. Available at: <https://doi.org/10.3390/microorganisms8101587>.
- Zhu, H. *et al.* (2015) 'Autoantigen Microarray for High-throughput Autoantibody Profiling in Systemic Lupus Erythematosus', *Genomics, Proteomics & Bioinformatics*, 13(4), pp. 210–218. Available at: <https://doi.org/10.1016/j.gpb.2015.09.001>.

7 Appendix



Appendix 1. Heatmaps of IgG responses to all 60 antigens on the antigen microarray platform across all serum samples.

These heatmaps were used to identify antigens with positive antibody responses in autoantigen, pathogen-associated and viral recall antigen subsets for further analysis. Dark to light blue indicates high and low IgG responses, respectively. The signal intensity values are log₂ transformed from arbitrary fluorescent units (AFUs).



Appendix 2. Heatmaps of IgA responses to all 60 antigens on the antigen microarray platform across all serum samples.

These heatmaps were used to identify antigens with positive antibody responses in autoantigen, pathogen-associated and viral recall antigen subsets for further analysis. Dark to light red indicates high and low IgA responses, respectively. The signal intensity values are log₂ transformed from arbitrary fluorescent units (AFUs).