INVESTIGATING THE GENETIC AETIOLOGIES OF SEVERE INFECTIOUS DISEASES USING WHOLE EXOME SEQUENCING AND FUNCTIONAL VALIDATION

Stephanie Najwa Hodeib

A thesis submitted for the degree of

Doctor of Philosophy

Department of Infectious Diseases, Faculty of Medicine

Imperial College London, United Kingdom

February 2023

Abstract

Human genetics can play a key role in determining susceptibility, severity, and outcome of infectious diseases resulting from otherwise common infections. Inborn errors of immunity (IEI) in invasive bacterial and viral infections, including *Neisseria meningitidis, Staphylococcal aureus, Streptococcus pneumoniae, Group A Streptococcus* and SARS-CoV-2, have been recognised to contribute to disease pathogenesis but do not explain all cases, highlighting the need for discovery of novel genetic aetiologies. This thesis aimed to use large existing well-defined disease cohorts, as well as patients recruited into a specialised paediatric IEI clinic, to discover novel causative variants associated with infectious disease. This was accomplished by employing whole exome sequencing to statistically identify significant and promising candidates that were subsequently characterised by functional validation assays. Firstly, gain-of-function *JAK2* variants were discovered in severe invasive meningococcal disease cases and were shown to constitutively activate the JAK/STAT signalling pathway. Secondly, heterozygous *KRT10* variants were enriched in an invasive bacterial disease cohort, including *Neisseria meningitidis, Staphylococcal aureus, Streptococcus pneumoniae,* and *Group A Streptococcus,* and hypothesised to confer predisposition of invasive bacterial infection by enhancing bacterial adherence and thereby, colonisation. Thirdly, *TICAM1* loss-of-function variants were characterised to impair type I IFN signalling resulting in severe COVID-19 pneumonia. Finally, genetic analysis and functional immunological techniques developed during this project was applied to a new specialist paediatric clinic. This facilitated novel genomic discoveries, in real time, and improved understanding of disease pathogenesis in children with severe or unusual infections which allowed for improved and more accurate patient care and clinical follow up. In this work, efforts were made to understand the impact of novel genetic aetiologies in susceptibility to and severity of infectious disease, discovered by whole exome sequencing and functionally validated.

Declaration of Originality

This thesis is submitted for the degree of Doctor of Philosophy to Imperial College London. The work presented here was undertaken in the Department of Infectious Disease, Faculty of Medicine at Imperial College London. The research was co-supervised by Dr Vanessa Sancho-Shimizu and Dr Jethro Herberg. I declare that none of the research results offered in this thesis has been previously submitted by me for any other degree or qualifications at Imperial College London or any other university. I certify that all the material included in this thesis is my own and those from contributors have been appropriately acknowledged and referenced.

Copyright Declaration

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

Acknowledgments

Firstly, I would like to thank my supervisor and mentor Dr Vanessa Sancho-Shimizu for her unwavering support and kindness. She has taught me so much and is an inspirational woman in the science community.

I thank my second supervisor Dr Jethro Herberg for his support and patience in helping me to navigate the clinical aspects of my thesis.

I would also like to extend gratitude to all the patients and family members who participated in the studies discussed in this thesis, and the teams involved in organising and running the studies.

Next, I would like to thank all the members of the Sancho-Shimizu Lab, past and present, for their support and friendship and I also extend that thanks to everyone in the Department of Infectious Disease who gave me support, guidance, and friendship throughout the years. Special thanks to my work and coffee buddies Ira, Dilys, and Heather who kept me sane throughout my PhD.

The biggest thank you to my family, including my brothers Mike and Chris and sister-in-law Stephanie and an extra special thanks to my parents, Hussein and Maria, for their unwavering support and constant love. بحبك Amo-te

All my love to my nephews and niece, Alexandre, Roberto, and Angela, you are the lights of my life.

A massive thank you to all my friends, for all the memories, for being my support network and for encouraging me through all the ups and downs of my PhD.

Finally, to me…you did it girl.

TO ALL THE GIRLS WHO TAKE UP SPACE

Table of Contents

Abstract 2
Declaration of Originality 3
Copyright Declaration 4
Acknowledgments 5
Abbreviations 14
List of Figures 17
List of Tables 19
Chapter 1 Introduction 22
1.1 Genetic basis of infectious disease 22
1.1.1 Inborn errors of immunity associated with infectious disease 22
1.2 Genomic investigations to determine to response to infectious disease 23
1.2.1 WES and inborn errors of immunity 24
1.2.2 WGS and inborn errors of immunity 25
1.2.3 Genome wide association studies 26
1.3 Invasive meningococcal disease (IMD) 27
1.3.1 *Neisseria meningitidis* (Nm) 27
1.3.2 Microbiology 27
1.3.3 Nm carriage and colonisation 28
1.3.4 Invasive meningococcal disease 29
1.3.5 Epidemiology 30
1.3.6 Treatments and vaccinations 32
1.3.7 Host immune response to *Neisseria meningitidis* infection 33
1.3.8 Innate immune response to Nm 34
1.3.9 Complement pathways 36
1.3.10 Adaptive immune response to Nm 38
1.4 Genetic basis underlying invasive meningococcal disease 39

1.4.1 Colonisation 39
1.4.2 Invasion 41
1.4.3 Inflammatory response 46
1.4.4 Acquired immunity 47
1.4.5 Coagulation and fibrinolysis 51
1.5 Other significant bacterial infections 55
1.5.1 *Staphylococcus aureus* 55
1.5.1.1 Microbiology 55
1.5.1.2 Staphylococcus disease and health impact 55
1.5.1.3 Genetic basis underlying invasive staphylococcus disease 57
1.5.2 *Streptococcus pneumoniae* 58
1.5.2.1 Microbiology 58
1.5.2.2 Pneumococcal disease and health impact 58
1.5.2.3 Genetic basis underlying invasive pneumococcal disease 59
1.5.3 *Group A Streptococcus* 60
1.5.3.1 Microbiology 60
1.5.3.2 Streptococcal disease and health impact 60
1.5.3.3 Genetic basis underlying invasive streptococcal disease 61
1.6 COVID-19 62
1.6.1 SARS-CoV2 62
1.6.2 Microbiology 63
1.6.3 Type I interferon response to SARS-CoV-2 infection 63
1.6.4 COVID-19 disease, treatments, and vaccination 64
1.6.5 Inborn errors of type I IFN immunity in viral infections 65
1.6.6 Role of TICAM1 in TLR3-IFN pathway 67
1.6.6.1 Association of *TICAM1* variants in viral disease 67
1.7 Summary 68

1.8 Aims of the studies 68
Chapter 2 Methods 70
2.1 Human tissue culture 70
2.1.1 Culturing human cell lines 70
2.1.2 Culturing primary human cells 70
2.1.2.1 Isolating primary human PBMCs 70
2.1.2.2 Isolating primary human polymorphonuclear granulocytes 71
2.1.3 Transfection of human cell lines 71
2.1.4 Generation of human fibro sarcoma cells expressing *JAK2* variants 71
2.2 Bacterial strains 72
2.2.1 Growth of cloning bacteria *Escherichia coli* 72
2.2.2 Growth of *Neisseria meningitidis* 73
2.2.3 Growth of Staphylococcus aureus 73
2.2.4 Growth of *Group A Streptococcus pyogenes* 73
2.2.5 Growth of *Streptococcus pneumoniae* 74
2.2.6 Heat killing *Neisseria meningitidis* 74
2.2.7 Formalin fixing of *Neisseria meningitidis* 74
2.2.8 Formalin fixing of *Staphylococcus aureus* 75
2.2.9 Formalin fixing of *Group A Streptococcus pyogenes* 75
2.2.9 Formalin fixing of *Streptococcus pneumoniae* 75
2.3 Ligands and Drugs 76
2.4 DNA editing 76
2.4.1 Plasmids 76
2.4.2 Oligonucleotides used in this thesis 76
2.4.3 Mutagenesis 88
2.4.4 Transformation 88
2.4.5 Isolation of plasmid DNA 88

2.4.6 Sanger sequencing 89
2.5 DNA extraction 89
2.5.1 DNA extraction from whole blood 89
2.5.2 DNA extraction from saliva 89
2.6 Whole exome sequencing 90
2.6.1 Whole exome sequencing 90
2.6.2 Gene burden analysis 90
2.7 Western blotting 91
2.7.1 Protein isolation and quantification 91
2.7.2 SDS polyacrylamide gel electrophoresis 91
2.7.3 Transfer to membrane 91
2.7.4 Protein identification 92
2.8 Adhesion and invasion assay 92
2.9 Flow cytometry 92
2.10 Static adhesion assay 93
2.11 Dual luciferase reporter assay 93
2.12 *CTLA-4* transendocytosis 93
2.13 Immunofluorescence 94
Chapter 3 Gain-of-function *JAK2* variants in invasive meningococcal disease. 96
3.1 Background 96
3.2 Whole exome sequencing cohort (EUCLIDS) 97
3.3 Bioinformatics and cohort analysis approach 98
3.4 Discovery of missense heterozygous *JAK2* variants in unrelated paediatric cases of IMD and other significant bacterial infections 101
3.4.1 Clinical presentation of EUCLIDS patients with *JAK2* variants 101
3.5 Review of *JAK2* variants discovered in EUCLIDS 106
3.6 Construction of *JAK2* variants in expression vectors 110

3.7 Overexpression of *JAK2* variants in *JAK2* deficient cells (γ2A) 112
3.7.1 Western blot analysis of baseline expression..... 112
3.7.2 Western blot analysis of JAK2 and STAT1 expression in stimulated cells 114
3.7 Flow cytometry analysis of phosphorylated STAT1 in stimulated γ2A cells..... 115
3.8 Flow cytometry analysis of phosphorylated STAT1 in stimulated PBMCs 118
3.8.1 Flow cytometry optimisation in stimulated control PBMCs 118
3.8.2 Flow cytometry in stimulated control vs V617F PBMCs..... 120
3.9 Creating stable mutant JAK2 expression in γ2A cells 122
3.9.1 Determination of stable JAK2 expression 125
3.10 Impact of *JAK2* on neutrophil activity 127
3.10.1 Optimisation of static adhesion assay with control neutrophils..... 128
3.10.2 Static adhesion of stimulated V617F neutrophils 130
3.10.3 Static adhesion of stimulated V617F neutrophils upon Rap1 inhibition 132
3.11 Discussion..... 133
Chapter 4 Impact of *KRT10* variants in invasive bacterial disease 142
4.1 Background 142
4.2 Discovery of heterozygous *KRT10* variants in unrelated paediatric cases of invasive bacterial disease 143
4.3 Clinical presentation of invasive bacterial disease patients with *KRT10* variants .. 145
4.4 Construction of *KRT10* variants in mammalian expression vector..... 148
4.5 Endogenous expression of *KRT10* in keratinocytes..... 149
4.6 Protein expression of *KRT10* variants in epithelial cells and keratinocytes 149
4.7 Assessing subcellular localisation of *KRT10* variants by immunofluorescence 152
4.8 Investigating the impact of *KRT10* variants in bacterial infection by adhesion and invasion assays..... 154
4.9 Discussion 156
Chapter 5 Loss-of-function *TICAM1* variants in SARS-CoV2 infection..... 161
5.1 Background 161

5.2 Identification of *TICAM1* variants in COVID-19 pneumonia patients 163
5.3 Construction of *TICAM1* variants in mammalian expression vectors 165
5.4 Optimisation of transient transfection in TICAM1-/- cells 167
5.5 Loss of expression *TICAM1* variants identified using an overexpression system 168
5.5.1 Impact of *TICAM1* variants on protein expression by immunoblotting 168
5.5.2 Visualisation of the impact of *TICAM1* variants with confocal microscopy 168
5.6 Functional characterisation of *TICAM1* loss of expression variants shows impaired type I IFN and NF-κB signalling 171
5.7 Discussion 172
Chapter 6 BRC Paediatric Clinic (INGENIOS) 176
6.1 Background 176
6.1.1 Aims 177
6.2 Summary of INGENIOS recruits 178
6.3 Genetic analysis of WES 181
6.3.1 Initial WES processing 181
6.3.2 Development of a universal analysis pipeline 181
6.4 Discovery of genetic variants in paediatric patients 184
6.4.1 De novo heterozygous *STAT3* in pneumococcal infection (1018) 184
6.4.2 Known AD *STAT1* in oral thrush (1028) 186
6.4.3 Novel X-linked *CFP* in IMD (1034) 188
6.4.4 AD *CTLA4* in HSE (1002) 189
6.3 Discussion 193
Chapter 7 Discussion 196
7.1 Overview 196
7.2 Main findings 197
7.3 Study limitations and future recommendations 202
7.4 Concluding remarks 205
References 206

Chapter 8 Appendix 228
8.1 Gene lists used to filter WES data 228
8.1.1 Mendelian susceptibility to mycobacterial diseases (MSMD) 228
8.1.2 Herpes simplex virus infection 228
8.1.3 Neisseria meningitidis infection 228
8.1.4 Candidiasis 228
8.2 INGENIOS patient details 229
8.3 Plasmid Maps 239
8.3.1 pTRIP-JAK2 wildtype plasmid map 239
8.3.2 pEFBos-TICAM1 wildtype plasmid map 240
8.3.3 pCMV-KRT10 wildtype plasmid map 241

Abbreviations

ACMG	American College of Medical Genetics
AD	Autosomal dominant
ANOVA	Analysis of Variance
AR	Autosomal recessive
ARDS	Acute respiratory distress syndrome
BHI	Brain heart infusion
BSA	Bovine Serum Albumin
cDNA	complementary DNA
CFP	Complement factor properdin
Clf	Clumping factor
CMC	Chronic mucocutaneous candidiasis
CMV	Cytomegalovirus
COVID19	Coronavirus disease
COVIDHGE	COVID Human Genetic Effort
CTLA4	Cytotoxic T-Lymphocyte Associated Protein 4
DAPI	4′,6-diamidino-2-phenylindole
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
dsRNA	double stranded RNA
EDTA	Ethylenediaminetetraacetic acid
ET	Essential Thrombocythemia
EUCLIDS	EU funded Childhood Life-threatening Infectious Diseases Study
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FF	Formalin fixed
FLAG	DYKDDDDK
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GAS	Group A streptococcus
GGTI	Geranylgeranyl transferase inhibitor
GoF	Gain of function
GWAS	Genome-wide association study
HA	Human influenza hemagglutinin

HaCaT	Human keratinocytes
HBSS	Hanks' Balanced Salt Solution
HEK293	Human embryonic kidney cells
HIES	Hyper IgE syndrome
HK	Heat killed
HRP	Horseradish peroxidase
HSE	Herpes simplex encephalitis
HSV	Herpes simplex virus
ICAM1	Intercellular Adhesion Molecule 1
ICHNT	Imperial College Healthcare NHS Trust
IEI	Inborn error of immunity
IFN	Interferon
Ig	Immunoglobulin
IL	Interleukin
IMD	Invasive meningococcal disease
INGENIOS	Investigating the GENetics of InfectiOn Study
IRF	Interferon regulatory factors
JAK	Janus kinase
KRT	Keratin
LB	Luria broth
LoF	Loss of function
LPS	Lipopolysaccharides
MAC	Membrane attack complex
MAF	Minor allele frequency
MDT	Multidisciplinary team
MFI	Median Fluorescence Intensity.
MOI	Multiplicity of infection
MPN	Myeloproliferative neoplasms
MSCRAMM	Microbial surface components recognizing adhesive matrix molecules
NET	Neutrophil extracellular traps
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NGS	Next generation sequencing
NHS	National Health Service

Nm	Neisseria meningitidis
NMDAR	N-methyl-D-aspartate receptor
NT	Non transfected
OD	Optical density
PAMP	Pathogen-associated molecular patterns
PBMC	Peripheral blood mononuclear cells
PBS	Phosphate-buffered saline
PCR	Polymerase Chain Reaction
PICU	Paediatric Intensive Care Unit
PID	Primary immunodeficiency
PMN	Polymorphonuclear neutrophils
PRR	Pattern Recognition Receptors
PsrP	Pneumococcal serine-rich repeat protein
PV	Polycythaemia vera
RFP	Red fluorescent protein
RNA	Ribonucleic acid
RPMI	Roswell Park Memorial Institute
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SNP	Single nucleotide polymorphism
STAT	Signal transducer and activator of transcription
TICAM1	TIR Domain Containing Adaptor Molecule 1
TIR	Toll/interleukin-1 receptor
TLR	Toll-like receptor
TNF	Tumour necrosis factor
TSBYE	Trypticase Soy Broth with Yeast Extract
VCAM1	Vascular cell adhesion protein 1
WES	Whole exome sequencing
WGS	Whole genome sequencing
WHO	World health organisation
WT	Wild-type
XL	Chromosome X-linked

List of Figures

Figure 1.1 *Neisseria meningitidis* pathogenesis 30
Figure 1.2 Toll-like receptor signalling in bacterial infection including *Nm* 35
Figure 1.3 Complement pathways 37
Figure 3.1 Pathogens identified in EUCLIDS 98
Figure 3.2 Overview of JAK2/STAT signalling 100
Figure 3.3 Lolliplot of *JAK2* variants and corresponding MAF 109
Figure 3.4 Sequence confirmation of site directed mutagenesis in *JAK2* plasmids 111
Figure 3.5 Impact of *JAK2* variants on protein expression 113
Figure 3.6 Impact of *JAK2* variants on protein expression upon stimulation 115
Figure 3.7 Expression of pSTAT1 in γ2A cells expressing JAK2 variants upon stimulation 117
Figure 3.8 Expression of pSTAT1 in PBMC's upon stimulation 119
Figure 3.9 Impact of V617F JAK2 variant on response to interferon and infection in PBMCs 121
Figure 3.10 Fluorescent microscopy of JAK2-lentiviral particles generation in HEK293 cells 123
Figure 3.11 Transduction efficiency of JAK2-lentiviral particles into γ2A cells 124
Figure 3.12 Fluorescent microscopy of stable transduced γ2A-JAK2 cells 125
Figure 3.13 Protein expression in JAK2 stably transduced human fibrosarcoma cells 126
Figure 3.14 JAK2-V617F promotes thrombus formation. 127
Figure 3.15 Optimisation of the static adhesion assay in control neutrophils 129
Figure 3.16 Static adhesion assay in control and V617F neutrophils 131
Figure 3.17 Static adhesion assay in control and V617F neutrophils upon stimulation 131
Figure 3.18 Inhibition of V617F neutrophil attachment upon Rap1 inhibition by treatment with GGTI-2147 132
Figure 4.1 Lolliplot of *KRT10* variants and corresponding MAF 145
Figure 4.2 Sequence confirmation of site directed mutagenesis in *KRT10* WT 148
Figure 4.3 Endogenous KRT10 expression in differentiating HaCaT keratinocytes 150
Figure 4.4 Protein expression of KRT10 variants in epithelial cells and keratinocytes 151
Figure 4.5 Impact of *KRT10* variants on protein expression by immunofluorescence 153
Figure 4.6 *KRT10* variants improve adhesion and invasion of epithelial cells by *S. aureus* 154

Figure 4.7 *KRT10* variants improve adhesion and invasion of keratinocytes by *S. aureus* 155
Figure 5.1 Lolliplot of TICAM1 variants and corresponding MAF 165
Figure 5.2 Sequence confirmation of site directed mutagenesis in *TICAM1* WT 166
Figure 5.3 Transient transfection optimisation in TICAM1-/- cells 167
Figure 5.4 Impact of TICAM1 variants on protein expression by immunoblotting 169
Figure 5.5 Impact of TICAM1 variants on protein expression by immunofluorescence 170
Figure 5.6 Impact of TICAM1 variants on type I IFN and NF-κB signalling 171
Figure 6.1 Demographic summary of INGENIOS patients 179
Figure 6.2 Clinical diagnoses summary of INGENIOS patients 180
Figure 6.3 Analysis pipeline used for INGENIOS patients 183
Figure 6.4 Pedigree of INGENIOS patient 1018 186
Figure 6.5 Pedigree of INGENIOS patient 1028 187
Figure 6.6 Pedigree of INGENIOS patient 1034 189
Figure 6.7 Pedigree of INGENIOS patient 1002 191
Figure 6.8 *CTLA4* transendocytosis assay 192

List of Tables

Table 1.1 Genetic aetiologies reported to be associated with IMD, involved in meningococci colonisation....40
Table 1.2 Genetic aetiologies reported to be associated with IMD, involved in meningococci invasion....42
Table 1.3 Genetic aetiologies reported to be associated with IMD, involved in the inflammatory response....48
Table 1.4 Genetic aetiologies reported to be associated with IMD, involved in acquired immunity....50
Table 1.5 Genetic aetiologies reported to be associated with IMD, involved in coagulation and fibrinolysis....52
Table 1.6 Monogenic inborn errors of immunity implicated in viral infections....66
Table 2.1 Drugs and ligands used in this thesis....76
Table 2.2 Plasmids used in this thesis....77
Table 2.3 Primers designed to amplify plasmid cDNA....80
Table 2.4 Oligonucleotides designed for site directed mutagenesis....82
Table 2.5 Antibodies used in this thesis....87
Table 3.1 Demographic and clinical data summary of EUCLIDS patients, and EUCLIDS patients with *JAK2* variants....102
Table 3.2 Detailed demographic and clinical details of EUCLIDS patients with *JAK2* variants....103
Table 3.3 Summary *JAK2* variants discovered in EUCLIDS....108
Table 3.4 Germline *JAK2* variants in disease....135
Table 4.1 Summary of *KRT10* variants discovered in EUCLIDS....144
Table 4.2 Demographic and clinical data summary of EUCLIDS patients, and all those with *KRT10* variants....146
Table 4.3 Demographic and clinical details of EUCLIDS patients with *KRT10* variants functionally tested in this thesis....147
Table 5.1 TICAM1 variants identified in life-threatening COVID-19 pneumonia patients....164
Table 6.1 *STAT3* variant summary discovered in INGENIOS patient 1018....186
Table 6.2 *STAT1* variant summary discovered in INGENIOS patient 1028....187
Table 6.3 *CFP* variant summary discovered in INGENIOS patient 1034....189

Table 6.4 *CTLA4* variant summary discovered in INGENIOS patient 1002.....191
Table 8.1 Recruited INGENIOS patient details.....229

Chapter 1

Chapter 1 Introduction

1.1 Genetic basis of infectious disease

Infectious disease remains one of the leading causes of morbidity and mortality globally, along with noncommunicable diseases like ischaemic heart disease which is the number one global killer (1). In 2019, prior to the SARS-CoV-2 viral pandemic, the World Health Organisation (WHO) reported over 7.7 million deaths caused by infectious disease worldwide (1). Infections can lead to extremely varied clinical outcomes, ranging from asymptomatic to fatal, even in those infected by the same pathogen. This scale can be attributed to the complex relationship between host and pathogen, including internal factors, like the genetic variability found in both host and microbe, and non-genetic external environmental factors. Genetic variability in humans has evolved over time, with infectious diseases and the host's adaptability to establish an effective immune response against a diverse range of pathogens as a key driving force in natural selection (2). This selective pressure from the burden of infectious disease, in particular those resulting in childhood mortality, has resulted in great genetic diversity in human immune genes in line with the Red Queen hypothesis, which describes an evolutionary arms race between competing species (3).

Traditionally, twin and multigeneration familial studies have been used to highlight high genetic variability and its contribution to immune response, with a median of 37% heritability seen in immune gene variation, particularly in genes involved in the cytokine pathway (4). The most genetically diverse location in the human genome is at the human leukocyte antigen (HLA) locus which is associated with multiple infectious and autoimmune disorders, including risk of death and progression of HIV disease (5, 6). Therefore, investigating host genetic factors in the predisposition or development of severe infectious disease in those who are infected by a common pathogen, and were otherwise healthy, is important to help understand this relationship.

1.1.1 Inborn errors of immunity associated with infectious disease

Mendelian single-gene defects of immunity, also known as monogenic inborn errors of immunity (IEI), are responsible for more than 485 known clinical syndromes, known as primary immunodeficiencies (PID) (7). Damaging or gain-of-function variants affecting a single gene are increasingly being shown to compromise innate and adaptive immunity, and often show incomplete clinical penetrance. There are many examples in which monogenic IEI's

have affected susceptibility or severity of infectious disease including *IRAK4* and *MYD88* deficiencies conferring predisposition to invasive pneumococcal disease (8, 9) and single gene variants in the IL-12/IFN-γ pathway underlying atypical mycobacterial disease (10). Autosomal recessive deficiencies in the terminal complement components and properdin factor are well studied to influence susceptibility to recurrent meningococcal disease by stunting the hosts ability to clear bacteria (11). Genetic defects in the TLR3 pathway have also been shown to predispose to viral herpes simplex encephalitis (12).

1.2 Genomic investigations to determine to response to infectious disease

Siblings and multigeneration families have traditionally been employed for genome-wide linkage studies to investigate susceptibility to infection by identifying associated genetic loci and were successful in discovering valid genetic variants. However, a lack of analysis power due to an inadequate number of study participants is a major limitation of this type of genetic investigation (13). Candidate gene studies are another early investigative approach and relies on a good understanding of pathogenesis as it curates genes to investigate that are plausibly relevant in infection and in most studies variant frequencies were compared between unrelated cases and controls. Candidate gene studies were successful in validating association of genetic variants to infectious disease (14). However, significant results from candidate gene studies yielded poor reproducibility from other independent sources due to investigating selected genes and small cohort sizes. These approaches tended to focus on known genes in pathways critical to the immune response, but these are rare with monogenic IEIs arising in less than 1-2% of the population, whereas polymorphisms occur with a higher frequency in the population (15).

Candidate gene studies are also incapable of discovering new genes that are associated with disease, therefore limiting the investigative opportunity. However, the development of next generation sequencing (NGS) allows for a broader scope of genetic investigations, with whole genome sequencing allowing researchers to consider the entire genome. This allows for a potential expansion of knowledge in pathogenesis, host susceptibility, and disease progression in infection, possibly leading to better developed therapies and preventative interventions.

1.2.1 WES and inborn errors of immunity

NGS was born in 2010 and has been constantly nurtured, developed, and improved since then. NGS is now an irreplaceable resource in genetic investigations and has facilitated the discovery of novel disease-causing variants and genes. Whole exome sequencing (WES) is currently the most cost-effective NGS method to use for research into and discovery of disease-causing variants, though whole genome sequencing (WGS) offers more advantages such as, having a more homogenous coverage and being able to call copy number variants more accurately (16, 17). Protein coding regions of the human genome constitutes of less than 2% of the total DNA but more than 85% of all known disease causing gene variants are located in the exome, making WES a cost-effective method for investigating disease that does not have a known genetic diagnosis (18). Another benefit to WES is the ease of variant processing and analysis that does not require complicated analytical tools, compared to WGS or RNA-sequencing (18). On average, around 20,000 – 30,000 variants per exome are called and a challenge for WES is filtering out disease-causing variants from benign common variants, which makes up around 95% of called variants (18).

Some different analysis strategies can be used to filter out disease-causing variants in WES, and it is important to consider the locus heterogeneity of the trait, the allelic segregation of the pedigree, and ethnicity of the affected individuals to account for differences in variant allele frequencies, as well as disease prevalence. If pedigrees are available, WES can be analysed to look for *de novo* or recessive variants in cases of one affected family member and unaffected parents or, if there are multiple family members affected, WES can be screened for either shared casual variants or help limit the search to a shared haplotype region (18).

Variants causative for severe disease can be filtered for in multiple ways, one such way that works particularly well when applied to investigating unrelated affected cases is to first search for variants with a low minor allele frequency (MAF), as rare disease is intuitively likely to be caused by rare pathogenic variants. Deciding on the MAF cut off should be determined by prevalence of the disease, clinical penetrance, and inheritance model (19). A MAF of less than 1 % is typically accepted in most disease models as a good place to start investigating for rare variants (20). Mode of inheritance is also an important factor to consider, as most IEI's are causative only if inherited in a particular manner, either autosomal dominant, autosomal recessive, or X-linked (7). An autosomal dominant disorder is when a variant is in only one allele of the gene, but this is enough to affect function and cause disease, typically either in a

dominant negative, haplo-insufficient, or gain-of-function manner. Autosomal recessive disorders are rarer as there needs to be a biallelic expression of the variants, and these disorders can result in more severe disease when the cause is due to a loss-of-function variation. Recessive disorders are of particular interest to investigate in consanguineous families, as both parents share a near common ancestor and this therefore leads to increased expression of autosomal recessive disorders (21). Variants can also be annotated using public databases such as gnomAD (http://gnomad.broadinstitute.org) and dbSNP (https://www.ncbi.nlm.nih.gov/snp/) to exclude synonymous SNPs and inform on known pathogenic variants. This filtering can be further stratified to prioritise variants predicted to have a significant functional effect on the protein, such as changes in a splice site, frameshifts, and nonsense variants, which indicates the presence of an early stop codon. However, in most scenarios where impact on protein function is predicted there is a bias towards variants that are protein damaging or lead to a loss-of-function, which disallows possible gain-of-function variants to be considered.

1.2.2 WGS and inborn errors of immunity

NGS can also be used to sequence entire genomes (WGS), though this is currently more costly than WES it is probable that in the near future with improvements in technology that the cost will be comparable. WGS processes DNA in a similar way to WES, with some notable differences, for example, the "capture" enrichment step used in WES is not needed for WGS allowing for a much more homogeneous coverage (22). WGS also has better overall uniform coverage of targeted bases compared to WES, despite WES having a higher mean read depth, and is more accurate identifying copy number variants, which often do not start or end within exons, and variants in non-coding RNA genes (16, 22). The human genome comprises of over 3 billion base pairs and the obvious advantage to WGS is, of course, its ability to detect intronic and intergenic variants in addition to those in the exome, but these produce massive data files that are costly to store and cumbersome to analyse. Though 85% of all known disease-causing gene variants are located in the exome, some are not, and this is where costly WGS can be valuable (23). For example, WGS allowed for the discovery of a causative intronic variant in X-linked reticulate pigmentary disorder that had been missed by prior exome investigations (24).

1.2.3 Genome wide association studies

WGS enables genome-wide association studies (GWAS), that have significantly developed research into genotype-phenotype associations. GWAS are used to identify genome loci associated with disease phenotype at a genome-wide significance level, and meta-analysis is a model method used to detect associations with increased statistical power. In association studies, disease-causing variants are typically not directly genotyped but are shown in linkage disequilibrium with the genotyped SNPs. Since its advent, over 52,000 significant SNP to disease associations have been reported according to the NHGRI-EBI GWAS Catalog. GWAS is able to utilise the vast public and private genomic databases to discover novel casual SNPs associated with disease, but it is important to carefully select cohorts and well-matched control groups to eliminate false-positive hits. One of the most effective GWAS was published in 2010 and described an impressive cohort of 475 invasive meningococcal disease cases analysed against 4703 UK population controls (Wellcome Trust Case Control Consortium). This study revealed two significant SNPs found in complement factor H (*CFH*) and 3 SNPs in CFH-related protein 3 (*CFHR3*), that were shown to be protective against invasive meningococcal disease and validated in two separate independent cohorts (25). There are multiple limitations to GWAS, the chief being the need of a high significance threshold to account for multiple independent statistical tests, however, this can be overcome by increasing the sample size included in the GWAS.

WGS and WES both have strengths and weaknesses, and both are powerful resources in genetic investigation and, increasingly, clinical diagnostics. WGS covers the entire genome, maximising the information there is to obtain from NGS and detecting variants that could be missed by WES, for example in intronic regions. However, this results in data that is an order of magnitude larger than WES data which can make interpretation challenging and requires more expertise in data processing. Due to massive data from WGS, operational costs can be high, as adequate storage and a high level of software resources are needed. WGS has also been shown to have a more uniform coverage than WES, especially in GC-rich regions, however a comparative study shows that WGS in clinical diagnostics is 0.81 % improved over WES (26), a modest increase that may not offset the increased cost of WGS for some. Though WES only captures ~2% of the entire genome, 85% of disease-causing variants occur in the coding regions (23). WES is also cost-effective in both actual sequencing costs and operational costs as the data is easier to store, handle, and process. All these factors are considered when deciding to perform WGS or WES. It is typical in cases of clinical diagnostics to use gene panels or WES

for patients, and only resort to costly WGS where WES is not able to identify an accurate diagnosis. For all the benefits discussed here, WES was the sequencing method of choice used in this project.

1.3 Invasive meningococcal disease (IMD)

1.3.1 *Neisseria meningitidis* (Nm)

Neisseria meningitidis (Nm) is a leading cause of bacterial septicaemia and meningitis worldwide. Nm belongs in the family *Neisseriaceae* and the wider genus *Neisseria*. Within this large genus are many Gram-negative, typically diplococci, bacteria that can colonise the mucosal surfaces of a range of animals, with 17 species that can be isolated from humans (27). However, only two of these species have caused any significant pathogenesis exclusively in humans. *Neisseria gonorrhoeae*, an obligate pathogen that typically colonies the urogenital tract and can result in sexually-transmitted gonorrhoea, and *Neisseria meningitidis,* a pathogen whose niche is the naso-oropharynx. Nm invasion can result in endemic and epidemic meningitis and/or sepsis disease, overall classed as invasive meningococcal disease (IMD). Nm was first identified by Anton Weichselbaum in 1887 in Austria, a Viennese pathologist who isolated the bacterium from the cerebrospinal fluid of patients during an epidemic of cerebrospinal meningitis occurring at the time (28). However, meningococcal disease was first clinically characterised earlier, in 1805, during a meningitis epidemic by Vieusseux Gaspard in Geneva, Switzerland (29). The first cases of IMD caused by meningococcus outside of Europe were reported West Africa, in the 1905 meningitis epidemic in northern Nigeria, and the 1906 epidemic in Ghana (30-32). These outbreaks were followed by other devastating outbreaks in African countries that are now collectively classified as a Sub-Saharan African region which stretches from Senegal to Ethiopia, known as the "African Meningitis Belt". Nm infection and subsequent IMD will be a focus in this project.

1.3.2 Microbiology

Nm is a common commensal bacterium that is paradoxically also an important and devasting invasive human pathogen. Nm can sometimes be described as an opportunistic pathogen. It is a gram-negative aerobic diplococcus β-proteobacterium, commonly described as resembling coffee beans, that selectively colonises the human nasopharynx, though in some rare individual case studies it has been seen to colonise the rectal and urogenital tracts (33-35). Nm is

encapsulated with thirteen capsular serogroups and of these, six major serogroups are typically associated with disease: A, B, C, Y, X, and W-135.

Meningococcal virulence and characterisation are mainly determined by the polysaccharide capsule, expression of surface adhesive proteins including Opacity proteins Opa and Opc, pili, porins PorA and PorB, endotoxin lipooligosaccharide, and virulent clonal complexes (36). Most meningococcal strains causing invasive disease will possess a polysaccharide capsule as this assists in evasion of the host immune response by preventing opsonisation and phagocytosis, therefore allowing the bacterium to survive in the bloodstream (37). In encapsulated Nm the cell envelope comprises an outer membrane composed of lipooligosaccharide (LOS/LPS), a peptidoglycan layer, and an inner membrane, surface pili and opacity proteins on the surface are used to attach to host non-ciliated epithelial cells, colonise, and invade (36).

1.3.3 Nm carriage and colonisation

Neisseria meningitidis is transmitted via mucus droplets and almost exclusively colonises the human nasopharynx. Initial adhesion is mediated by type IV pili and then facilitated by Opacity-associated adhesion (Opa) proteins interactions with host cell surface proteins on the nasopharyngeal epithelium, including carcino-embryonic antigen cell adhesion molecule (CEACAM) proteins (38, 39). Carriage of Nm refers to the commensal colonisation of the bacterium, with invasion through the mucosal layer resulting in bacterial dissemination throughout the body causing IMD with complications such as meningitis and sepsis. Carriage rates vary depending on factors including age, living conditions, and geographical location. The Stonehouse survey conducted in the UK, in a non-epidemic setting, concluded an overall 10 % carriage rate of meningococci with a peak seen in 0–4-year-olds and a second peak in 15-19 year olds with a decline then seen in increasing age (40, 41). Also, greatly increased carriage rates have been reported among military recruits from multiple countries including 37.4 % in Greece, 70 % in Norway, and up to 61 % in Poland suggesting that crowded environments can increase the carriage rate of *Neisseria meningitidis* in the population (41-43). In landmark publications in the 1960s, Goldschneider and colleagues reported that carriage of Nm led to the production of protective antibodies and development of acquired immunity, with bacterial colonisation very rarely leading to invasive disease (44, 45). Carriage of Nm in human mucosal nasopharyngeal membranes often leads to asymptomatic colonisation and can last for days or months with an increase of serum bactericidal response seen (46). Majority of *Neisseria*

meningitidis carriers remain asymptomatic or can develop low grade bacteraemia (47, 48). However, there is a minority of the population that suffers from meningococcal invasion resulting in IMD.

1.3.4 Invasive meningococcal disease

IMD is rare is non-epidemic settings but can be life-threatening, with rapid onset and significant mortality at an overall rate of over 50% in untreated individuals and 10% in those treated, with up to 19% of survivors suffering from life-long severe sequelae with a reduced quality of life (49, 50). As mentioned, in majority of the population carriage of *Neisseria meningitidis* is asymptomatic, however, in susceptible hosts, using mechanisms still unclear, meningococci can invade and cross the nasopharyngeal mucosal epithelium to gain access to the blood stream (Figure 1.1) (38). Once inside the bloodstream the meningococci grow in number and are disseminated throughout the host, this can lead to septicaemia and can develop into the often-fatal purpura fulminans (50, 51). Shock is also a typical feature seen with meningococcal sepsis (52). In some cases, the meningococci can breach the blood brain barrier (BBB), at vulnerable sites such as the choroid plexus, leading to infection in the cerebrospinal fluid, which can spread along the meninges, developing into meningitis (Figure 1.1) (50, 53). Sepsis and meningitis phenotypes in IMD can occur independently or in combination.

IMD incidence rates typically average 1-3/100,000 in Europe and North America (54). However, the incidence of IMD greatly increases during outbreaks, with the highest rates of IMD seen at 1000/100,000 in sub-Saharan Africa during large epidemics which occur every 5-10 years (54, 55). During a ten year period in sub-Saharan Africa between 1992-2012, before the introduction of any vaccine programme, 947,000 cases of meningitis predominantly caused by Nm serogroup A (~85% of cases) were reported with an estimated 100,000 deaths occurring (56).

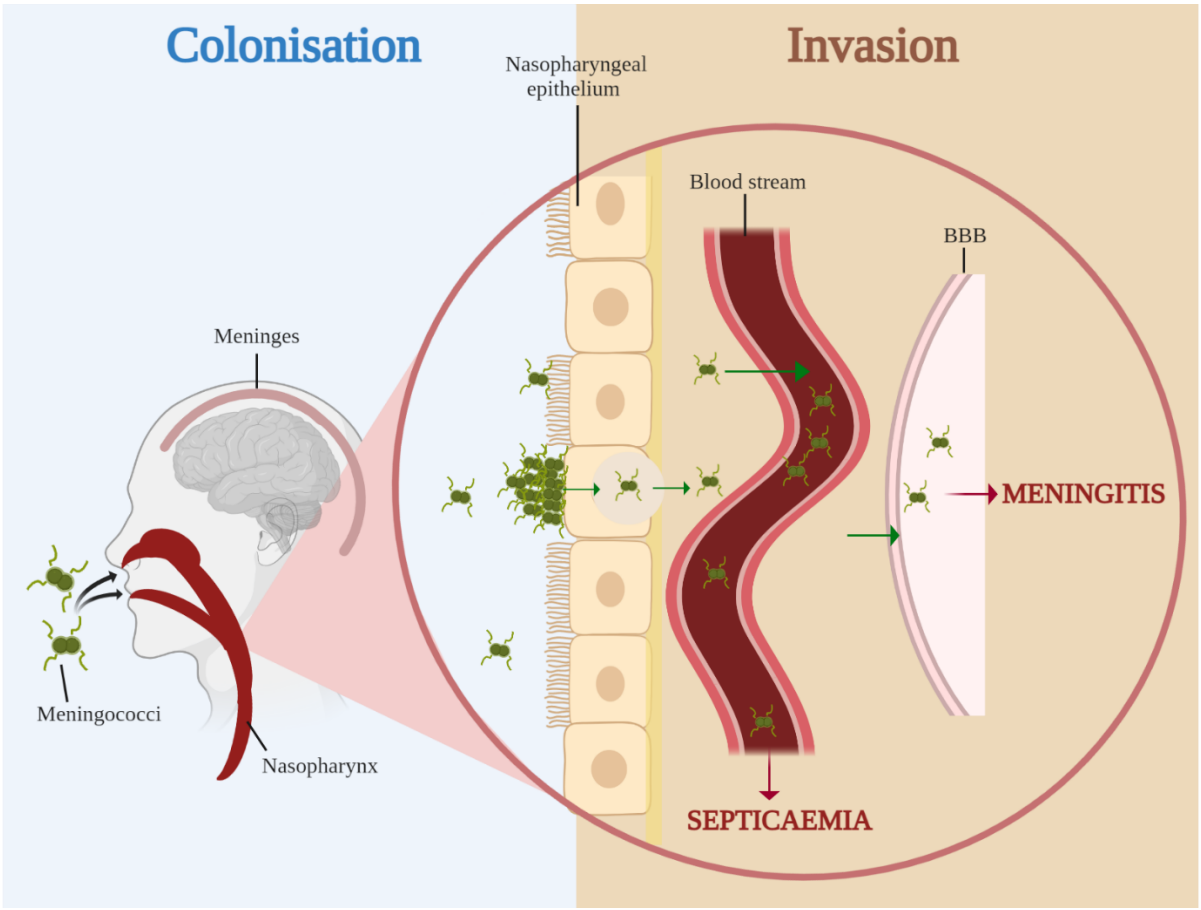


Figure 1.1 ***Neisseria meningitidis*** **pathogenesis**

Representation of the pathogenesis route undertaken by Nm. Meningococci colonises the nasopharynx after transmission on mucus droplets and in a rare subset of people, with mostly unknown mechanisms, the meningococci can invade the epithelium, gaining access to the blood stream. From here, meningococci disseminates throughout the host leading to septicaemia which can develop into the often-fatal purpura fulminans. In some cases, the meningococci can breach the blood brain barrier (BBB), at vulnerable sites like the choroid plexus, leading to infection in the cerebrospinal fluid, which can spread along the meninges, developing into meningitis. Adapted from Hodeib et al. (2020) (11) and created using biorender.com.

1.3.5 Epidemiology

The six major Nm serogroups typically associated with disease are serogroups A, B, C, Y, X, and W-135 (57). Nm is found globally, however, there are geographical hotspots traditionally associated with certain disease causing serogroups. Serogroups B and C are the most common serogroups found to cause IMD in developed countries in the Americas, Europe, Asia, and Australia with an incidence rate of 1-3 cases per 100,000 (58, 59). Serogroup A had posed a recurrent pandemic threat in sub-Saharan Africa affecting as many as 1000 cases per 100,000 since the early 20th Century, mostly in the African “meningitis belt” and in Asia (54, 60, 61).

However, Nm epidemics in Africa in recent years have geographically extended outside the historic boundaries of the meningitis belt (54, 62). Serogroup W-135 is conventionally associated with IMD in Saudi Arabia, specifically in individuals undertaking the annual Hajj pilgrimage (63). However, in recent years there have been surges of MenW incidence globally and it is now a dominant causative serogroup of IMD, with a predilection to a sepsis phenotype, in countries in Europe, South America, Australia, and along the African meningitis belt (64). Incidence of IMD caused by serogroup Y is scattered, mainly found the United States of America (USA) and Sweden, but also reported in the United Kingdom (UK), Israel, and South Africa (65-67). Finally, serogroup X is typically seen in outbreaks in Sub-Saharan Africa with some sporadic cases seen in Italy and China (68-70). A key determinant in the distribution of serogroups in countries causing IMD in recent years is due to the health strategies each country has undertaken, with particular reference to differing national vaccine strategies implemented.

The UK has had the highest number of clinically confirmed IMD cases in Europe (71). Prior to the national introduction of the MenC vaccination programme, C was the most dominant of the disease-causing serogroups accompanied with a high case fatality rate. Following the widespread implementation of MenC immunisation in the UK, serogroups B, W, and Y are now the predominant disease-causing meningococcal serogroups with an IMD case fatality rate recorded between 2008-2015 of 64 %, 17 %, and 14 % respective to the serogroups (72). However, in 2015, the 4CMenB and MenACWY vaccination programme for university students was rolled out in the UK, with good uptake, causing a consistent decrease of IMD since, with ~1/100,000 cases clinically confirmed in 2019 (73).

Nm can be phenotypically separated into serogroups and genetically separated by sequence types (STs) and clonal complexes (CCs). A minority of Nm clonal complexes, termed temporally stable hyperinvasive virulent lineages, are responsible for the majority of IMD cases seen in outbreaks and epidemics. In recent years, serogroup C CC10217 has caused an outbreak of IMD in Sub-Saharan Africa and serogroup C and W CC11 has been associated with outbreaks in Saudi Arabia, South America, the USA, and Italy (74-77).

Age is a major risk factor for the development of IMD with young children being at particular risk of developing IMD due to the absence or waning of protective maternal antibodies and the lessened ability of complement-mediated phagocytosis of bacteria compared to older persons. Therefore, disease rates are high in those aged between 3 months and 5 years of age, with other

peaks of IMD incidence seen in adolescents and in the elderly, understandably coinciding with increased carriage rates (48, 49, 78).

1.3.6 Treatments and vaccinations

If left untreated, IMD can reach a mortality rate of 50 %, however, if symptoms are correctly identified early in infection and effective parenteral antimicrobial therapy is administered the mortality rate drops to 10 % (52, 56). Treatment has varied over the decades and continues to vary between resource rich countries and middle to low income developing countries. IMD onset is rapid with 85 % of IMD-related deaths occurring within a day of diagnosis (72). Clinical management of shock and intracranial pressure are key in effective treatment of IMD, and an antibiotic regiment should be started as soon as IMD is recognised in the patient. Antibiotic treatment helps to reduce the levels of circulating endotoxins in the bloodstream, which is associated with the development of more severe disease that can lead to death such as multiple organ failure and severe sepsis (79). A one week antibiotic regime is recommended of cefotaxime, ceftriaxone, and penicillin (or benzylpenicillin in countries where penicillin resistance is rare, like the UK), these are typically used for initial antibiotic therapy in IMD cases to clear colonising bacteria and reduce the risk of transmission (80). If the patient is allergic to penicillin, then chloramphenicol and meropenem can be used instead (80).

However, administration of prophylactic vaccines is the best management approach to reducing IMD fatalities. This is because there is still a 10 % fatality rate in treated IMD cases and it is not uncommon for there to be late diagnoses or misdiagnoses of IMD, particularly in rural areas and low to middle income developing countries with limited medical resources and testing facilities.

There are multiple vaccinations available against *Neisseria meningitidis*. Currently, the most popular are the quadrivalent conjugate vaccine (MenACWY), the MenA conjugate vaccine (MenAfriVac) which has had widespread success particularly in the sub-Saharan region, and the recombinant protein vaccines targeting serogroup B (4CmenB, also known as Bexsero, and Trumeba). Conjugate vaccines successfully target the polysaccharide capsule of serogroups A, C, W, and Y however, serogroup B capsule is covered with polysialic acid rendering it non-immunogenic (81). Therefore, a newer technique, reverse vaccinology pioneered by Dr Rino Rappuoli and colleagues (82), was tasked with discovering novel vaccine candidates in Nm serogroup B and has led to recombinant protein vaccines that successfully target highly immunogenic protein antigens and the outer membrane vesicles of MenB (37).

In 2015, the UK was the first country to implement 4CMenB nationwide in the childhood vaccination regimen, with two doses given at 8 and 16 weeks of age and a booster at 1 year (83). Vaccine uptake in the UK was high with 88 % receiving all three doses, and it has been estimated that this vaccine scheme has prevented 75 % of IMD cases caused by MenB (84). However, this vaccine does not protect against all disease causing MenB strains and vaccine effectiveness was calculated at 59.1 % when all three doses were administered (84). This study shows the positive effect this vaccine has had on the vaccine eligible cohort in the UK but also shows that more work needs to be done to discover more robust and widespread vaccine targets in MenB. Also, studies have shown that the MenB recombinant protein vaccines have no significant effect on meningococcal carriage and is unable to provide herd immunity in the community, unlike its conjugate vaccine partners (85).

MenAfriVac was created to limit the cost of the expensive quadrivalent conjugate vaccines to include only the dominant disease-causing serogroup in Africa, MenA. According to the World Health Organisation (WHO) (April 2021), due to the mass vaccination efforts from African countries in the meningitis belt from 2010 with MenAfriVac, incidence of IMD due to MenA has fallen ~99 % with no MenA IMD cases confirmed since 2017. However, IMD from other serogroups, mainly C, W-135, and X are still occurring as there is no wide accessibility to multivalent meningococcal conjugate vaccines in these African countries yet.

1.3.7 Host immune response to *Neisseria meningitidis* infection

Host evasion and a high rate of replication are key determinants for Nm to cause severe disease, though the exact reason why this occurs in only a small subset of the population is not fully understood (86). As with most infectious agents, the human adaptive and innate immune responses play key roles in protection and clearing infection. As mentioned previously, in the late 1960s Goldschneider and colleagues published landmark epidemiological studies that were important for reporting Nm carriage and responsible for establishing that serum bactericidal antibodies acquired against circulating invasive strains was the most important host defence mechanism (44, 45). Risk of IMD is particularly high in infants between 3 months and 2 years of age, when maternal humoral antibodies start to wane, suggesting that innate immunity is of equal importance in response to Nm infection, especially during this infant period (45, 86).

1.3.8 Innate immune response to Nm

Nasopharyngeal colonisation by Nm can trigger the host innate immune response, in which two main outcomes can be achieved by the host, 1) the activation of pro-inflammatory cytokines, and 2) inducing the complement cascade to clear meningococci via opsonophagocytosis. Nm expresses various pathogen-associated molecular patterns (PAMPs) that are recognised by host pattern recognition receptors (PRRs) as non-self molecules on antigen-presenting or phagocytic cells. A main class of PRRs are the toll-like receptor (TLR) family which can recognise evolutionary conserved PAMPs (Figure 1.2). Nm LPS, expressed on the outer polysaccharide capsule, is a main driver of host inflammation and is recognised by TLR4 (87). Lipoteichoic acid, peptidoglycan, and PorA are key Nm adhesion molecules that are involved in initial pili attachment and colonisation and are recognised by TLR2 (88). Thirdly, it is believed that TLR9, which is localised to the intracellular vesicles of the endoplasmic reticulum, lysosomes, and endosomes detects and binds to Nm unmethylated CpG DNA (89). TLR recognition of and binding to microbial PAMPs triggers a signalling cascade via the main adaptor molecules TIRAP, and MyD88. This leads to activation and nuclear translocation of NF-kB and IRFs which induces gene transcription of complement components, interferons (IFN), and pro-inflammatory chemokines and cytokines such as IL-6, TNF-α, IL-8, IL-1β, and CXCL1-3 (Figure 1.2). These pro-inflammatory cytokines are essential in clearing invading meningococci and preventing IMD, however, abnormal high levels of circulating IL-1, IL-6, and TNF-α have been associated with poor outcome in meningococcal septic shock and septicaemia (90-93).

These secreted cytokines/chemokines can activate the Janus kinase-signal transducer and activator of transcription (JAK-STAT) pathway, which can also be activated by circulating interferons, cytokines, hormones, and growth factors. In the canonical JAK-STAT pathway, ligand binding to membrane receptors phosphorylates JAKs that can then phosphorylate latent STATs located in the cytoplasm, which dimerise and translocate to the nucleus to regulate the transcription of genes (Figure 1.2). This pathway is involved in mediating multiple processes including induction of anti-microbial proteins, T cell differentiation, B cell maturation and development, apoptosis, cell differentiation, cell proliferation, and many more (94).

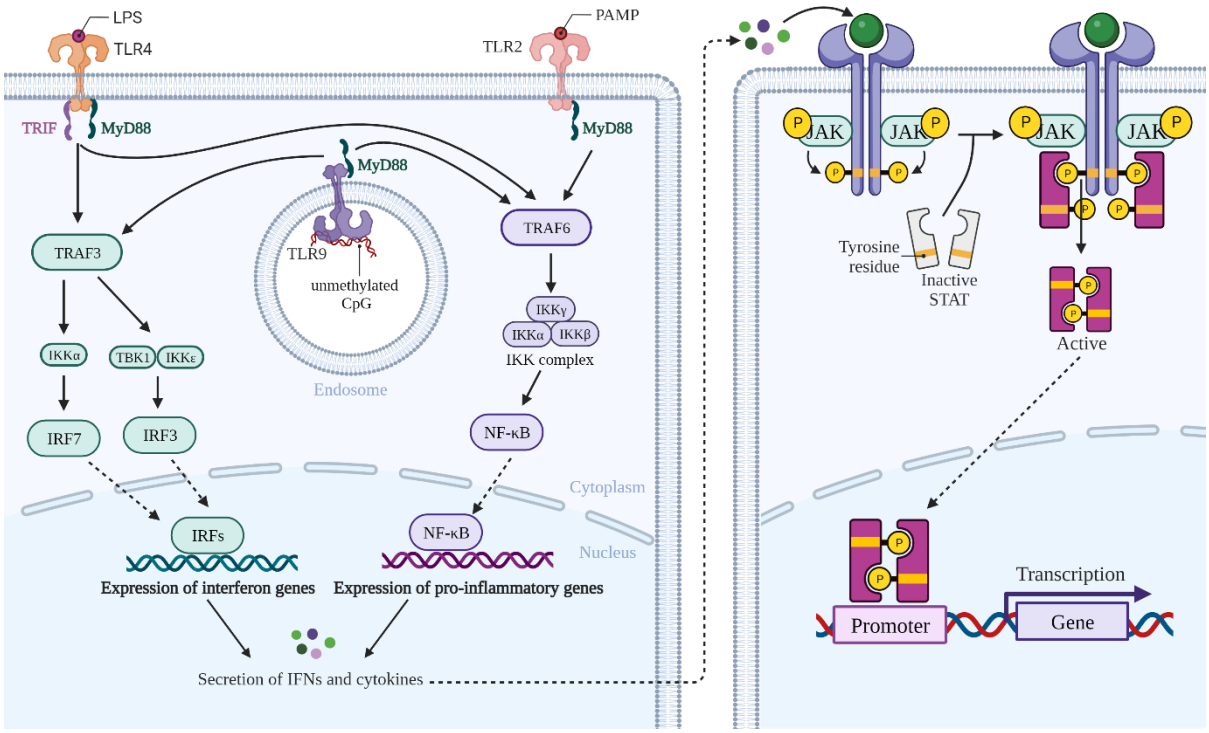


Figure 1.2 Toll-like receptor signalling in bacterial infection including *Nm*

Representation of the TLR signalling pathway induced by bacterial PAMPs. Nm LPS is recognised by TLR4 and lipoteichoic acid, peptidoglycan, and PorA are recognised by TLR2 on the cell membrane. TLR9, found in intracellular endosome compartments detects and binds to Nm unmethylated CpG DNA. TLR recognition triggers a signalling cascade via the main adaptor molecules MyD88 and TICAM1 which activates and translocate NF-kB and IRFs to the nucleus to induce gene transcription of complement components, IFNs, and pro-inflammatory chemokines and cytokines, which are then secreted extra-cellularly. In the canonical pathway, ligand binding to membrane receptors phosphorylates JAK tyrosine residues which recruit and phosphorylate latent STATs located in the cytoplasm. Active STATs dimerise and translocate to the nucleus to regulate the transcription of cytokine/IFN stimulated genes by binding to their promoters. This results in the induction of anti-microbial proteins, T cell differentiation, B cell maturation and development, apoptosis, cell differentiation, cell proliferation, and many more. Image created using biorender.com.

Neutrophils are polymorphonuclear leukocytes (PMNs) that are recruited to infection sites, and can secrete large amounts of pro-inflammatory cytokines, which can, among other interactions, go on to trigger the JAK-STAT pathway and activate the inflammatory pathway (95). Detection of high levels of neutrophils in the cerebrospinal fluid, where neutrophils are typically absent, is a hallmark of Nm infection (96). Neutrophils harbour specialised granules containing antimicrobial enzymes and molecules in their cytoplasm and are bactericidal. They release these granules in internal phagolysosomes where phagocytosed pathogens are engulfed into or via extracellular granule secretion via exocytosis or release of neutrophil extracellular traps

(NETs) (97). NETs are large, extra-cellular, web-like constructions assembled from neutrophil proteins, chromatin, and DNA and, much like a fishing net, can trap pathogens and neutralise them (98). In response, Nm has developed virulence factors to evade NET mediated killing, including modification of its LPS to prevent digestion, expression of high-affinity zinc uptake receptor ZnuD allowing Nm to escape NET-mediated nutritional immunity, and binding of Nm's spontaneously released outer membrane vesicles to NETs, therefore allowing the meningococci to escape the subsequent bacteriostatic effect (99). These evasion measures by Nm in response to NET induction have been postulated to aggravate thrombosis, likely leading to disseminated intravascular coagulation, and therefore contributing to rapid progression of severe sepsis in IMD (99).

1.3.9 Complement pathways

The human complement pathways play a chief role in innate immunity by successful management of invasive pathogens, including Nm, via complement-dependent lysis and opsonophagocytosis. Flexner and colleagues first established the importance of host complement-dependent bactericidal activity in meningococcal infections in the early 20th century and this was then built on by Goldschneider and colleagues, who clearly elucidated the protective roles and complement and antibodies in invading Nm (44, 45, 100, 101).

Complement is activated by three main pathways in which component 3 (C3) plays a central role (Figure 1.3). The classical antibody–antigen interaction, mediated by C1q, and triggered by IgG and IgM antibody interactions. The mannose-binding lectin (MBL) pathway in which MBL and associated serine proteases can interact with microbial cell walls. The alternative pathway (AP), regulated by complement factor P (CFP) and activated by C3 interacting with complement Factor B (CFB) and complement Factor D (CFD). The alternative pathway can also uniquely amplify C3 deposits in the other pathways in a positive loop manner. All three pathways converge into the cleavage of C3 into C3a, a proinflammatory mediator, and C3b an opsonin that can bind to meningococci and facilitate phagocytosis (Figure 1.3) (102). Secondly, C3b also triggers the downstream formation of the membrane attack complex (MAC) by assembling terminal complement components (C5 – C9). The MAC can clear meningococci by creating pores in bacterial cell membranes leading to cell death (103, 104). The importance of complement in protection against invading meningococci is highlighted by deficiencies in C3 and the terminal complement components, C5 - C9, conferring susceptibility to IMD (11, 105, 106). The complement pathway is also important in bridging the divide to adaptive

immunity by assisting in modulating lymphocytes and mediating antigen-specific responses (97).

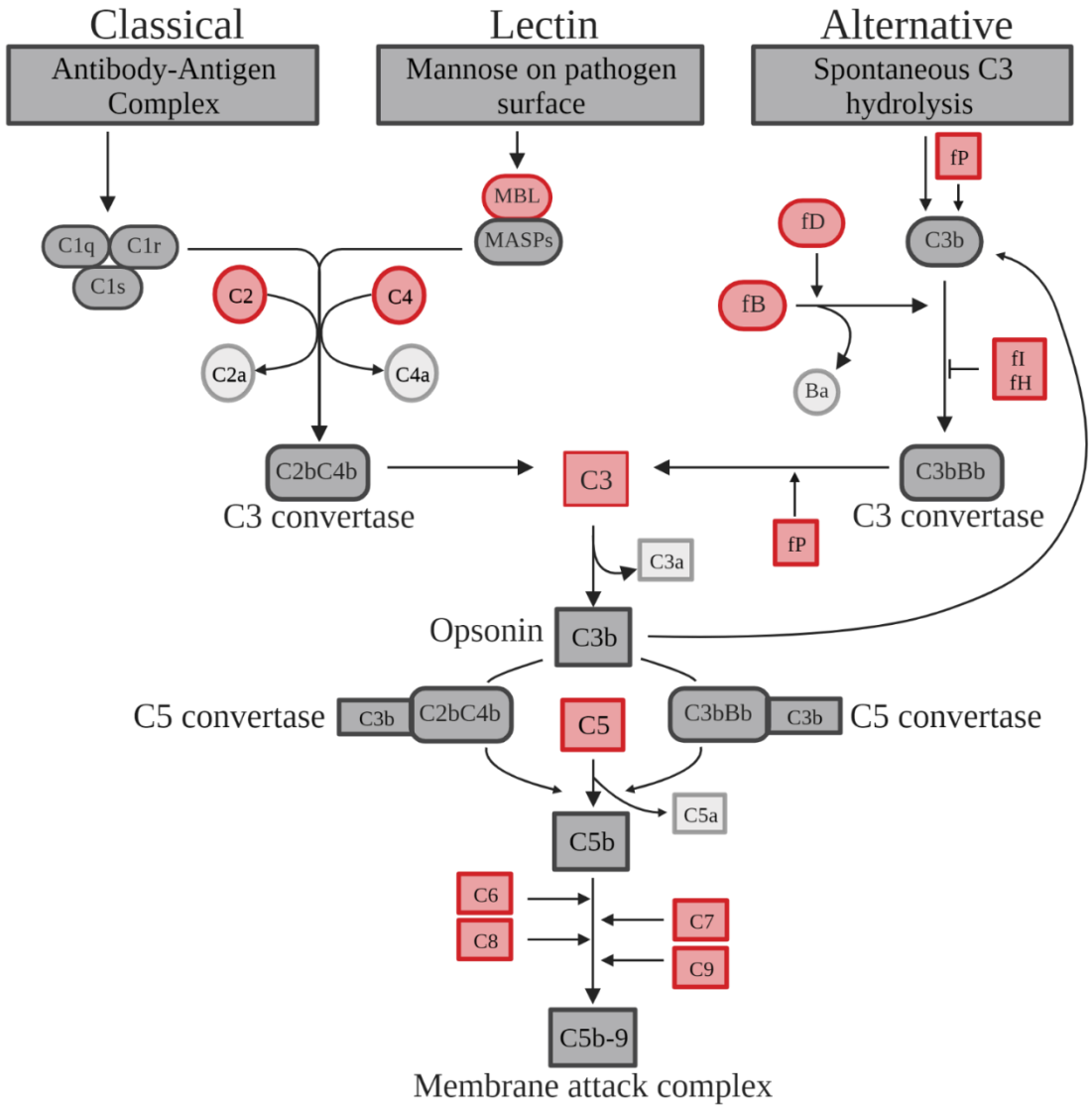


Figure 1.3 Complement pathways

Overview of three main complement pathways. The classical antibody–antigen interaction, mediated by C1q and performed by IgG and IgM antibodies. Secondly, interaction of mannose-binding lectin and associated serine proteases with microbial cell walls. Lastly, the alternative pathway, regulated by complement factor P (CFP) and activated by C3 interacting with complement CFB and complement CFD. CFI, and CFH act as negative regulators of the alternative pathway and this pathway also amplifies C3 deposits in the other pathways in a positive loop manner. All three pathways converge into the cleavage of C3 into C3a, an inflammation effector, and C3b an opsonin that can facilitate phagocytosis. Secondly, C3b also triggers the downstream formation of the membrane attack complex by assembling terminal complement components (C5 – C9). The MAC can clear bacteria by creating pores in bacterial cell membranes leading to cell death. Those in red symbolises factors that have reported loss-of-function mutations that are associated with either chronic meningococcaemia or IMD. Adapted from Hodeib et al. (2020) (11) and created using biorender.com.

1.3.10 Adaptive immune response to Nm

Adaptive immunity can be activated upon immune cells directly encountering meningococci or through cross-reactive immunity with non-pathogenic neisserial strains, such as *Neisseria lactamica* which shares some lipooligosaccharide epitopes with Nm (54, 107). Antibodies, also known as immunoglobulins (Ig), are produced by B cells and are important in the fight against invading Nm. They assist in bactericidal killing, activating complement mediated opsonophagocytosis, and interacting with T helper cells leading to long-term memory (108). Out of the five known immunoglobulin classes, IgA, IgD, IgE, IgM, and IgG, the main driver of the human humoral response to Nm by activating the classical complement pathway resulting in complement-mediated killing by the MAC complex and phagocytosis of meningococci through C3 opsonisation is IgG (105). IgG is also the only antibody class able to traverse the placenta and is the main source of protective maternal antibodies provided to neonates via transplacental transfer and through ingesting breast milk (109). Anti-meningococcal IgG importance is highlighted further in age-related incidence of IMD inversely correlated to serum IgG levels (110). IgA and IgM have also been shown to have roles in protection against invading meningococci. IgA is mainly expressed at mucosal surfaces including the respiratory and gastrointestinal tracts and are known more for their neutralising activity by binding to meningococci surfaces, as their ability to induce opsonisation and phagocytosis of bacteria is weaker compared to IgG (111, 112). IgM antibodies do not opsonise or phagocytose pathogens, but they are able to highly stimulate the classical complement pathway and induce formation of a high level of MACs. Some neisserial strains have evolutionary developed evasion methods for the host antibody response, for example Nm produces IgA1 proteases that have able to cleave IgA antibodies, at the Fc region, therefore stunting its ability to perform phagocytosis and signal further downstream immune responses (113).

1.4 Genetic basis underlying invasive meningococcal disease

There have been decades of research investigating the role of human genetics in predisposition to Nm infection, stretching back to the discovery of complement deficiencies increasing risk of IMD in the 1970s (114, 115). Both rare monogenic variants and common polymorphisms have been associated with IMD, and their discoveries have assisted in elucidating the host susceptibility, pathogenesis of Nm, and aetiology of IMD. This section will briefly outline variants that have been associated with IMD with respect to the stages of Nm infection, with a more exhaustive description available in a review we recently published (11).

1.4.1 Colonisation

During Nm colonisation of the epithelial surface of the human nasopharynx, bacterial surface proteins like pili and Opa proteins adhere to the host cell and facilitate successful attachment. In a recent study, our lab described a novel dominant negative heterozygous variant in short palate, lung, and nasal epithelial clone 1 gene (*SPLUNC1*, also known as *BPIFA1*) found in two siblings suffering from meningococcal sepsis (116). *In vitro* studies showed the *SPLUNC1* variant resulted in reduced antibiofilm activity thereby increasing Nm adhesion and invasion of cells (116) (Table 1.1). Polymorphisms in human *CEACAM3* and *CEACAM6* proteins, known binding partners of Nm bacterial attachment proteins, have been associated with IMD with *CEACAM3* haplotype C and *CEACAM6* haplotype B associated with protection against IMD and *CEACAM6* Haplotype C associated with increased susceptibility to IMD (117). Surfactant protein A2, encoded by *SFTPA2,* is expressed in the nasopharynx, and is involved in innate immunity, as it binds to microbial cell walls and can activate the phagocyte and inflammatory responses. A single candidate gene study showed various polymorphisms in *SFTPA2* resulting in either protective or increased susceptibility risk in IMD (118) (Table 1.1).

Table 1.1 Genetic aetiologies reported to be associated with IMD, involved in meningococci colonisation

Gene	Variant	Genetic model	Impact	Gene specific phenotype	Infection Phenotype	Disease `outcome	Reference
SPLUNC1	c.65G>A, p.G22E	AD	Monogenic	Increased bacterial adhesion	IMD	Increased susceptibility	(119)
CEACAM3	Haplotype C	Additive	Polymorphism	N/A	IMD	Protective	(117)
CEACAM6	Haplotype B	Additive	Polymorphism	N/A	IMD	Protective	(117)
	Haplotype C	Additive	Polymorphism	N/A	IMD	Susceptibility	(117)
SFTPA2	$1A^1$ genotype	Recessive	Polymorphism	N/A	IMD	Susceptibility	(118)
	$1A^5$ genotype	Dominant	Polymorphism	N/A	IMD	Protective	(118)
	rs1965708	Recessive	Polymorphism	N/A	IMD	Susceptibility	(118)

1.4.2 Invasion

Successful invasion of Nm is poorly understood but results in dissemination of meningococci in the bloodstream. It is well established that complement mediated bactericidal activity plays a central protective role against Nm infection, and complement deficiencies leave a hole in an important host defence mechanism and can lead to IMD. Complement deficiencies can be acquired or inherited, with frequencies depending on complement component and ethnicity, inherited complement deficiencies are rarer than acquired, occurring in 0.03% of the general population (120). Impaired complement function that increases risk to IMD was first characterised functionally in terminal complement deficiencies (*C5, C6, C7, C8A, C8B,* or *C9*), in the 1970s/80s, and later genetically as autosomal recessive monogenic disorders (Table 1.2). Patients with these deficiencies are clinically distinct customarily presenting with recurrent meningococcal infection, with lower mortality rates per incidence (121, 122). As all complement pathways feed into the terminal complement components, deficiencies in these can stunt all complement response to infection, however there are other deficiencies that affect the separate top branches of the different complement pathways.

In the alternative complement pathway functional deficiencies in all the related factors have been associated with IMD. The most well studied are those affecting properdin (CFP), a positive regulator of the AP as it can bind to and stabilise C3b thereby prolonging its functional activity (120). *CFP* is encoded on the X chromosome and genetic recessive deficiencies are typically connected with non-recurrent and rapidly progressive fatal meningococcaemia, with around 50% of patients infected by Nm serogroups W and Y (Table 1.2) (121). Autosomal recessive complement deficiencies associated with IMD have also been identified in AP complement factors D, B, and I, typically discovered from familial, or small patient number studies (Table 1.2). Alternatively, GWAS have been utilised recently and polymorphisms in a broad genetic region across complement factor H (*CFH*) and complement factor H-related 3 (*CFHR3*) was identified to be highly significantly associated with IMD, and this finding was validated in different cohorts (Table 1.2) (25, 123). CFH acts as a negative regulator of the AP and competes with CFB resulting in inactive C3b (104). CFH and CFHR proteins, which share partial homology to CFH, are also able to bind to Nm factor H-binding protein (fHBP) which is expressed on the bacteria's surface, this mechanism is typically used by Nm for immune evasion and deficiencies in CFH and CFHR3 could antagonise this evasion technique (124).

Table 1.2 Genetic aetiologies reported to be associated with IMD, involved in meningococci invasion

Gene	Variant	Genetic model	Impact	Gene specific phenotype	Infection Phenotype	Disease outcome	Reference
CFP	c.481C>T, p.R161X	XR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(125)
	c.1240T>G, p.Y414D	XR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(126)
	c.617C>G, pS206X	XR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(127)
	c.893G>T pG298V	XR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(127)
	c.1164G>A, p.W388X	XR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(128)
C5	c.1055A>G, p.Y352C	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(129)
	c.754G>A, p.A252T	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(130)
	c.55C>T, p.Q19X; c.4444C>T, p.R1482X	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(131)
	c.1115A>G, p.K372R	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(132)
	c.4890-4891delinsG, p.L1631fs	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(133)
C6	c.878delA	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(134)
	c.1599T > A, p.Y493X ; IVS3+3A>C	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(134)

	c.1936delG	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(135)
	c.879delG	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(136)
	c.1195delC	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(137)
C7	c.2107C>T, p.Q681X	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(138)
	c.2184T>A, p.C728X	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(139)
	c.281-1G>T c.1-?-2350+?del	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(140)
	c.1135G>C, p.Gly379Arg	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(141)
	c.1922delAG	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(142)
	c.633_643del c.1922delAG [a]	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(143)
C8B	c.1282C>T, p.R428X	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(144)
	c.271C>T, p.Q91X; c.820C>T, p.R274X	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(145)
	c.1041_1047dup, p.L350fs c.271C>T, p.Q91X [a]	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(146)
C9	c.346C>T, p.R116X	AR	Monogenic	Reduced complement function	Meningitis	Increased susceptibility	(147)
	c.162C>A, p.C54X	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(148)
CFD	c.638T>G, p.V213G;	AR	Monogenic	Reduced complement function	Sepsis	Increased susceptibility	(149)

	c.640T>C, p.C214R						
	c.125C>A, p.S42X	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(150)
CFB	c.766C>T, p.Q256X c.1894-1897delTTTG, p.F632C**fs**X8 [a]	AR	Monogenic	Reduced complement function	Meningitis	Increased susceptibility	(151)
CFI	c.1282A>T p.H400L	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(152)
	c.1282A>T p.H400L c.801G>A	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(152)
	del-exon 5 [a]						
	c.266_?_536 + ?del c.1420C>T, p.R474X	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(153)
	c.485G>A, p.G162D c.1176_1177dupAT, p.W393Yfs*5	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(153)
	c.772G>A, p.A258T	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(153)
C3	c.1716G>A, p.K552X	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(154)
C2	28-bp deletion at the 3′ end of exon 6	AR	Monogenic	Reduced complement function	IMD	Increased susceptibility	(155)
CFHR3	rs426736	Additive	Polymorphism	N/A	IMD	Increased susceptibility	(25)
CFH	c.-496C>T	Recessive (REF)	Polymorphism	High fH levels and reduced bactericidal activity	IMD	Increased susceptibility	(156)

	rs1065489	Additive	Polymorphism	N/A	IMD	Increased susceptibility	(25)
	rs1061170	Dominant	Polymorphism	N/A	IMD	Increased susceptibility	(157)
	rs3753396	Dominant	Polymorphism	N/A	IMD	Protective	(157)
MBL2	c.154C>T, p.R52C; c.161G>A, p.G54D; c.170G>A, p.G57E	Dominant	Polymorphism	N/A	IMD	Increased susceptibility	(158)
	c.154C>T, p.R52C; c.161G>A, p.G54D; c.170G>A, p.G57E	Dominant	Polymorphism	N/A	IMD	Increased susceptibility (hospital cohort)	(159)
	c.154C>T, p.R52C; c.161G>A, p.G54D; c.170G>A, p.G57E	Dominant	Polymorphism	N/A	IMD	Increased susceptibility (community-based study)	(159)

(REF), denotes referent allele

Deficiencies in C3 are rarely identified, most likely as it is fundamental to healthy survival due to the chief role it plays in innate immunity, but some patients with C3 autosomal recessive deficiencies have been identified to suffer from recurrent bacterial infections, including Nm (Table 1.2) (154). Early complement components, C1, C2, and C4, are involved in the classical complement path and autosomal recessive deficiencies are typically associated with autoimmune disease, the chief of these being systemic lupus erythematosus (SLE) and their impact in IMD is heavily questioned with very few IMD patients having deficiencies in early complement components. Variants in the lectin complement pathway have also been investigated with candidate gene studies which highlighted three polymorphisms in exon 1 of mannose-binding lectin, encoded by *MBL2,* to be associated with reduced plasma protein concentrations and increased incidence of IMD with younger age (Table 1.2) (158, 159). However, other studies have showed that *MBL2* polymorphisms are not significantly associated with IMD (160, 161).

1.4.3 Inflammatory response

Autosomal recessive monogenic disorders in *IRAK4*, and *IKBKG* (NEMO) have been associated with IMD, with these deficiencies shown to dysregulate cytokine production, including an abolished IL-6 response, therefore compromising the function of some of the main signalling molecules in the innate immune response (Table 1.3) (8, 162). A variant in the hematopoietic transcription factor *GATA2,* leading to haploinsufficiency, in a cytopenic patient infected with Nm has also been reported (Table 1.3) (163). Recently, a monoallelic missense variant in *NFKB1*, has been reported in two siblings presenting with meningococcal meningitis with sepsis, and functional studies showed that this resulted in impaired nuclear translocation (164) (Table 1.3).

Polymorphisms in genes involved in the inflammatory response have also been associated with IMD, specifically in toll-like receptors and key cytokines involved in Nm infection, these polymorphisms are typically discovered by candidate gene studies. Polymorphisms in *TLR4* have shown to mediate a hypo-responsiveness to LPS and a *TLR9* polymorphism indicates a protective effect against meningococcal meningitis in a paediatric cohort (Table 1.3) (165, 166). A rare *TNF* polymorphism found in the promoter region has been shown to be associated with IMD though studies are conflicting. Two studies show the single nucleotide polymorphism (SNP) resulted in increased TNF-α levels that may predispose individuals to IMD (167, 168), though other studies disagree and show no association between IMD and the

rare *TNF* allele (Table 1.3) (169-172). TNF-α is a key activator of the inflammatory response and mediates septicaemia, with high levels of circulating TNF-α and IL-6 correlated to increase risk of severe IMD, and death (90, 91). IL-6 is another key cytokine involved in haematopoiesis and the inflammatory response to Nm infection, it is a pro-coagulant secreted by T cells and macrophages and an *IL6* SNP has been associated with increased IMD fatality (Table 1.3) (169, 170, 173). SNPs in pro-inflammatory cytokines *IL1RN* and *IL1B* and anti-inflammatory cytokine *IL10*, have been associated with poor outcome in IMD (Table 1.3).

1.4.4 Acquired immunity

Antibody-mediated phagocytosis is a key immune defence against Nm infection and Fc receptors for IgG (FcγR) can be found on phagocytes, this includes three FcγR subclasses IIa, IIIa, and IIIb encoded by *FCGR2A*, *FCGR3A*, and *FCGR3B*, respectively that are involved in Ig-G binding efficiency. FcγRIIa binds IgG2 on poly-morphonuclear cells and is effective at IgG2-mediated phagocytosis of encapsulated bacteria, including Nm (174, 175), FcγRIIIa binds IgG1, IgG3, and IgG4 on monocytes, macrophages and natural killer cells, and FcγRIIIb binds IgG1, and IgG3 on neutrophils. Candidate gene studies have shown that allotypes within these three subclasses are associated with IMD either in a protective or antagonistic role depending on the binding efficiency of the certain allotypes (Table 1.4). A combination of polymorphisms in all three subclasses in conjunction have also been reported to be associated with IMD (176, 177). Again, similar to other polymorphisms identified by candidate gene studies that have been discussed thus far, there are conflicting studies that show no significant correlation between FcγR polymorphisms and IMD (176, 178).

Table 1.3 Genetic aetiologies reported to be associated with IMD, involved in the inflammatory response

Gene	Variant	Genetic model	Impact	Gene specific phenotype	Infection Phenotype	Disease outcome	Reference
IRAK4	c.877C > T, p.Q293X	AR	Monogenic	Impaired IL-6 production	IMD	Increased susceptibility	(179)
IKBKG	c.1249T>C, p.C417R	XR	Monogenic	N/A	IMD	Increased susceptibility	(180)
NFKB1	c.691C>T, p.R230C	AD	Monogenic	Impaired p50 nuclear translocation	IMD	Increased susceptibility	(164)
TLR4	c.896A>G, p.D299G	Dominant	Polymorphism	N/A	Death	Increased susceptibility	(181)
	c.896A>G, p.D299G	Dominant	Polymorphism	N/A	IMD	Increased susceptibility	(182)
TLR9	rs352140	Dominant	Polymorphism	N/A	Meningitis	Protective	(166)
TNF	rs1800629	Dominant	Polymorphism	High TNF-α levels	Death	Increased susceptibility	(167)
	rs1800629	Dominant	Polymorphism	High TNF-α levels	IMD	Increased susceptibility	(167)
	rs1800629	Dominant (REF)	Polymorphism	N/A	IMD	Increased susceptibility	(169)
	rs1800629	Recessive (REF)	Polymorphism	N/A	IMD	Increased susceptibility	(169)

	rs1800629	Recessive	Polymorphism	High TNF-α levels	IMD	Increased susceptibility	(168)
IL1B	rs16944	Recessive (REF)	Polymorphism	N/A	Death	Increased severity	(172)
	rs16944	Recessive	Polymorphism	N/A	Death	Increased severity	(172)
	rs16944	Dominant (REF)	Polymorphism	N/A	Death	Protective	(183)
IL1B/ IL1RN	rs16944/ rs419598 (REF)	Dominant/ Recessive	Polymorphism	N/A	Death	Protective	(172)
IL1B/ IL1RN	rs16944 / rs419598	Dominant	Polymorphism	N/A	Death	Increased susceptibility	(183)
IL1RN	86-basepair variable number tandem repeat (VNTR)	Recessive	Polymorphism	N/A	IMD	Increased susceptibility	(170)
	rs419598	Recessive	Polymorphism	N/A	IMD	Increased susceptibility	(184)
IL-6	rs1800795	Recessive (REF)	Polymorphism	N/A	Death	Increased susceptibility	(170)
	rs1800795	Recessive (REF)	Polymorphism	N/A	IMD	Increased susceptibility	(169)
IL10	rs1800896	Recessive	Polymorphism	N/A	IMD	Increased susceptibility	(170)

(REF), denotes referent allele

Table 1.4 Genetic aetiologies reported to be associated with IMD, involved in acquired immunity

Gene	Variant	Genetic model	Impact	Gene specific phenotype	Infection Phenotype	Disease outcome	Reference
FCGR2A	c.500A>G, p.H131R	Recessive	Polymorphism	Reduced phagocytosis	Sepsis	Increased susceptibility	(185)
	c.500A>G, p.H131R	Recessive	Polymorphism	N/A	IMD	Increased susceptibility	(175)
	c.500A>G, p.H131R	Recessive (REF)	Polymorphism	N/A	IMD	Protective	(175)
	c.500A>G, p.H131R	Dominant	Polymorphism	N/A	IMD	Increased susceptibility	(174)
	c.500A>G, p.H131R	Recessive	Polymorphism	N/A	IMD	Increased susceptibility	(186)
	c.500A>G, p.H131R	Recessive	Polymorphism	N/A	IMD	Increased susceptibility	(186)
	c.500A>G, p.H131R	Recessive	Polymorphism	N/A	IMD	Increased susceptibility	(171)
FCGR2A/ FCGR3B	c.500A>G, p.H131R/ NA2 allotype	Recessive	Polymorphism	Reduced phagocytosis	IMD	Increased susceptibility	(177)

(REF), denotes referent allele

1.4.5 Coagulation and fibrinolysis

Meningococcal endotoxins, chiefly LPS/LOS, are expressed on the outer bacterial membrane and circulate in the host bloodstream during Nm infection, strongly activating the coagulation pathway and generating thrombin (51). Severe IMD is often characterised by coagulopathy, which can result in the most severe clinical thrombotic complication, purpura fulminans, developing in up to 20 % of cases and can progress into skin necrosis requiring grafting or amputations (51, 187, 188). Aberrated regulation of coagulation by the fibrinolytic system due to polymorphisms in genes involved in these processes, discovered by candidate gene studies, have been linked to manifestation of the most severe IMD phenotypes (Table 1.5) (51). A SNP in the *F5* gene, named the factor V Leiden mutation (FV^L), is associated with thrombotic events development of severe purpura fulminans in IMD (Table 1.5) (187, 189). FV^L also shows resistance to activated protein C, a key anti-coagulant activated by thrombin that can inhibit plasminogen activator inhibitor (PAI) and deactivate factor V, and factor VIII, to downregulate a pro-coagulation signalling cascade. Polymorphisms in the *PROC* promoter region, which encodes protein C, resulting in haplotypes producing low protein C plasma levels have been associated with meningococcal septicaemia and increased IMD severity (Table 1.5) (190-192).

Thrombin activatable fibrinolysis inhibitor (TAFI), encoded by *CPB2* gene, attenuates the fibrinolytic response and a polymorphism has been shown to increase its anti-fibrinolytic potential and increase the risk of severe IMD and death (Table 1.5) (193, 194). PAI-1, encoded by *SERPINE1*, inhibits plasminogen activators like tPA and uPA and is responsible for blood clotting (haemostasis). However, rising levels of PAI-1 in the blood stream can shift the balance from haemostasis to thrombosis. A 4G/5G insertion/deletion polymorphism in the *SERPINE1* promoter region has been shown to increase PAI-1 levels and is associated with disseminated intravascular coagulation, severe meningococcal septic shock, and death in IMD (Table 1.5) (195-198). To note, a meta-analysis report shows no association between the 4G/5G *SERPINE1* polymorphism and sepsis susceptibility (199)

Table 1.5 Genetic aetiologies reported to be associated with IMD, involved in coagulation and fibrinolysis

Gene	Variant	Genetic model	Impact	Gene specific phenotype	Infection Phenotype	Disease outcome	Reference
F5	c.1601G>A, p.R534Q (FV^{L})	Dominant	Polymorphism	Increased thrombosis	Purpura fulminans	Increased severity	(187)
	c.1601G>A, p.R534Q (FV^{L})	Recessive	Polymorphism	Increased thrombosis	Purpura fulminans	Increased severity	(189)
SERPINE1	rs1799889	Recessive	Polymorphism	Higher PAI-1 concentration	Death	Increased severity[a]	(200)
	rs1799889	Recessive	Polymorphism	N/A	Death	Increased severity[a]	(195)
	rs1799889	Recessive	Polymorphism	N/A	Sepsis	Increased severity[a]	(195)
	rs1799889	Recessive	Polymorphism	N/A	Vascular complications	Increased severity[a]	(195)
	rs1799889	Recessive	Polymorphism	N/A	Sepsis	Increased severity[a]	(196)
	rs1799889	Recessive	Polymorphism	N/A	Meningitis	Increased severity[b]	(196)
	rs1799889	Recessive	Polymorphism	N/A	Death	Increased severity[a]	(197)
	rs1799889	Recessive	Polymorphism	N/A	Sepsis	Increased severity[a]	(197)

	rs1799889	Recessive	Polymorphism	N/A	Disseminated intravascular coagulation	Increased severity[a]	(198)
PROC	rs1799808 (REF), rs1799809	Recessive	Polymorphism	Low protein C levels	IMD	Increased susceptibility	(190)
	rs1799808, rs1799809	Dominant	Polymorphism	N/A	Sepsis	Increased severity	(190)
	rs1799808, rs1799809 (REF)	Recessive	Polymorphism	N/A	Sepsis	Protective	(190)
CPB2	c.1040C>T, p.T325I	Recessive	Polymorphism	N/A	Death	Increased severity	(194)
	c.1040C>T, p.T325I	Recessive	Polymorphism	Increased antifibrinolytic activity	Sepsis	Increased severity	(193)

(REF), denotes referent allele. [a]4G/4G genotype; [b]5G/5G genotype

Alternatively, polymorphisms, which are more common, are typically discovered via candidate gene studies in cohorts, or within large studies such as a GWAS. As described here, there have been multiple rare monogenic (IEI) disorders and common polymorphisms associated with IMD, however, many of these are either the result of a single candidate gene study or have multiple studies with conflicting results, and therefore many of these polymorphisms would need further investigation to prove definite host genetic contribution to disease. IEIs have also been implicated in other important infectious diseases (7), including viral and other significant bacterial infections, some of which are relevant in this project and will be described in the following sections.

1.5 Other significant bacterial infections

Nm is a leading cause of bacterial septicaemia and meningitis worldwide, but there are other invasive bacterial infections that can cause significant health impact, including *Staphylococcus aureus, Streptococcus pneumoniae*, and *Group A Streptococcus.*

1.5.1 *Staphylococcus aureus*

1.5.1.1 Microbiology

Staphylococcus aureus (*S. aureus*) is a spherical gram-positive opportunistic bacterium that, similarly to Nm, is both a human commensal and invasive pathogen. *S. aureus* is a catalase-positive, facultative anaerobe and are coagulase-variable as though most are coagulase-positive, some may atypically not secrete coagulase, a bacterial enzyme that induces coagulation in the bloodstream by binding and activating prothrombin nonproteolytically and is a main driver of disease alongside von Willebrand factor binding protein (201, 202). Coagulase testing is a primary method used to detect *S. aureus* infection. *S. aureus* has a thick cell wall peptidoglycan layer that covalently anchors proteins at their C-terminus, labelled cell wall-anchored (CWA) proteins. These surface CWA proteins perform various functions for successful infection establishment, and some are key virulence factors in infection. The largest group of CWA proteins in *S. aureus* are the microbial surface components recognizing adhesive matrix molecules (MSCRAMM) family, structurally defined by two adjacent IgG-like folded domains (203). The MSCRAMM family consists primarily of adhesion proteins including fibronectin and collagen binding proteins and clumping factors A (ClfA) and B (ClfB) that assist in binding to different host tissues and driving colonisation (203). ClfA and serine aspartate repeat protein E (SdrE), are also vital for host immune evasion as ClfA binds to complement factor I, acting as a co-factor and increasing cleavage of C3b to iC3b, thereby degrading C3b and limiting its bactericidal opsonin activity (204). SdrE binds to complement factor H, a factor essential for the downregulation of the alternative complement pathway, via a 'dock-lock-latch' mechanism to evade the host immune response (205).

1.5.1.2 Staphylococcus disease and health impact

S. aureus colonises skin and mucous membranes and ~30 % of the human population persistently carries *S. aureus* commensally on their anterior nares (206), and other areas such as skin, throat, and along the GI tract (207). Carriage is asymptomatic but invasion can occur, often due to damaged skin (i.e. a wound) and damaged mucous membranes (i.e. from viral

infection) providing opportunity, leading to either mild skin and soft tissue focal infections or more serious invasive life-threatening disease like septicaemia, septic arthritis, pneumonia, and endocarditis (208). *S. aureus* infections has a case fatality rate of ~25 % (209) and involves a complex relationship with the host and pathogenesis can be acute, recurrent, or chronic and persistent. As briefly discussed above, *S. aureus* has a wide repertoire of virulence factors and immune evasion mechanisms that it uses to establish successful bacterial dissemination, including sequestering inside host cells (210, 211).

S. aureus pathogenesis is multi-factorial from adherence to host cells, immune evasion, and damage to host cells via secreted toxins that can be characterised into three major groups, pore-forming toxins, exfoliative toxins, and superantigens. *S. aureus* α-toxin is one of the more potent toxins and once bound forms pores in target cells, causing them to leak their cellular contents and cytokines leading to the production of pro-inflammatory molecules that influence sepsis and development of septic shock (212). Host platelets and monocytes are particularly susceptible to α-toxin. Panton-Valentine leukocidin (PVL) is a pore-forming cytotoxin secreted by *S. aureus* and is chiefly responsible for tissue necrosis i.e. necrotic skin lesions and severe necrotizing pneumonia, and leukocyte destruction (213).

S. aureus is a major source of hospital acquired and community-acquired infections and treatment largely depends on the type of infection and if the strain is antibiotic resistance. Drug resistant or sensitive strains are largely classified into two major groups, Methicillin-Sensitive *Staphylococcus aureus* (MSSA), and Methicillin-Resistant *Staphylococcus aureus* (MRSA). Typically, penicillin is used in MSSA infections and vancomycin or daptomycin is used for MRSA strains, and in some cases additional therapy is necessary depending on the disease development (214). MRSA is widespread in hospitals and patients are typically exposed to MRSA during or after surgery, by medical devices like intravenous lines, or exposure to other patients infected with MRSA (214). Community-associated MRSA has also been recorded and is typically the result of recent stays in a healthcare facility, uncovered wounds in sports or crowded living spaces, sharing personal items, and injection drug use (215). There is currently no vaccine licensed against *S. aureus* thus far, though some are in the development and the increasing presence of MRSA makes the need for a vaccine, or new therapeutic targets, even more important.

1.5.1.3 Genetic basis underlying invasive staphylococcus disease

There are risk factors associated with development of invasive staphylococcal disease. Non-genetic risk factors include age with rates of infection high in those at either end of life, the very young or the elderly with an incidence rate of more than 100 per 100,000 in the over-70s age group (214, 216). Ethnicity is also a risk factor with double the incidence rate of MRSA seen in the black population in the USA compared to the white population, and seen in up to 20 times more in the indigenous population in Australia compared to the non-indigenous population of European ethnicity (217-219). However, it is arguable that this ethnicity risk factor is not a representation of genetic susceptibility and more likely the result of socioeconomic differences in ethnicity groups within those countries (219).

Genetic variation is also a risk factor with inborn errors of immunity associated with invasive staphylococcal disease. IEIs are rarely exclusively associated with *S. aureus* infection alone and more so associated to susceptibility of overall pyogenic bacterial infections. Autosomal recessive *IRAK4* and *MyD88* deficiencies are associated with predisposition to invasive pyogenic bacterial infections with *S. aureus* seen in 14 % of *IRAK4* deficiencies and 21 % of *MyD88* deficiencies (8). Bacterial infection in *IRAK4* and *MyD88* deficiencies are often seen before two years of age, but clinical presentation improves with age with no invasive bacterial infection seen in adulthood (8, 9). Both *IRAK4* and *MyD88* deficiencies also pre-dispose to non-invasive staphylococcal infection seen in 43 % and 53 % of patients, respectively (220). Biallelic variants resulting in autosomal recessive *HOIL1* deficiency, also known as *RBCK1,* and a component of the linear ubiquitin chain assembly complex (LUBAC), has been seen in individuals suffering from reoccurring pyogenic bacterial infections, including *Staphylococcus aureus* (221). Autosomal recessive *TIRAP* deficiency is associated with an incomplete penetrance of invasive staphylococcal infection and was primarily described in a case study involving a clinical phenotype of *Staphylococcus aureus* pneumonia and sepsis and an absent/reduced cytokine production level in response to TLR2 and TLR4 ligand in fibroblasts, monocytes and granulocytes was recorded (222).

1.5.2 *Streptococcus pneumoniae*

1.5.2.1 Microbiology

Streptococcus pneumoniae (*S. pneumoniae*) is a lancet shaped gram-positive, facultative anaerobic, extracellular bacterium often found in pairs (diplococci). *S. pneumoniae* is surrounded by an antiphagocytic polysaccharide capsule composed of peptidoglycan and teichoic acid, decorated with surface proteins. The capsule is a large contributor to the virulence due to its involvement in immune evasion and there are currently 46 known serogroups characterised by the capsules biochemical structure and serologic (immunologic) properties, and the most common disease-causing serogroups are 6, 14, 18, 19, and 23 (223). *S. pneumoniae* pili, LPxTG motif proteins, and choline binding protein A (CbpA) is key for cell adhesion as it binds to the polymeric immunoglobulin receptor on nasopharyngeal cells and along with other bacterial surface proteins promotes attachment and bacterial uptake (224). Bacterial aggregation and biofilm formation is mediated by the pneumococcal serine-rich repeat protein (PsrP) (225). Another vital protein is pneumococcal surface protein A (PspA) which plays a role in immune evasion as it inhibits complement deposition and activation, inhibits phagocyte engulfment, and can bind lactoferrin (224, 226).

1.5.2.2 Pneumococcal disease and health impact

S. pneumoniae is traditionally known to be the most common cause of community-acquired pneumonia (CAP) worldwide, today accounting for 27% of cases (227). It is a major cause of mortality and morbidity worldwide and is a significant cause of bacteraemia and meningitis, i.e. invasive pneumococcal disease. Development of disease is determined by *S. pneumoniae* ability to adhere and colonise the alveolar epithelium of the lower respiratory tract and invade. Invasive pneumococcal disease is typically the result of significant activation of the host inflammatory response that causes significant cell damage driven by bacterial antagonists. Almost 50 % of healthy children and 30 % of healthy adults are carriers of *S. pneumoniae* in their nasopharynx and a major risk factor for invasive disease incidence is age, with an incidence rate of 36.4 per 100,000 in over 65 years of age and 34.2 per 100,000 in infants less than one year of age, whereas the overall rate in adults is ~6 per 100,000 (228, 229). Upon successful colonisation, *S. pneumoniae* secretes large quantities of cytotoxins including pore-forming toxin pneumolysin which disrupt the alveolar epithelium and edema fluid builds up in the alveolar space leading to lobar pneumonia (230). Pneumolysin can be recognised by TLR4, inducing the inflammatory response, and it also activates the classical complement pathway

(231, 232). Release of pro-inflammatory cytokines, predominantly IL-6 and IL-1β, mediated by NF-κB signalling upon PAMP recognition, is necessary for infection clearance as they induce neutrophil recruitment and macrophage activation. However, this inflammation response can also contribute to lung damage and IEI's that decrease activation of this inflammation pathway have shown a protective effect against invasive pneumococcal disease (233). Pneumococcal pneumonia, if not immediately treated, can evolve into life-threatening complications such as lung abscess, empyema, necrotizing pneumonia, parapneumonic effusion, and septicaemia.

Treatment for invasive pneumococcal disease is an antibiotic regime, typically penicillin or a broad-spectrum of antibiotics if antibiotic resistance is suspected, with supportive therapy if necessary, such as mechanical ventilation. Fortunately, a pneumococcal conjugate vaccine (PCV, brand name Prevenar 13), covering serotypes 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 19A, 19F, 18C, and 23F, is available and in the UK it is part of the national vaccination schedules for children under 2 years old. A pneumococcal polysaccharide vaccine (PPV), covering 23 serotypes, is administered to people aged 65 and over and to immunocompromised individuals. The use of vaccines has reduced overall incidence of severe disease and hospitalisations but has caused an increase of invasive pneumococcal disease caused by serogroups not covered in the currently available vaccines (234, 235).

1.5.2.3 Genetic basis underlying invasive pneumococcal disease

Similarly to invasive staphylococcal disease, autosomal recessive inborn errors of immunity resulting in *IRAK4* and *MyD88* deficiencies are associated with predisposition to invasive pyogenic bacterial infections with *S. pneumoniae* seen in 54 % of *IRAK4* deficient infections and 45 % of *MyD88* deficient infections (8). *IKBKG* encodes for IKK-γ/NEMO, an essential protein needed to activate NF-κB signalling. Hypomorphic *IKBKG* mutations in over 100 patients have been identified with encapsulated pyogenic bacterial infections by *Haemophilus influenzae* and *Streptococcus pneumoniae* characterising the infectious clinical phenotype (8, 236, 237). One third of NEMO deficient patients have been reported to have died from invasive infections, and 72 % had an inability to generate pneumococcus-specific antibodies (238). Polymorphisms in *NFKBIA, and NFKBIE* are also associated with susceptibility to invasive pneumococcal disease. *NFKBIA* and *NFKBIE* belong to the IκB kinase family and are responsible for regulation of the inflammatory response by inhibiting NF-κB until the IκB kinase complex is degraded and NF-κB is free to activate pro-inflammatory gene transcription.

SNPs in *NFKBIA* promotor region are associated with protection from invasive pneumococcal disease, and more specifically pneumococcal empyema (233). Conversely, SNPs in *NFKBIE* were associated with increased risk of pneumococcal meningitis but not bacteraemia in a paediatric cohort (239).

1.5.3 *Group A Streptococcus*

1.5.3.1 Microbiology

Group A Streptococcus (GAS), also known as *Streptococcus pyogenes*, are gram-positive, facultative anaerobic, beta-haemolytic, extracellular ubiquitous bacteria that exclusively colonises humans and causes a wide range of invasive disease. GAS virulence is closely linked to its antiphagocytic hyaluronic capsule and its surface antigen, M protein, which are central in host immune evasion (240). Lipoteichoic acid and protein F also decorate the cell surface and are key bacterial adhesins molecules necessary for successful colonisation. The M-protein surface antigen is used to serotype GAS and there are currently more than 100 M serotypes, with the M1 serotype the predominant invasive disease-causing GAS serotype (240, 241). Possible mechanisms proposed that renders encapsulated GAS antiphagocytic is M protein binding to CFH and fibrinogen, which interferes with the alternative pathway activation of complement and reduces bound C3b (242, 243). This reduces leukocyte mediated phagocytosis, allowing GAS to survive longer in host tissues (242, 243). GAS colonises epithelial layers and its cell surface protein adhesins allow colonisation of distinct tissue sites, with the principal colonisation sites being the nasopharynx and skin but GAS has also been found elsewhere such as in the vagina and rectum (244). Colonisation of the nasopharynx and skin often results in asymptomatic or mild disease such as such as pharyngitis and impetigo, with invasion into the deeper layers of the body such as blood and organs resulting in severe life-threatening disease.

1.5.3.2 Streptococcal disease and health impact

GAS serotype M1T1 clone is found worldwide and is the primary disease-causing serotype. GAS infection peaks in winter and early spring with ~700 million cases of pharyngitis recorded annually worldwide, and is particularly prevalent in causing bacterial pharyngitis in school-aged children (245). Invasive GAS (iGAS) is rare but can result in life-threatening disease, with streptococcal toxic shock syndrome (STSS) and necrotizing fasciitis as some of the most

severe manifestations (240). iGAS is responsible for more than 500,000 deaths annually worldwide and STSS has an incidence rate of ~3 cases per 100,000 people per year in Europe. The USA has a predicted necrotizing fasciitis incidence rate of 2.5 per 1,000,000 people per year (246-248). Acute rheumatic fever is an autoimmune disease that can be triggered by GAS infection and, in low to middle income countries, can lead to high mortality rates from rheumatic heart disease (244). Concerningly, a higher than usual incidence rate of iGAS and related deaths has been seen in children in the UK and France in the past year (248). GAS remains globally sensitive to penicillin, and in situations of penicillin allergy, clindamycin and macrolides can be used as treatment, however, there have been reports of GAS resistant to clindamycin and macrolides (240). Even with treatment, severe toxic symptoms can result in a mortality rate ~50 % (249). There is no currently licensed vaccine against GAS, but potential antibiotic resistant strains in the future may drive the need for an effective vaccine to be developed.

1.5.3.3 Genetic basis underlying invasive streptococcal disease

Though iGAS is rare it carries a severe disease burden and high case fatality rate, the reason why healthy individuals can develop iGAS is still unknown and host genetic variability could explain susceptibility to iGAS, though IEIs in relevance to GAS infections remains ill-studied. A recent study showed an association of iGAS to the *HLA-DQA1**01:03 haplotype allele and postulates an altered binding interaction between HLA and GAS superantigens (250). However, this study had a small sample size, though that can be typical of rare disease cohorts, and most of their cases had a specific clinical phenotype of necrotizing fasciitis making it difficult to determine if the genetic association is specific to the necrotizing fasciitis phenotype or can be generalised to susceptibility to all invading GAS (250). Independent studies would need to be done to validate this finding.

1.6 COVID-19

1.6.1 SARS-CoV2

Infectious diseases have a great impact in all aspects of human society, highlighted more so by the recent devasting severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic. SARS-CoV-2 is an extremely transmissible pathogenic coronavirus that emerged in 2019/2020 and is responsible for coronavirus disease (COVID-19) which typically presents as a respiratory illness that can develop into severe morbidity and/or mortality. SARS-CoV-2 belongs to the *Coronaviridae* family and the betacoronavirus genus. There are four different genus in the *Coronaviridae* family, alpha, beta, gamma, and delta. Alpha- and betacoronviruses are known to infect only mammals, causing respiratory illness in humans, while gamma- and deltacoronaviruses typically infect birds (251). Coronaviruses are common pathogens that infect humans and animals, with four coronaviruses that are endemic to humans (human coronavirus NL63 (HCoV-NL63), HCoV-229E, HCoV-OC43, and HCoV-HKU1) and infect the upper respiratory tract, contributing to 15 % – 30 % of cases of seasonal common colds in adults (252, 253). Though, severe disease can arise from infection in at risk groups such as the immunocompromised, infants, and the elderly. Furthermore, coronaviruses can be zoonotic and spill over from animal reservoirs have resulted in significant pathogenic coronavirus strains that have had a substantial impact of human heath, well-being, and society. The most notable coronaviruses being severe acute respiratory syndrome coronavirus (SARS-CoV), Middle East respiratory syndrome coronavirus (MERS-CoV) and SARS-CoV-2.

The high chance of recombination and mutation rate also make coronaviruses important pathogens to monitor. SARS-CoV emerged in China in 2002 to cause two major but self-limiting epidemics between 2002-2004 that was mainly contained in Asia, with a major outbreak in Toronto, Canada (254). SARS-CoV-2, the most recent pathogenic zoonotic coronavirus, emerged in the Wuhan province in China late in 2019 and shares 86.85 % sequence similarity with SARS-CoV (255). Currently, the closest relative of SARS-CoV-2 is the bat coronavirus RaTG13 sharing 96.75 % sequence similarity. The exact origin or animal reservoir of SARS-CoV-2 is not known, but due to the high sequence similarity with RaTG13 it is highly suspected for SARS-CoV-2 to have emerged from a bat coronavirus potentially facilitated by an intermediate host species (256). SARS-CoV-2 is transmitted mainly via respiratory droplets and has an enhanced rate of transmission compared to SARS-CoV, a key

reason for the quick global spread of SARS-CoV-2 that resulted in the first worldwide pandemic for in generations.

1.6.2 Microbiology

SARS-CoV-2 is a positive sense, single-stranded enveloped RNA virus. SARS-CoV-2 is enveloped with a lipid bilayer membrane decorated with trimeric spike proteins that are used for cellular receptor binding, trigging virion fusion to the membrane and cell egress, with angiotensin-converting enzyme 2 (*ACE2*) identified as the principal host cell receptor (257, 258). It has also been shown that secondary host factors such as transmembrane protease serine protease 2 (*TMPRSS2*) are necessary for viral entry, and can also explain SARS-CoV-2 tissue tropism to airway cells and both *ACE2* & *TMPRSS2* are high highly co-expressed in nasal epithelial, alveolar epithelial and bronchial transient secretory cells (259, 260). Once in the cytoplasm, the viral positive single strand RNA can be directly translated into initial viral polyproteins by hijacking the host cells ribosomes as it mimics messenger RNA (mRNA) (261). Proteolysed viral polyproteins then form the replication-transcription complex, which generates both an antisense negative-strand template from viral RNA for genomic replication and continues mRNA synthesis and splicing. The latter produces viral structural proteins, and double-membrane vesicles are also formed from the endoplasmic reticulum and Golgi to segregate viral replication and transcription to avoid host detection (261). Structural viral membrane (M) proteins complex with viral RNA and nucleocapsid (N) proteins to form new infectious virions that are released extracellularly via exocytosis (261). Though SARS-CoV-2 has evolved certain mechanisms to avoid host detection, the process of entry and replication exposes viral PAMPs and provides opportunity for host sensing via PRRs which triggers the first lines of defence, the innate immune and inflammation responses.

1.6.3 Type I interferon response to SARS-CoV-2 infection

SARS-CoV-2 PAMPs can typically trigger the host immune response via PRRs found on innate immune cells like macrophages, monocytes, neutrophils, and natural killer cells. Several PRRs have been shown to activate during SARS-CoV-2 infection including, toll-like receptors (TLRs), retinoic acid-inducible gene I (RIG-I)-like receptors (RLRs), nucleotide-binding oligomerization domain (NOD)-like receptors (NLRs), and inflammasomes (262). This project will focus on the TLR response, specifically the TLR-3 pathway. TLRs are found on the host cell surfaces, and endosomal membranes can respond to viral PAMPs and trigger the immune signalling pathways. TLR-1, -2, -4, and -6 are found on the cell surface and recognise viral

proteins. This recognition triggers signalling through the MyD88 adaptor molecule and induces transcription of proinflammatory cytokines via NF-κB and MAPK signalling. TLR3 is intracellularly expressed in endosomes and lysosomes and recognises viral dsRNA triggering a TICAM1-dependent, MyD88-independent, signalling pathway. This pathway is mediated by NF-κB and IRF3 activation, and results in the induction of type I and type III IFNs, proinflammatory cytokines, and chemokines. TLR4 is also capable of activating this TICAM1-dependent signalling pathway in addition to the MyD88-dependent pathway (263). TLR3 is widely expressed in different cell types but is notable absent from plasmacytoid dendritic cells and neutrophils, with minimal expression seen in T cells (264).

Induction of IFNs, cytokines, and chemokines aid in clearing infection by signalling the immune and inflammation response and creating an anti-viral environment. For example, type I IFNs upregulate induction of IFN-stimulated genes (ISGs) such as dsRNA-binding protein kinase R and 2′–5′ oligoadenylate synthetase, within virus-infected cells and halt protein translation and induce RNA degradation (265, 266). However, dysregulated excessive induction of pro-inflammatory molecules can lead to a cytokine storm resulting in inflammatory cell death, a life-threatening condition (267). Cytokine storms, involving high levels of IL-1, IL-6, IL-4, IL-10, and IFN-γ, are often seen in symptomatic COVID-19 cases (268).

1.6.4 COVID-19 disease, treatments, and vaccination

COVID-19 pandemic caused by SARS-CoV-2 infection seized the world and continues to cause a major global health crisis. Its significant impact is mainly due to its high transmissibility rate and ability to cause extremely severe disease in those infected with an overall case fatality rate of ~1 % (267). As of January 2023, the WHO has reported over 750,000,000 confirmed cases of COVID-19, including over 6.8 million deaths (https://covid19.who.int/). Up to 20 % of COVID-19 patients require hospitalisation, of which up to 30 % need intensive care (267). This highlights the high disease burden of SARS-CoV-2 infection, which during pandemics and outbreaks can overwhelm national health systems. Some SARS-CoV-2 infections are asymptomatic, ~17 - 33 % (269, 270) but, in some cases, it can result in mild to severe respiratory disease. The most severe outcomes of SARSCoV-2 infection are acute respiratory distress syndrome (ARDS) that can lead to life-threatening pneumonitis (271). Extrapulmonary diseases have also been reported in COVID-19 cases

including multi-organ failure, coagulopathy, cardiac arrhythmias, rhabdomyolysis, and shock (271).

A variety of therapeutic options are available to treat COVID-19 including antiviral drugs, anti-SARS-CoV-2 monoclonal neutralising antibodies, anti-inflammatory drugs, and immunomodulators agents (272). However, as with most infectious diseases, the most crucial step to contain and limit the high health impact is development and implementation of effective vaccines. With world-wide collaboration multiple effective vaccines against SARS-CoV-2 were created and approved in record time, some notable ones used in the UK are: BNT162b2 an mRNA-based vaccine from BioNTech/Pfizer, mRNA-1273 vaccine another mRNA-based vaccine from Moderna, and ChAdOx1 nCoV-19 a viral vector vaccine from Oxford/AstraZeneca. As of January 2023, the WHO has reported administration of over 13 billion vaccine doses worldwide (https://covid19.who.int/).

1.6.5 Inborn errors of type I IFN immunity in viral infections

To date, there are 21 IEI's in the type I IFN pathway associated with major viral infections, including COVID-19, HSE, Influenza, Yellow fever virus (YFV), varicella zoster virus (VZV), and Measles, Mumps, Rubella (MMR), these are summarised in Table 1.6, Deficiencies in the TLR3-IFN pathway have shown to confer susceptibility to herpes virus encephalitis including autosomal recessive (AR) *UNC93B1*, a transmembrane protein that transports TLR3 from the endoplasmic reticulum (ER) to the endosome (273). Additionally, autosomal dominant (AD) *IRF3*, an IFN transcription factor that localises to the nucleus after TLR3-TICAM1 interaction, *TBK1,* a non-canonical Ikβ-related kinase that mediated IRF3 activation, and *TRAF3*, a facilitator for TBK1-TICAM1 interactions, and both AR and AD *TLR3,* and *TICAM1* deficiencies (12, 274-279). *TLR3* deficiency has also been reported in influenza A associated ARDS (280). *TICAM1* will be a focus in this project.

Table 1.6 Monogenic inborn errors of immunity implicated in viral infections

Gene	Inheritance	Viral pathogen					
		COVID-19	Influenza	HSE	MMR	YFV	VZV
IFNAR1	AR	✓		✓	✓	✓	
IFNAR2	AR				✓	✓	
TYK2	AR	✓		✓	✓		✓
STAT1	AR		✓	✓	✓		✓
STAT2	AR	✓	✓		✓		
IRF9	AR	✓	✓				✓
IRF7	AR	✓	✓				
TLR3	AR, AD	✓	✓	✓			✓
TICAM1	AR, AD			✓			
UNC93B1	AR			✓			
TBK1	AR, AD	✓		✓			
IRF3	AD			✓			
NEMO	XR			✓	✓		
TLR7	XR	✓					
MYD88	AR	✓					
IRAK4	AR	✓					
MDA5	AR, AD		✓				
POLR3A	AD						✓
POLR3C	AD						✓
MX1	AD		✓				

Adapted from Casanova and Anderson (2023) (281)
HSE; herpes simplex encephalitis, MMR, Measles, Mumps, Rubella; YFV, yellow fever virus; VZV, varicella zoster virus.

1.6.6 Role of TICAM1 in TLR3-IFN pathway

TICAM1, also known as TRIF, (TIR domain-containing adaptor inducing interferon-β/TIR domain-containing adaptor molecule 1), is the key adaptor molecule for the TLR3 signalling pathway that is chiefly activated by the dsRNA PAMP, from invading viruses. TICAM1 is used in signalling pathways independent of MyD88 and once bound to activated TLR3, TICAM1 will recruit TRAF3 which undergoes self- ubiquitination to then complex with TBK1 and IKKi/IKKε kinases, leading to phosphorylation of the kinases which triggers phosphorylated activation of IRF3 (282-285). Activated IRF3 then forms homodimers and heterodimers with IRF7, which facilitates translocation into the nucleus and targeted transcription of IFN and ISGs (282-284). TICAM1 can also recruit TRAF6, triggering complex formation with NEMO and IKKα/IKKβ kinases which facilitates downstream activation of NF-κB (282). TICAM1s C-terminal TIR domain is crucial for TIR:TIR domain interactions, chiefly facilitating signalling for IFN induction via IRF3 whereas the N-terminal region of TICAM1 coordinates the signalling cascade downstream of TRAF6 which activates NF-κB (285-287).

1.6.6.1 Association of *TICAM1* variants in viral disease

As TICAM1 is the sole mediator for the TLR3 signalling pathway, genetic variants that affect the function of TICAM1 have proved to be consequential in severe viral disease. The first account was reported in 2011 of two separate variants from unrelated patients resulting in *TICAM1* deficiencies and associated with paediatric HSE (12). AD pS186L, and AR p.R141X variants were described to impair TLR3 signalling (12). Fibroblasts isolated from TICAM1 deficient patients showed impaired IFN-β, IFN-λ, and IL-6 production after TLR3 stimulation, reasonably due to compromised activation of IRF-3 and NF-κB. AD p.A568T, and AD p.S160F were discovered in two unrelated men over the age of 60 with adult-onset HSE, and functional testing showed significantly impaired IFN-β, CXCL10, and TNF-α responses to poly(I;C) stimulation (a synthetic dsRNA molecule) and HSV-1 compared to controls (288). Loss-of-function heterozygous *TICAM1* variants have also been reported in severe COVID-19 pneumonia patients and subsequent functional data showed impaired IFN-β induction (289), this work will be expanded upon in this project in a subsequent chapter.

1.7 Summary

This chapter serves to highlight the current knowledge on host genetic variability in human infectious disease with a focus on IMD, COVID-19, and other significant invasive bacterial disease resulting from infection by *Staphylococcus aureus, Streptococcus pneumoniae,* and *Group A Streptococcus.* It also summarises the current methodologies and technologies being used in investigating novel genetic aetiologies of severe infectious diseases and reviews current known genetic variants associated with predisposition to and development of severe disease. However, this chapter also reveals gaps in our current knowledge on the genetic basis of life-threatening infectious disease, and this project will aim to answer some of these questions.

1.8 Aims of the studies

- To identify rare disease-causing variants that confer susceptibility to and/or severity of invasive infectious disease, including IMD, COVID-19, and other significant invasive bacterial disease, using the WES approach.
- To functionally validate the role of the variants in the identified genes in the context of invasive infectious diseases.
- To distinguish the mechanism by which the identified genes, and their associated variants, may lead to susceptibility or severity of invasive disease.
- To apply the methodologies developed in the previous aims to new patients recruited through the Imperial College Biomedical Research Council Paediatric Infection Clinic.

Chapter 2

Chapter 2 Methods

2.1 Human tissue culture

2.1.1 Culturing human cell lines

HEK293 cells are derived from human embryonic kidney cells (ATCC: CRL1573). HaCaT cells are immortalised human keratinocytes and were kindly gifted by Professor Peter O'Hare (Imperial College London). TICAM1 deficient human fibroblast cells were immortalised using the SV40 virus and previously described (12). HEK293, HaCaT, and TICAM1(-/-) fibroblast cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) GlutaMAX (Invitrogen) supplemented with 10 % of heat-inactivated foetal bovine serum (FBS) (Gibco) and incubated at 37°C, 5% CO2. Additionally, HaCaT cells were cultured in collagen coated vessels, in which plates and flasks were pre-treated at 37°C for 30 minutes with rat tail Collagen type I (Sigma) diluted 1:400 in PBS before cells were cultured.

γ2A are human fibrosarcoma cells that are JAK2 deficient and were a kind gift from Dr Ana Costa- Pereira (Imperial College London) and previously described (290). γ2A cells were maintained in DMEM GlutaMAX supplemented with 10% FBS, 1x L-Glutamine (Gibco), and 400 µg Geneticin (Gibco) at 37°C, 5% CO2.

2.1.2 Culturing primary human cells

Venous blood sampling was performed by a medical professional from both fully informed and consenting patients and healthy donors.

2.1.2.1 Isolating primary human PBMCs

For isolation of human peripheral blood mononuclear cells (PBMCs) from blood, whole blood was collected in BD Vacutainer® Heparin Tubes and processed with a sterile density gradient centrifugation method using Ficoll-Paque™ (Cytiva) within three days of blood draw. Briefly, collected whole blood was centrifuged at 400 g for 10 minutes and separated plasma was removed and replaced with Phosphate Buffered Saline (PBS) (Gibco), after plasma replacement the blood was diluted in a 1:1 ratio with PBS and overlaid carefully on top of 1 volume of Ficoll-Paque™ medium. The sample was then centrifuged at 600 g for 30 minutes with the centrifuge brakes turned off, so the gradients do not mix. Post-centrifugation, a clear four-layer gradient can be seen with a cloudy middle layer containing PBMCs, this layer was

carefully collected in PBS and washed twice by centrifuging for 10 minutes at 600 g and resuspending the cell pellet in PBS. Isolated and washed PBMCs were manually counted with a haemocytometer using Trypan Blue (Invitrogen) and maintained in Roswell Park Memorial Institute (RPMI) 1640 Medium (Gibco) supplemented with 10 % FBS and incubated at 37°C, 5% CO2.

2.1.2.2 Isolating primary human polymorphonuclear granulocytes

Isolation of human polymorphonuclear granulocytes (also known as polymorphonuclear leukocytes) were performed on whole undiluted blood collected using BD Vacutainer® K2E EDTA tubes and processed immediately upon blood draw with a sterile density gradient centrifugation method using PolymorphPrep™ (Progen) according to manufacturer's instructions. Collected granulocytes were resuspended in Hanks' Balanced Salt Solution (HBSS) no calcium, no magnesium, no phenol red (Gibco) supplemented with 1% FBS and used immediately for experiments.

2.1.3 Transfection of human cell lines

HEK293, HaCaT, γ2A, and TICAM1(-/-) fibroblast cells were seeded in 24-well plates as described in 2.1.1 to achieve an ~80% confluent monolayer the next day, culture media was then removed and replaced with 450 µL of fresh culture media, additionally with TICAM1 transfections resulting in Western blots, apoptosis was inhibited by adding 20 µM z-VAD-fmk (Promega) to each well. Transient transfections were performed using Lipofectamine™ 3000 Transfection Reagent (Invitrogen) according to manufacturer's instructions. Briefly, in one tube, 1 µL of P3000™ Enhancer Reagent was mixed with plasmid DNA diluted to an appropriate concentration in 25 µL of Opti-MEM medium (Invitrogen). Total amount of DNA (500ng/well) was kept constant by addition of empty plasmid vectors. In a separate tube, 1.5 µL of Lipofectamine™ 3000 Reagent was added to 23.5 µL of Opti-MEM medium. The contents of both tubes were then mixed, incubated at room temperature for 15 minutes and added in a drop wise manner to the cells. The plates were incubated at 37°C, 5% CO2 for at least 24 hours. Where applicable, transfection efficiency was analysed by fluorescent microscopy using an EVOS™ FL Auto 2 (Invitrogen).

2.1.4 Generation of human fibro sarcoma cells expressing *JAK2* variants

The mutagenized *JAK2* constructs in a lentiviral pTRIP backbone (see Table 2.2), were used to create *JAK2* stable expressing γ2A cells. As lentiviral vectors can not replicate on their own,

pseudoparticles were generated in HEK293 cells in a 6-well plate by co-transfecting 1 µg of pTRIP-JAK2 plasmids with 0.5 µg of pCMV-VSV and 0.4 µg of pGag-Pol plasmids using Lipofectamine™ 3000 as detailed in 2.1.3. Transfection media was replaced with fresh DMEM containing 10 % FBS 6 hours after transfection. The pseudoparticles from cell supernatants were harvested at 24 hours and 48 hours post transfection, and filtered through a 0.45 µm filter to be used for transduction or stored at -80°C.

For transduction, γ2A cells were seeded in 6-well plates at $5x10^5$ cells/mL and grown overnight to reach a confluency of ~70% the next day. Culture media was replaced with fresh media containing polybrene at a final concentration of 8 µg/mL and varying amounts of harvested lentiviral particles (0 µL , 250 µL, 500 µL, and 1000 µL) and incubated for 24 hours at 37°C, 5% CO2 where viral particles were removed, and cells were placed in fresh culture media and incubated for a further 48 hours. At 72 hours post transduction, positive RFP expression was checked by fluorescent microscopy and then cells were trypsinised, washed once with PBS, and harvested in PBS with 1% FBS 1x EDTA at 3 million cells/500 µL in sterile fluorescence-activated cell sorting (FACS) tubes. FACS was performed by Dr Parisa Amjadi (ICL) where samples were sorted to select for RFP positive cells excited by 565 nm laser. Successfully sorted cells were washed once with PBS, to remove residual EDTA, recovered in fresh DMEM containing 20 % FBS, seeded into an appropriate culture vessel, and incubated at 37°C, 5% CO2. New stably transduced JAK2-γ2A cells were expanded and cell stocks were made.

2.2 Bacterial strains

All bacteria strains, except *E.coli*, were a kind gift from Dr Fadil Bidmos (ICL).

2.2.1 Growth of cloning bacteria *Escherichia coli*

XL10-gold ultracompetent *E.coli* cells (Agilent) was grown in Luria-Brentani (LB) broth shaking or on solid LB agar plates in a static incubator at 37°C, 5% CO2. Growth mediums were prepared as follows: 20 g LB powder (Sigma), plus additional 15 g of agar (Sigma) for solid LB agar plates, were dissolved in 1 L of ddH_2O and autoclaved. Where necessary, growth mediums were supplemented with antibiotics at the following final concentrations: 25ug/mL of kanamycin; 100ug/mL of ampicillin.

2.2.2 Growth of *Neisseria meningitidis*

Bacterial stocks were stored in 1 mL aliquots at -80°C in BHI + 25% glycerol. The clinical isolate encapsulated serogroup B *N. meningitidis* (Nm) strain MC58 was cultured on solid blood agar plates and incubated overnight (approximately 16 hours) at 37°C, 5% CO_2. Solid blood agar plates were prepared as follows: 7.5 g of agar (Sigma) and 18.5 g of Bacto™ brain heart infusion (BHI, BD biosciences) were dissolved in 500 mL ddH_2O, autoclaved and supplemented with 5% defibrinated horse blood (Thermo-Fisher).

2.2.3 Growth of Staphylococcus aureus

Bacterial stocks were stored in 1 mL aliquots at -80°C in LB + 25% glycerol. *Staphylococcus aureus* strain PVL+MSSA was cultured in liquid LB broth and incubated overnight (approximately 16 hours) in a shaking incubator, 220 rpm, at 37°C, 5% CO2. Liquid LB broth was prepared by dissolving 10 g of LB in 500 mL ddH2O, and autoclaved. A 1:10 dilution of the overnight growth was made with fresh LB broth and incubated for 4 hours in a shaking incubator, 220 rpm, at 37°C, 5% CO_2 to allow the bacteria to enter a log phase of growth. Post incubation, the bacteria were then centrifuged at 2000 rpm for 5 minutes and resuspended in PBS, twice. Bacteria was then diluted in PBS to an OD_{600} 0.5. OD_{600} was measured to determine the colony forming units (CFU). The calculation was made based on the estimation that OD_{600} 1 equates to ~10^9 CFU/mL.

2.2.4 Growth of *Group A Streptococcus pyogenes*

Bacterial stocks were stored in 1 mL aliquots at -80°C in TSBYE + 25% glycerol. *Group A Streptococcus* (GAS) used in this project was grown from inhouse clinical isolate strain M1UK, emm type 1. GAS was cultured in liquid Trypticase Soy Broth with 0.5% Yeast Extract (TSBYE) and incubated overnight (approximately 16 hours) in a shaking incubator, 220rpm, at 37°C, 5% CO2. Liquid TSBYE broth was prepared by dissolving 30 g of TSB (Sigma) and 5 g Yeast Extract (Sigma) in 1L ddH2O, and autoclaved. A 1:10 dilution of the overnight growth was made with fresh TSBYE and incubated for 4 hours in a shaking incubator, 220 rpm, at 37°C, 5% CO_2 to allow the bacteria to enter a log phase of growth. Post incubation, the bacteria were then centrifuged at 2000 rpm for 5 minutes and resuspended in PBS, twice. Bacteria was then diluted in PBS to an OD_{600} 0.2. OD_{600} was measured to determine the colony forming units (CFU). The calculation was made based on the estimation that OD_{600} 0.2 equates to ~10^8 CFU/mL.

2.2.5 Growth of *Streptococcus pneumoniae*

Bacterial stocks were stored in 1 mL aliquots at -80°C in TSBYE + 25% glycerol. *Streptococcus pneumoniae* strain D39 was cultured in liquid TSBYE and incubated overnight (approximately 16 hours) in a static incubator, at 37°C, 5% CO_2. A 1:10 dilution of the overnight growth was made with fresh TSBYE and incubated for 4 hours in a static incubator, at 37°C, 5% CO_2 to allow the bacteria to enter a log phase of growth. Post incubation, the bacteria were then centrifuged at 2000 rpm for 5 minutes and resuspended in PBS, twice. Bacteria was then diluted in PBS to an OD_{600} 0.2. OD_{600} was measured to determine the colony forming units (CFU). The calculation was made based on the estimation that OD_{600} 0.2 equates to ~10^8 CFU/mL.

2.2.6 Heat killing *Neisseria meningitidis*

Single colonies were collected from an overnight growth of Nm using a sterile loop, replated on solid blood agar plates and incubated for 4 hours at 37°C, 5% CO_2 to allow the bacteria to enter a log phase of growth. Post incubation, colonies were resuspended in PBS and diluted to an OD_{600} between 0.4-0.6. Dilutions were transferred into a sealed tube and heated in a water bath at 56°C for 30 minutes. OD_{600} was measured to determine the colony forming units (CFU). The calculation was made based on the estimation that OD_{600} 1 equates to ~10^9 CFU/mL. Heat killing was confirmed by streaking bacteria on a BHI + 5% defibrinated horse blood solid agar plate and incubating overnight at 37°C, 5% CO_2.

2.2.7 Formalin fixing of *Neisseria meningitidis*

Single colonies were collected from an overnight growth of Nm using a sterile loop, replated on solid blood agar plates and incubated for 4 hours at 37°C, 5% CO_2 to allow the bacteria to enter a log phase of growth. Post incubation, colonies were resuspended in 0.5% formalin in PBS and incubated at room temperature for a minimum of 4 hours. The bacteria were then centrifuged at 2000 rpm for 5 minutes and resuspended in PBS, twice. The bacteria were then diluted to an OD_{600} between 0.4-0.6 in PBS. OD_{600} was measured to determine the colony forming units (CFU). The calculation was made based on the estimation that OD_{600} 1 equates to ~10^9 CFU/mL. Fixing was confirmed by streaking bacteria on a BHI + 5% defibrinated horse blood solid agar plate and incubating overnight at 37°C, 5% CO_2.

2.2.8 Formalin fixing of *Staphylococcus aureus*

Staphylococcus aureus was grown as described above (section 2.2.3). Bacteria formalin fixed by centrifuging cells, resuspending in 0.5% formalin in PBS and incubated at room temperature for a minimum of 4 hours. Fixing was confirmed by incubating bacteria in LB broth at in a shaking incubator 37°C, 5% CO_2. The fixed bacteria were then centrifuged at 2000 rpm for 5 minutes and resuspended in PBS, three times, and diluted in PBS to an OD_{600} 0.5.

2.2.9 Formalin fixing of *Group A Streptococcus pyogenes*

Group A Streptococcus pyogenes was grown as described above (section 2.2.4). Bacteria was formalin fixed by centrifuging cells, resuspending in 5% formalin in PBS and incubated at room temperature for a minimum of 4 hours. Fixing was confirmed by incubating bacteria in Liquid TSBYE broth at in a shaking incubator 37°C, 5% CO_2. The fixed bacteria were then centrifuged at 2000 rpm for 5 minutes and resuspended in PBS, three times, and diluted in PBS to an OD_{600} 0.2.

2.2.9 Formalin fixing of *Streptococcus pneumoniae*

Streptococcus pneumoniae was grown as described above (section 2.2.5). Bacteria was formalin fixed by centrifuging cells, resuspending in 5% formalin in PBS and incubated at room temperature for a minimum of 4 hours. Fixing was confirmed by incubating bacteria in liquid TSBYE broth at in a static incubator 37°C, 5% CO_2. The fixed bacteria were then centrifuged at 2000 rpm for 5 minutes and resuspended in PBS, three times, and diluted in PBS to an OD_{600} 0.2.

2.3 Ligands and Drugs

The drugs and stimulation ligands used are listed below (Table 2.1).

Table 2.1 Drugs and ligands used in this thesis

Name	Source	Working concentration
Recombinant human interleukin 6 (IL-6)	R&D systems	100 U/mL
Recombinant human interferon gamma (IFN-γ)	R&D systems	10 ng/mL
Human interferon-Alpha 2A (2a) (IFN-α)	Pbl assay science	100 U/mL
Lipopolysaccharide, purified from *S. minnesota* (LPS)	Invivogen	100/150 ng/mL
Recombinant human tumour necrosis factor α (TNF-α)	R&D systems	20/200 ng/mL
Resiquimod (R848)	Invivogen	3 µg/mL
GGTI-2147	Sigma	10 µM

2.4 DNA editing

2.4.1 Plasmids

For a list of all plasmids used and constructed during this thesis see Table 2.2. Wildtype plasmid maps can be found in the appendix (8.3 Plasmid Maps)

2.4.2 Oligonucleotides used in this thesis

Primers used to amplify genomic *JAK2* or for plasmid sequence confirmation of plasmids expressing *TICAM1, JAK2,* and *KRT10* were designed using the online PrimerQuest™ Tool (Integrated DNA Technologies IDT) and shown in Table 2.3. Overlapping primer pairs used to introduce single nucleotide substitutions at specific sites of interest for site directed mutagenesis were manually designed (Table 2.4).

Table 2.2 Plasmids used in this thesis

Plasmids	Details	Origin
pEF-Bos-TICAM1	Expression vector, C-terminally tagged with HA, ampicillin resistant cassette	(12)
pEF-Bos-TICAM1-T4I	Cloned from pEF-BOS-TICAM1, T4I introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-T32I	Cloned from pEF-BOS-TICAM1, T32I introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-S60C	Cloned from pEF-BOS-TICAM1, S60C introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-R71Q	Cloned from pEF-BOS-TICAM1, R71Q introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-V80M	Cloned from pEF-BOS-TICAM1, V80M introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-A111T	Cloned from pEF-BOS-TICAM1, A111T introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-R141X	Cloned from pEF-BOS-TICAM1, R141X introduced using site directed mutagenesis	(12)
pEF-Bos-TICAM1-T157M	Cloned from pEF-BOS-TICAM1, T157M introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-S160F	Cloned from pEF-BOS-TICAM1, S160F introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-L162V	Cloned from pEF-BOS-TICAM1, L162V introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-S186L	Cloned from pEF-BOS-TICAM1, S186L introduced using site directed mutagenesis	(12)
pEF-Bos-TICAM1-R195E	Cloned from pEF-BOS-TICAM1, R195E introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-G245S	Cloned from pEF-BOS-TICAM1, G245S introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-V302L	Cloned from pEF-BOS-TICAM1, V302L introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-T377I	Cloned from pEF-BOS-TICAM1, T377I introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-L386P	Cloned from pEF-BOS-TICAM1, L386P introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-Q392K	Cloned from pEF-BOS-TICAM1, Q392L introduced using site directed mutagenesis	This thesis

pEF-Bos-TICAM1-S465N	Cloned from pEF-BOS-TICAM1, S465N introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-D557N	Cloned from pEF-BOS-TICAM1, D557N introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-M595L	Cloned from pEF-BOS-TICAM1, M595L introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-G598W	Cloned from pEF-BOS-TICAM1, G598W introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-P620S	Cloned from pEF-BOS-TICAM1, P620S introduced using site directed mutagenesis	This thesis
pEF-Bos-TICAM1-Q702R	Cloned from pEF-BOS-TICAM1, Q702R introduced using site directed mutagenesis	This thesis
pEF-Bos-empty	Cloning vector, C-terminally tagged with HA, ampicillin resistant cassette	(12)
pTRIP-JAK2	Lentiviral expression vector, C-terminally tagged with RFP, ampicillin resistant cassette	(291)
pTRIP-JAK2-L113V	Cloned from pTRIP-JAK2, L113V introduced using site directed mutagenesis	This thesis
pTRIP-JAK2-L393V	Cloned from pTRIP-JAK2, L393V introduced using site directed mutagenesis	This thesis
pTRIP-JAK2-F556V	Cloned from pTRIP-JAK2, F556V introduced using site directed mutagenesis	This thesis
pTRIP-JAK2-R564L	Cloned from pTRIP-JAK2, R564L introduced using site directed mutagenesis	Rahmeen Rahman (ICL)
pTRIP-JAK2-N782K	Cloned from pTRIP-JAK2, N782K introduced using site directed mutagenesis	This thesis
pTRIP-JAK2-F844L	Cloned from pTRIP-JAK2, F844 introduced using site directed mutagenesis	Rahmeen Rahman (ICL)
pTRIP-JAK2-E846D	Cloned from pTRIP-JAK2, E846D introduced using site directed mutagenesis	Rahmeen Rahman (ICL)
pTRIP-JAK2-L871V	Cloned from pTRIP-JAK2, L871V introduced using site directed mutagenesis	This thesis
pTRIP-JAK2-R1063H	Cloned from pTRIP-JAK2, R1063H introduced using site directed mutagenesis	Rahmeen Rahman (ICL)
pTRIP-JAK2-N1108S	Cloned from pTRIP-JAK2, N1108S introduced using site directed mutagenesis	This thesis
pTRIP-JAK2-V617F	Cloned from pTRIP-JAK2, V617F introduced using site directed mutagenesis	This thesis
pTRIP-JAK2-G127A	Cloned from pTRIP-JAK2, G127A introduced using site directed mutagenesis	This thesis
pTRIP-JAK2-M187I	Cloned from pTRIP-JAK2, M187I introduced using site directed mutagenesis	This thesis

pTRIP-JAK2-V392M	Cloned from pTRIP-JAK2, V392M introduced using site directed mutagenesis	This thesis
pTRIP-JAK2-S523T	Cloned from pTRIP-JAK2, S523T introduced using site directed mutagenesis	This thesis
pTRIP-JAK2-G571S	Cloned from pTRIP-JAK2, G571S introduced using site directed mutagenesis	This thesis
pTRIP-JAK2-I724T	Cloned from pTRIP-JAK2, I724T introduced using site directed mutagenesis	This thesis
pTRIP-empty	Lentiviral cloning vector, ampicillin resistant cassette	(291)
pCMV6-KRT10	Expression vector, kanamycin resistant cassette, C-terminally tagged with MYC-DDK	Origene
pCMV6-KRT10-R4P	Cloned from pCMV6-KRT10, R4P introduced using site directed mutagenesis	This thesis
pCMV6-KRT10-S42F	Cloned from pCMV6-KRT10, S42F introduced using site directed mutagenesis	This thesis
pCMV6-KRT10-S53N	Cloned from pCMV6-KRT10, S53N introduced using site directed mutagenesis	This thesis
pCMV6-KRT10-R86H	Cloned from pCMV6-KRT10, R86H introduced using site directed mutagenesis	This thesis
pCMV6-KRT10-F102S	Cloned from pCMV6-KRT10, F102S introduced using site directed mutagenesis	This thesis
pCMV6-KRT10-G124V	Cloned from pCMV6-KRT10, G124V introduced using site directed mutagenesis	This thesis
pCMV6-KRT10-K334Q	Cloned from pCMV6-KRT10, K334Q introduced using site directed mutagenesis	This thesis
pCMV6-KRT10-G575V	Cloned from pCMV6-KRT10, G575S introduced using site directed mutagenesis	This thesis
pCMV-empty	Cloning vector, kanamycin resistant cassette	Origene
pRL-TK Renilla luciferase	Control reporter vector, used in luciferase reporter assay	(12)
IFN-β promoter-luciferase	Firefly reporter vector, used in luciferase reporter assay	(12)
NF-κB luciferase	Firefly reporter vector, used in luciferase reporter assay	(12)

Table 2.3 Primers designed to amplify plasmid cDNA

Primer	Sequence 5'-3'	Details
TICAM1-31F	GCGCCTTCGACATTCTAGG	TICAM1 cDNA sequencing
TICAM1-351	TACCAGGAAGCCGTCCGCA	TICAM1 cDNA sequencing
TICAM1-537F	CTGGCTGAGGAGAAGCTGTG	TICAM1 cDNA sequencing
TICAM1-681F	AGCAAGCTCTGTGACGACC	TICAM1 cDNA sequencing
TICAM1-853F	AAATCAGCCAGTCCCCTACC	TICAM1 cDNA sequencing
TICAM1-1021F	ACCACCTCTCCAAATACCAAGC	TICAM1 cDNA sequencing
TICAM1-1681F	ACTGTTGTCATCCCCTTCTG	TICAM1 cDNA sequencing
TICAM1-1945F	CAGCCAGCAGCCTTTCCA	TICAM1 cDNA sequencing
TICAM1-2201R	GAAGGGCAGTGACTGTGGA	TICAM1 cDNA sequencing
TICAM1-1945R	TGGAAAGGCTGCTGGCTG	TICAM1 cDNA sequencing
TICAM1-1298R	ACCTGGAAATCCTCGCAGAA	TICAM1 cDNA sequencing
TICAM1-1134R	CTGGGTAGTTGGTGCTGGTT	TICAM1 cDNA sequencing
TICAM1-853R	GGTAGGGGACTGGCTGATTT	TICAM1 cDNA sequencing
TICAM1-407R	AGTTCCCCCAGCCGGTGGTC	TICAM1 cDNA sequencing
JAK2-9F	GGCCTGCCTTACGATGACAG	JAK2 cDNA sequencing
JAK2-2022F	TGAAAGCCGAGTTGTAAC	JAK2 cDNA sequencing
JAK2-1381F	GAGTACAACCTCAGTGGGAC	JAK2 cDNA sequencing
JAK2-2225F	CTTTGTGGGAAATCTGCAGTGG	JAK2 cDNA sequencing
JAK2-3059F	GGTATGCTCCAGAATCACTGAC	JAK2 cDNA sequencing

JAK2-549R	CACTGCCATCCCAAGACAT	JAK2 cDNA sequencing
JAK2-1381R	GTCCCACTGAGGTTGTACTC	JAK2 cDNA sequencing
JAK2-1973R	GCCAACTGTTTAGCAACTTC	JAK2 cDNA sequencing
JAK2-2720R	GCTGTAGGGATTTCAGGAT	JAK2 cDNA sequencing
JAK2-3394R	CAGCCATGTTATCCCTTATTTGATC	JAK2 cDNA sequencing
KRT10-10F	CGATACAGCTCAAGCAAGCACTACTC	KRT10 cDNA sequencing
KRT10-457F	CTGAATGACCGCCTGGCTTCCTAC	KRT10 cDNA sequencing
KRT10-906F	GAATGTGGAAATGAATGCTGCCCCG	KRT10 cDNA sequencing
KRT10-1355F	GCTAGAAGGAGAGGGAAGTTCCGG	KRT10 cDNA sequencing
KRT10-1740R	CTTAGATGAAGACTCGCCCAC	KRT10 cDNA sequencing
KRT10-1353R	GGCTGCGGTAGGTTTGAATTTC	KRT10 cDNA sequencing
KRT10-900R	CCAGTGGACACATTTCGAAGG	KRT10 cDNA sequencing
KRT10-350R	CCAAAGCTGCCCCCACCAAAGC	KRT10 cDNA sequencing

Table 2.4 Oligonucleotides designed for site directed mutagenesis

Primer	Sequence 5'-3'	T_{Anneal} (X ˚C)	Details
TICAM1-T4I_F	tggtaccatggcctgcaTAggcccatcacttcct	75	Generation of T4I in pEF-bos-TICAM1 wildtype
TICAM1-T4I_R	aggaagtgatgggccATtgcaggccatggtacca		
TICAM1-T32I_F	agcacaaactgaagaTcccacgcccaggct	63	Generation of T32I in pEF-bos-TICAM1 wildtype
TICAM1-T32I_R	agcctgggcgtgggAtcttcagtttgtgct		
TICAM1-S60C_F	gaggccaggatctGtctagaggcattg	68	Generation of S60C in pEF-bos-TICAM1 wildtype
TICAM1-S60C_R	caatgcctctagaCagatcctggcctc		
TICAM1-R71Q_F	ccgatgcggtggcccAgctggtggcccgccagt	71	Generation of R71Q in pEF-bos-TICAM1 wildtype
TICAM1-R71Q_R	actggcgggccaccagcTgggccaccgcatcg		
TICAM1-V80M_F	gccagtgggctggcAtggacagcaccgagga	67	Generation of V80M in pEF-bos-TICAM1 wildtype
TICAM1-V80M_R	tcctcggtgctgtccaTgccagcccactggc		
TICAM1-A111T_F	gagaagctgtgccccAcctcgctgcgggacgtg	68	Generation of A111T in pEF-bos-TICAM1 wildtype
TICAM1-A111T_R	cacgtcccgcagcgaggTggggcacagcttctc		
TICAM1-T157M_F	atccagggagcatccggaTgctccagtccaatctgggct	66	Generation of T157M in pEF-bos-TICAM1 wildtype
TICAM1-T157M_R	agcccagattggactggagcAtccggatgctccctggat		
TICAM1-S160F_F	atccggacgctccagtTcaatctgggctgcct	64	Generation of S160F in pEF-bos-TICAM1 wildtype
TICAM1-S160F_R	aggcagcccagattgAactggagcgtccggat		
TICAM1-R195E_F	tgctccctgcAatccactggc	71	Generation of R195E in pEF-bos-TICAM1 wildtype
TICAM1-R195E_R	gccagtggatGgcagggagca		

TICAM1-G245S_F	tgcccgagcctgtccccAgtggctgccaggagcc	71	Generation of G245S in pEF-bos-TICAM1 wildtype
TICAM1-G245S_R	ggctcctggcagccacTggggacaggctcgggca		
TICAM1-P271S_F	tgccaagcagcTcacctcctgggc	78	Generation of P271S in pEF-bos-TICAM1 wildtype
TICAM1-P271S_R	gcccaggaggtgAgctgcttggca		
TICAM1-V302L_F	agcaccaactacccaCtggagtgcaccgaggggt	66	Generation of V302L in pEF-bos-TICAM1 wildtype
TICAM1-V302L_R	acccctcggtgcactccaGtgggtagttggtgct		
TICAM1-T377I_F	gctcacctgaAcccctcctccctg	61	Generation of T377I in pEF-bos-TICAM1 wildtype
TICAM1-T377I_R	cagggaggaggggTtcaggtgagc		
TICAM1-L386P_F	ccctgttcccttcctcccCggaatcatcatcggaac	75	Generation of L386P in pEF-bos-TICAM1 wildtype
TICAM1-L386P_R	gttccgatgatgattccGgggaggaagggaacaggg		
TICAM1-Q392K_F	catcatcggaaAagaaattctataactttgtg	63.1	Generation of Q392K in pEF-bos-TICAM1 wildtype
TICAM1-Q392K_R	cacaaagttatagaatttctTttccgatgatg		
TICAM1-S465N_F	cttcgactgtcgcctgaAcctgcaccaggtg	75	Generation of S465N in pEF-bos-TICAM1 wildtype
TICAM1-S465N_R	cacctggtgcaggTtcaggcgacagtcgaag		
TICAM1-D557N_F	agagccaacacctgAacggtgagcggatgc	63	Generation of D557N in pEF-bos-TICAM1 wildtype
TICAM1-D557N_R	gcatccgctcaccgtTcaggtgttggctct		
TICAM1-M595L_F	gtggcttttgggagccacTtgtcatttgggactg	76	Generation of M595L in pEF-bos-TICAM1 wildtype
TICAM1-M595L_R	cagtcccaaatgacaAgtggctcccaaaagccac		
TICAM1-G598W_F	ccacatgtcatttTggactggggcgccctat	61	Generation of G598W in pEF-bos-TICAM1 wildtype
TICAM1-G598W_R	atagggcgccccagtccAaaatgacatgtgg		

TICAM1-P620S_F	ggagccccgccaTcctttcccacttgg	77.6	Generation of P620S in pEF-bos-TICAM1 wildtype
TICAM1-P620S_R	ccaagtgggaaaggAtggcggggctcc		
TICAM1-Q702R_F	accagagagggtcccGggcgcccgaggacaa		
TICAM1-Q702R_R	ttgtcctcgggcgccCgggaccctctctggt	69	Generation of Q702R in pEF-bos-TICAM1 wildtype
JAK2-L113V_F	gatgagtcaaccaggcataatgtaGtctacagaataagattttac	74	Generation of L113V in pTRIP-JAK2 wildtype
JAK2-L113V_R	gtaaaatcttattctgtagaCtacattatgcctggttgactcatc		
JAK2-L393V_F	gtagcacctccagccgtgGttgaaaatatacaaagc	74	Generation of L393V in pTRIP-JAK2 wildtype
JAK2-L393V_R	gctttgtatattttcaaCcacggctggaggtgctac		
JAK2-F556V_F	cttggccaaggcactGttacaaagatttttaaaggc	71	Generation of F556V in pTRIP-JAK2 wildtype
JAK2-F556V_R	gcctttaaaaatctttgtaaCagtgccttggccaag		
JAK2-N782K_F	gtgggcagaattagcaaaActtataaataattgtatgg	69	Generation of N782K in pTRIP-JAK2 wildtype
JAK2-N782K_R	ccatacaattatttataagTtttgctaattctgcccac		
JAK2-L871V_F	cggtatgaccctGtacaggacaacactggg	71	Generation of L871V in pTRIP-JAK2 wildtype
JAK2-L871V_R	cccagtgttgtcctgtaCagggtcataccg		
JAK2-N1108S_F	atgacagaatgctggaacaGtaatgtaaatcaacgcccc	71	Generation of R1063H in pTRIP-JAK2 wildtype
JAK2-N1108S_R	ggggcgttgatttacattaCtgttccagcattctgtcat		
JAK2-G127A_F	ggtattgcagtgCcagcaacagagcctatcggc	74	Generation of G127A in pTRIP-JAK2 wildtype
JAK2-G127A_R	gccgataggctctgttgctgGcactgcaatacc		
JAK2-M187I_F	gggatggcagtgttagatatgatAagaatagcc	68	Generation of M187I in pTRIP-JAK2 wildtype
JAK2-M187I_R	ggctattctTatcatatctaacactgccatccc		

JAK2-V392M_F	gcacctccagccAtgcttgaaaatatac	65	Generation of V392M in pTRIP-JAK2 wildtype
JAK2-V392M_R	gtatattttcaagcaTggctggaggtgc		
JAK2-S523T_F	gtaccaaccAcaccaacattacagaggcctactc	64	Generation of S523T in pTRIP-JAK2 wildtype
JAK2-S523T_R	gagtaggcctctgtaatgttggtgTggttggtac		
JAK2-G571S_F	taggagactacAgtcaactgcatgaaacag	64	Generation of G571S in pTRIP-JAK2 wildtype
JAK2-G571S_R	ctgtttcatgcagttgacTgtagtctccta		
JAK2-I724T_F	gggtaccacctgaatgcaCtgaaaatcc	70	Generation of I724T in pTRIP-JAK2 wildtype
JAK2-I724T_R	ggattttcaGtgcattcaggtggtaccc		
JAK2-V617F_F	ggttttaaattatggagtatgtTtctgtggagacgaga	68	Generation of V617F in pTRIP-JAK2 wildtype
JAK2-V617F_R	tctcgtctccacagaAacatactccataatttaaaacc		
KRT10-R4P_F	tgttcCatacagctcaagcaag	63.8	Generation of R4P in pCMV6-KRT10 wildtype
KRT10-R4P_R	cttgcttgagctgtatGgaaca		
KRT10-S42F_F	gcagcaaaggctTccttggtggagg	72	Generation of S42F in pCMV6-KRT10 wildtype
KRT10-S42F_R	cctccaccaaggAagcctttgctgc		
KRT10-S53N_F	gctcagggggttcaAtggtggctc	73	Generation of S53N in pCMV6-KRT10 wildtype
KRT10-S53N_R	gagccaccaTtgaacccccctgagc		
KRT10-R86H_F	ggtggaggtagctttcAtggaagc	64.5	Generation of R86H in pCMV6-KRT10 wildtype
KRT10-R86H_R	gcttccaTgaaagctacctccacc		
KRT10-F102S_F	gttatggaggcatctCtggagggggcagtttc	74	Generation of F102S in pCMV6-KRT10 wildtype
KRT10-F102S_R	gaaactgccccctccaGagatgcctccataac		

KRT10-G124V_F	ggcggctttggtgTaggcggctttgg	76	Generation of G124V in pCMV6-KRT10 wildtype
KRT10-G124V_R	ccaaagccgcctAcaccaaagccgcc		
KRT10-K334Q_F	gaacaaaaccgcCaagatgctgaag	67	Generation of K334Q in pCMV6-KRT10 wildtype
KRT10-K334Q_R	cttcagcatcttGgcggttttgttc		
KRT10-G575V_F	ccgggtccgtggTcgagtcttcatctaag	72.5	Generation of G575V in pCMV6-KRT10 wildtype
KRT10-G575V_R	cttagatgaagactcgAccacggacccgg		

Table 2.5 Antibodies used in this thesis

Antibody target	Application	Dilution	Buffer used	Source
pJAK2 (Y1007)	Western blot, primary	1:500	5% BSA in 1x TBST	CST 4406
JAK2	Western blot, primary	1:1000	5% BSA in 1x TBST	Abcam ab39636
pSTAT1 (Y701)	Western blot, primary	1:1000	5% BSA in 1x TBST	CST 9167
STAT1	Western blot, primary	1:1000	5% BSA in 1x TBST	CST 14994
GAPDH	Western blot, primary	1:1000	5% milk in 1x TBST	CST 2118
KRT10	Western blot, primary	1:500	5% milk in 1x TBST	Biolegend 905404
FLAG	Western blot, primary	1:1000	5% milk in 1x TBST	CST D6W5B
HA	Western blot, primary	1:1000	5% BSA in 1x TBST	Sigma H6908
Polyclonal goat anti-rabbit IgG HRP	Western blot, secondary	1:10,000	1% BSA/milk* in 1x TBST	Dako P0448
CD14-FITC Clone M5E2	Flow cytometry	2 µL/1x10^6 cells	1x PBS 1% FBS	Biolegend 301804
CD4-PerCP/Cy5.5 Clone OKT4	Flow cytometry	2 µL/1x10^6 cells	1x PBS 1% FBS	Biolegend 317428
pSTAT1-BV421 (Y701) Clone 4a	Flow cytometry	5 µL/1x10^6 cells	1x PBS 1% FBS	BD 562985
FLAG	Immunofluorescence	1:100	1x PBS 1% BSA	Sigma F1804
Alexa Fluor 488 secondary antibody	Immunofluorescence	1:1000	1x PBS 1% BSA	Invitrogen A11017
HA[b]	Immunofluorescence	1:100	1x PBS 1% BSA	Sigma H6908
DyLight™ 488 secondary antibody [b]	Immunofluorescence	1:1000	1x PBS 1% BSA	Invitrogen 35552

*Same buffer used as for the primary antibody incubation. CST; Cell signalling technology, BD; biosciences

[a] Used together, concurrently, in the same immunofluorescence experiment [b] Used together, concurrently, in the same immunofluorescence experiment

2.4.3 Mutagenesis

Site directed mutagenesis was performed using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent) according to manufacturer's instructions. Mutagenesis was performed to introduce single nucleotide substitutions into pEF-bos-TICAM1, pTRIP-JAK2, and pCMV6-KRT10 wildtype expression vectors (Table 2.2). Mutant alleles were generated using PCR: 95˚C for 10 minutes, then 18 cycles of [95 ˚C for 20 seconds, X ˚C for 10 seconds, 68 ˚C for 4.5 minutes when using pEF-bos-TICAM1, 7.5 minutes when using pTRIP-JAK2, or 3.5 minutes when using pCMV6-KRT10] and a final extension of 68 ˚C for 5 minutes. Annealing temperatures and extension times varied depending on the primer pairs and plasmids used, respectfully, see Table 2.4 for details of annealing temperatures. Custom designed oligonucleotides were used for mutagenesis (Table 2.4) and parental plasmids present in the PCR products were digested with DpnI for 15 minutes at 37 ˚C before transformation occurred. All isolated plasmids were verified by Sanger sequencing.

2.4.4 Transformation

To increase plasmid yield, purified plasmids or PCR products from site directed mutagenesis were transformed into XL10-Gold ultracompetent cells (Agilent). Briefly, either 10 ng of purified plasmids or 2 µL of PCR mutagenesis products were added to 30 µL of competent cells and incubated on ice for 30 minutes. The cells were then heat shocked for exactly 30 seconds at 42 ˚C in a water bath and then placed immediately back on ice for a further 2 minutes and then 400 µL of SOC media (Invitrogen) was added to the cells and incubated at 37 ˚C, 5% CO_2 in a shaking incubator (200 rpm) for 30-60 minutes. Cells were then plated onto LB agar plates, supplemented with appropriate antibiotics, and incubated overnight at 37 ˚C, 5% CO_2. Single individual overnight colonies were sterilely picked and grown in LB liquid broth supplemented with appropriate antibiotics overnight at 37 ˚C, 5% CO_2 in a shaking incubator (200 rpm) before plasmids could be isolated.

2.4.5 Isolation of plasmid DNA

Plasmids were isolated from overnight LB broth growth using QIAprep® Spin Miniprep Kit (QIAGEN), according to manufacturer's instructions and eluted isolated plasmids were measured by a Nanodrop™ 2000/2000c spectrophotometer (Thermo Fisher) before being sent for sequencing.

2.4.6 Sanger sequencing

To confirm successfully single nucleotide substitution within the plasmid cDNA target regions, isolated plasmids were sent for sanger sequencing. Sanger sequencing was performed offsite by Genewiz, Azenta Life Sciences (https://www.genewiz.com/en-GB/) using primers designed in Table 2.3 and any full length plasmid sequencing was performed by Plasmidsaurus (https://www.plasmidsaurus.com/).

2.5 DNA extraction

2.5.1 DNA extraction from whole blood

DNA extraction from whole blood was performed using QIAamp® DNA Blood Midi Kit (QIAGEN) according to manufacturer's instructions and performed on whole undiluted blood collected using BD Vacutainer® EDTA tubes. Concentration of eluted DNA was measured using a Qubit 4 fluorometer (Thermo Fisher).

2.5.2 DNA extraction from saliva

DNA extraction from saliva were performed on samples collected using Oragene DNA Mailable Tubes (DNA genotek), within three months of sample collection. Briefly, samples are incubated at 50 °C in a water bath for 1 hour and then transferred to a sterile Corning™ Falcon™ tube where 1/25th volume of PT-L2P reagent (DNA genotek) is added to the sample, mixed by vortexing, and incubated on ice for 10 minutes. Samples are then centrifuged for 10 minutes at top speed and the clear supernatant is transferred to a fresh tube and mixed with 1.2x volume of room temperature ethanol (95-100%), gently inverted to mix, and allowed to stand for 10 minutes at room temperature to allow the DNA to fully precipitate. Sample are then centrifuged for 10 minutes at top speed and supernatant is discarded. The pellet is resuspended in 1 mL of 70% ethanol and centrifuged for 10 minutes at top speed, as much as possible of the supernatant is removed and the pellet is allowed to airdry. The DNA pellet is rehydrated with TE buffer, vortexed, and incubated at room temperature overnight. The next day, samples are incubated at 50 °C for 1 hour with occasional vortexing and then transferred to a labelled sterile tube and DNA concentration is measured using a Qubit 4 fluorometer.

2.6 Whole exome sequencing

2.6.1 Whole exome sequencing

Library preparation and whole exome sequencing on isolated patient and control donor DNA extracted from either whole blood or saliva is performed by Novogene (UK, Cambridge centre). Whole exome sequencing is performed on an Illumina NovaSeq 6000 platform generating paired end reads of 150 base pairs in length. This generated exomes with an effective mean coverage of 50x. Raw reads in FASTQ format were then processed by Dr Evangelos Bellos for de-duplication and sequencing quality assessment using the Picard tools (http://broadinstitute.github.io/picard) and then aligned to the current human reference genome (build UCSC hg19) using bwa v0.7. Post-alignment, the exomes were jointly genotyped using GATK v3.8 (https://github.com/broadinstitute/gatk) (292). Initial variant annotation is achieved using Ensembl's Variant Effect Predictor (VEP) database.

2.6.2 Gene burden analysis

A gene-burden test was designed and performed by Dr Evangelos Bellos to identify genes carrying more variants than expected in our patients. To achieve this, comparator groups were generated using data from the gnomAD project (https://gnomad.broadinstitute.org/) (293), the largest and most ethnically diverse collection of exomes currently available. Exonic variants with allele frequencies below 1% were aggregated at the gene level, and statistical significance of the burden in cases was empirically evaluated using random sampling of gnomAD. To account for multiple testing, the significance threshold was set to $3x10^{-6}$. Genes identified as significantly burdened using this approach, were then incorporated into the subsequent analysis pipelines as potential additional candidate genes.

2.7 Western blotting

2.7.1 Protein isolation and quantification

Cells were washed with PBS and adherent cells were mechanically detached, collected and placed immediately on ice. Cells were then centrifuged at 1500 rpm, for 5 minutes at 4 °C and supernatant was discarded. For lysis, cell pellets were then resuspended in RIPA buffer (Sigma), supplemented with 1x Halt™ Protease and Phosphatase Inhibitor Cocktail (Thermo Scientific) and Pierce™ Phosphatase Inhibitor (Thermo Scientific), and incubated on ice for a minimum of 30 minutes. Samples were then centrifuged at 13000 rpm, for 15 minutes at 4°C and protein supernatant was placed back on ice and either used immediately for downstream applications or stored at -80 °C. Protein quantification was performed using the Pierce™ Coomassie (Bradford) Protein Assay Kit (Thermo Fisher) according to manufacturer's Standard Microplate Protocol, and absorbance was measured using a Versamax microplate reader (Molecular Devices) at 595nm.

2.7.2 SDS polyacrylamide gel electrophoresis

Protein concentrations were equalised amongst samples to a total volume of 6.5 µL to which 1 µL of NuPAGE Sample Reducing Agent (Thermo Fisher) and 2.5 µL of NuPAGE LDS Sample Buffer (Thermo Fisher) was added, total volume 10 µL. Prepared protein samples were then denatured for 10 minutes at 70 °C and then placed back on ice. Meanwhile, precast 10% polyacrylamide NuPAGE™ Bis-Tris, 1.0 mm, Midi Protein Gels (Thermo Fisher) were set up using adaptors and placed into a Criterion™ Cell tank (Bio-Rad). Tank and gel reservoirs were filled with 1x XT MES Running Buffer (Bio Rad) and 10 µL of each sample were loaded into each well alongside a Precision Plus Protein Dual Colour Standards ladder (Bio Rad). Gel electrophoresis was then executed at 110 Volts for at least 1.5 hours to achieve good protein separation.

2.7.3 Transfer to membrane

Post gel electrophoresis, resolved proteins were dry transferred from the polyacrylamide gel to a polyvinylidene difluoride (PVDF) membrane using the iBlot™ Transfer Stacks and iBlot™ 2 Gel Transfer Device system (Invitrogen) according to manufacturer's instructions. Dry blotting transfer method: 20 V for 1 minute, 23 V for 4 minutes, then 25 V for 2 minutes. Membranes with transferred resolved proteins were then cut to size with a sterile scalpel and

placed in a blocking buffer of 5% Bovine Serum Albumin (BSA) in 1X Tris-Buffered Saline, 0.1% Tween (Sigma) (TBST) for at least 1 hour with gentle rocking.

2.7.4 Protein identification

Antibody details are listed in Table 2.5. Membranes were incubated in either 5% BSA or 5% milk (Pierce) in 1x TBST buffer with primary antibodies diluted to an appropriate concentration (Table 2.5) overnight at 4 °C with gentle rocking. Excess antibody was removed by washing the membranes three times with fresh 1x TBST for 5 minutes each with gentle rocking. Next, the membranes were incubated in either 1% BSA or 1% milk in 1x TBST buffer with secondary anti-rabbit IgG antibodies conjugated to horseradish peroxidise (HRP) (Table 2.5) at room temperature for 1 hour with gentle rocking. Protein bound antibodies were developed using the ECL Western Blotting kit (GE Healthcare), visualised with Azure BioSystems c600, and analysed using ImageJ software.

2.8 Adhesion and invasion assay

Human cell lines, 16HBE14, HEK293, and HaCaT, grown in appropriate complete medium and transfected for 24 hours with KRT10 constructs (Table 2.2) as described previously. Bacteria were suspended in PBS to achieve an multiplicity of infection (MOI) of 1 for both adherence and invasion assays, as previously described (119). Briefly, for the adherence assay, the confluent monolayer of transiently transfected cells was infected with bacteria for 90 minutes, then the wells were washed 3 times with PBS and incubated with 1% saponin (Sigma) for 10 minutes at 37°C. For the invasion assay, 150 µg/ml of gentamicin (Sigma) was added to each well and incubated for a further 1 hour at 37°C. Wells were then processed in the same manner as the adhesion assay and viable counts were determined by serial dilution and CFU counting.

2.9 Flow cytometry

Flow cytometry was performed on PBMCs and γ2A cells. 1×10^6 cells were stimulated for 20 min at 37 °C with IFN-α, IFN-γ, IL-6, LPS, (see Table 2.1 for details). Cells were also live infected with MC58 Nm, *S. aureus*, *Group A Streptococcus pyogenes*, and *Streptococcus pneumoniae* at either MOI 1 or 10. Stimulation was stopped by adding 1 volume of fixation buffer (4% PFA in PBS) and incubating in a 37°C water bath for 10 minutes. Cells were centrifuged at 250 g for 5 mins and washed once with 1x PBS. Cells were permeabilized for 30 min on ice with pre-chilled Phosflow PermIII buffer (BD biosciences) prior to staining.

Cells were stained with fluorochrome-conjugated antibodies (Table 2.5) in the dark for 30 minutes at room temperature. Data acquisition was done using BD LSR II (BD Biosciences) and analysed with FlowJo v10.8.1.

2.10 Static adhesion assay

Static adhesion assays were performed as previously described (294). To summarise, 96-well tissue culture plates were coated overnight at 4°C with 0.85 µg/well of goat anti-human immunoglobulin G (IgG) Fc fragment (Dianova). Wells were then washed with twice with 1x PBS (-Ca2+, -Mg2+) and coated with 0.3 µg/well of recombinant Human VCAM-1 and/or ICAM-1 protein (R&D Systems) or 1% BSA in PBS as a control and incubated for 90mins at room temperature. Wells were then washed with three times with 1x PBS and blocked with 200 µL/well of 1% BSA in PBS for 2 hours. Isolation of human polymorphonuclear leukocytes were performed as described previously (see 2.1.2.2) and cells were incubated with either mock or GGTI-2147 (10 µM, 30 minutes), at 37°C, 5% CO2 and then plated at $1.5x10^5$ cells/well and incubated at 37°C, 5% CO2 for 1 hour. Non-adherent cells were removed by washing the wells 3 times with HBSS and cell adhesion were determined by microscopy and processed using ImageJ Software.

2.11 Dual luciferase reporter assay

TICAM1-/-fibroblast cells were plated in 24-well plates and transfected with the expression vectors in the manner described previously (see 2.1.5). 24 hours after transfection, cell lysates were mechanically collected and transferred to a 96-well white polystyrene tissue culture treated plate (Corning Costar) and analysed with the Dual-Luciferase Reporter Assay System (Promega Corporation), according to the manufacturer's instructions. Microplate measurements were read using FLUOstar® Omega (BMG Labtech). Firefly luciferase values were normalized against the Renilla luciferase activity and expressed as fold change with respect to cells transfected with the wild-type vector.

2.12 *CTLA-4* transendocytosis

This assay was performed by Dr Bodo Grimbacher's team (Freiburg, Germany). CTLA-4 function was measuring by investigating the transendocytosis of CD80 (a CTLA-4 ligand) by flow cytometry. Briefly, primary human CD4+ T cells were purified from PBMCs by negative selection and activated with CD3/CD28 beads for 16 hours in the presence of mouse embryonic

fibroblast (MEF) cells expressing CD80-mScarlet. After 16 hours cells were harvested and gated using flow cytometry on CD4+CD45RO+FOXP3+ and analysed for mScarlet uptake.

2.13 Immunofluorescence

Cells were seeded on ethanol-sterilised coverslips (VWR, USA) in 24-well plates and transfected as previously described for 24 hours and incubated at 37°C, 5% CO2. Cells were washed once with 1x PBS and fixed for 15 minutes on ice with pre-chilled 100% methanol (Sigma). The methanol was then aspirated off and cells were washed 3x with PBS and permeabilised with 0.1% Triton-PBS (BDH, UK) for 4 minutes at room temperature. Samples were washed with PBS and blocked with 1% BSA in PBS for 10 minutes at room temperature. Coverslips were then incubated in a dark humidified chamber with the anti-flag primary antibody (Table 2.5) for 45 minutes at room temperature. Coverslips were then washed 7 times with 1x PBS and incubated with anti-mouse secondary antibody (Table 2.5) in a dark humidified chamber for 45 minutes at room temperature. Coverslips were washed 7 times with 1x PBS and mounted on SuperFrost Microscope Slides (VWR) in ProLong™ Gold Antifade Mountant with DAPI (Invitrogen) and sealed using clear varnish. Slides were incubated at 4ºC overnight, imaged on the LSM 5 Pascal Confocal microscope (Zeiss, Germany), and processed using ImageJ Software.

Chapter 3

Chapter 3 Gain-of-function *JAK2* variants in invasive meningococcal disease

3.1 Background

Genetic variability underlying invasive meningococcal disease (IMD) has been extensively described in the introduction however, most of these studies have been heavily based on GWAS and candidate gene studies (Chapter 1, section 1.4). While they have provided valuable knowledge of Nm pathogenesis, many of the significant findings have not been independently replicated successfully, and have inappropriate control groups or inadequate sample sizes (11). In this thesis, we opted to carry out WES and genetic burden analysis to identify variants that could then be functional characterised. WES, when appropriately powered, can be effective in discovering rare variants, in an unbiased non-candidate driven approach, in both an individual case study (295) and in groups of affected unrelated cases (296).

Despite effective vaccinations available, not all countries have mandated their use in national vaccination programmes and there are still cases of IMD (297). IMD is rare but has significant mortality with a case fatality rate of 10%, with up to 19% of survivors suffering from life-long severe sequelae with a reduced quality of life (49, 50). Left untreated, IMD case fatality rate can rise to 50% in non-epidemic settings and climb further to 80% during meningococcal outbreaks (50, 298). The identification of causative variants underlying IMD will allow for greater insight into disease pathogenesis and could therefore facilitate more accurate clinical care and diagnosis. To identify rare and novel variants associated with IMD this project utilised WES on a cohort of unrelated patients, with a clearly defined phenotype, and significant results were followed by functional characterisation.

3.2 Whole exome sequencing cohort (EUCLIDS)

In 2012, the FP7 European Union funded childhood life-threatening infectious diseases study (EUCLIDS) was founded to understand the genetic basis underlying susceptibility and outcome of life-threatening childhood bacterial infections. The consortium aimed to undertake large scale genomic investigations and bioinformatic studies to identify novel biomarkers to predict susceptibility and outcome of childhood infection. The study recruited infections caused by meningococcus, pneumococcus, streptococcus, staphylococcus, *E. coli*, and *Salmonella* alongside controls, via the EUCLIDS consortium. The cohort of patients with a range of childhood infections and no known predisposing risk factors, including exclusion of known primary immunodeficiencies, were recruited in collaboration with 11 study sites based in 7 countries including the UK (mainly through St Mary's Hospital, London and the Meningitis Research Foundation charity network), Spain, Holland, Germany, Italy, Switzerland and The Gambia. Due to the pre-existing legacy cohorts of IMD patients, meningococcal disease investigations were prioritized for further genetic investigations. Meningococcal cases were confirmed by microbiological culture at the individual study sites, as well as by PCR amplification at the diagnostic Micropathology laboratory based in Warwick, UK. Over a 5-year period, a total of 2,844 patients and controls were recruited into EUCLIDS with the meningococcal cohort contributing 9% (n=261) (Figure 3.1).

A sub-cohort of 750 paediatric patients with a clear pathogen diagnosis of meningococcus, pneumococcus, streptococcus, or staphylococcus invasive disease were selected for WES to contribute to subsequent genomic investigations (Figure 3.1). Within this group, a meningococcal disease cohort (n=261) were included in the WES cohort if the following criteria were met. The inclusion criteria consisted of, 1) children under the age of 18, 2), laboratory confirmation of Nm as the causative pathogen in IMD cases detected either by blood or CSF culture/PCR, and 3), presentation of severe IMD outcomes, including death, skin graft, amputation, meningitis and sepsis.

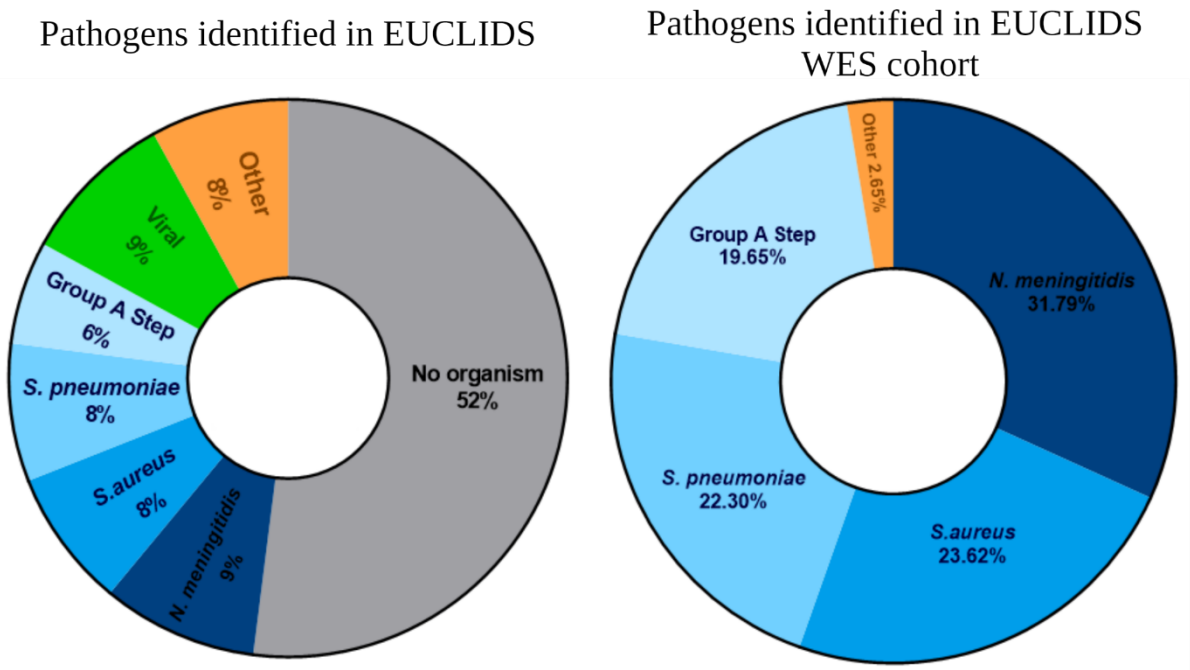


Figure 3.1 Pathogens identified in EUCLIDS

A summary of causative pathogens identified in patients recruited into EUCLIDS, n=2,844 (left) and in the subset of patients whose DNA was selected for whole exome sequencing n=750 (right).

3.3 Bioinformatics and cohort analysis approach

Due to the studies clearly defined phenotype a gene burden analysis of WES was performed to reveal rare and novel genes enriched in the IMD cases. The gene burden test was designed and performed by Dr Evangelos Bellos (Imperial College London) to identify genes carrying more variants than expected in our patient cohort compared to a healthy control cohort. To achieve this, comparator groups were generated using data from the gnomAD project (https://gnomad.broadinstitute.org/) (293), the largest and most ethnically diverse collection of sequences (exomes and genomes) currently available. This control group are largely adults who are unlikely to have suffered from invasive bacterial infections or harbour any IEIs, though this cannot be confirmed with certainty as it is derived from an anonymised public database. Non-synonymous genetic variants with allele frequencies below 1% were aggregated at the gene level, and statistical significance of the burden in IMD cases was empirically evaluated using random sampling of gnomAD. To account for multiple testing, the significance threshold was set to $3x10^{-6}$. A list of significantly enriched genes was produced using this approach and

interrogated to determine biological relevance to continue with functional characterisation. Among this list, *JAK2*, encoded on chromosome 9p24, was identified and found to carry statistically significantly more non-synonymous rare variants in severe IMD cases vs controls, and was originally chosen as a promising candidate for further investigations due to its role in cytokine signalling, innate host immunity (299, 300), and essential thrombocythemia, a rare blood disease in which platelets are over produced by the bone marrow (301-303).

JAK2 is a multifunctional tyrosine kinase that acts as a critical intermediate between membrane receptors and downstream signalling molecules such as STATs (Signal Transducer and Activator of Transcription). It belongs to the Janus family of cytoplasmic tyrosine kinases which also include JAK1, JAK3, and TYK2 in mammals (299). When a membrane receptor binds to a particular cytokine, hormone, or growth factor, including interferon-γ (IFN-γ), IL-6, IL-3, IL-5, IL-12, IL-13, IL-23, erythropoietin (EPO), thrombopoietin (TPO), growth hormones, leptin, oncostatin M (OSM), and leukaemia inhibitory factor, JAK2 undergoes trans-phosphorylation at its tyrosine 1007–1008 activation loop sites, located in the tyrosine kinase domain (also known as JH1), activating JAK2 signalling ability (Figure 3.2) (299, 300). This activates downstream signalling pathways involved in a wide range of different processes such as cell proliferation, cell survival, apoptosis, cell metabolism, cell differentiation, cytokine signalling, neutrophil recruitment, and haematopoiesis (Figure 3.2).

Human *JAK2* mutations are typically associated with adult onset acquired somatic mutations causing myeloproliferative diseases. One particular variant, p.V617F, is a well-known constitutive activator and is the most clinically relevant variant, found in over 95% of individuals with polycythaemia vera (PV), over 50% of those with essential thrombocythemia (ET), and around 65% of primary myelofibrosis cases (301-303). Myeloproliferative neoplasms driven by clonal expansion and over-proliferation of bone marrow cells of the myeloid lineage, can lead to clinical complications such as thrombosis, secondary fibrosis, or leukemic transformation (304-307). According to a literature search, genetic variability in *JAK2* has not yet been clearly associated with susceptibility to or severity of infectious disease.

My role in this project was to further investigate the *JAK2* variants discovered in the gene burden analysis, conducted by Dr Evangelos Bellos, involving severe IMD cases. Some of the work at the beginning of this project was conducted by a colleague in the group, Mr Rahmeen Rahman, as I will clearly state throughout this chapter where appropriate.

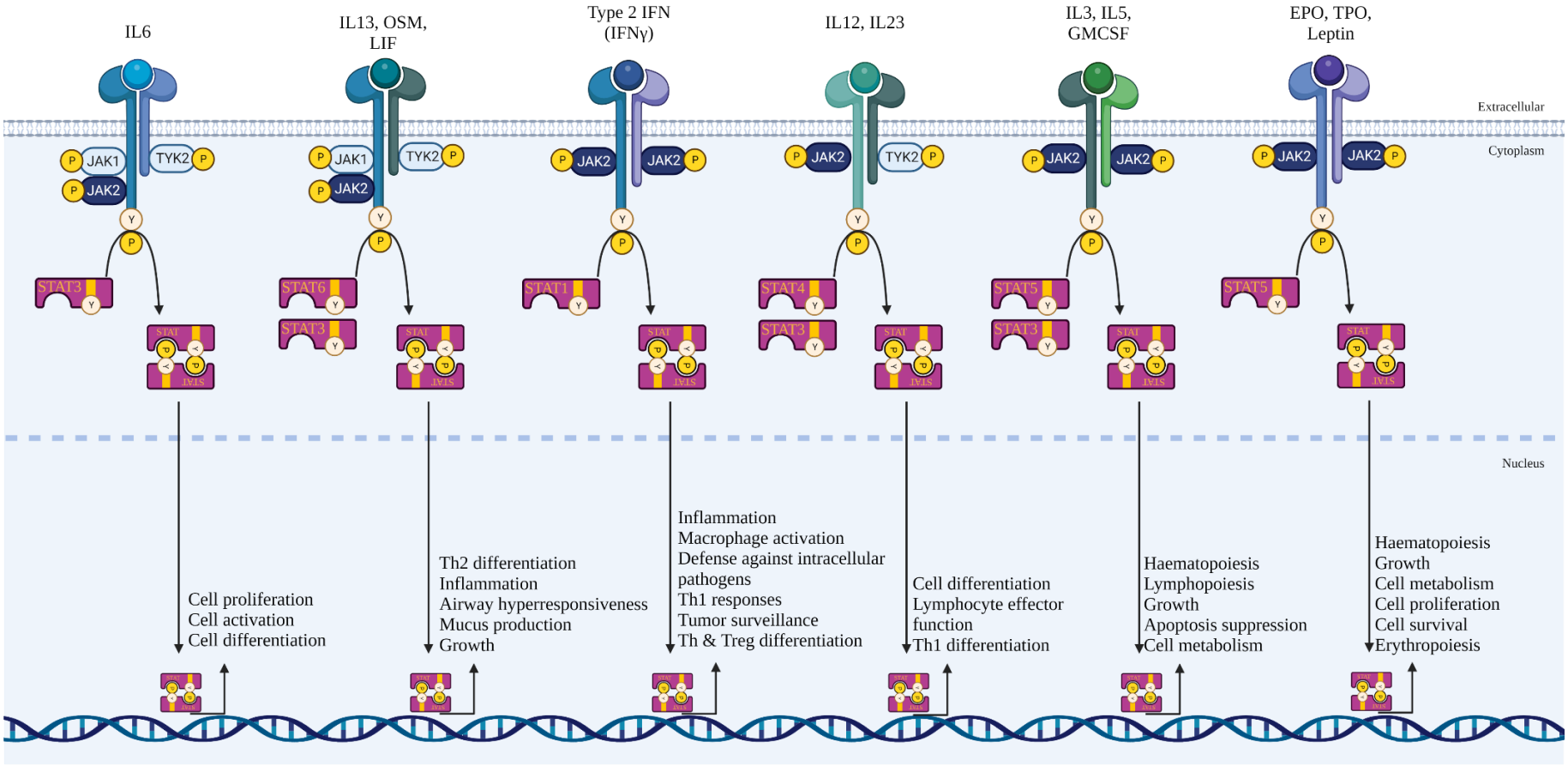


Figure 3.2 Overview of JAK2/STAT signalling

Circulating interferons, cytokines, hormones, and growth factors bind to membrane receptors which activates and phosphorylates (P) tyrosine residues (Y) on JAK2 that can then phosphorylate tyrosine residues in latent STATs located in the cytoplasm, which dimerise and translocate to the nucleus to regulate the transcription of genes involved in mediating multiple processes including induction of anti-microbial proteins, T cell differentiation, B cell maturation and development, apoptosis, cell differentiation, cell proliferation, haematopoiesis, erythropoiesis, and cell metabolism. Image created using biorender.com.

3.4 Discovery of missense heterozygous *JAK2* variants in unrelated paediatric cases of IMD and other significant bacterial infections

3.4.1 Clinical presentation of EUCLIDS patients with *JAK2* variants

A total of 10 heterozygous mutations were identified in 18 IMD cases, and from these 10 variants I discovered three of the variants were also present in other cases of non-meningococcal invasive bacterial infections caused by staphylococcus, Group A streptococcus, and pneumococcus (Table 3.3). Further scrutiny of the EUCLIDS WES cohort also revealed 6 new *JAK2* variants found in invasive staphylococcal, Group A streptococcal, and pneumococcal infections but not IMD (Table 3.3). The gene burden analysis revealed *JAK2* to be significantly enriched, with a p value $< 3x10^{-6}$, in only severe IMD cases when compared to controls and not when overall invasive bacterial disease was compared to controls. I generated both an overall summary of demographic and clinical data seen within EUCLIDS as a whole and in the cases where *JAK2* variants are found (Table 3.1) and a more detailed table containing individual data of all EUCLIDS patients where *JAK2* variants were discovered (Table 3.2).

Sex and age demographics in the cohorts with *JAK2* variants are evenly spread and similar to the overall EUCLIDS cohort (Table 3.1). 5/18 IMD patients with *JAK2* variants (27.78%) presented with meningitis, similar to the overall rate seen in EUCLIDs patients (21.63%) and in the IMD subset (20.83%). Notably, the cohort of *JAK2* variants present in IMD cases contain some of the most extreme manifestations of IMD, including 4 of the 18 patients (22.22%) dying, a higher percentage than is seen in both the overall EUCLIDS cohort (3.97%) and larger IMD cohort (5.56%) (Table 3.1). IMD patients with *JAK2* variants also had a higher proportion of sepsis incidence seen in 13/18 patients (72.22%), than the overall IMD cohort (48.61%). 3/13 IMD sepsis patients with *JAK2* variants (16.67%) received a skin graft and one of these patients subsequently died while another patient went on to have a four-limb amputation (Table 3.1, Table 3.2). 3/18 IMD patients with *JAK2* variants developed septic shock, of which two patients died (Table 3.1, Table 3.2). Table 3.2 details the demographic and clinical data from severe bacterial disease patients in which *JAK2* variants were discovered. All the IMD cases were caused by serogroup B and C, and most of them required PICU admission (72.22%) (Table 3.2).

Table 3.1 Demographic and clinical data summary of EUCLIDS patients, and EUCLIDS patients with *JAK2* variants

	EUCLIDS All *n=453*	**EUCLIDS Men** *n=144*	**EUCLIDS Staph** *n=107*	**EUCLIDS Pneumo** *n=101*	**EUCLIDS GAS** *n=89*	***JAK2* variants in IMD** *n=18*	***JAK2* variants in IMD cases in OBD** *n=12*	***JAK2* variants seen in OBD, but not IMD** *n=6*
Demographic characteristics								
Female	212 *46.80%*	68 *47.22%*	47 *43.93%*	45 *44.55%*	49 *55.06%*	8 *44.44%*	6 *50%*	4 *66.67%*
Male	231 *50.99%*	70 *48.61%*	60 *56.07%*	56 *55.45%*	40 *44.94%*	10 *55.56%*	6 *50%*	2 *33.33%*
Median age, months	39	38	109	23.5	43.5	38	88	51.5
Clinical syndrome								
Sepsis	258 *56.96%*	70 *48.61%*	62 *57.94%*	52 *51.48%*	64 *71.91%*	13 *72.22%*	1 *8.33%*	2 *33.33%*
Septic shock	64 *14.13%*	27 *18.75%*	6 *5.61%*	14 *13.86%*	15 *16.85%*	3 *16.67%*	1 *8.33%*	2 *33.33%*
Meningitis	98 *21.63%*	30 *20.83%*	5 *4.67%*	52 *51.48%*	10 *11.24%*	5 *27.78%*	2 *16.67%*	1 *16.67%*
Clinical data								
PICU admission	217 *47.90%*	46 *31.94%*	37 *34.58%*	69 *68.32%*	64 *71.91%*	13 *72.22%*	8 *66.67%*	4 *66.67%*
Death	18 *3.97%*	8 *5.56%*	3 *2.80%*	7 *6.93%*	0 *0.00%*	4 *22.22%*	1 *8.33%*	0 *0.00%*
Skin graft/ amputation	47 *10.38%*	31 *21.53%*	2 *1.87%*	0 *0.00%*	13 *14.61%*	3 *16.67%*	1 *8.33%*	1 *16.67%*

Values that ≠ 100% are due to mathematical rounding or an absence of data. OBD; Other bacterial disease (exclusive of IMD), IMD; invasive meningococcal disease, PICU; Paediatric intensive care unit.

Table 3.2 Detailed demographic and clinical details of EUCLIDS patients with *JAK2* variants

Patient	Sex	Ethnicity	Age of Onset (months)	Causative pathogen	Clinical phenotypes	PICU Admission	Severe clinical outcomes	Identified *JAK2* amino acid variant
IMD patients with *JAK2* variants								
102-1229	Female	European	9	Men B	Sepsis	Yes		L113V
318-1013	Female	European	193	Men	Septic Shock	Yes	Septic shock, death	L393V
MEN_2673	Female	Caucasian	NR	Men	Meningitis	Yes		L393V
101-5341	Male	European	NR	Men	Sepsis	NR		L393V
440492	Female	Caucasian	35	Men B	Sepsis	Yes	Skin graft	F556V
440582	Male	Caucasian	9	Men C	Sepsis	Yes	Death	R564L
ZH236	Male	Caucasian	15	Men C	Sepsis - *purpura fulminans*	Yes	Septic shock, death	N782K
439952	Male	Caucasian	10	Men B	Sepsis	Yes	Death	F844L
431052	Male	Caucasian	10	Men B	Sepsis	Yes	Skin graft, 4-limb amputation	E846D
PRS486	Male	Caucasian	10	MenC	Meningitis			E846D
SG112	Female	Caucasian	54	Men B	Sepsis			L871V
1956	Male	Caucasian	117	Men	Meningitis	Yes		R1063H

433692	Male	Caucasian	55	Men B	Sepsis	Yes	Unknown	R1063H
101-5611	Female	European	NR	Men	Meningitis			R1063H
437952	Female	Caucasian	210	Men C	Sepsis	Yes	Skin graft	N1108S
443662	Male	Caucasian	186	Men B	Sepsis	Yes	Unknown	N1108S
101-5501	Female	European	NR	Men	Meningitis		Unknown	N1108S
318-0001	Male	European	38	Men	Septic shock	Yes	Septic shock	N1108S
Other invasive bacterial disease patients that share *JAK2* variants present in IMD patients								
305-0002	Male	European	88	*S. pneumoniae*	Pneumonia with pleural empyema	Yes		L393V
401-1060	Female	Other/mixed	3	*S. pneumoniae*	Meningo/encephalitis	Yes	Hearing loss	L393V
101-1309	Male	Asian	12	GAS	Septic arthritis, Septic shock	Yes	Septic shock	L393V
401-1041	Female	African/North African	76	GAS	Meningo/encephalitis	No		L393V
201-1178	Male		126	GAS	Soft tissue infection, necrotising fasciitis, peritonitis	Yes		L393V
402-1037	Male	European	89	*S. aureus*	Soft tissue infection, pyomyositis	No		L393V
101-1267	Male	Other/mixed	35	*S. aureus*	osteomyelitis	No		R1063H
101-0022	Female	Other/mixed	144	GAS	Soft tissue infection, TSS	Yes	Skin graft, death	R1063H

308-1013	Male	African/North African	140	*S. aureus*	Soft tissue infection, osteomyelitis	No		N1108S
301-1129	Female	European	83	GAS	Pneumonia with pleural empyema, soft tissue infection, TSS	Yes		N1108S
402-1003	Female		143	GAS	TSS	Yes		N1108S
Invasive bacterial disease patients with *JAK2* variants, exclusive of IMD								
401-1106	Female	Other/mixed	9	*S. aureus*	Pneumonia with pleural empyema, septic shock	Yes	Septic shock, severe overall disability	G127A
101-0112	Female	European	56	*S. aureus*	Soft tissue infection, pyomyositis	No		M187I
101-0050	Female	Asian	24	*S. pneumoniae*	Pneumonia with pleural empyema, septic shock	Yes	Septic shock	V392M
101-0041	Female	Asian	47	GAS	Sepsis, TSS	Yes		S523T
301-0002	Male	European	157	*S. aureus*	Meningo/encephalitis, Mastoiditis and cerebellar abscess	Yes		G571S
308-1013	Male	African/North African	140	*S. aureus*	Soft tissue infection, osteomyelitis	No		I724T

NR; Not Recorded, IMD; Invasive Meningococcal Disease, TSS; Toxic Shock Syndrome, GAS; Group A Streptococcus.

3.5 Review of *JAK2* variants discovered in EUCLIDS

Table 3.3 summaries the 16 *JAK2* variants discovered in the EUCLIDS cohort in patients with significant bacterial disease. I have also included references from a brief literature review to determine if a particular *JAK2* variant has been described before and in what context. The two key parameters taken into consideration were whether they have been observed in myeloproliferative neoplasm (MPN) patients, and if there has been any functional work published that has characterised the variants. Nine of the overall 16 *JAK2* variants have been previously sequenced in MPN patients, L393V, F556V, R564L, E846D, R1063H, N1108S, V392M, G571S, and I724T (Table 3.3). However, L393V, I724T, and some MPN patients with R1063H, have been discovered only in co-existence with *JAK2* variant, V617F (Table 3.3). Additionally, E846D has only been co-inherited with R1063H in MPN patients (308). Adult onset somatically acquired *JAK2* V617F is the foremost clinically relevant variant strongly associated with myeloproliferative diseases (301-303). V617F is a gain-of-function variant that constitutively activates JAK2 and subsequent signalling. In myeloproliferative disorders, gain-of-function *JAK2* mutations drives over-proliferation of bone marrow cells which underlies the clinical manifestations of myeloproliferative neoplasm (305). Due to its well-established characteristics, I decided to use V617F as a positive control in my experiments to assess gain-of-function effects.

Four of the *JAK2* variants discovered in the EUCLIDS cohort have been previously characterised in published literature as gain-of-function mutations, L393V, F556V, E846D, and R1063H (Table 3.3). *JAK2* F556V, and E846D were clear gain-of-function variants with F556V shown to activate JAK-STAT5 signalling (309), promote cell proliferation and constitutively activate JAK2 in γ2A cells, JAK2 deficient human fibrosarcoma cells, and JAK2 mutants stably expressed in Ba/F3-EpoR cells (310). *JAK2* E846D was shown to lead to constitutive JAK2 activation and hyperactivate JAK2/STAT5 signalling in the specific context of the erythropoietin receptor in Ba/F3-EPOR-IL-3–dependent cells transduced to stably express E846D (308). The same study discovered E846D and R1063H in co-existence in MPN patients but showed R1063H to individually have a very weak effect on JAK2/STAT5 signalling, therefore hypothesising that R1063H augments JAK2 activity caused by E846D (308). This was also shown in functional studies where R1063H co-exists with V617F and further amplifies the gain-of-function effect of V617F but not independently (311). *JAK2* L393V alone has been reported to weakly activate JAK2, shown by STAT1/3/5 signalling luciferase assays, but though it was discovered in co-existence with V617F it did not synergise

V617F JAK2 activating activity (312). The *JAK2* G571S variant has also been sequenced in MPN patients, however functional studies so far have shown this variant to have no effect on JAK2 activity (309). The three *JAK2* variants with the highest minor allele frequency (MAF), p.L393V, pR1063H, and p.N1108S with MAFs of 0.007883, 0.004699, 0.001962, respectively are seen in both IMD cases and other invasive bacterial disease patients and have also been identified in MPN patients, (Table 3.2, Figure 3.3).

Figure 3.3 is a representative image of all the *JAK2* variants discovered in our EUCLIDS cohort, and colour coded to differentiate in which EUCLIDS subset they are found in, either solely in IMD cases (black), in both IMD and other invasive bacterial infections (blue), or in invasive bacterial infections exclusive of IMD (green). Two identified *JAK2* variants are novel, L871V, and S523T (Table 3.3, Figure 3.3). JAK2 has many key phosphorylation sites used to positively activate JAK2 activity, with Y1007/1008 as the chief activating site, and JAK2 also self-regulates via auto-phosphorylation (Figure 3.3). Only one *JAK2* variant found in the EUCLIDS cohort falls on one of these phosphorylation sites, S523T, a novel variant found in a female GAS patient presenting with sepsis and toxic shock syndrome (Figure 3.3, Table 3.2). *JAK2* S523 has been shown to act as a negative regulator in the context of growth hormone stimulation (313).

Table 3.3 Summary *JAK2* variants discovered in EUCLIDS

JAK2 variant	dbSNP ID	Global MAF (gnomAD)	Functional GoF reported?	Reported in patients with MPNs?	Present in EUCLIDS bacterial disease group
L113V	rs143103233	0.0006471			IMD
L393V	rs2230723	0.007883	Yes (312)[c]	Yes[a] (312, 314, 315)	IMD, *S. aureus*, GAS, *S. pneumoniae*
F556V	rs764634461	0.00001455	Yes (309, 310)	Yes (309)	IMD
R564L	rs368927897	0.00006051		Yes (316)	IMD
N782K	rs778634399	0.00001436			IMD
F844L	rs368688124	0.000007957			IMD
E846D	rs150221602	0.0004492	Yes (308)	Yes (308, 315)[b]	IMD
L871V	Novel	.			IMD
R1063H	rs41316003	0.004699	Yes[d] (308)[b] (311)[a]	Yes (315, 317) (308)[b] (311)[a]	IMD, *S. aureus*, GAS
N1108S	rs142269166	0.001962		Yes (317)	IMD, *S. aureus*, GAS
G127A	rs56118985	0.000003995			*S. aureus*
M187I	rs770883479	0.000008			*S. aureus*
V392M	rs200018153	0.0006015		Yes[a] (318)	*S. pneumoniae*
S523T	Novel	.			GAS
G571S	rs139504737	0.0004612	Predicted no effect (309)	Yes (309, 316, 319)	*S. aureus*
I724T	rs372254348	0.00008489		Yes[a] (318)	*S. aureus*

[a] co-exist with V617F [b] E846D & R1063H in co-existence [c] weak GoF [d] R1063H is not GoF alone but appears to be an amplifier when found in co-existence with other GoF variants.
IMD; Invasive Meningococcal Disease, GAS; *Group A Streptococcus*, MPN; myeloproliferative neoplasms, GoF; Gain-of-function, MAF; minor allele frequency

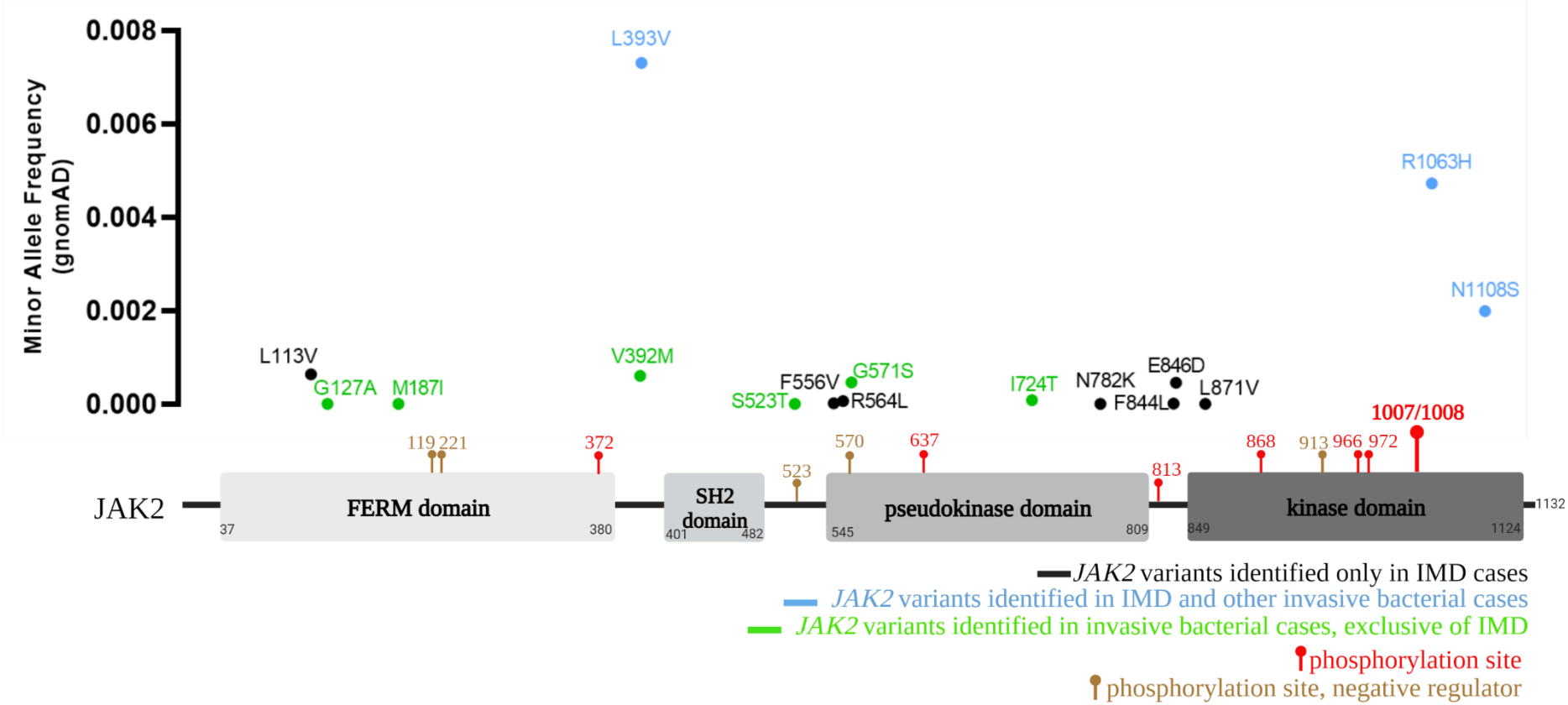


Figure 3.3 Lolliplot of *JAK2* variants and corresponding MAF

JAK2 amino acid variants discovered in EUCLIDS patients plotted against their reported global minor allele frequency (MAF) according to gnomAD, and along a schematic of *JAK2* protein with significant domains and phosphorylation sites. 4.1 protein, ezrin, radixin and moesin (FERM), Src Homology 2 (SH2). Known key JAK2 phosphoyrlation sites are also marked, including auto-phosphorylation that negatively regulates JAK2 activity. Image created using biorender.com.

3.6 Construction of *JAK2* variants in expression vectors

All 16 *JAK2* variants identified in the EUCLIDS cohort, plus an extra *JAK2* variant V617F, were constructed for use in experiments, (4 of the 17 *JAK2* variants, R564L, F844L, E846D, and R1063H, were constructed by Mr Rahmeen Rahman). All other plasmids were generated by me. Firstly, the pTRIP expression vector containing wild-type (WT) JAK2 C-terminally tagged with RFP, previously described (291), was used to introduce the missense variants by QuikChange Lightning Site-Directed Mutagenesis Kit from Agilent and overlapping primer pairs containing the single nucleotide substitution (detailed protocol found in section 2.4 and primer details found in Table 2.4). Confirmation of successful mutant constructs were determined by Sanger sequencing and then aligned to *JAK2* WT (Figure 3.4). Whole plasmid sequencing confirmed no second-site variants were introduced during the mutagenesis process.

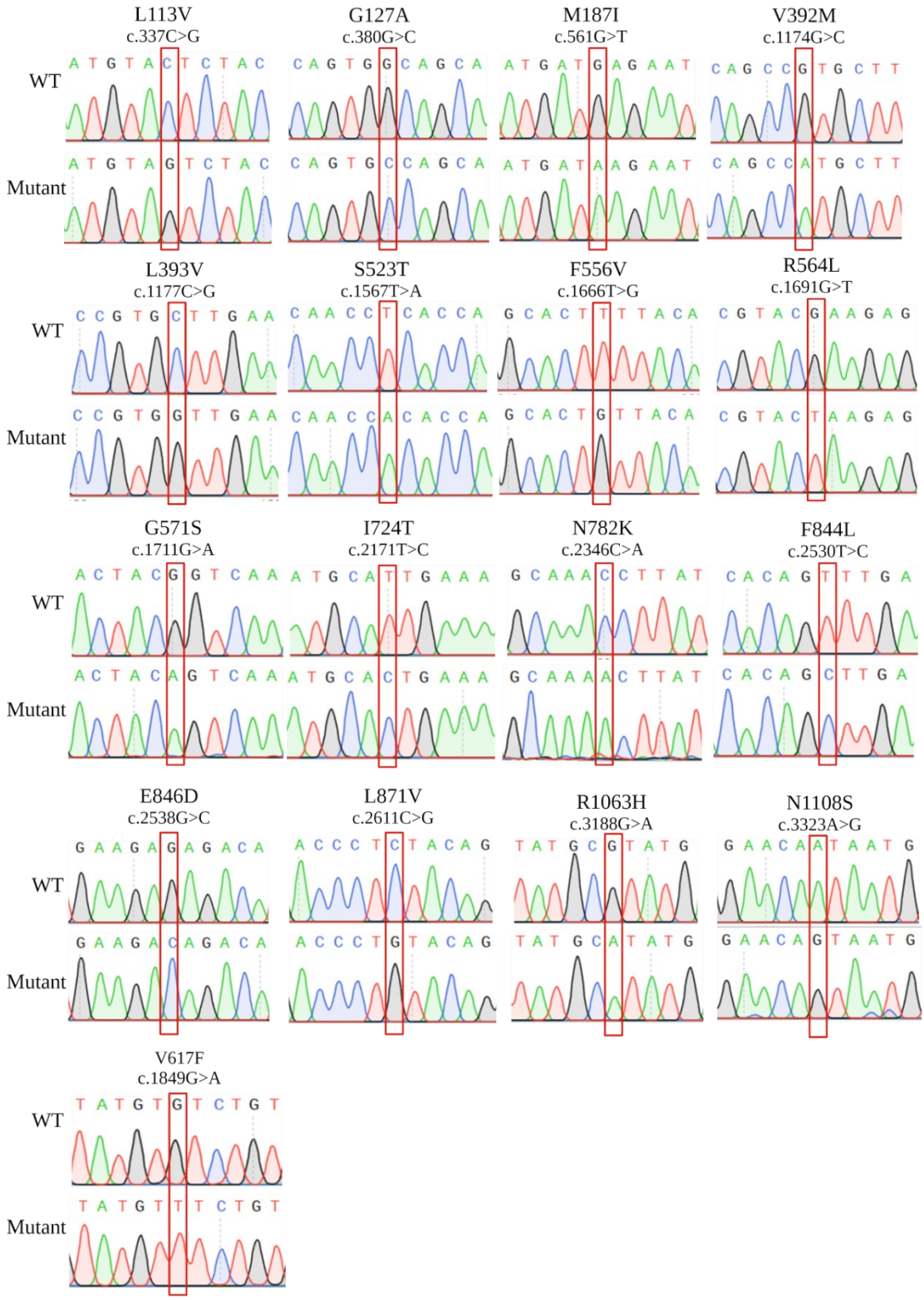


Figure 3.4 Sequence confirmation of site directed mutagenesis in *JAK2* plasmids

Sanger sequence confirmation of successfully integrated *JAK2* variants in pTRIP-JAK2 WT plasmid. Alignment of pTRIP-JAK2 WT and mutant plasmid, visualised using SnapGene (by Dotmatics). Nucleotide changes are marked with a red box. R564L, F844L, E846D, and R1063H, were constructed by Mr Rahmeen Rahman, all others were generated by me.

3.7 Overexpression of *JAK2* variants in *JAK2* deficient cells (γ2A)

3.7.1 Western blot analysis of baseline expression

To identify the impact of *JAK2* variants on protein expression, pTRIP-JAK2 constructs were transiently transfected into JAK2-/- human fibrosarcoma cells (γ2A), previously described (290). Western blots were then performed using antibodies recognising whole JAK2, activated phospho-JAK2 (Y1007/Y1008), STAT1, and activated phospho-STAT1 (Y701), with GAPDH used as a loading control, (detailed protocol outlined in section 2.6). The pTRIP-JAK2 construct V617F was used as a gain-of-function control. Three variants showed a clear gain-of-function with clear detectable bands in phosphorylated JAK2 immunoblots, F556V, E846D, and L871V (Figure 3.5). Densitometry of at least three independent experiments showed a significant increase in pJAK2 for F556V, E846D, and L871V when compared to JAK2-WT. Phosphorylated STAT1 was also investigated, as STAT1 is one of the latent STATs recruited downstream of phosphorylated JAK2 upon ligand binding as is a generally accepted read out indicative of JAK2 activation (Figure 3.2). Also, *JAK2* variants that may not appear as GoF in direct investigations of phosphorylated JAK2 may still prove to activate the JAK2/STAT pathways upon examining activity of downstream molecules. Activated STAT1 mirrors activated JAK2 immunoblots with F556V, E846D, and L871V being statistically significantly increased compared to JAK2 WT and other variants including, R564L, and R1063H also show a clear increase in activated STAT1 though these did not reach statistical significance (Figure 3.5). Other *JAK2* variants, L113V, L393V, N782K, F844L, N1108S also showed weak increase in phosphorylated STAT1 compared to JAK2-WT but this was not found to be statistically significant (Figure 3.5). All gain-of-function *JAK2* variants seen to activate JAK2 and STAT1, shown by immunoblotting, F556V, E846D, and L871V, are to a significant level but lesser than what is seen with V617F (Figure 3.5). No *JAK2* variants found in other invasive bacterial infection, exclusive of IMD were shown to be activating variants by immunoblotting (Figure 3.5).

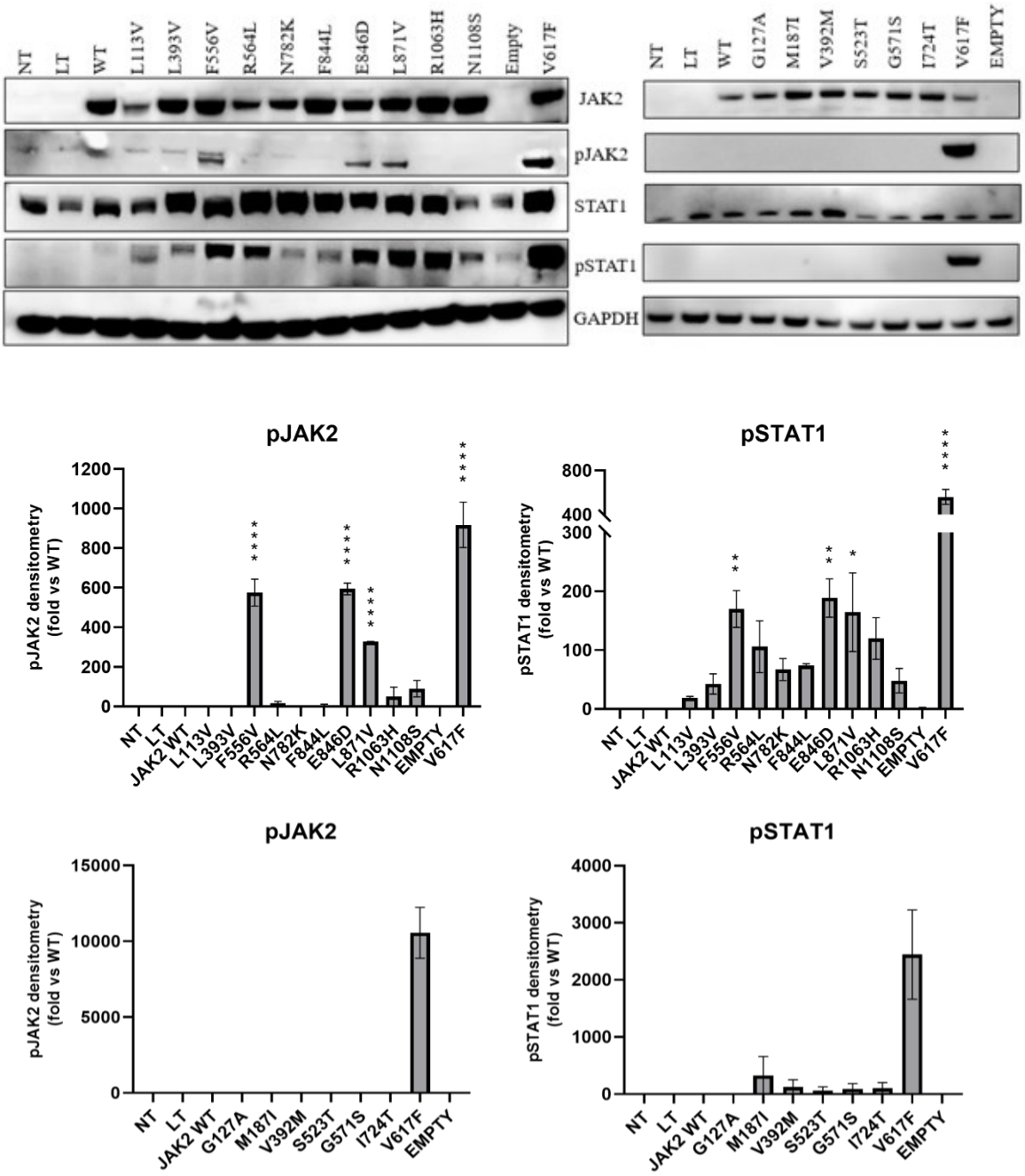


Figure 3.5 Impact of *JAK2* variants on protein expression

γ2A cells were transiently transfected with JAK2 constructs, immunoblots show resolved JAK2, pJAK2(Y1007), STAT1, and pSTAT1(Y701) proteins, with GAPDH as a loading control. N=3 and immunoblots are representative of at least three independent experiments. Densitometry analysed with ImageJ, data shown as mean ± SEM and the differences between variants and wild type were tested in one-way ANOVA, significance shown with *p<0.05, **p<0.01, ****p<0.0001. NT; non-transfected, LT; lipofectamine control, WT; wild type.

3.7.2 Western blot analysis of JAK2 and STAT1 expression in stimulated cells

Next, I wanted to investigate the impact of the *JAK2* variants in the context of stimulation with ligands that trigger the JAK/STAT pathway (Figure 3.2). Due to time restrictions, this experiment was performed with 4 of the JAK2 variants F556V, E846D, L871V, and R1063H. IFN-γ, and IL-6 were chosen as ligands for this experiment as they have been shown to signal downstream of JAK2 (Figure 3.2), and IFN-α is used as a control as it signals via JAK1 and TYK2 to phosphorylate STAT1 but it does not signal through JAK2 (320). Formalin-fixed *N. meningitis* MC58 strain, serogroup B (MenB) (MOI 1) was chosen due to its relevance in the context of this project. At Imperial College London, live *N. meningitis* is a category level 3 pathogen, and its use is extremely restrictive, therefore we chose to formalin fix the bacteria to allow for more flexibility in my experiments. This experiment has been performed once, and therefore no statistics was performed. In γ2A cells pTRIP-JAK2 constructs, were transiently transfected then stimulated with either mock, IFN-α, IFN-γ, IL-6, or FF-MenB and immunoblotting were performed. There is a clear trend that the highest pJAK2 and pSTAT1 expression is observed upon stimulation with IFN-γ in F556V, E846D, and V617F and only in pSTAT1 in JAK2-WT, L871V, and R1063H, when compared to JAK2-WT mock (Figure 3.6). Phosphorylated JAK2 was seen in F556V, E846D, and V617F to a higher degree in all stimulated conditions, compared to JAK2-WT (Figure 3.6). Additionally, increased expression of pSTAT1 was seen for all E846D stimulation conditions compared to WT (Figure 3.6).

In this experiment no activated JAK2 was seen in any conditions of L871V (Figure 3.6), in contrast to previous seen activation at base level (Figure 3.5). This is likely due to the normal subtle expression of phosphorylated JAK2 that is experimentally difficult to investigate, also a higher dilution of antibody recognising pJAK2 was also likely used as expression of pJAK2 in the V617F positive control is to a significantly lower degree (Figure 3.6) than what has been seen previously (Figure 3.5). The marked 10-fold increase in pSTAT1 expression upon IL-6 stimulation in cells expressing E846D is interesting, as this cytokine typically signals through STAT3 downstream of JAK1, JAK2, and TYK2 (321). However, it has been shown that though STAT3 is the dominant signalling molecule, STAT1 may be involved in the IL-6 signalling cascade and previous studies show IL-6 can induce phosphorylation of both STAT1 and STAT3 in γ2A cells (321, 322). However, as stated, this is the result of one biological experiment, and it would need to be repeated before any confident conclusionary remarks can be formulated. In future experiments, I also would stimulate with FF-MenB for at least 2 hours

at an MOI of 10 to replicate a more comparative infection scenario, as 20 minutes may not be enough time for bacterial PAMPs to activate the signalling pathway.

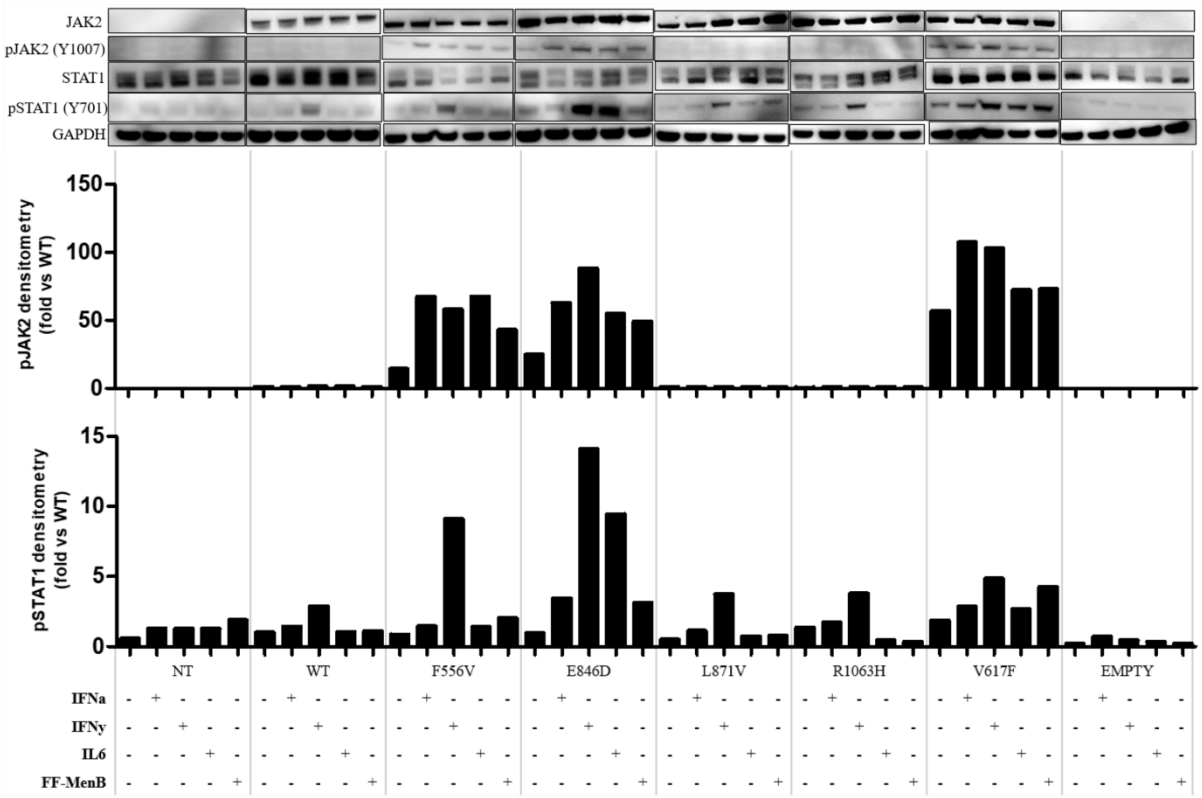


Figure 3.6 Impact of *JAK2* variants on protein expression upon stimulation

γ2A cells were transfected with JAK2 constructs then stimulated either with mock, IFN-α, IFN-γ, IL-6 , or formalin-fixed (FF) MenB. Immunoblots representative of n=1 and shows resolved JAK2, pJAK2(Y1007), STAT1, and pSTAT1(Y701) proteins, with GAPDH used as a loading control. Densitometry analysed with ImageJ. NT; non-transfected, WT; wild type.

3.7 Flow cytometry analysis of phosphorylated STAT1 in stimulated γ2A cells

Working with phosphorylated proteins is experimentally difficult as they are sensitive to time and temperature, and phosphorylated JAK2 antibodies created for immunoblotting does not always pick up the sensitive JAK2 phosphorylated state. Due to these limitations, I decided to attempt the same stimulation experiment as shown in Figure 3.6 but using a flow cytometry output, with the hope that flow cytometry's ability to measure at the single cell level would allow for more sensitive measurements and reproducible results. As time was restrictive, this

experiment was performed once with two of the *JAK2* variants shown to be strongly gain-of-function, F556V, and E846D and one JAK2 variant that thus far has not shown any gain-of-function activity, L113V, by immunoblotting (Figure 3.5). Activated JAK2 (pJAK2) was not included in this experiment as there are no commercially available antibodies. Briefly, γ2A cells were transiently transfected for 24 hours and then stimulated with either mock, IFN-α, or IFN-γ. Cells were stained with pSTAT1 (Y701) and *JAK2* E846D showed an increase in baseline pSTAT1 expression compared to JAK2-WT and showed the highest expression of pSTAT1 among all the variants upon IFN-γ stimulation (Figure 3.7). *JAK2* F556V also showed a marked increase in pSTAT1 upon IFN-γ stimulation, ~3-fold increase (Figure 3.7). All *JAK2* variants, showed a similar increase in pSTAT1 upon stimulation with IFN-α, when compared to JAK2-WT (Figure 3.7). This data is the result of one biological experiment and would need to be repeated with all the *JAK2* variants before any confident conclusionary remarks could be formulated.

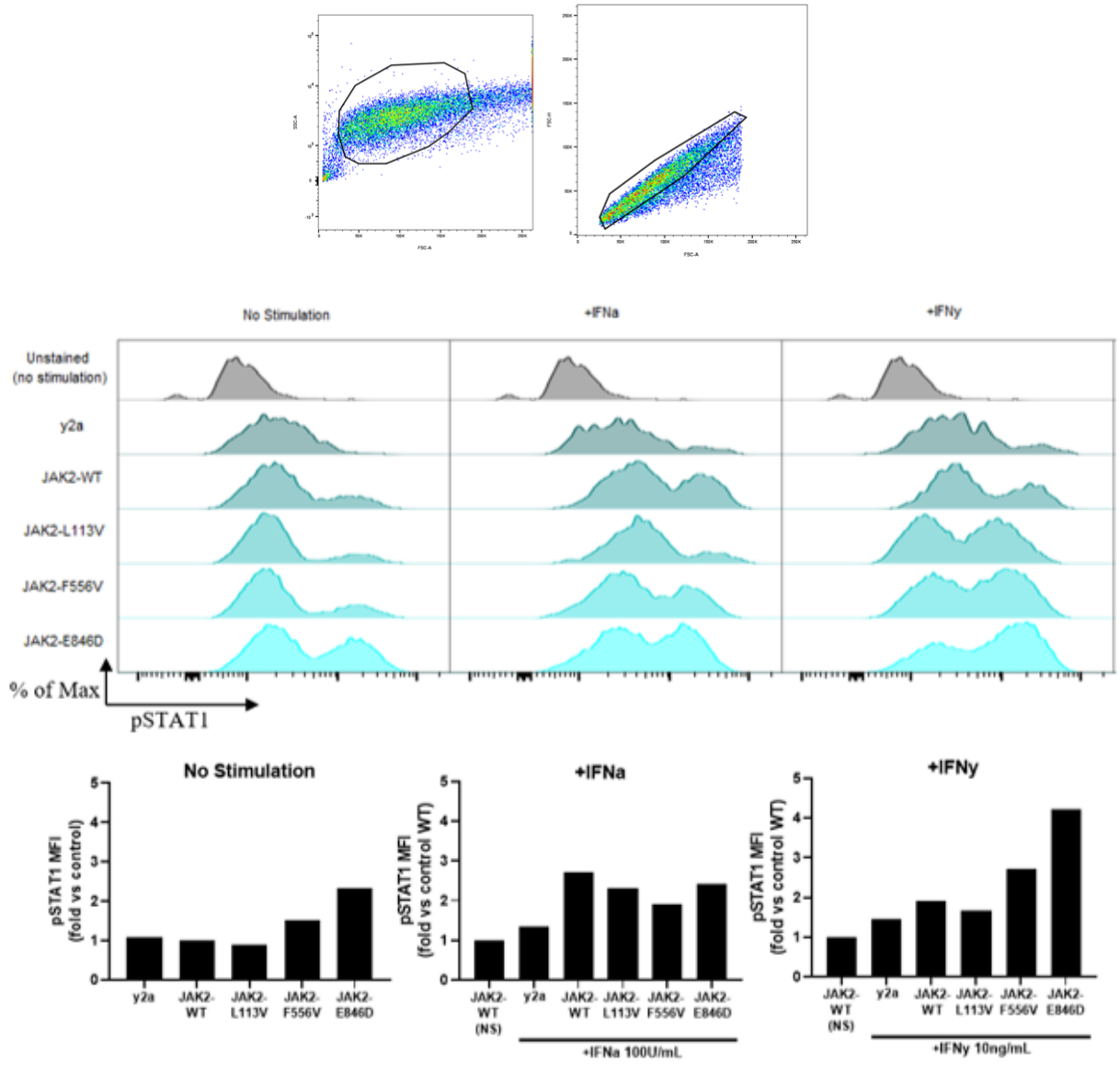


Figure 3.7 Expression of pSTAT1 in γ2A cells expressing JAK2 variants upon stimulation

Graphs show expression of intracellular pSTAT1 (Y701) in γ2A cells expressing JAK2 variants and stimulated with mock, IFN-α (1000 U/mL), or IFN-γ (100ng/mL). Gating strategy for single cell population is shown (top) and data is representative of n=1. Graphs showing MFI fold change are compared to non-stimulated JAK2-WT.

3.8 Flow cytometry analysis of phosphorylated STAT1 in stimulated PBMCs

3.8.1 Flow cytometry optimisation in stimulated control PBMCs

During this project, I was presented with the possibility of being able to successfully recall some of the EUCLIDS patients with *JAK2* variants, therefore, I began optimising experiments that could take advantage of fresh patient samples. Though I would like to investigate the activation of JAK2 and STAT1 by immunoblotting, the amount of blood needed to isolate enough protein to successfully run the experiment is high and therefore prohibitive. Therefore, the first experiment would be to investigate activation of STAT1 in patient peripheral blood mononuclear cells (PBMCs) using flow cytometry as this requires a lower volume of blood. To optimise this experiment, control PBMCs were stimulated with mock, IFN-α, IFN-γ or were live infected for 20 minutes with *N. meningitidis* (serogroup B strain MC58), *S. aureus, S. pneumoniae, or Group A Streptococcus* (GAS). Cells were stained with CD14-FITC, to gate for monocytes, CD4-PerCP/Cy5.5, to gate for T cells, and pSTAT1-BV421 (Y701). Cells were gated for a single cell population using forward scatter and then CD14+ and CD4+ single stained cells were used to positively gate for these populations (Figure 3.8).

CD14+pSTAT1+ cells were significantly increased upon stimulation with IFN-α (1000 U/mL), and IFN-γ (100ng/mL), all bacterial infections also showed an increased expression in CD14+pSTAT1+ cells, though this did not reach significance (Figure 3.8). In all PBMCS, there was a significant increase in pSTAT1 only upon IFN-α stimulation and no increase in pSTAT1+ cells were seen in any CD4+ cells (Figure 3.8).

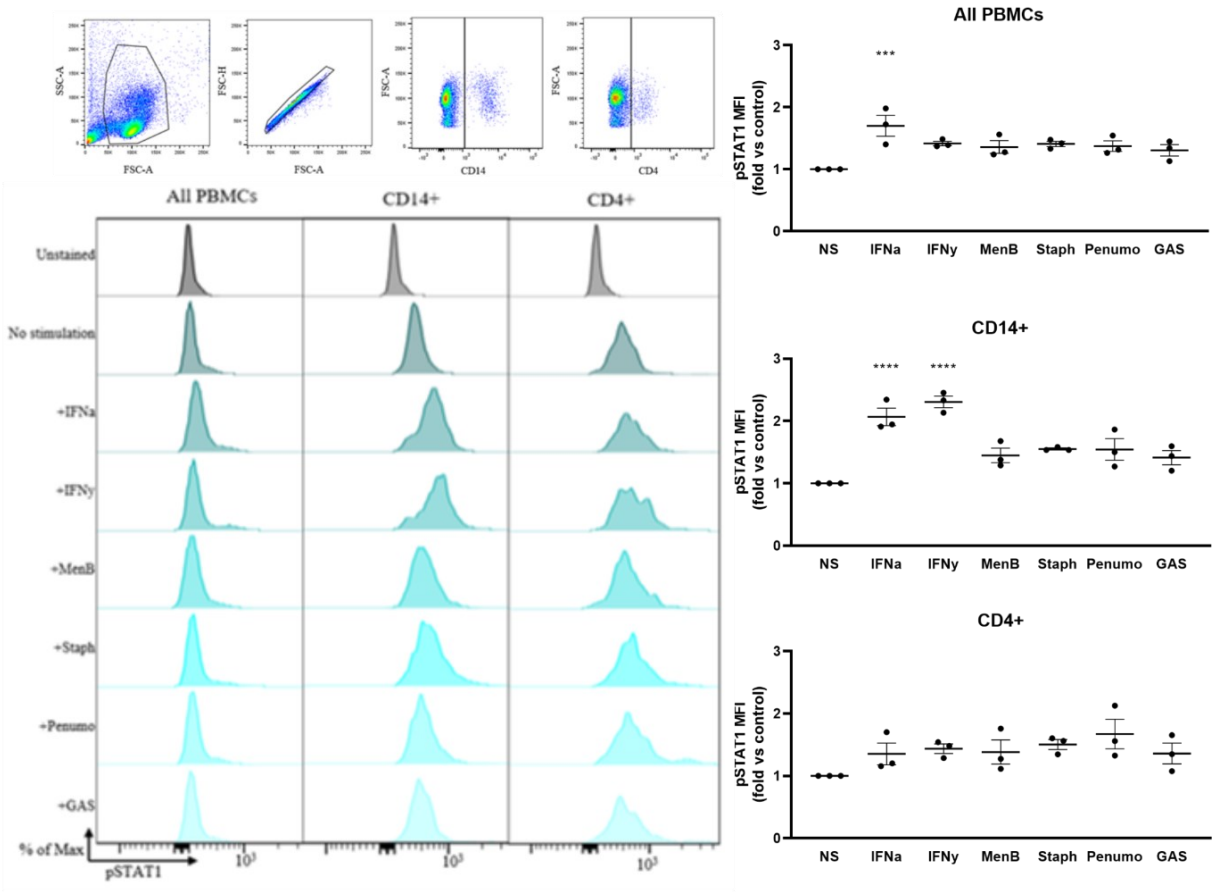


Figure 3.8 Expression of pSTAT1 in PBMC's upon stimulation

Graphs show expression of intracellular pSTAT1 (Y701) in PBMCs cells stimulated with mock, IFN-α, IFN-γ, or live infection with *N. meningitidis*, *S. aureus, S. pneumoniae,* or *Group A streptococcus*. Gating strategy for single cell population, CD14+ population, and CD4+ population is shown (top). Data is representative of n=3 biological replicates. Graph shows mean ± SEM and MFI fold change is compared to non-stimulated. The differences were statistically tested in one-way ANOVA, and significance is shown with ***p<0.001, ****p<0.0001.

3.8.2 Flow cytometry in stimulated control vs V617F PBMCs

I shall be receiving EUCLIDS patient samples shortly, however I was recently presented with an opportunity to obtain blood samples from a myeloproliferative neoplasm patient positive for the *JAK2* V617F variant. This was possible through collaboration with Dr Simone Claudiani (Imperial College London). I was able to run my optimised flow cytometry experiment measuring the expression of pSTAT1 in PBMCs upon IFN-stimulation and bacterial infection described above (section 3.7.1) using blood from a patient positive for V617F and a control donor (Figure 3.9). Only one modification was made post-optimisation altering infection time with live bacteria from 20 minutes to 2 hours. This data is derived from one independent experiment, and more repeats will be needed to form any conclusory statements.

From this experiment, V617F shows a constitutive activation of pSTAT1 at baseline without any stimulation in all PBMCs and CD14+ cells, however this was not seen in the CD4+ cell population (Figure 3.9). Levels of pSTAT1 in control cells then reach similar levels to V617F cells upon stimulation in all PBMCs and CD14+ cells (Figure 3.9). CD4+pSTAT1+ cell populations show some interesting trends with an absence of pSTAT1 expression in V617F CD4+ cells upon stimulation with IFN-γ, MenB, and *S. pneumoniae,* compared to CD4+ control cells (Figure 3.9). An increased pSTAT1 response to all bacterial infections in the CD4+ population in control cells is also seen here and had not been present in previous experiments (Figure 3.8), however this could be attributed to the longer infection time employed for this experiment (Figure 3.9).

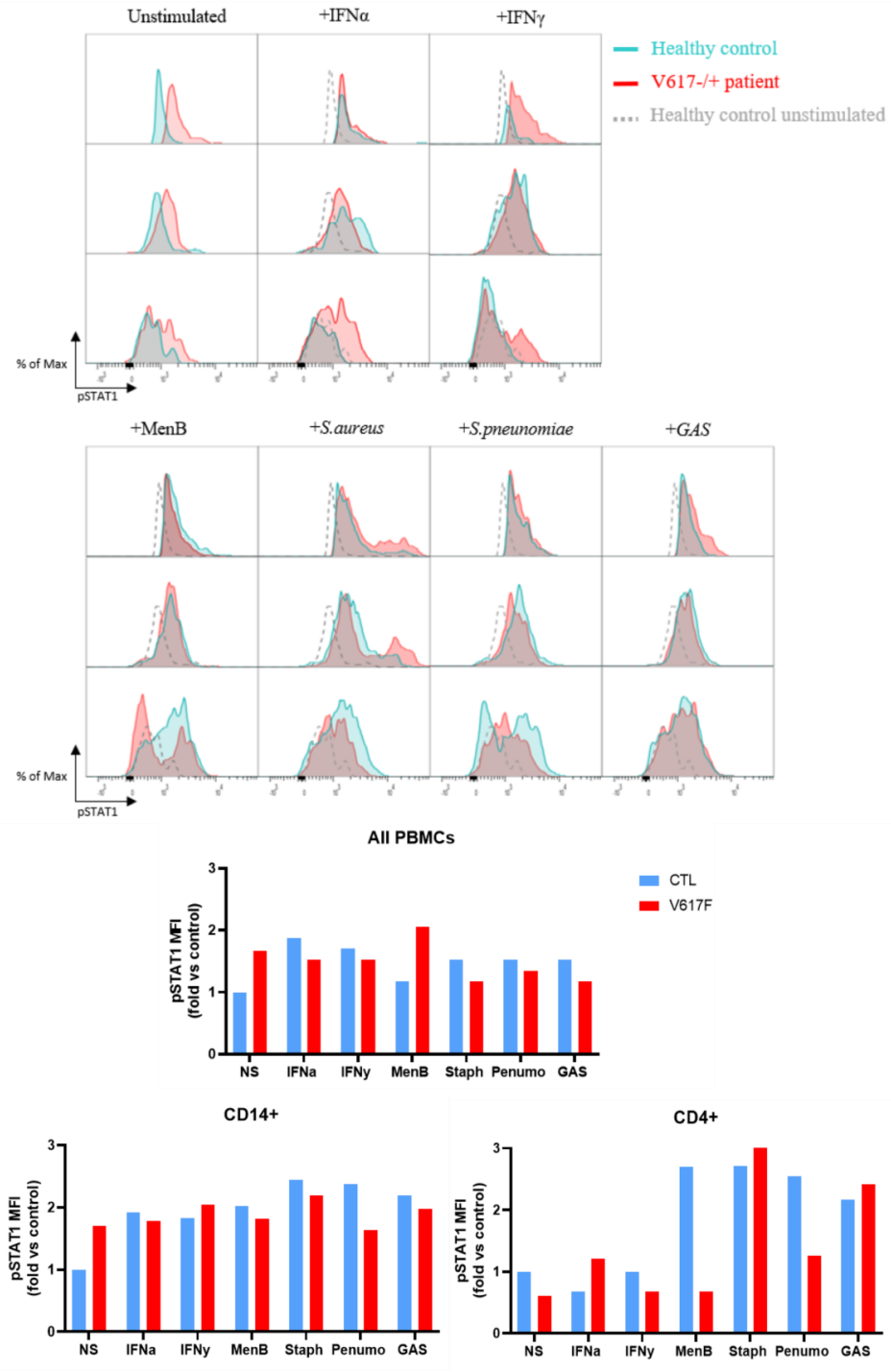


Figure 3.9 Impact of V617F JAK2 variant on response to interferon and infection in PBMCs

Graphs show expression of intracellular pSTAT1 (Y701) in human PBMCs from control and V617F whole blood and stimulated with mock, IFN-α, IFN-γ, or live infection with *N. meningitidis*, *S. aureus, S. pneumoniae,* or *Group A Streptococcus* (GAS). Cells were gated for CD14+ and CD4+ populations and data was plotted as mean ± SEM with MFI fold change determined by comparison to non-stimulated control PBMCs. Data is representative of n=1 biological replicate.

3.9 Creating stable mutant JAK2 expression in γ2A cells

I decided to create stable cell lines permanently expressing the *JAK2* variants identified in invasive bacterial infections from the EUCLIDs cohort, as transient transfections do not result in 100% transfection efficiency and can introduce variability in results. Also working with phosphorylated proteins can pose a difficulty in reproducible data and using stable cell lines could, hopefully, circumnavigate this issue.

The pTRIP-JAK2 lentiviral constructs were used to transduce JAK2 stable expressing γ2A cells. As lentiviral vectors cannot replicate on their own, pseudoparticles were generated in HEK293 cells first and successful expression of RFP+ particles was visualised using fluorescent microscopy. All JAK2 variants successfully produced lentiviral particles in HEK293 cells 48 hours post co-transfection, and as expected, significant cell death is also seen in the microscope images (Figure 3.10).

The pseudoparticles were transduced into γ2A cells and then fluorescence-activated cell sorting (FACS) was performed by Dr Parisa Amjadi (Imperil College London) to positively sort only RFP+ cells (Figure 3.11). All JAK2 transduced γ2A cells had positive RFP populations, with efficiency ranging from 13% in V617F transduced cells to >70% in M187I transduced cells (Figure 3.11).

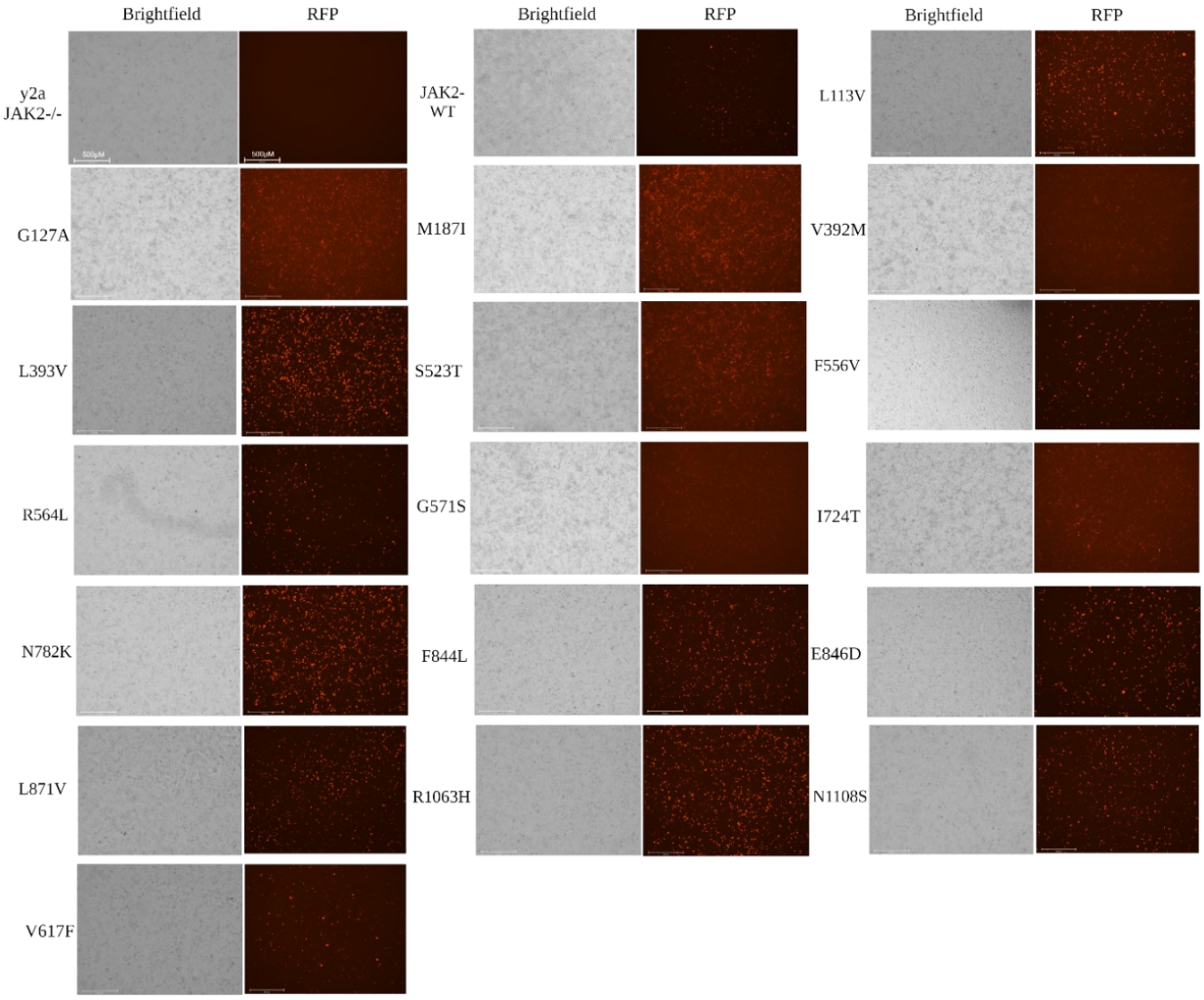


Figure 3.10 Fluorescent microscopy of JAK2-lentiviral particles generation in HEK293 cells

RFP expression in HEK293 cells at 48 hours post co-transfection of pTRIP-JAK2-RFP plasmids with pCMV-VSV and pGag-Pol to generate JAK2-lentiviral particles. Images captured with an EVOS™ FL Auto 2 Imaging System. Scale bar represents 500 µM. RFP; red fluorescent protein. Scale bar represents 500 µM.

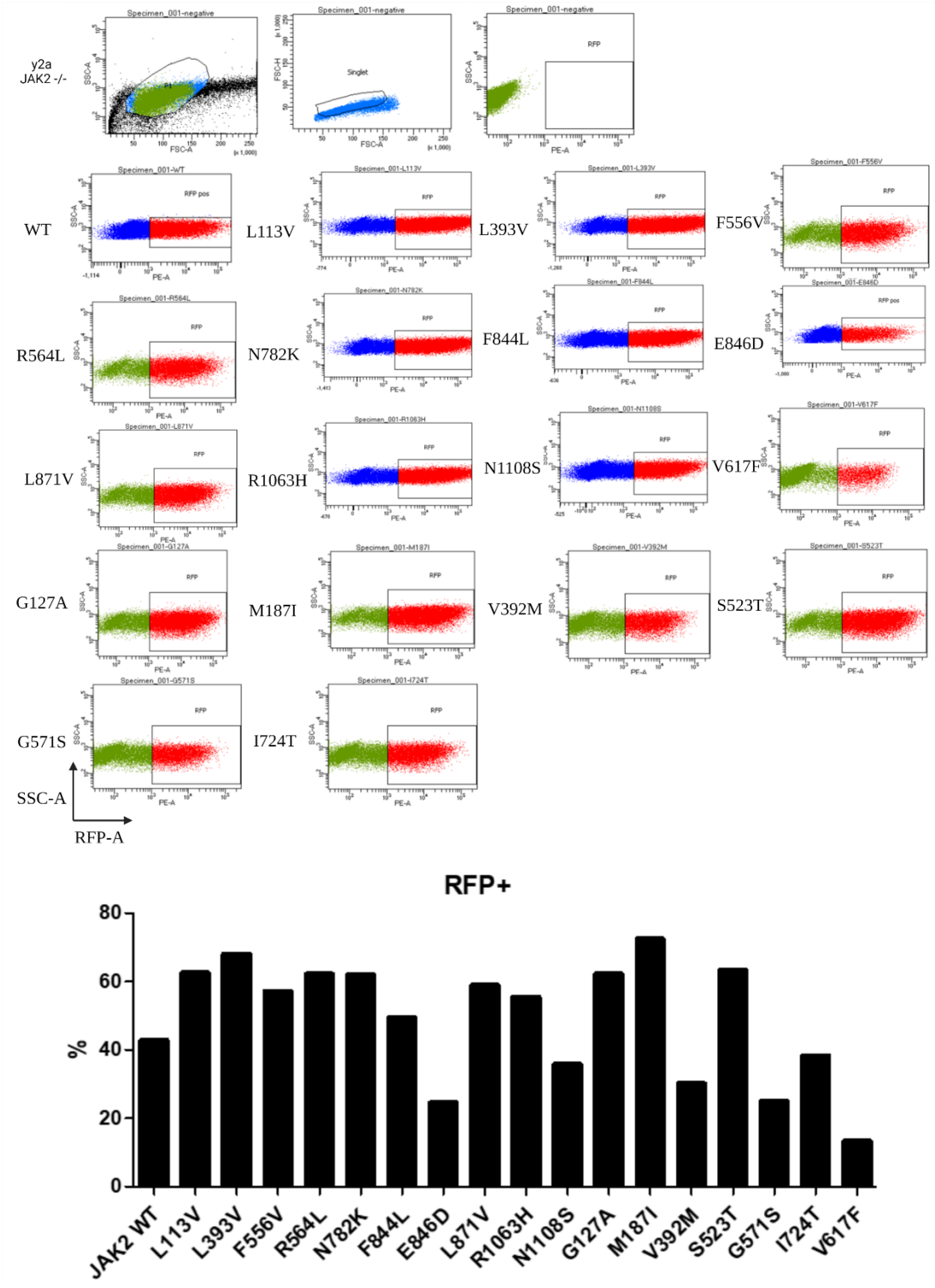


Figure 3.11 Transduction efficiency of JAK2-lentiviral particles into γ2A cells

JAK2-RFP-lentiviral particles were transduced into γ2A cells and RFP was selected for by fluorescence-activated cell sorting (FACS), performed by Dr Parisa Amjadi. The gating strategy shown selected for single cells positive for RFP expression (top row). The graph (bottom) represents the number of RFP+ sorted cells as a percentage of the single cell population., n=1.

3.9.1 Determination of stable JAK2 expression

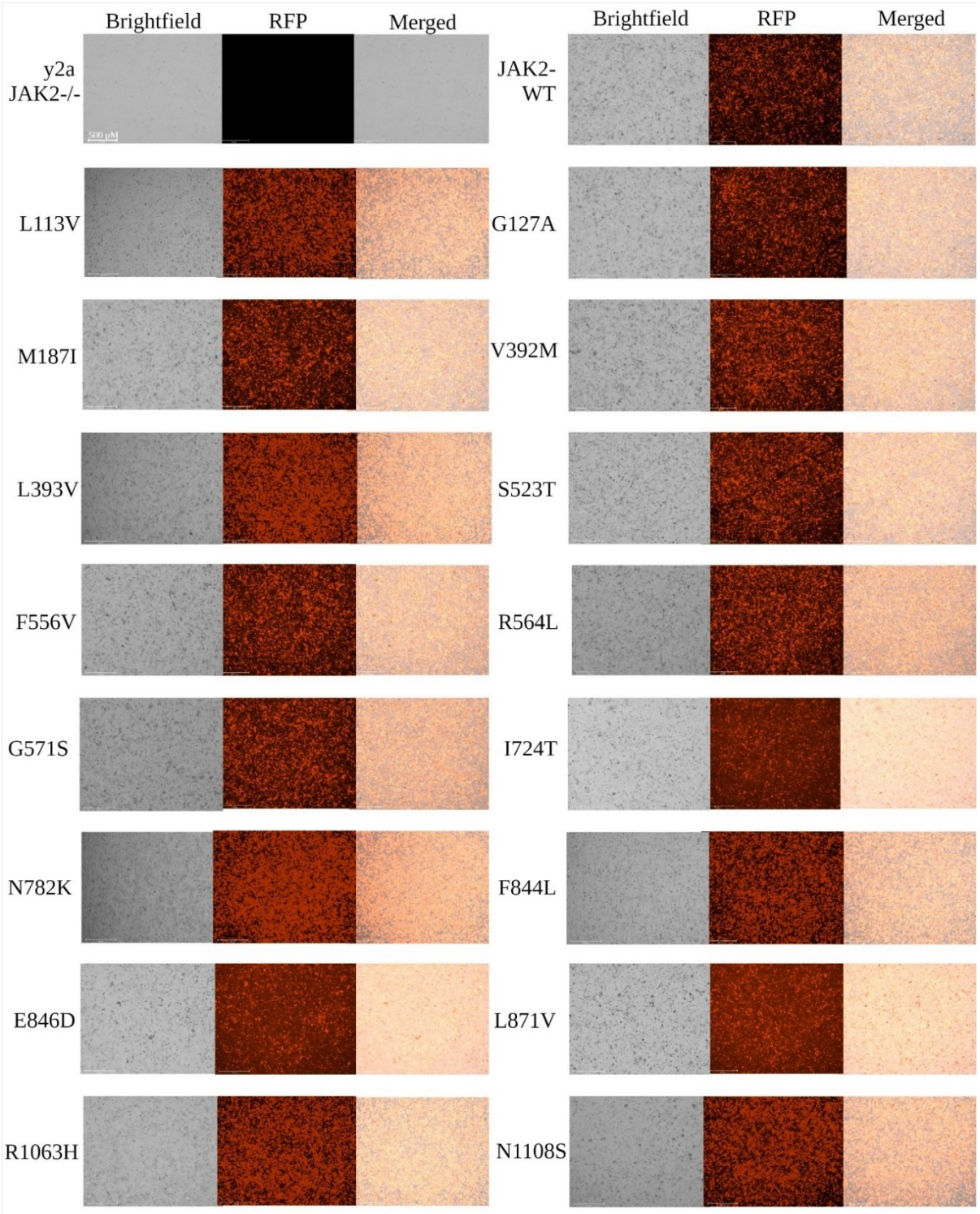


Figure 3.12 Fluorescent microscopy of stable transduced γ2A-JAK2 cells

RFP expression in γ2A cells post JAK2-lentiviral transduction and Fluorescence-activated cell sorting (FACS). Cells were cultured in complete media and incubated at 37°C, 5% CO_2. RFP; red fluorescent protein. Figure shows pseudo-coloured images of brightfield, RFP channel and merged captured with an EVOS™ FL Auto 2 Imaging System, scale bar represents 500 µM, n=1.

All γ2A-JAK2 transduced cells survived recovery, except for those transduced with JAK2 V617F which failed to grow and spontaneously died shortly after transduction and FACS. Post sorting and culture growth, positive RFP expression was captured with an EVOS™ FL Auto 2 Imaging System to determine successful transduction (Figure 3.12). Immunoblotting was performed to determine stable JAK2 expression in stably transduced γ2A-JAK2 cells, using V617F transfected cells as a control. All variants show stable expression of JAK2, though no γ2A-JAK2 transduced cells showed phosphorylated JAK2 or STAT1 (Figure 3.13) despite transiently transfected F556V, E846D, and L871V showing high expression of phosphorylated JAK2 and STAT1 (Figure 3.5). This may be due to the more subtle nature of transduced cells compared to overexpression in transient transfection systems. However, as V617F protein lysates isolated from transiently transfected cells failed to show pJAK2 by immunoblotting (Figure 3.13), this is likely the result of inadequate antibody dilution. Figure 3.12 is representative of one experiment and will be repeated to know this for sure. Sequence confirmation of *JAK2* in the stable γ2A-JAK2 transduced cells will also be performed imminently to successfully confirm integration of target JAK2 variants.

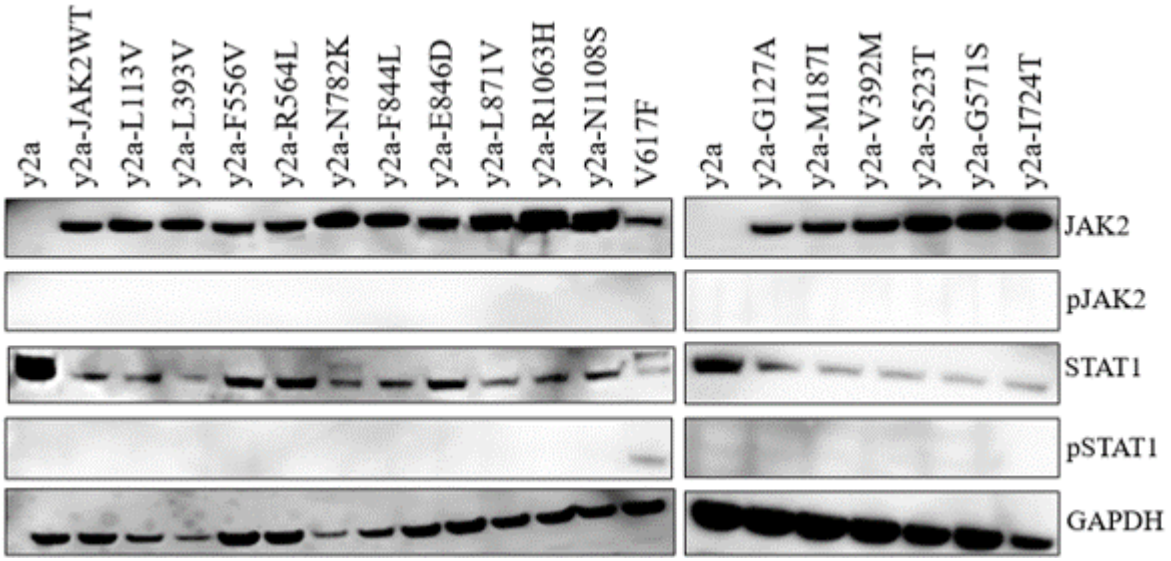


Figure 3.13 Protein expression in JAK2 stably transduced human fibrosarcoma cells

Immunoblots of protein isolated from γ2A-JAK2 transduced cells showing resolved JAK2, pJAK2(Y1007), STAT1, and pSTAT1(Y701), with GAPDH used as a loading control. Immunoblots representative of n=1.

3.10 Impact of *JAK2* on neutrophil activity

To characterise a possible mechanism that could explain an increased susceptibility to or severe outcome of IMD I decided to investigate the role of JAK2 in neutrophil recruitment. Neutrophil recruitment plays a major part in the development of sepsis and are recruited to the endothelial site of inflammation during infection (323). Neutrophil β1 and β2 integrins bind respectively to vascular cell adhesion molecule 1 (VCAM1) and (intercellular Adhesion Molecule 1) ICAM1 typically expressed on endothelial and immune cells (324). Neutrophils expressing V617F have shown increased adhesion to VCAM1 & ICAM1 by constitutive activation of β1 and β2 integrins by enhancing Rap1 activation which signals the β integrins to their open high-affinity conformation (Figure 3.14) (294, 325). I hypothesised that the gain-of-function *JAK2* variants found in IMD cases within the EUCLIDS cohort will modify neutrophil activity, which could cause the development of severe disease in IMD, particularly severe sepsis including disseminated coagulation which can lead to death, amputations and skin grafts. The high representation of severe sepsis and death from sepsis in the cohort of IMD patients with *JAK2* variants and reports of V617F association with venous thrombosis helped to inform this hypothesis (294).

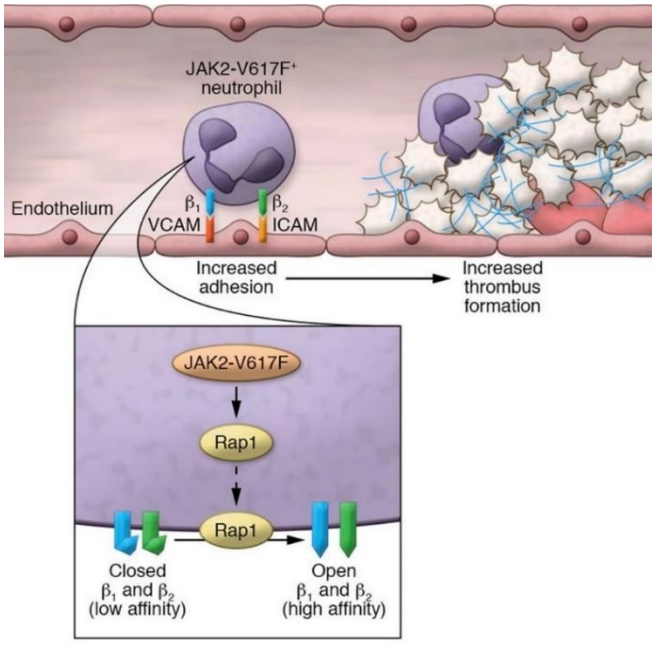


Figure 3.14 JAK2-V617F promotes thrombus formation.

Visual representation of JAK2-V617F impact on neutrophils which promotes thrombus formation. Neutrophils expressing JAK2 V617F have shown increased adhesion to VCAM1 & ICAM1 on the endothelium by constitutive activation of β1 and β2 integrins to their open high-affinity conformation mediated by enhanced Rap1 activation. Figure adapted from Oh. S (2018) (325).

3.10.1 Optimisation of static adhesion assay with control neutrophils

I decided to employ a static adhesion assay to investigate this hypothesis. Briefly, neutrophils were isolated, stimulated, and plated onto a 96-well pre-coated plate with either BSA (control), ICAM1 (3 µg/mL), or VCAM1 (3 µg/mL) and number of adherent cells were measured. I optimised this experiment using control donors and initially stimulated isolated neutrophils with TNF-α (20 ng/mL), shown to activate and prime neutrophils (326-329), LPS, Heat-killed (HK) MenB, or formalin-fixed (FF) MenB, at a MOI 1 for 1 hour. This produced a significant increase in attachment of neutrophils adhered to both ICAM1, VCAM1, and both ICAM1 & VCAM co-coated together upon stimulation with TNF-α, HK-MenB, and FF-MenB, but not LPS when compared to non-stimulated control (Figure 3.15A). I then re-optimised the experiment using increased doses of stimuli, TNF-α (200 ng/mL), HK- and FF-MenB (MOI 10) and adding formalin-fixed *S. aureus, S. pneumoniae,* and *Group A Streptococcus* (GAS) at a dose of MOI 10 (Figure 3.15B). This resulted in neutrophil attachment of TNF-α and MenB stimulated cells at a ten-fold higher than seen at the lower dose (Figure 3.15), with FF *S. aureus, S. pneumoniae,* and *Group A Streptococcus* also showing similar levels of attachment to FF-MenB (Figure 3.15B).

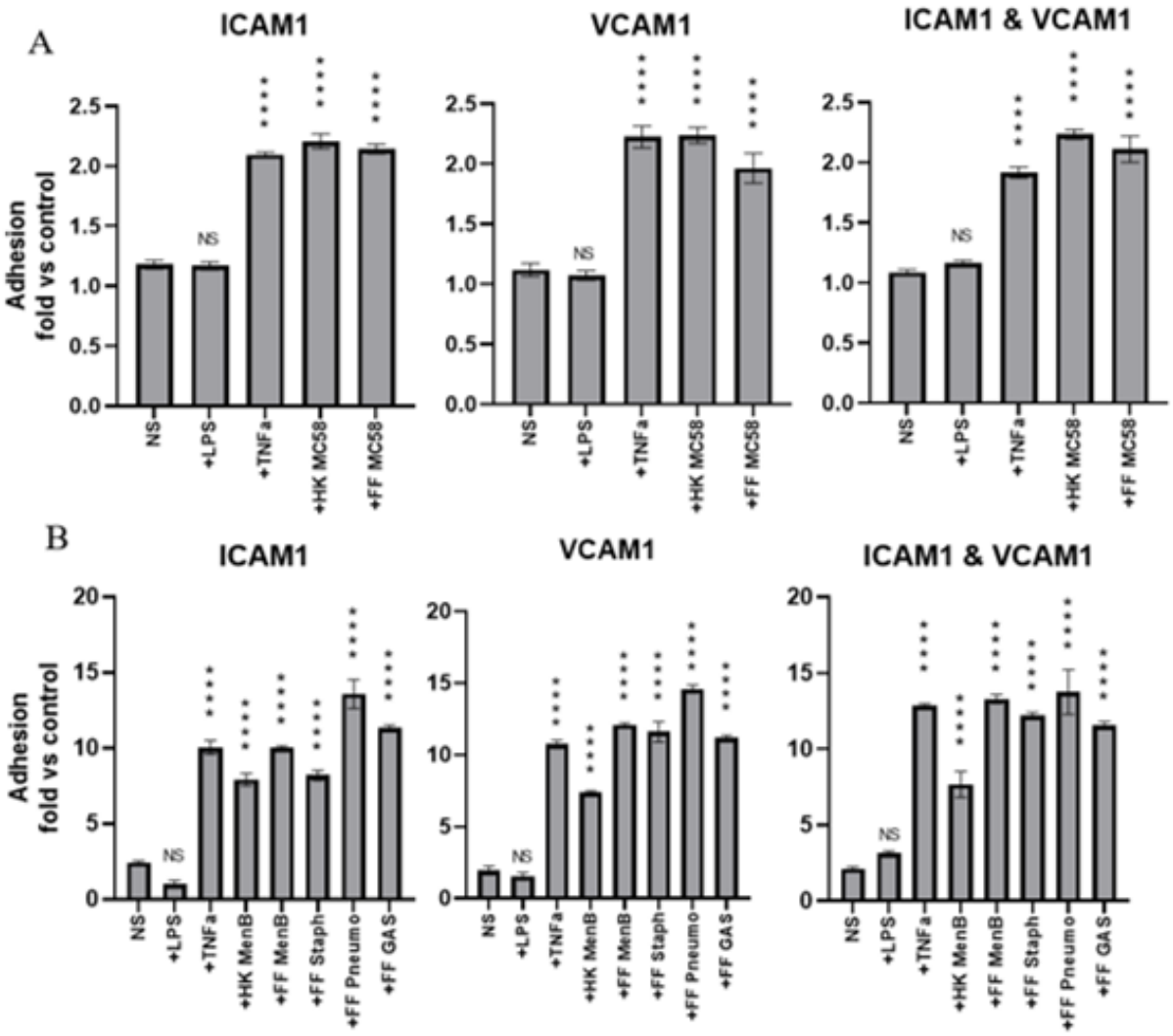


Figure 3.15 Optimisation of the static adhesion assay in control neutrophils

Bar graphs representing number of neutrophils attached to VCAM1 and ICAM1 post-stimulation, compared to control non-stimulated conditions. Adhered cells were measured by microscopy and processed using ImageJ. Neutrophils stimulated with A) mock, LPS (100 ng/mL), TNF-α (20 ng/mL), Heat-killed (HK) *N. meningitidis* serogroup B (MenB), or formalin-fixed (FF) MenB at MOI 1, or B) mock, LPS (100 ng/mL), TNF-α (200 ng/mL), HK-MenB, or FF-MenB, *S. aureus, S. pneumoniae,* or *Group A Streptococcus* (GAS) at MOI 10. Data is representative of n=2 independent experiments with three technical replicates each. Graphs display mean ± SEM as a fold change compared to non-stimulated. Any increased differences were tested with a one-way ANOVA, statistical significance shown with ****p<0.0001.

3.10.2 Static adhesion of stimulated V617F neutrophils

As explained in section 3.7.2 I shall be receiving EUCLIDS patient samples shortly, however the experiment was carried out on *JAK2* V617F blood samples via collaboration with Dr Simone Claudiani (ICL). Because of this opportunity, I was able to conduct my neutrophil static adhesion experiment with a gain-of-function JAK2 sample alongside a control donor. This experiment has only been conducted independently once, though I hope to repeat it, and all the data in this section is representative of one biological repeat with three technical repeats. The experiment was conducted using the optimised conditions shown in Figure 3.15B however during neutrophil isolation of V617F patient blood, there was poor separation of cell layers, leading to a reduced number of collected neutrophils and platelet contamination which can compete with neutrophils for ICAM1 and VCAM1 binding. This is expected of V617F patients as they have thrombocytosis and to minimise this in future experiments, blood will first be centrifuged, and plasma removed to reduced platelet contamination. However due to the reduced number of collected neutrophils, cells number were reduced from an optimised $1.5x10^5$ cells/well to $4x10^4$ cells/well, which can account for the lower fold change values see in the following data compared to optimised experimental results (Figure 3.15B). Also, co-coating with both VCAM1 & ICAM1, as well as stimulating with HK-MenB were removed as investigative variables in order to preserve reagents and conserve neutrophils. These variables were chosen to be excluded as FF-MenB was processed the same as the other bacteria, thereby removing the extra variable and VCAM1 and ICAM1 separately showed similar attachment levels to co-coating so I deemed it unnecessary to have all three conditions.

Figure 3.16 shows that at base level, with no stimulation, V617F neutrophils show significant increased attachment to both VCAM1 and ICAM1 when compared to controls. Upon stimulation with LPS (100 ng/mL), TNF-α (200 ng/mL), FF-MenB, FF-*S. aureus*, FF-*S. pneumoniae*, or FF-GAS at a MOI 10 for 1 hour, neutrophil attachment levels out between V617F and control cells and were similar (Figure 3.17). All stimulation, with the exception of LPS, showed increased V617F and control neutrophil attachment to both VCAM1 and ICAM1 (Figure 3.17).

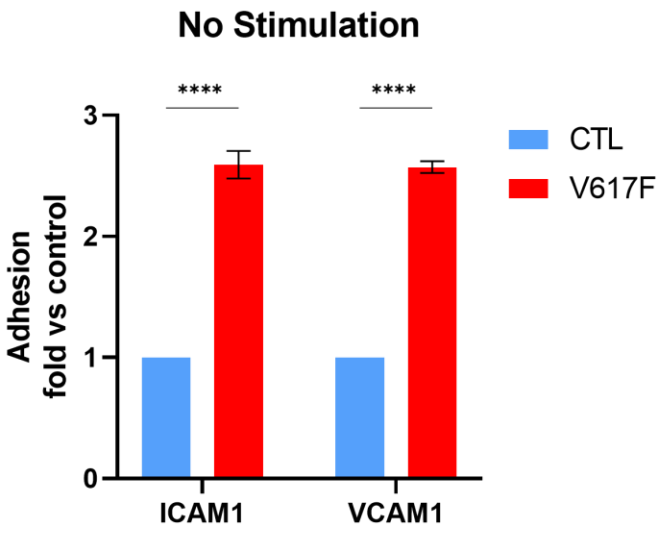


Figure 3.16 Static adhesion assay in control and V617F neutrophils

Bar graph representing number of V617F neutrophils attached to VCAM1 and ICAM1 at baseline with no stimulation compared to control neutrophils. Adhered cells were measured by microscopy and processed using ImageJ. Data is representative of n=1 independent experiment with three technical replicates. Graphs display mean ± SEM as a fold change compared to control. Any increased differences were tested with a one-way ANOVA, and statistical significance shown with ****p<0.0001.

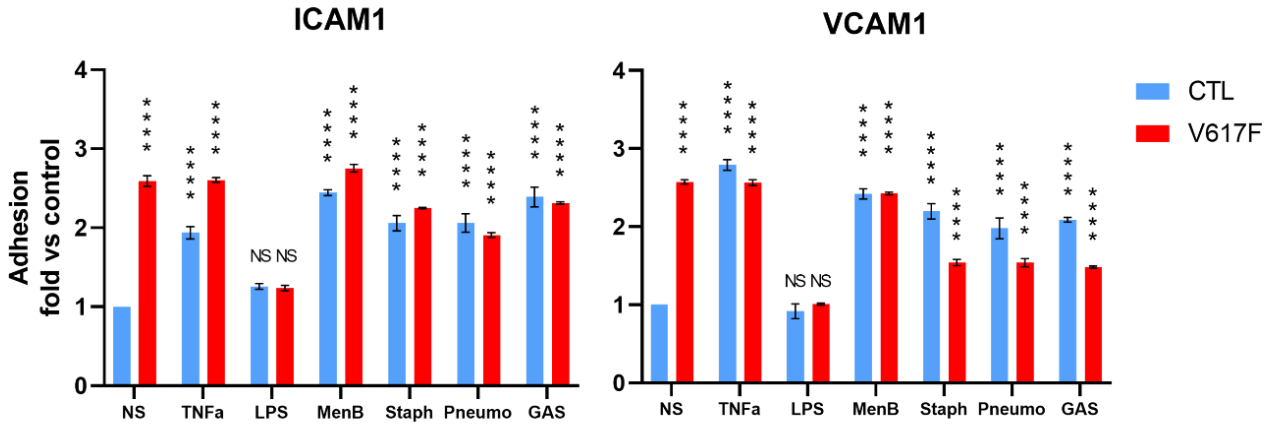


Figure 3.17 Static adhesion assay in control and V617F neutrophils upon stimulation

Bar graph representing number of V617F stimulated neutrophils attached to VCAM1 and ICAM1, compared to control non-stimulated conditions. Stimulated with mock, LPS (100 ng/mL), TNF-α (200 ng/mL), HK-MenB, or FF-MenB, *S. aureus, S. pneumoniae,* or *Group A Streptococcus* (GAS) at MOI 10. Adhered cells were measured by microscopy and processed using ImageJ. Data is representative of n=1 independent experiment with three technical replicates. Graphs display mean ± SEM as a fold change compared to control. Any increased differences were tested with a one-way ANOVA, and statistical significance shown with ****p<0.0001.

3.10.3 Static adhesion of stimulated V617F neutrophils upon Rap1 inhibition

Next, I wanted to test if JAK2 activation and gain-of-function in neutrophil attachment was Rap1-dependent. I investigated this by utilising a cell permeable drug, GGTI-2147 (Sigma), a potent molecule that selectively inhibits geranylgeranyl transferase I (GGTase I), thereby blocking the activation of Rap1. Prior to stimulation and plating, I pre-treated the isolated neutrophils with GGTI-2147 (10 μM). Results showed that upon Rap1 inhibition, neutrophil attachment was significantly impeded in V617F cells compared to controls (Figure 3.17). Significantly inhibited V617F neutrophil attachment to both VCAM1 and ICAM1 was also seen, regardless of stimulation, when compared to control (Figure 3.18). This shows that increased neutrophil attachment to VCAM1 and ICAM1 in JAK2 gain-of-function samples is Rap1 dependent, and by extension caused by JAK2 activation. These data of V617F and control neutrophil attachment, with no stimulation, replicate results previously reported (294), and though needs to be independently repeated shows a robust methodology for investigating *JAK2* gain-of-function in neutrophil activity, that can be repeated with recalled EUCLIDS patients with *JAK2* variants.

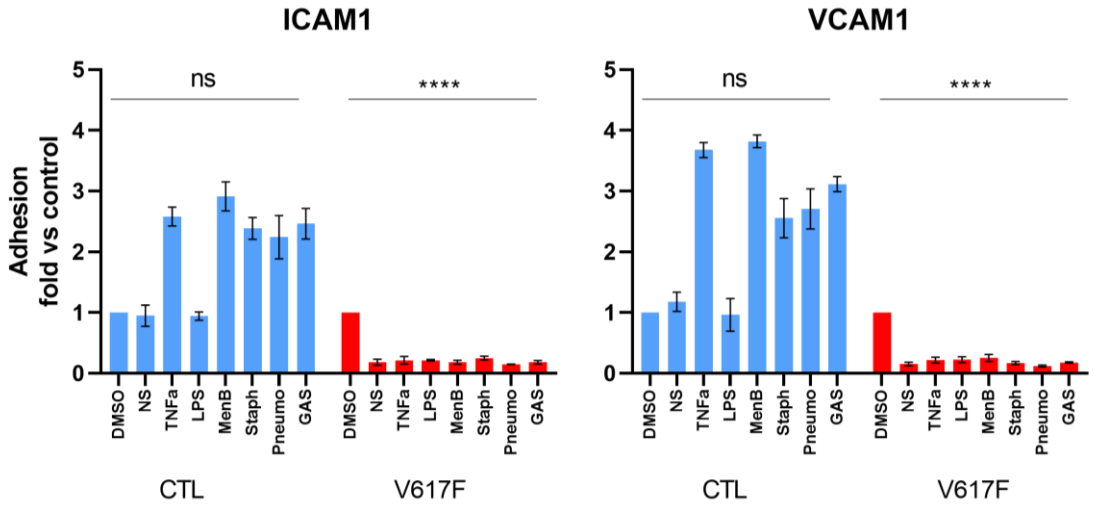


Figure 3.18 Inhibition of V617F neutrophil attachment upon Rap1 inhibition by treatment with GGTI-2147

Bar graph representing number of GGTI-2147 treated V617F stimulated neutrophils attached to VCAM1 and ICAM1, compared to non-stimulated conditions. Stimulated with mock, LPS (100 ng/mL), TNF-α (200 ng/mL), HK-MenB, or FF-MenB, *S. aureus, S. pneumoniae,* or *Group A Streptococcus* (GAS) at MOI 10. Adhered cells were measured by microscopy and processed using ImageJ. Data is representative of n=1 independent experiment with three technical replicates. Graphs display mean ± SEM as a fold change compared to control. Any decreased differences were tested with a one-way ANOVA, and statistical significance shown with ****$p<0.0001$.

3.11 Discussion

JAK2 genetic mutations in myeloproliferative neoplasms and non-communicable human disease has been well reported, however, thus far, no clear associations of *JAK2* variants have been reported to be involved in predisposition to invasive bacterial infection (330). In this thesis, I aimed to investigate the impact of *JAK2* variants on protein expression and function in the context of invasive bacterial infection, with a particular focus on IMD. A gene burden analysis of a large cohort of IMD cases recruited into EUCLIDS, an international collaborative recruitment of paediatric patients presenting with invasive bacterial infections (Figure 3.1), revealed variants in *JAK2* to be significantly enriched in IMD cases when compared to controls, with a p value less than $3x10^{-6}$. Further analysis of the clinical data of EUCLIDs IMD patients with *JAK2* variants showed a high representative of severe manifestations of IMD including sepsis (72.22%), septic shock (16.67%) and extreme clinical outcomes of skin grafts and amputations (16.67%), and death (22.22%) (Table 3.1, Table 3.2).

JAK2 is a multifunctional tyrosine kinase that acts as a critical intermediate between membrane receptors and downstream signalling molecules such as STATs. JAK2 activity is chiefly driven by phosphorylation triggered by membrane receptor bind to a particular cytokine, hormone, or growth factor (Figure 3.2) (299, 300). This binding activates downstream signalling pathways involved in a wide range of different processes such as cell proliferation, cell survival, apoptosis, cell metabolism, cell differentiation, cytokine signalling, neutrophil recruitment, and haematopoiesis (Figure 3.2). Phosphorylation of tyrosine sites is the main driving force of JAK2 activity, in both activating and inhibitory roles. Upon activation JAK2 becomes phosphorylated on as many as 20 tyrosine residues. In addition to the main tyrosine 1007-1008 location, phosphorylation of tyrosine sites 372, 637, 813, 868, 966 and 972 enhance JAK2 activity, whereas phosphorylation of Ser523, and tyrosine sites 119, 221, 570, and 913 negatively regulate JAK2 (Figure 3.3) (313, 331-335).

Adult onset acquired somatic mutations of *JAK2* are known to result in myeloproliferative diseases. One particular variant, p.V617F, is a well-known constitutive activator and is the most clinically relevant variant, found in over 95% of individuals with polycythaemia vera (PV), over 50% of those with essential thrombocythemia (ET), and around 65% of primary myelofibrosis cases (301-303). Due to this, *JAK2* is now one of the first diagnostic markers tested upon diagnosis with a myeloproliferative disorder and JAK inhibitor, ruxolitinib, is an approved therapeutic drug used to treat myelofibrosis (301). In addition to the main V617F

variant identified in the majority of MPN cases there are other variants identified in exon 12 that affect a region adjacent to V617F and are present in polycythaemia vera cases that test negative for *JAK2* V617F (336, 337).

JAK2/STAT signals downstream of multiple hematopoietic cell surface receptors, and in myeloproliferative disorders, *JAK2* somatic variants originating in hematopoietic stem cells and myeloid progenitor cell compartment can constitutively hyperactivate cytokine/ligand-independent JAK-STAT signalling (304, 305). This activation drives clonal expansion and over-proliferation of bone marrow cells of the myeloid lineage, which underlies the clinical manifestations of myeloproliferative neoplasms, with incomplete penetrance, and can lead to complications such as thrombosis, secondary fibrosis, or leukemic transformation (304-307). A higher *JAK2* V617F allele burden has been associated with a higher risk of development of arterial and venous thrombotic events (338). *JAK2* genetic variability associated with MPNs typically occur as heterozygous somatic mutations in hematopoietic stem cells that then gain homozygosity through loss of heterozygosity on chromosome 9p due to mitotic recombination during disease progression (339). Differences in JAK2 V617F allelic burden can also inform on disease manifestations in MPNs, with a lower allele burden seen in ET compared to PV (340).

Myeloproliferative disease-causing variations in *JAK2* are traditionally characterised as adult-onset somatic variants, however, recently reported case studies have also identified germline *JAK2* variants that have complete penetrance, with low risk of significant disease (Table 3.4) (341-345). These germline variants are polyclonal and selectively affects the megakaryocytic lineage and presents with thrombocytosis that typically has a benign clinical course with minimal complications and no risk of disease progression (Table 3.4). Notably, most *JAK2* variants associated with human disease appear to be clustered around the pseudokinase domain (also known as JH2) which is a protein kinase that stretches from amino acid 545 to 809, and negatively regulates cytokine signalling by auto-inhibiting JAK2 kinase activity (Table 3.4) (346, 347). Only one patient with a reported germline *JAK2* variant, E846D, was reported to have a significant bacterial infection episode (Table 3.4). Interestingly, this is the same variant seen in the IMD EUCLIDs cohort and that had been shown to be a GoF variant (Figure 3.5and Figure 3.7). No other reports of significant infection in other *JAK2* germline patients could be due to incomplete penetrance, a lack of reporting, or a lack of pathogen challenge during the vulnerable childhood years.

Table 3.4 Germline *JAK2* variants in disease

JAK2 variant	Disease presentation	Infection episodes?	Ref
G571S	Essential thrombocythemia, but also seen in non-affected family members – assumed to be non-causative	Unknown	(348)
L393V	V617F polycythaemia vera	Unknown	(312)
G335D	Essential thrombocytosis	Unknown	(309)
E846D	Familial erythrocytosis in adults	One case of bacterial endocarditis with pneumococcal meningitis	(349)
E846D/R1063H	Erythrocytosis and splenomegaly with megakaryocytic atypia in an adolescent	Unknown	(308)
V617F	Susceptibility to myeloid neoplasms and essential thrombocytosis	Unknown	(350)
R564Q	Essential thrombocytosis in child siblings and mother	Unknown	(342)
H608N	Essential thrombocythemia in adults	Unknown	(344)
L611S	Chronic thrombocytosis in a 3-month-old child	Unknown	(341)
V617I	Thrombocytosis with vascular events	Unknown	(343, 351)
V625F	Essential thrombocytosis	Unknown	(309)
R867Q	Thrombocytosis in adults	Unknown	(345)
S755R/R938Q	Thrombocytosis in adolescents and adults	Unknown	(345)
T875N	Essential thrombocytosis in an adult	Unknown	(352)
Haplotype GGCC_46/1[a]	Predisposition to the development of V617F myeloproliferative neoplasms	Unknown	(353, 354)

[a] rs3780367, rs10974944, rs12343867, rs1159782

All *JAK2* variants discovered within EUCLIDS patients are assumed to be germline due to their presence in young children. Ideally, confirmation of germline variants would be performed by sequencing DNA from cheek swab and/or hair follicles, and in circumstances where patients are recalled this will be completed. However, from existing WES data one can estimate the likelihood of a non-germline by assessing the allelic fraction of the variant. Allelic fraction refers to the percentage of a sample represented by an allele and is determined by the number of times a variant base is observed, divided by the total number of times any base is observed at the locus (355). Typically, germline variants have an allele fraction around 50% and 100% for heterozygous and homozygous variants, respectively, but an allelic fraction range of 30-70% is typically accepted to characterise germline heterozygous variants (356-358). All *JAK2* variants discovered in the EUCLIDS cohort falls within the allelic fraction range for germline heterozygous variants.

Functional assessment of 16 JAK2 variants by phosphorylated JAK2 and STAT1 immunoblotting revealed 3 variants to be GoF, F556V, E846D, and L871V (Figure 3.5). F556V, and E846D have both been previously reported to be GoF (Table 3.3), in the context of cell proliferation and EPO, MPL, and G-CSF stimulated activation of STAT5, as most *JAK2* published data is in relevance to cell proliferation associated with MPNs (308, 310). *JAK2* L871V is a novel variant, not previously described. Two other JAK2 variants identified in the EUCLIDs cohort, L393V, and R1063H have also been previously prescribed as gain-of-function however, upon closer inspection of the literature there variants were seen in co-operation with other more dominant activating variants such as V617F and E846D and are unlikely to be independently GoF variants but rather augment the JAK2 activity of other GoF variants (Table 3.3) (308, 311, 312). JAK2 F556V and E846D showed activation of JAK2 and STAT1 upon IFN-γ stimulation to a higher degree compared to WT, while L871V showed similar levels to WT (Figure 3.6). Similar results were also seen with a flow cytometry output investigating expression of pSTAT1 (Figure 3.7) and which revealed that JAK2 F556V and E846D variants were hyper-responsive to cytokine signalling, which could contribute to disease pathogenesis.

Phosphorylated proteins are difficult to work with experimentally, as they are sensitive to time and temperature, and many of my experiments had to be performed diligently and continuously with no optional experimental stops. Because of this, along with transient transfection not reaching 100% transfection efficiency, I decided to create stable cell lines of all *JAK2* variants via lentiviral transduction. Stable JAK2 expression in γ2A cells was shown by

immunofluorescence (Figure 3.12), and immunoblot (Figure 3.13), These JAK2 stable expressing cells will be used in future experiments to investigate the impact of variants on JAK2 expression, cytokine signalling, and function.

In preparation of JAK2 patient samples, from a patient harbouring the activating E846D variant two main experiments were optimised. As immunoblotting requires a lot of protein, and therefore a high volume of blood, for the detection of phosphorylated proteins, I decided to pursue a flow cytometry output experiment that uses a lower starting volume of blood while remaining sensitive to detection of phosphorylated proteins. I optimised the effect of IFN stimulation and live bacterial infection on activated STAT1 (pSTAT1) in control PBMCs and showed that IFN-α and IFN-γ stimulation resulted in a significant increase phosphorylated STAT1 in the CD14+ monocyte population, which is part of the myeloid lineage, when compared to non-stimulated cells (Figure 3.8). No difference in pSTAT1 was observed in any condition of the CD4+ T cell population (Figure 3.8). Previously published studies have reported an increase in CD4+ pSTAT1 upon IFN-α stimulation but used 10-100-fold higher concentrations of IFN-α than what was used in my experiment (359). JAK2 V617F experiment has only been performed once and will need to be repeated but pSTAT1 was observed to be constitutively active with no stimulation compared to control in the CD14+ cell population (Figure 3.9), this is expected as acquisition of V617F is a myeloid lineage-specific event (360). Upon live bacterial infection, IFN-α, and IFN-γ stimulation V617F and control cells show similar levels of pSTAT1 expression in CD14+ cells but there appears to be an inhibition of pSTAT1 expression in CD4+V617F cells live infected with *N. meningitidis* and *S. pneumoniae*, this is an unusual result as V617F mutations are typically absent from T cell populations (360) and I would therefore expect data from this cell population to be similar to control cells. The experiment will need to be repeated to determine if this is a real consequence (Figure 3.9).

The final experiment sought to characterise the potential mechanism in which GoF *JAK2* variants exacerbate invasive bacterial infections. Neutrophils are polymorphonuclear leukocytes (PMNs) that are recruited to infection sites, and can secrete large amounts of pro-inflammatory cytokines, which can, among other interactions, go on to trigger the JAK-STAT pathway and activate the inflammatory pathway (95). Studies have shown that the JAK2 GoF V617F mutation contributes to increased neutrophil adhesion (294, 361) which may account for the severity in the IMD patients with JAK2 variants as prolonged or excessive neutrophil activity has been documented to lead to cellular and tissue damage of the endothelium and can result in sepsis and thrombosis (323, 362, 363). In addition to the inflammatory cytokine

signalling pathway, increased adhesion to VCAM1 and ICAM1, ligands for β1 and β2 integrin, respectively has been observed with granulocytes isolated from mice knocked in with JAK2-V617F (294). JAK2-V617F expression was demonstrated to shift β integrins from a closed, low-affinity confirmation to an open, high-affinity conformation, via Rap-1 dependent translocation signalling (294). This research also showed that increased endothelial adhesion and β1/β2 integrin activation in a JAK2-V617F mice inferior vena cava ligation model promoted venous thrombosis (294). Furthermore, they found increased granulocyte binding to VCAM1 and increased Rap1 activation in V617F MPN patients compared to healthy controls (361). A static adhesion assay was successfully optimised and showed significantly increased neutrophil binding to VCAM1 and ICAM1 upon stimulation with TNF-α and formalin fixed *N. meningitidis* (MenB), *S. aureus, S. pneumoniae,* or *Group A Streptococcus* (GAS) (Figure 3.13). This is expected as it is indicative of the known mechanism of neutrophil recruitment that occurs during bacterial infection, which sees neutrophils recruited from the bone marrow to sites of infection where they play a crucial role in the inflammatory pathway to help clear infection (364-366).

Performing the static adhesion assay with V617F patient samples, showed significantly increased attachment of neutrophils to both VCAM1 and ICAM1 at baseline, with no stimulation, compared to controls (Figure 3.16), replicating published work (294, 361). Upon stimulation with formalin fixed bacteria and TNF-α, neutrophil attachment did not differ between V617F and control cells (Figure 3.17). The absence of neutrophil attachment upon LPS stimulation is interesting as circulating LPS is known to induce neutrophil chemotaxis in inflammation via the p38-membrane receptor pathway (367), however our experiments show that LPS is not responsible for inducing binding of β-integrins to VCAM1 and ICAM1 (Figure 3.15 and Figure 3.17). Consistent with previous reports , neutrophil attachment is significantly inhibited when cells are treated with GGTI-2147, a drug that inhibits Rap1 activation and transmembrane localisation, when compared to non-treated and control cells (Figure 3.18). Upon stimulation with all bacteria and TNF-α, GGTI-2147 pre-treated V617F cells show inhibited neutrophil attachment compared to pre-treated control cells which showed adhesion levels similar to non-treated (Figure 3.18). This indicates that JAK2 V617F neutrophils bind to VCAM1 and ICAM1 predominantly via a Rap-1 dependent pathway, whereas JAK2 WT neutrophils may potentially use other pathways independent of Rap1, such as Rho PAR-2 mediated signalling (368). This experiment will be repeated with V617F cells to confirm these findings and with patient E846D cells which will be donated by a EUCLIDs IMD patient in

the process of being recalled. Investigating JAK2-dependent Rap1 activation can also be further explored in future experiments by immunoblotting for Rap1-GTP, the activated form of Rap1, and investigating localisation of Rap1 in cells possessing *JAK2* variants with immunofluorescence.

Thus far, 3 of the 10 JAK2 variants found exclusively in the IMD EUCLIDS cohort were found to be GoF, F556V, E846D, and L871V. I hypothesise that GoF variants in *JAK2* exacerbate IMD outcomes by increasing neutrophil adherence to its endothelial binding partners, VCAM1 and ICAM1. This increased adhesion could lead to cellular and tissue damage of the endothelium and can promote sepsis via hyperactivation of the coagulation pathway (323, 362, 363). If proved, this could result in use of *JAK2* inhibitors, such as ruxolitinib an already approved therapeutic drug used to treat myelofibrosis, to treat patients with meningococcal infection to prevent the development of severe sepsis. Interestingly, MPN patients with GoF *JAK2* variants are not known to be at risk for IMD. As *JAK2* GoF variants in MPN are typically somatic mutations acquired in adulthood, this lack of infectious disease risk could be due to incomplete penetrance, a lack of reporting, or a lack of pathogen challenge during the vulnerable childhood years. Further functional work, detailed below, will need to be done to further characterise GoF JAK2 variants associated with IMD.

This work is and will be continued, with immediate efforts put into confirming stably transduced γ2A-JAK2cells, making sure all experiments have statistical significance with at least three independent biological repeats, and conducting experiments with new blood samples from patients possessing *JAK2* variants. Another imminent goal is to investigate phosphorylated STAT1 expression with all *JAK2* variant expressions upon stimulation, using flow cytometry, as it will be informative in revealing if additional *JAK2* variants have GoF properties during cytokine signalling. Some future experiments also in progress includes optimising the static adhesion assay to utilise PLB-985 cells, pluripotent hematopoietic cells that can be differentiated to neutrophil-like cells, as these cells can be genetically manipulated to express the *JAK2* variants of patients we are not able to recall (369). However, this will take significant optimising as I ran a pilot experiment that showed no specific binding of PLB-985s to either VCAM1 or ICAM1 regardless of stimulation condition (data not shown), which is not in accordance with what is seen with real neutrophils (Figure 3.13). I suspect that the PLB-985s need to be differentiated first, which runs into significant experimental difficulties as PLB-985 differentiation causes significant cell death (370). I, alongside a colleague in the lab, are also beginning to investigate the impact of orthologs of the *JAK2* variants identified in the

EUCLIDs cohort in an established zebrafish inflammation model in collaboration with Dr Milka Sarris (University of Cambridge) (371). The premise of this pilot work is to prepare *JAK2* zebrafish ortholog (jak2a) RNA that will be used for transgenesis into zebrafish embryos whose lineage have undergone transgenic modification to express fluorescently tagged neutrophils (371, 372). The larvae expressing the jak2a ortholog variants can then be challenged with molecules such as LPS and/or bacteria and real-time imaging of microbe-leukocyte interactions can be performed, including measurement of neutrophil recruitment to the challenge site. I aligned human JAK2 and zebrafish jak2a and identified six of the *JAK2* variants in EUCLIDs patients have orthologs in zebrafish jak2a, JAK2 L113V, L393V, F556V, R564L, F844L, and L871V to jak2a L90V, L359V, F520V, R528L, F807L, and L834V respectively. I created these jak2a orthologs using site directed mutagenesis (data not shown) and my colleague is currently subcloning the jak2a ortholog inserts into vectors to be used for transgenesis. This is an exciting opportunity to investigate the impact of *JAK2* ortholog variants *in vivo*.

This chapter has described the work I have undertaken to investigate the impact of *JAK2* variants on protein expression and function and attempt to characterise the potential mechanism used in the context of bacterial infection. Finalising the data produced so far will provide the first instance of genetic variability in *JAK2* being associated with severe outcome in invasive bacterial disease.

Chapter 4

Chapter 4 Impact of *KRT10* variants in invasive bacterial disease

4.1 Background

Nm is a leading cause of bacterial septicaemia and meningitis worldwide, but there are other invasive bacterial infections that can cause significant health impact, including *Staphylococcus aureus* (*S. aureus*/Staph), *Streptococcus pneumoniae* (*S. pneumoniae*/Pneumo), and *Group A Streptococcus* (GAS). *S. pneumoniae* infection is the leading worldwide cause of bacterial pneumonia, accounting for 27 % of community-acquired pneumonia (227) and *S. aureus* invasive infections can result in severe clinical manifestations such as like septicaemia, septic arthritis, pneumonia, and endocarditis with a case fatality rate of ~25 % (208, 209). Mild GAS-pharyngitis is recorded in ~700 million cases/year but invasive GAS disease resulting in streptococcal toxic shock syndrome (STSS) and necrotizing fasciitis are responsible for 500, 000 deaths annually with an incidence rate of ~3 cases per 100,000 people per year in Europe (246-248). Bacterial infection contributes to ~20 % of childhood mortality worldwide and *N. meningitidis, S. aureus, S. pneumoniae,* and *GAS* infections are some of the main causes of most invasive bacterial disease worldwide (373, 374).

The FP7 EU-funded 5-year study, EUCLIDS, described previously in section 3.2, recruited cases of childhood bacterial infections that resulted in severe disease such as meningitis, septicaemia, and focal infections. The bacterial infections recruited into EUCLIDs included meningococcal, pneumococcal, streptococcal, staphylococcal, *E.coli*, and *Salmonella* infections (Figure 3.1). Recruitment criteria included children under the age of 18 with severe or unusual illness with no known predisposing risk factors. A gene burden analysis was performed by Dr Evangelos Bellos (as described previously in section 3.3) of rare (MAF <1%) non-synonymous variants investigating enrichment of variants in genes in the severe invasive bacterial disease patients from the EUCLIDS WES cohort vs controls. Keratin, type I cytoskeletal 10 (*KRT10*) was found significantly enriched in all invasive bacterial diseases, with a significant p value $<3x10^{-6}$.

KRT10 is a member of the type I cytokeratin family belonging to the intermediate filament protein superfamily and is widely expressed in different tissues but is the most highly expressed keratin in skin, particularly suprabasal keratinocytes (375, 376). KRT10 typically dimerises with KRT1, via a disulphide bond at KRT10 amino acid site C401, and two heterodimers of KRT1 and KRT10 form a heterotetrametric complex that continues to further assemble and

composes part of the epithelial cell cytoskeletons (375). In addition to structural support, KRT10 also has a critical role in keratinocyte differentiation, growth, and proliferation. Genetic variants that interfere with these roles can contribute to clinical manifestations of disease (377). Autosomal dominant genetic variations in *KRT10* are well-known to be associated with epidermolytic hyperkeratosis and ichthyosis with confetti, skin disorders that are typically present at birth or shortly after and can be clinically characterised by the presence of very red skin (erythroderma) and severe blisters (12, 13).

I initiated the work of this project including preliminary literature reviews, collection of clinical data, and construction of *KRT10* variants in a mammalian expression vector. Subsequent experimental work was performed by two Master of Sciences students, Ms Afnan Alahmadi and Ms Laia Rigat Nogareda who I attentively supervised during my project. Their experimental work was decided by discussion and performed under my direction. I shall make it transparent during this chapter what work was performed by who.

4.2 Discovery of heterozygous *KRT10* variants in unrelated paediatric cases of invasive bacterial disease

A total of 17 heterozygous variants, 12 missense and 5 in frame indel variants, were identified across all bacterial disease groups within the EUCLIDS cohort (Table 4.1). None of these variants have been reported in epidermolytic hyperkeratosis, ichthyosis with confetti, or in association with any human disease. *KRT10* harbours a Gly-Ser rich repeat region towards its C-terminal tail and several polymorphisms have been reported in this repeat region, according to gnomAD. However, due to the difficulties in sequencing Gly-Ser rich regions, sequencing artifacts masquerading as polymorphisms are common and expected, and should be confirmed by sanger sequencing. Online resources report high sequence conflict in this tail region, meaning differences between the default canonical sequence and different independent sequence submissions have been reported. Because of this, and the higher MAF and prevalence of some *KRT10* variants found in this cohort, 8 of 17 variants were initially excluded from further investigations (Table 4.1). The remaining 9 heterozygous missense *KRT10* variants, G575V, Y499H, K344Q, G124V, F102S, R86H, S53N, S42F, and R4P were taken forward for functional analysis.

Table 4.1 Summary of *KRT10* variants discovered in EUCLIDS

KRT10 amino acid variant	dbSNP ID	MAF (gnomAD)	Reported in patients with EH or IC	Present in patients from which EUCLIDS bacterial disease group	Reason for variant not further investigated
G575V	rs199785115	0.001074	No	2x Men, 1x Pneumo	
G548_S52dup	rs752563839	0.001796	No	2x Men	Potential sequencing artifact
S547Y	rs752817112	0.010692	No	40x evenly spread across disease groups	Potential sequencing artifact and high MAF
Y542S	rs780363296	0.001084	No	40x evenly spread across disease groups	Potential sequencing artifact
G541S	rs773883770	0.000852	No	40x evenly spread across disease groups	Potential sequencing artifact
H522_G529del	rs775424605	0.000218	No	39x evenly spread across disease groups	Potential sequencing artifact
G516_G524dup	rs769510355	0.018886	No	43x evenly spread across disease groups	Potential sequencing artifact and high MAF
S508_G512del	rs570853437	0.020752	No	26x evenly spread across disease groups	Potential sequencing artifact and high MAF
Y499H	rs571434577	0.000798	No	1x Pneumo	
H487_G498del	rs770927933	0.000969	No	1x Men	Potential sequencing artifact
K344Q	rs371899202	0.00000199	No	1x Pneumo	
G124V	rs200060640	0.000108	No	1x Men	
F102S	rs114467326	0.000672	No	1x Staph	
R86H	rs117610737	0.001496	No	2x Pneumo, 3x Men, 1x Staph	
S53N	rs146835683	0.000259	No	1x Pneumo, 1x Men	
S42F	rs142050024	0.000395	No	1x GAS	
R4P	rs142158041	0.000223	No	1x Staph	

EH; epidermolytic hyperkeratosis, IC; ichthyosis with confetti, GAS; *Group A Streptococcus,* MAF; Minor allele frequency

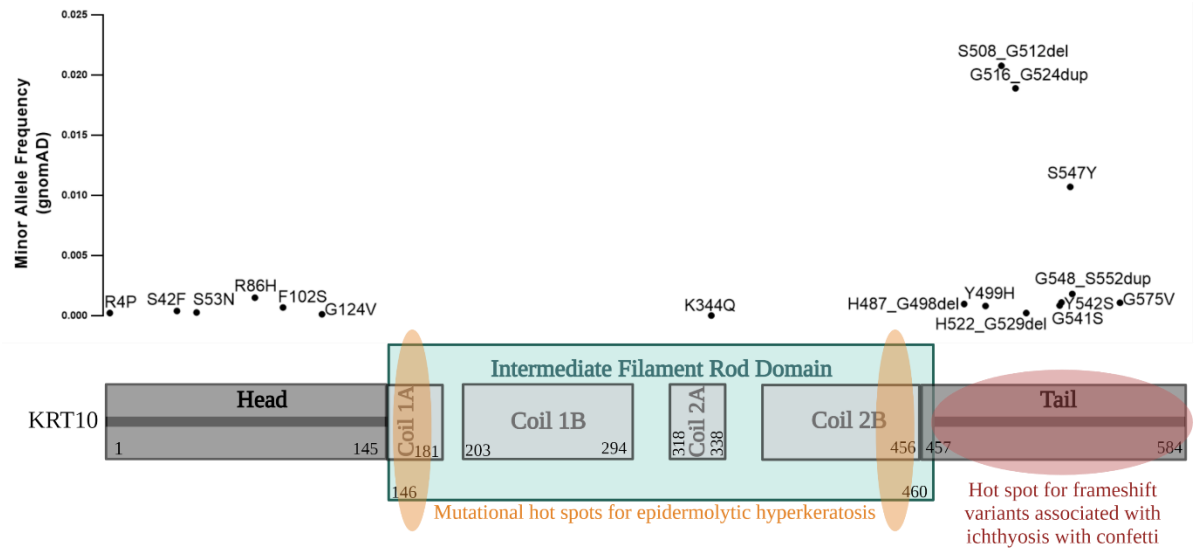


Figure 4.1 Lolliplot of *KRT10* variants and corresponding MAF

KRT10 amino acid variants discovered in EUCLIDS patients plotted against their reported global minor allele frequency (MAF) according to gnomAD, and along a schematic of *KRT10* protein with significant domains and regions. Key mutational hot spots where variants associated with epidermolyitc hyperkeratosis and ichthyosis with confetti are also marked. Image created using biorender.com.

4.3 Clinical presentation of invasive bacterial disease patients with *KRT10* variants

I generated both an overall summary of demographic and clinical data seen within EUCLIDS compared to the cases with *KRT10* variants (Table 4.2), and a more detailed table containing individual data of all EUCLIDS patients with *KRT10* variants (Table 4.3). Sex and age demographics in the cohorts with *KRT10* variants were similar to the overall EUCLIDS cohort, though median age was slightly lower in *KRT10* patients (Table 4.2). 9/17 *KRT10* patients presented with sepsis (52.94%), 4/17 pneumonia with pleural effusion/empyema (23.53%), and 3/17 with severe meningo/encephalitis disease (17.65%) (Table 4.2, Table 4.3). There were zero deaths in the *KRT10* patients, though the rate of PICU admission (82.35%) was almost double that of the overall EUCLIDS cohort (47.90%), indicating a more severe clinical presentation (Table 4.2, Table 4.3).

Table 4.2 Demographic and clinical data summary of EUCLIDS patients, and all those with *KRT10* variants

	EUCLIDS All *n=453*	**EUCLIDS Men** *n=144*	**EUCLIDS Staph** *n=107*	**EUCLIDS Pneumo** *n=101*	**EUCLIDS GAS** *n=89*	**EUCLIDS with all *KRT10* variants** *n=17*
Demographic characteristics						
Female	212 *46.80%*	68 *47.22%*	47 *43.93%*	45 *44.55%*	49 *55.06%*	7 *41.18%*
Male	231 *50.99%*	70 *48.61%*	60 *56.07%*	56 *55.45%*	40 *44.94%*	10 *58.82%*
Median age, months	39	38	109	23.5	43.5	21
Clinical syndrome						
Sepsis	258 *56.96%*	70 *48.61%*	62 *57.94%*	52 *51.48%*	64 *71.91%*	9 *52.94%*
Pneumonia pleural effusion/empyema	*51* *11.26%*	*0* *0.00%*	3 *2.80%*	30 *29.70%*	16 *17.98%*	4 *23.53%*
Severe meningitis	98 *21.63%*	30 *20.83%*	5 *4.67%*	52 *51.48%*	10 *11.24%*	3 *17.65%*
Clinical data						
PICU admission	217 *47.90%*	46 *31.94%*	37 *34.58%*	69 *68.32%*	64 *71.91%*	14 *82.35%*
Death	18 *3.97%*	8 *5.56%*	3 *2.80%*	7 *6.93%*	0 *0.00%*	0 *0.00%*
Skin graft/ amputation	47 *10.38%*	31 *21.53%*	2 *1.87%*	0 *0.00%*	13 *14.61%*	2 *11.76%*

Table 4.3 Demographic and clinical details of EUCLIDS patients with *KRT10* variants functionally tested in this thesis

Patient	Sex	Ethnicity	Age of Onset (months)	Causative pathogen	Significant severe phenotype	PICU Admission	Severe clinical outcomes	Identified *KRT10* mutation
305-1027	Female	Caucasian	115	*N. meningitidis*	Meningo/encephalitis	Yes		G575V
401-1016	Male	Caucasian	9	*S. pneumoniae*	Severe sepsis	Yes	Hearing loss	G575V
401552	Male	Caucasian	40	*N. meningitidis*	Severe sepsis	Yes		G575V
201-1136	Male	Caucasian	2	*S. pneumoniae*	Severe sepsis	Yes		Y499H
101-0090	Male	Mixed	65	*S. pneumoniae*	Severe sepsis	No		K344Q
308-1023	Female	Caucasian	4	*N. meningitidis*	Severe sepsis	Yes		G124V
102-1295	Male	Mixed	6	*S. aureus*	Severe sepsis	Yes		F102S
102-1250	Male	Caucasian	9	*S. pneumoniae*	Meningo/encephalitis, LRTI, subdural empyema, cerebral abscess	Yes	Severe overall disability	R86H
311-1015	Female	Caucasian	41	*N. meningitidis*	Severe sepsis	Yes		R86H
102-1535	Male	Caucasian	152	*S. pneumoniae*	Pneumonia pleural effusion	Yes	Moderate overall disability	R86H
204-1005	Male	Caucasian	1	*S. aureus*	Osteomyelitis	Yes		R86H
437222	Female	Caucasian	204	*N. meningitidis*	Severe sepsis	Yes	Skin graft	R86H
440302	Female	Caucasian	7	*N. meningitidis*	Severe sepsis	Yes	Skin graft	R86H
305-1016	Female	Caucasian	3	*S. pneumoniae*	Pneumonia with empyema	Yes		S53N
101-1623	Female	Caucasian	Unknown	*N. meningitidis*	Meningo/encephalitis	Yes		S53N
107-1074	Male	Caucasian	33	GAS	Pleural empyema, septic arthritis	Yes		S42F
201-1072	Male	Caucasian	74	*S. aureus*	Osteomyelitis	No		R4P

LRTI; lower respiratory tract infection

4.4 Construction of *KRT10* variants in mammalian expression vector

I constructed nine missense *KRT10* variants G575V, Y499H, K344Q, G124V, F102S, R86H, S53N, S42F, and R4P. *KRT10* variants were introduced into the pCMV6-KRT10 WT plasmid by site directed mutagenesis and confirmed by sanger sequencing (detailed protocol described in section 2.4 DNA editing) (Figure 4.2). Whole plasmid sequencing confirmed no second-site variants were introduced during the mutagenesis process. The Glycine-Serine rich C-terminal tail region proved to be a challenge as this was a highly mutagenic area and many clones would procure large indels during the mutagenesis process. This meant that large numbers of clones had to be picked and processed to locate successful clones with an absence of second-site variants.

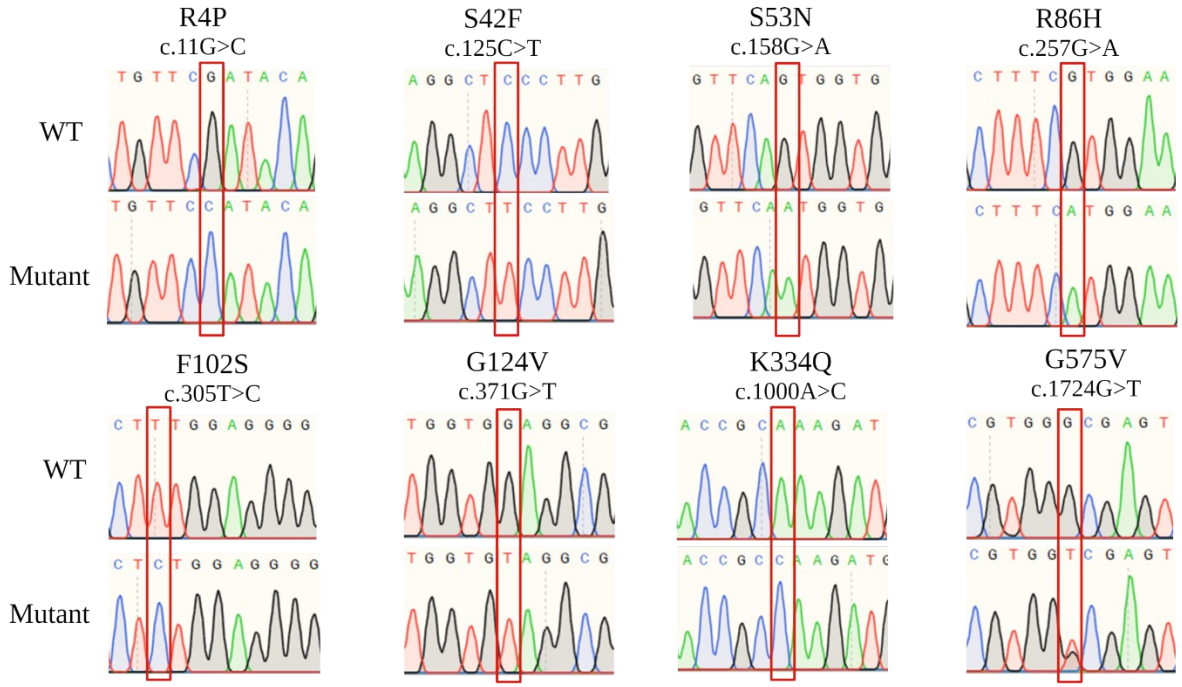


Figure 4.2 Sequence confirmation of site directed mutagenesis in *KRT10* WT

Sanger sequence confirmation of successfully integrated *KRT10* variants in pCMV6-KRT10 WT plasmid. Alignment of pCMV6-KRT10 WT and mutant plasmids, visualised using SnapGene (by Dotmatics). Nucleotide changes are marked with a red box.

4.5 Endogenous expression of *KRT10* in keratinocytes

Keratinocytes are particularly relevant for *S. aureus* infections as *S. aureus* internalises itself within human keratinocytes during live infection as a method of immune evasion (378). To make use of keratinocytes in future experiments, endogenous KRT10 expression had to be assessed first. In undifferentiated HaCaT cells, no endogenous KRT10 was detected, which was expected as KRT10 is a differentiation marker signalling a mature human epidermis layer (Figure 4.3). To explore the full expression kinetics undergone by KRT10 during differentiating, HaCaT cells were differentiated by confluency, as described previously (379) and proteins were isolated over a period of 12 days and stored at -80°C, by Ms Laia Rigat Nogareda. A subsequent immunoblot was undertaken in which I analysed the densitometry and this experiment revealed an increase in KRT10 expression with increasing time in an almost linear relationship after day 4 with expression of KRT10 after 12 days of differentiating ~45 fold higher than that seen as day 0, with day 0 representing HaCaT cells prior to seeding (Figure 4.3).

4.6 Protein expression of *KRT10* variants in epithelial cells and keratinocytes

To identify the impact of *KRT10* variants on protein expression, HEK293 and HaCaT cells were transiently transfected with the *KRT10* variants and immunoblotting was conducted by Ms Laia Rigat Nogareda and Ms Afnan Alahmadi. Proteins were probed with primary antibodies targeting both an epitope on the N-terminus of KRT10 and FLAG, which is tagged on the C-terminus of KRT10. I performed the densitometry from the western blots images and confirmed that the *KRT10* S42F, R86H, and K334Q reduced expression compared to WT when probed with KRT10 antibody targeted to the N-terminus in HEK239 cells, however this was not replicated when the same variants were probed with C-terminus tagged FLAG (Figure 4.4A). HaCaT cells showed reduced expression of S53N, and G575V compared to WT when probed with KRT10 antibody targeted to the N-terminus, however only G575V mirrored this reduced expression when probed with C-terminal tagged FLAG (Figure 4.4B).

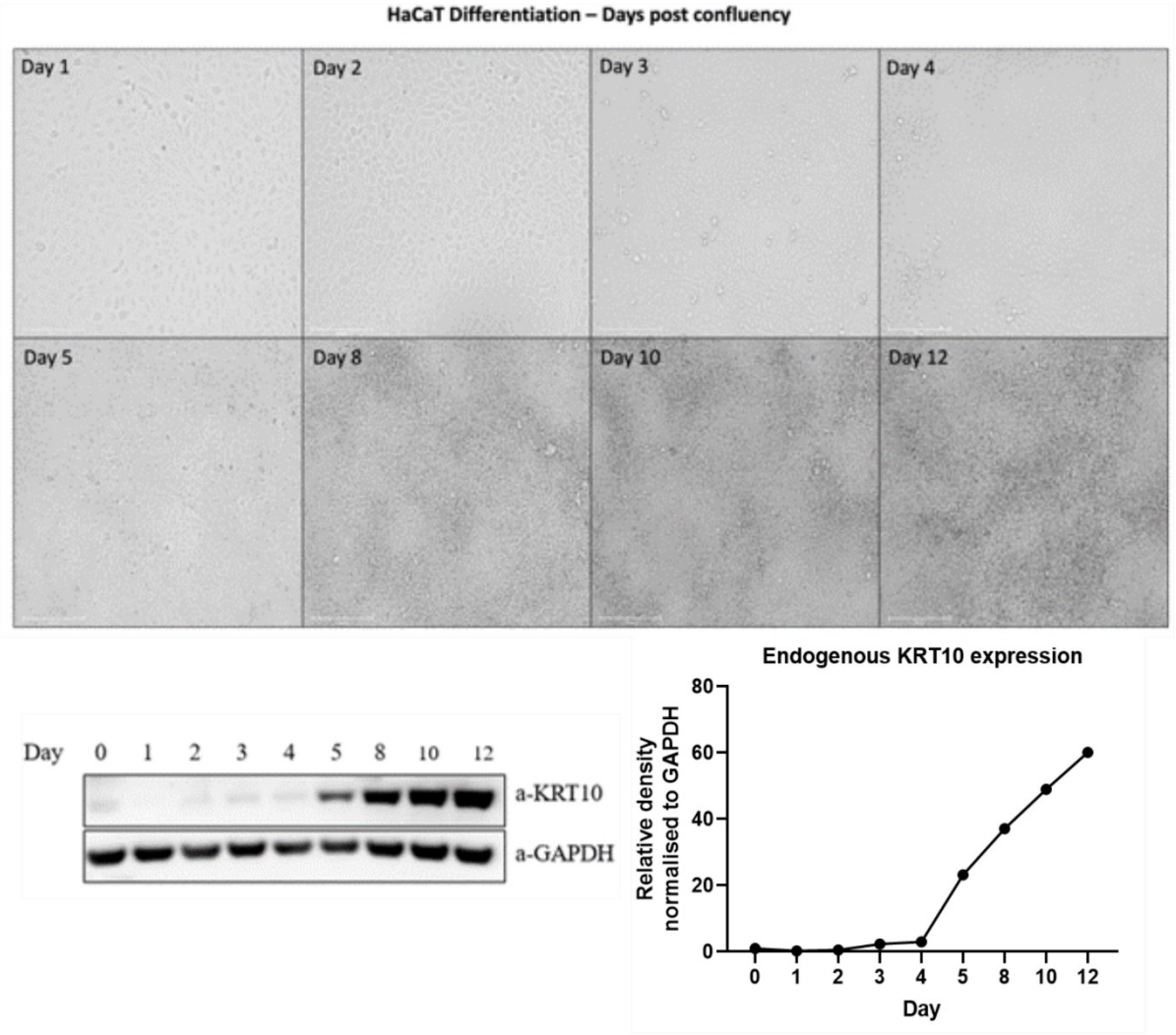


Figure 4.3 Endogenous KRT10 expression in differentiating HaCaT keratinocytes

The brightfield microscopy images (top) show HaCaT morphology over a 12-day period of confluency, cells were not split or removed during this time but left at confluency over time to induce differentiation. Western blotting (performed by Ms Laia Rigat Nogareda) of HaCaT protein lysates over twelve days shows resolved KRT10 and GAPDH proteins (bottom right), day 0 signifies HaCaT cells prior to seeding. Densitometry of immunoblots were analysed by me with ImageJ, N=1 and immunoblots are representative of one independent experiment.

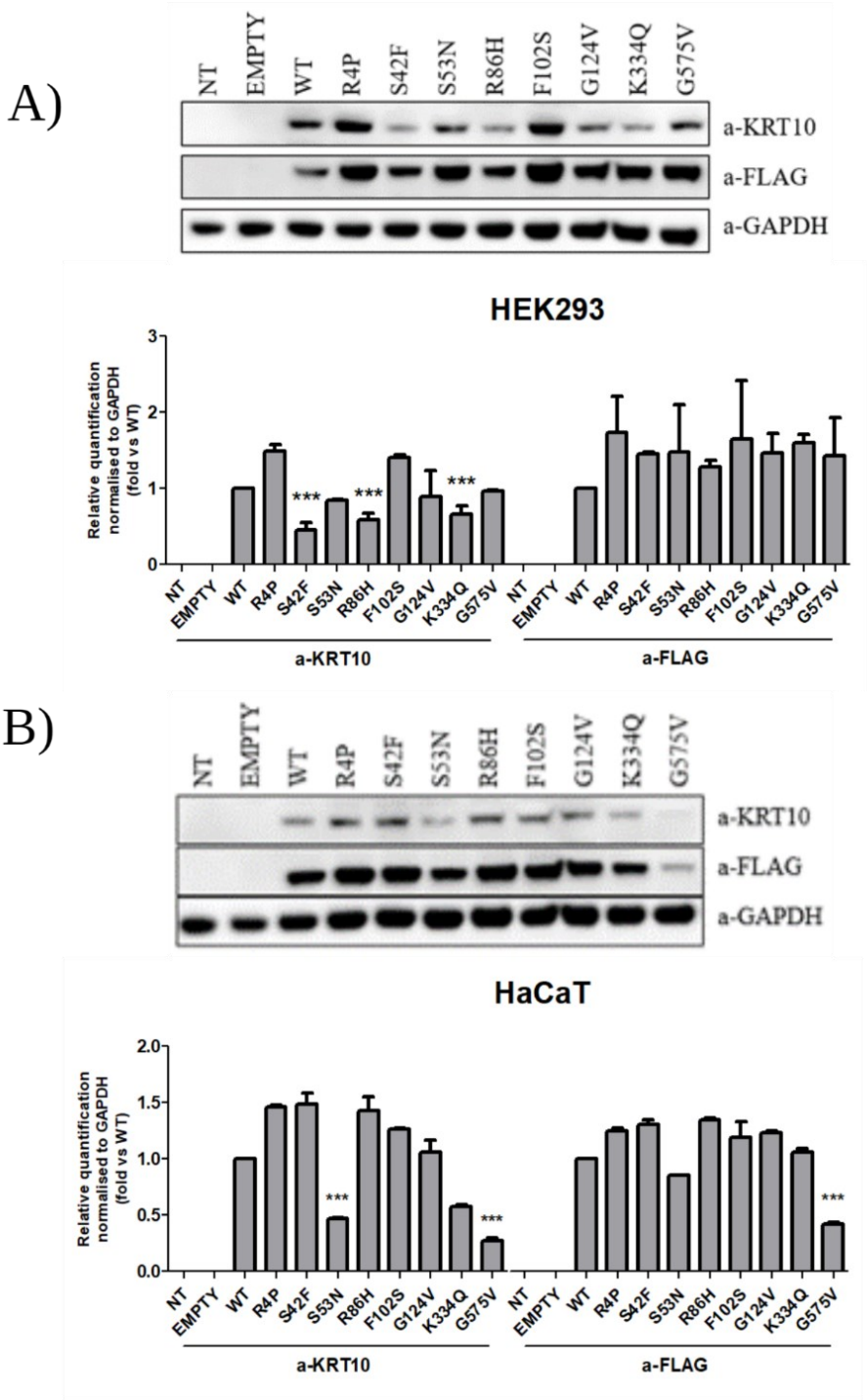


Figure 4.4 Protein expression of KRT10 variants in epithelial cells and keratinocytes

A) HEK293 or B) undifferentiated HaCaT cells were transiently transfected with KRT10 constructs. Immunoblots show resolved KRT10-FLAG proteins using antibodies against an epitope on the N-terminal of KRT10 and the C-terminal tag FLAG, with GAPDH as a loading control. N=2 and immunoblots are representative of at two independent experiments. Densitometry analysed with ImageJ, data shown as mean ± SEM and the differences between variants and wild type were tested in a two-way ANOVA. significance shown with ***p<0.001. NT; non-transfected, WT; wild type. Immunoblots were conducted by Ms Laia Rigat Nogareda and Ms Afnan Alahmadi, I performed the densitometry analysis.

4.7 Assessing subcellular localisation of *KRT10* variants by immunofluorescence

Genetic variability in *KRT10* has been reported to cause aberrant filament formation and mis-localisation of KRT10 in keratinocytes (380). With this information, I decided to investigate the impact of the *KRT10* variants overexpressed in undifferentiated keratinocytes, where there is minimal to no expression of endogenous KRT10, with experimental work and analysis conducted by Ms Laia Rigat Nogareda (Figure 4.5). Immunostaining with an antibody recognising the KRT10 C-terminally tagged FLAG motif showed a well-defined intermediate filament network spread uniformly throughout the cell with KRT10 WT. Interestingly= S53N, R86H, F102S, K334Q, and G575V KRT10 appeared to aggregate in the cytosol with intracellular punctate appearance (Figure 4.5). KRT10 R4P, S42F, and G124V appear to have similar KRT10 expression and filament network appearance to WT (Figure 4.5). In future experiments it will be necessary to co-stain with other subcellular markers to determine exact location of KRT10 expression and localisation.

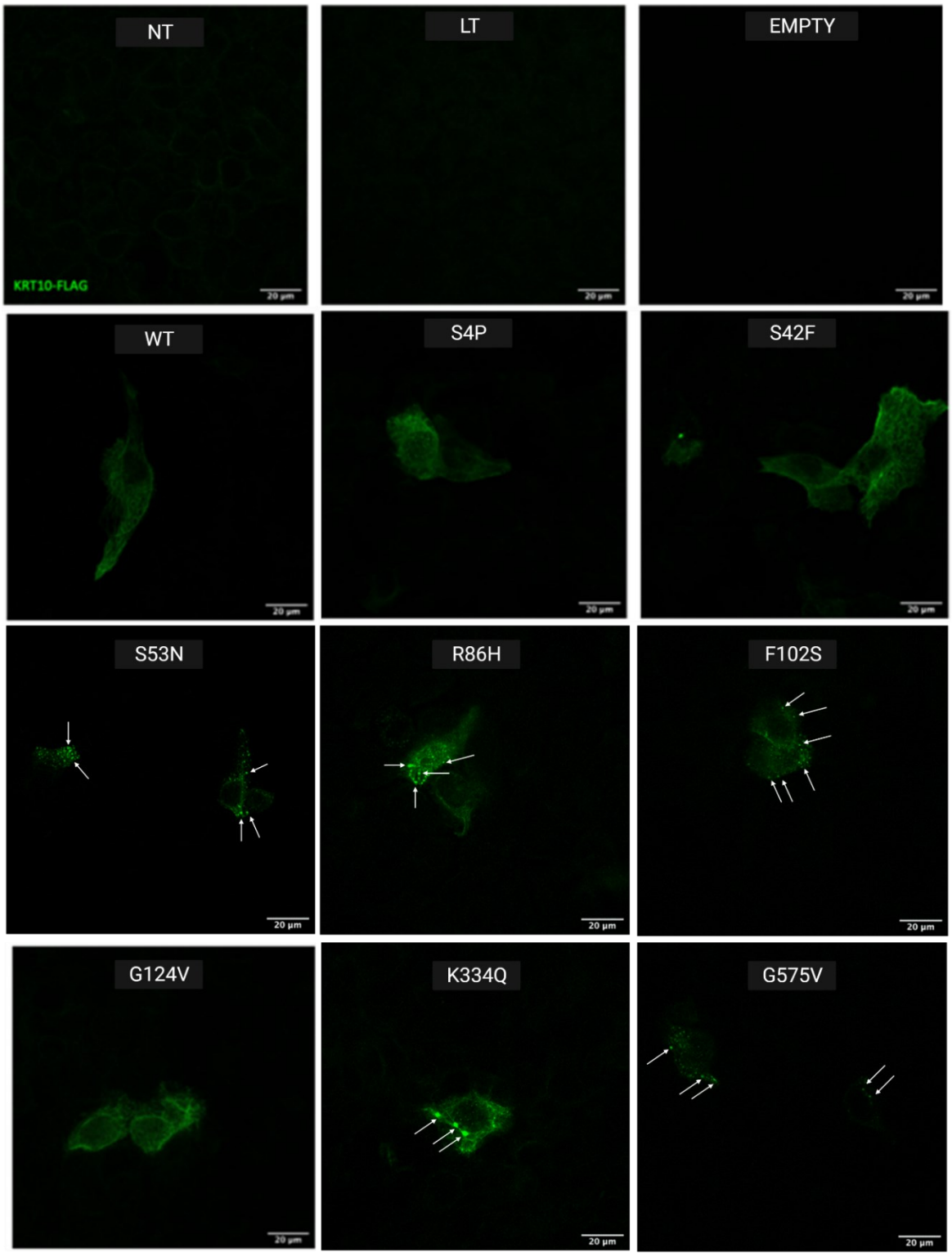


Figure 4.5 Impact of *KRT10* variants on protein expression by immunofluorescence

Undifferentiated HaCaT cells were transiently transfected with *KRT10*. A rabbit FLAG primary antibody (1:100) followed by a goat anti-rabbit IgG secondary antibody conjugated to AlexaFluor™ 488 (1:1000) was used to stain and visualisation of immunofluorescence was achieved with confocal microscopy at 63X. Arrows show aggregates of KRT10-FLAG. Scale bar, 20 µm. Figure shows a representative replicate of n=2 and was performed by Ms Laia Rigat Nogareda.

4.8 Investigating the impact of *KRT10* variants in bacterial infection by adhesion and invasion assays

KRT10 has been shown to bind to *S. aureus* clumping factor B (ClfB) (381-383) and *S. pneumoniae* pneumococcal serine-rich repeat protein (PsrP) (384). To understand if *KRT10* variants have an impact on this relationship a short time course adhesion and invasion assay was carried out using *S. aureus* in epithelial cells and keratinocytes. This assay has been described previously (116) and is detailed in section 2.7. Briefly, cell monolayers were transiently transfected for 24 hours and then infected with *S. aureus* for 90 minutes. Bacteria not associated with cells were washed away and the cell-associated bacteria were serially diluted, plated on solid growth medium and colony forming unit (CFU) values were determined. To measure invasion, a gentamicin killing of the surface adherent cells was added after the infection and wash step followed by a saponin mediated release of intracellular bacteria. Undifferentiated HaCaT cells were used as they express undetectable levels of endogenous *KRT10* (Figure 4.3).

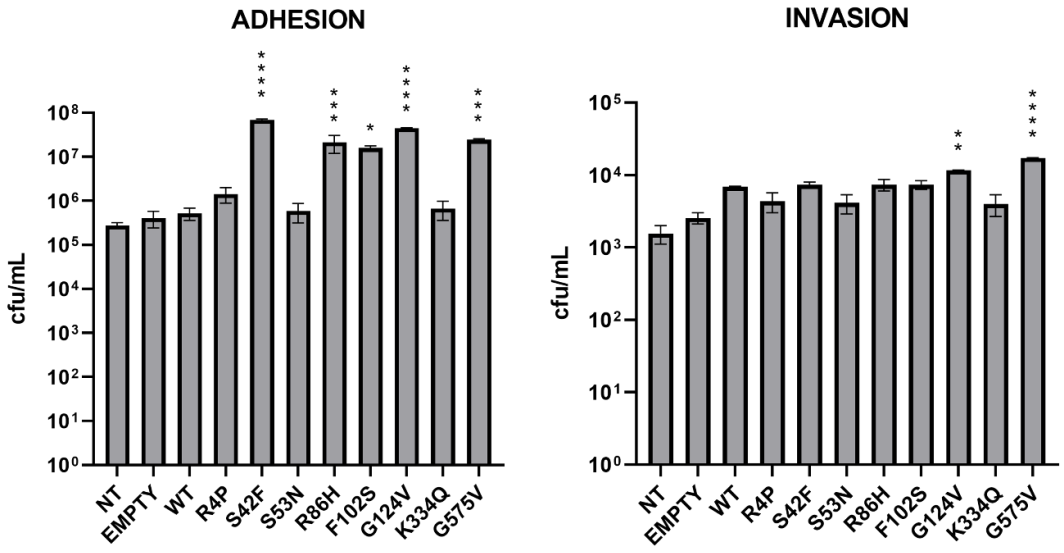


Figure 4.6 *KRT10* variants improve adhesion and invasion of epithelial cells by *S. aureus*

Bar graphs show HEK293 monolayers were transfected with *KRT10* variants and then infected with *S. aureus* for 90 minutes at a MOI of 10 and bacteria adherence (left) or invasion (right) were determined by colony counts (CFU/mL). Experiments were conducted by Ms Afnan Alahmadi and data analysed by me. Figure shows representative data of n=3 independent biological experiments. Data shows mean ± SEM and the differences between variants and wild type were tested in a two-way ANOVA with multiple comparisons corrected with Tukey, significance shown with *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. NT; non-transfected, WT; wild type.

Both epithelial cells (Figure 4.6) and keratinocytes (Figure 4.7) show a significant marked increase in both adhesion and invasion of *S. aureus* in the presence of *KRT10* variants. In epithelial cells, *KRT10* S42F, R86H, F102S, G124V, and G575V show a significant 1.5-2 log increase of staphylococci adhesion compared to WT, and only two variants, G124V, and G575V carried this significant difference through to staphylococci invasion (Figure 4.6). In keratinocytes, *KRT10* variants S42F, R86H, G124V, and G575V also show a significant 1.5-2 log increased staphylococci adhesion with S42F, and G575V displaying a significant log increase of intracellular staphylococci (Figure 4.7).

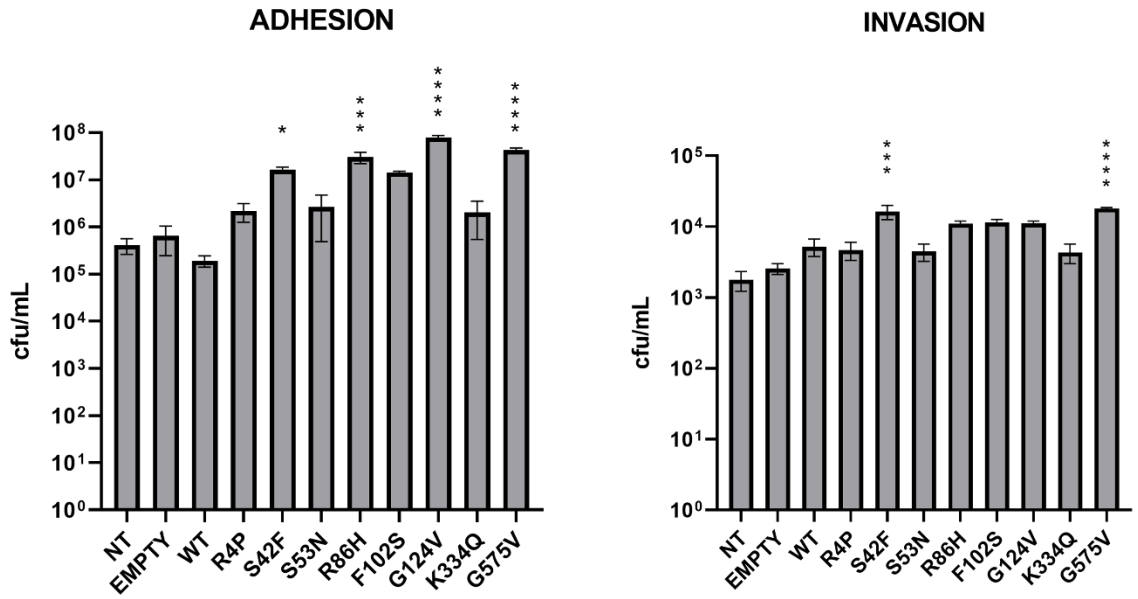


Figure 4.7 *KRT10* variants improve adhesion and invasion of keratinocytes by *S. aureus*

HaCaT monolayers were transfected with *KRT10* variants and then infected with *S. aureus* for 90 minutes at a MOI of 10 and bacteria adherence (left) or invasion (right) were determined by colony counts (CFU/mL). Experiments were conducted by Ms Afnan Alahmadi and data analysed by me. Figure shows representative data of n=3 independent biological experiments. Data shows mean ± SEM and the differences between variants and wild type were tested in a two-way ANOVA with multiple comparisons corrected with Tukey, significance shown with *p<0.05, ***p<0.001, ****p<0.0001. NT; non-transfected, WT; wild type.

4.9 Discussion

The *KRT10* gene is found in chromosome 17q21 and has two splice variants, the canonical transcript being 584 amino acids long (Accession number P13645). Genetic variations in *KRT10* are well-established to be associated with epidermolytic hyperkeratosis and ichthyosis with confetti, skin disorders that are typically present at birth or shortly after and can be clinically characterised by the presence of very red skin (erythroderma) and severe blisters (380, 385). These skin disorders caused by *KRT10* variants are often autosomal dominant but have also been seen to be inherited in an autosomal recessive manner, with the disease causative nature arising from the absence of KRT10 function, and/or the mis-localisation of KRT10 (380, 385). *KRT10* variants discovered in epidermolytic hyperkeratosis are commonly found in two "mutational hot spots" at beginning of the 1A rod domain, encoded by exon 1, and the other at the end of the 2B rod domain, encoded by exon 6 (Figure 4.1) (386, 387). Variation in *KRT10* associated with epidermolytic hyperkeratosis was first established in 1992 with the identification of a heterozygous p.R10H in an adolescent, later validated in two unrelated families with epidermolytic hyperkeratosis (385, 388). SNPs p.N8H (388), p.Y14D (388), p.R10C (386, 388), p.R10L (386), p.R15H, p.L15S (385), p.L103Q (388), p.R156H (389), p.R156C (390, 391), p.M150R (390, 391), p.Q434X (392), p.C427X (393), located in the mutational hotspots have also been reported in cases of epidermolytic hyperkeratosis. Heterozygous frameshift variants at the C-terminal tail of *KRT10* have been reported to cause aberrant filament formation and mis-localisation resulting in a replacement of the Gly-Ser rich tail with arginine- or alanine-rich motifs, causing KRT10 mis-localisation in keratinocytes and leading to a clinical presentation of ichthyosis with confetti (Figure 4.1) (380, 394-398).

Keratin proteins have been seen to interact with invading pathogens by binding to them, and potentially playing a role in their attachment, such as *Burkholderia cepacia* with KRT13 (399) and *Salmonella enterica* with KRT18 (400). In relevance to this project, *KRT10* has previously been shown to interact with *S. aureus* and *S. pneumonia.* Studies have showed preferential binding of the glycine/serine repeat motif located at the C-terminal tail region of *KRT10* to clumping factor B (ClfB), a surface binding protein of *S. aureus*, on the surface of desquamated nasal epithelial cells and keratinocytes, an interaction that may contribute to adherence and initial colonisation of the host (382, 383). Studies also show that *KRT10* binds to the pneumococcal serine-rich repeat protein (PsrP) of *S. pneumoniae*, on the endothelial lung cells, specifically at the PsrP basic binding region which is flanked by serine-rich repeats and promotes biofilm formation through self-oligomerisation (225, 384). A more recent

crystallography study further clarifies the interaction, by revealing that the PsrP basic binding region undergoes a conformational change, resulting in a novel MSCRAMM fold (401).

A gene burden analysis was performed by Dr Evangelos Bellos of whole exome sequences from 750 paediatric patients suffering from severe bacterial and viral infections recruited in EUCLIDS (Table 4.2). Using >140,000 exomes from gnomAD non-infectious non-cancer adult dataset as a control, *KRT10* was revealed as a top hit enriched in the invasive bacterial infection cohort which included invasive IMD, pneumococcal infection, invasive staphylococcal infection, and group A streptococcal infection. A range of clinical presentations and causative pathogens were seen in these patients, with severe sepsis the leading significant phenotype in 9/17 patients (Table 4.3). A total of 17 heterozygous variants, 12 missense and 5 in frame indel variants were identified in invasive bacterial infection patients in EUCLIDS (Table 4.1). Notably, KRT10 harbours a Glycine-Serine rich repeat region towards its C-terminal tail where there is reported high sequence conflict. In our cohort 5 in-frame indels and 3 missense variants found in high prevalence in this region were put aside for further investigations until their validity could be confirmed (Figure 4.1) (Table 4.1). The remaining 9 missense *KRT10* variants were introduced by site directed mutagenesis in a mammalian expression vector (Figure 4.2) and overexpressed in epithelial cells and keratinocytes (Figure 4.4).

KRT10 S42F, R86H, and K334Q had reduced expression compared to WT when probed with KRT10 antibody targeted the N-terminal in HEK293 cells, however this was not replicated when the same variants were probed with C-terminal tagged FLAG (Figure 4.4A). In HaCaT cells, there is a reduced expression of *KRT10* S53N, and G575V compared to WT when probed with the KRT10 antibody targeted the N-terminal, however only G575V mirrored this reduced expression when probed with C-terminal tagged FLAG (Figure 4.4B). These inconsistent results between a-KRT10 and a-FLAG immunoblots may be explained by difference in antibody binding affinity, as a-KRT10 binding likely does not bind as strongly to its epitope as a-FLAG, also evidenced by stronger bands seen in the a-FLAG immunoblots. However, reduced expression in G575V in HaCaT cells is likely a real result as this also objectively seen in immunofluorescence data (Figure 4.5). No other *KRT10* variants that show reduced expression via immunoblot is supported by immunofluorescence data. Immunofluorescence of KRT10 variant localisation was carried out in undifferentiated HaCaT keratinocytes, which were shown to produce undetectable levels of endogenous KRT10 (Figure 4.3). Immunofluorescence confirms that there is no loss of KRT10 expression caused by variants

discovered in the EUCLIDS cohort and 5/9 *KRT10* variants, S53N, R86H, F102S, K334Q, and G575V appear to aggregate in the cytosol with intracellular punctate appearance (Figure 4.5).

As discussed earlier, KRT10 preferentially binds to both *S. aureus* and *S. pneumonia* and this binding has been shown to encourage bacterial colonisation and biofilm formation, because of this, I hypothesised that *KRT10* variants associated with invasive bacterial disease may have increased or higher affinity binding of KRT10 to bacteria. To begin to characterise the impact of *KRT10* variants on function, I commissioned a bacterial assay, performed by Ms Afnan Alahmadi and Ms Laia Rigat Nogareda, that measured number of cell-associated bacteria and intracellular bacteria in cells expressing the *KRT10* variants. Results from this assay supports my hypothesis as multiple *KRT10* variants show a significant marked increase in staphylococci adhesion and invasion of HEK293 (Figure 4.6) and HaCaT (Figure 4.7) cells. In both cell types G575V shows significant increase in both adhesion and invasion of staphylococci, interestingly *S. aureus* is known to bind to the C-terminal glycine and serine motifs where G575 is located and this variant could result in improved binding (383). Adhesion of staphylococci was also shown to be significantly increased in HEK293 cells expressing *KRT10* S42F, R86H, F102S, (Figure 4.6) and G124V and HaCaT cells expressing S42F, R86H, and G124V (Figure 4.7). Invasion of staphylococci was also improved in HEK293 cells expressing G124V (Figure 4.6), and HaCaT cells expressing S42F (Figure 4.7). However, these experiments were performed again by a separate student and there were difficulties replicating the significance of this data (data not shown). The lack of reproducibility could be due to technical reasons to be investigated. For example, *S. aureus* ClfB is an MSCRAMM only expressed on the cell surface during the exponential phase of bacterial growth (402), and it could be that the bacteria in the second set of repeated experiments were grown insufficiently. Secondly, the transfection efficiency in the second set of repeated experiments was also lower, visualised by co-transfection with GFP and microscopy, (data not shown) then the original three independent replicates performed by the first student, which could result in decreased *KRT10* expression and therefore insignificant interactions with the bacteria. These will need to be further explored and the experiment repeated to determine true significance. Alternatively, KRT10 may not be significantly involved in bacterial adhesion or invasion, and this would have to be confirmed with additional investigations.

Additionally, I would also test the impact of *KRT10* variants on bacterial adhesion and invasion using the other pathogens discovered in the KRT10 patients, including *S. pneumoniae*, GAS, and *N. meningitides* in appropriate airway and lung cell lines. Other experiments to investigate the impact of *KRT10* variants in invasive bacterial infections are biofilm formation assays, which could be of particular interest as KRT10 binds to PsrP which self- oligomerises to promote biofilm formation and it would be interesting to investigate if *KRT10* genetic variations enhances this virulence mechanism (401). Another valuable experiment to utilise are pull down assays, which will allow investigations into the impact of *KRT10* variants on known bacterial binding partners, such as *S. aureus* ClfB. In conclusion, this work is still in its infancy but so far describes *KRT10* genetic variability as a promising candidate to investigate in susceptibility to invasive bacterial infection as preliminary data shows *KRT10* variants to modulate bacterial adhesion and invasion.

Chapter 5

Chapter 5 Loss-of-function *TICAM1* variants in SARS-CoV2 infection

The work presented in this chapter was carried out during the height of the SARS-CoV-2 pandemic lockdowns and due to the sudden nature of the project and tight deadline timely work was imperative, thus, all work carried out in this chapter was in direct collaboration with a lab colleague, Ms Ira Sabli.

5.1 Background

Monogenic inborn errors of type I IFN immunity have been implicated in respiratory viral infections. Severe influenza disease, including critical pneumonia and acute respiratory distress syndrome (ARDS), have been associated with AR *IRF-7*, AR *IRF-9*, AR and AD *TLR-3* deficiencies (280, 403, 404). It was shown *TLR-3* deficiencies result in enhanced susceptibility to influenza virus, which could be rescued by IFN-α2b or IFN-λ1 therapy (280). *IRF9* deficiency marked an inability to respond effectively to type I IFN in all cell types whereas *IRF7* deficiency effected the ability of pulmonary epithelial cells and plasmacytoid dendritic cells (pDCs), or both, to induce type I or type III IFNs (403, 404). This suggested the IRF-7-dependent type I and III amplification and the following IRF-9 response to be essential in protection from influenza virus. Notably, though GATA Binding Protein 2 (*GATA2*) is not directly involved in the type I IFN pathway, AD haploinsufficiency in *GATA2* have resulted in rare influenza associated ARDS, and might be due to a lack of pDCs in GATA2-deficient patients, which have a high baseline expression of IRF-7 and therefore produces large amounts of type I and type III IFNs that can signal induction of anti-viral defences (405-407).

SARS-CoV-2 is a single stranded positive RNA virus whose infection can manifest into severe COVID-19 disease. Around 30% of SARS-CoV-2 infections are asymptomatic with around 20% of COVID-19 patients requiring hospitalisation, of which, up to 30% needing intensive care (267). COVID-19 currently has a case fatality rate ~1% worldwide, but this can increase in pandemic settings and is variable depending on reporting countries and resources available in these countries to treat disease. For example, Peru currently has a COVID-19 case fatality rate of ~5% (267) (https://covid19.who.int/region/amro/country/pe). Mild symptoms of COVID-19 include common cold like symptoms such as coughing, high temperature, unexplained tiredness, muscle aches etc. However, serious disease can further develop and one of the more severe manifestations of COVID-19 is life-threatening pneumonia.

Risk factors for developing severe COVID-19 include older age (over 50), being male, and pre-existing co-morbidities, particularly conditions affecting the pulmonary system such as chronic obstructive pulmonary disease, cardiovascular diseases, and hypertension (267, 408-410). In 2020, the disease incidence rate of COVID-19 was 1 in 1000 in the young healthy population which rises to 1 in 10 in the over 80 population with underlying health conditions (267, 408-410). Genetic risk factors discovered early in the pandemic included discovery of types A and B blood groups conferring greater risk to COVID-19 compared to type O, while type AB, conferred the least amount of risk decrease (411). Another genetic risk factor identified early to be associated with increasing risk of hospitalisation from COVID-19 was a region on chromosome 3, spanning 6 genes, that were in high linkage disequilibrium (412). Around 5% of severe COVID-19 cases, predominantly presented with severe pneumonia but other presentation including encephalitis and cardiovascular disease are also reported to a lesser degree, are in young healthy adults that have no known pre-existing health conditions (410). Thus, the COVID Human Genetic Effort (COVIDHGE, https://www.covidhge.com) proposed that causative inborn errors of immunity (IEI) may be involved in development of COVID-19 pneumonia with complete or incomplete penetrance in otherwise healthy patients (413). IEIs in the TLR3 (280), IRF7 (404), and IRF9 (403) pathways have been shown to confer susceptibility to life-threatening influenza pneumonia, by impairing the response to and production of type I IFNs. Due to these established roles, it was decided to investigate IEIs in these pathways in life-threatening COVID-19 pneumonia (289, 414).

This work can contribute to understanding the wide clinical variability that occurs in an infected population, hypothesising that monogenic inborn errors of immunity can help explain the development of severe disease in otherwise healthy low-risk individuals. COVIDGHE is an international consortium, led by Professor Jean-Laurent Casanova and Dr Helen Su, that aims to discover the human genetic and immunological bases of the various clinical forms of SARS-CoV-2 infection. Our lab participates in this massive collaborative effort.

The work described in this chapter was published in Zhang *et al* (2020) "*Inborn errors of type I IFN immunity in patients with life-threatening COVID-19*" on which I am a second co-author (289). The overall project was to investigate *TLR3*, *IRF7*, *IRF9* and loci associated within these pathways, including *TICAM1, UNC93B1, TRAF3, TBK1, IRF3, NEMO/IKBKG, IFNAR1, IFNAR2, STAT1, and STAT2,* in the context of COVID-19 pneumonia. Due to our labs previous experience investigating *TICAM1* in herpes simplex encephalitis (12), we were tasked to elucidate the role *TICAM1* variants may play in SARS-CoV-2 infection and COVID-19. I

personally constructed all the *TICAM1* variants identified in the COVID-19 pneumonia cohort and performed the immunoblotting experiments alongside Ms Ira Sabli and this contribution was critical to establish pathogenicity of *TICAM1* variants.

5.2 Identification of *TICAM1* variants in COVID-19 pneumonia patients

In 2020, the COVIDHGE consortia internationally collected whole-exome (n=295) and whole-genome (n=364) sequences from 659 patients presenting to intensive care with critical COVID-19 pneumonia. Severe life-threatening COVID-19 pneumonia was characterised as patients with pneumonia who developed critical disease, whether pulmonary with mechanical ventilation (CPAP, BIPAP, intubation, hi-flow oxygen), septic shock, or with any other organ damage requiring hospital admission to the intensive care unit (ICU). Within this cohort, 74.5 % were men and 25.5 % were women, 13.9 % died, and patients were comprised from various ethnicities aged between 1 month and 99 years, with a median age of 51.8 years (289). As a control group, 534 individuals who tested positive for SARS-CoV-2, either via PCR or serological test, and who either remained asymptomatic or had self-resolving mild illness were enrolled.

Rare (MAF <0.01) monoallelic and biallelic nonsynonymous variants, predicted to be loss-of-function, were found to be enriched in the TLR3/IRF7/IRF9 pathway associated loci, using a principal component analysis (PCA)-adjusted burden test (289). This included *TLR3, IRF7, IRF9, TICAM1, UNC93B1, TRAF3, TBK1, IRF3, NEMO, IFNAR1, IFNAR2, STAT1*, and *STAT2* loci. From this original analysis 19 heterozygous *TICAM1* variants were identified and communicated to us. Table 5.1 summarises these variants, along with their MAF and clinical disease severity.

Table 5.1 *TICAM1* variants identified in life-threatening COVID-19 pneumonia patients

Protein variant	MAF (Global gnomAD)	Phenotype, number of cases
T4I	0.0003874	Severe, 1
T32I	0.003	Severe, 1
S60C	0.000294	Severe, 1
R71Q	0.000529	Severe, 3
V80M	0.000642	Severe, 1; Asymptomatic, 1
A111T	0.0001305	Critical, 1
T157M	0.000107	Critical, 1
S160F	0.00216	Severe, 1
R195Q	0.0000641	Severe, 1
G245S	0.00125	Severe, 1
V302L	0.0	Severe, 1
T377I	0.0000119	Severe, 1
L386P	0.0	Severe, 1
Q392K	0.0	Severe, 1
S465N	0.0	Severe, 1
D557N	0.000173	Severe, 1
M595L	0.000012	Severe, 1
G598W	0.000553	Severe, 1
P620S	0.0000114	Severe, 1; Asymptomatic, 1
Q702R	0.0	Severe, 1

MAF; minor allele frequency

Figure 5.1 visualises the identified *TICAM1* variants on a lolliplot, and shows localisation of variants along the TICAM1 protein and within certain domains. Previously reported *TICAM1* variants were used during this research as controls, including a loss-of-function variant, p.R141X, and a hypomorphic variant, p.S186L (12).

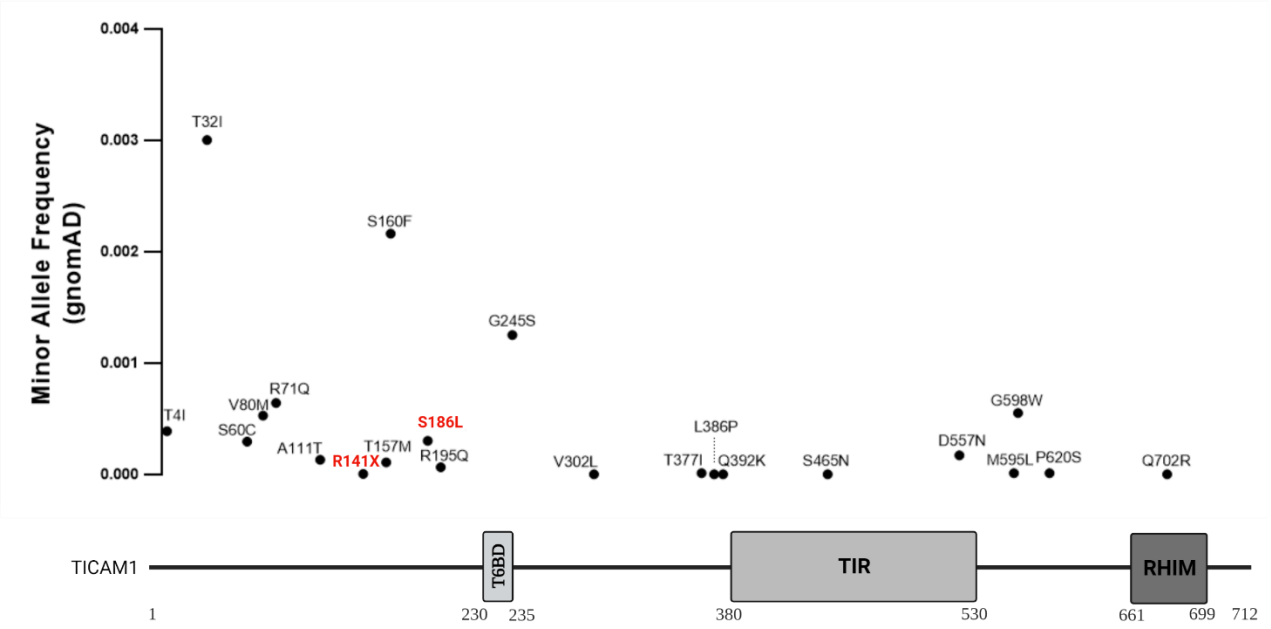


Figure 5.1 Lolliplot of *TICAM1* variants and corresponding MAF

TICAM1 variants discovered in COVID-19 pneumonia plotted along a schematic of TICAM1 protein against their reported minor allele frequency (MAF), according to gnomAD. Known *TICAM1* loss-of-function variant, p.R141X, and hypomorphic variant, p.S186L, included as experimental controls, highlighted in red. TRAF6 binding domain (T6BD), Receptor-interacting protein homotypic interaction motif (RHIM). Image created using biorender.com.

5.3 Construction of *TICAM1* variants in mammalian expression vectors

Firstly, the pEF-Bos mammalian expression vector containing wild-type (WT) *TICAM1* C-terminally tagged with HA, previously described (12), was used to introduce the missense variants by site-directed mutagenesis and overlapping primer pairs containing the single nucleotide substitution (detailed protocol found in section 2.4 and primer details found in Table 2.4). Plasmids generated from site-directed mutagenesis were sanger sequenced to confirm successful integration of target variants by aligning the constructed plasmid sequences to *TICAM1* WT (Figure 5.2). Whole plasmid sequencing confirmed no second-site variants were introduced during the mutagenesis process.

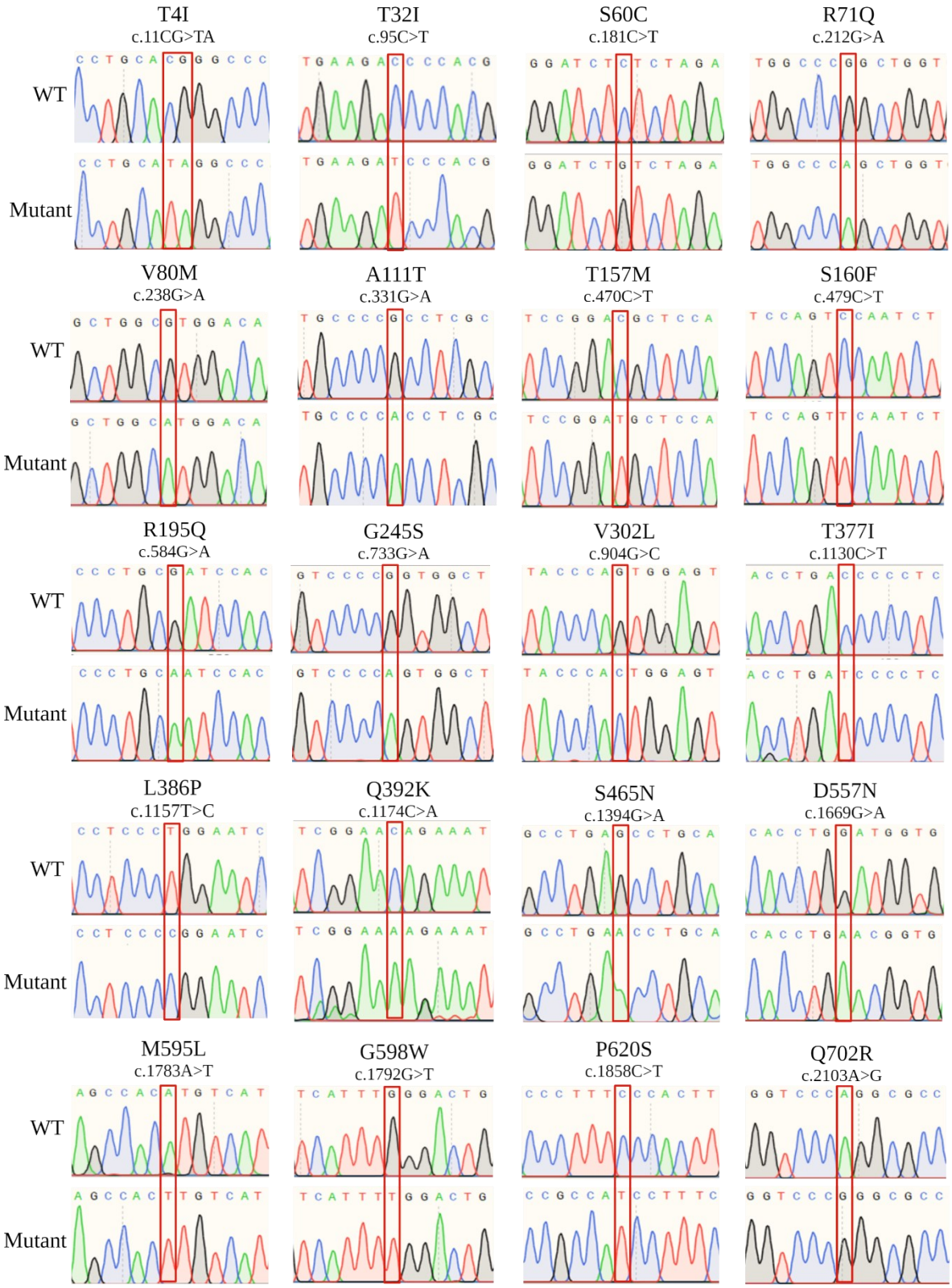


Figure 5.2 Sequence confirmation of site directed mutagenesis in *TICAM1* WT

Sanger sequence confirmation of successfully integrated *TICAM1* variants in pEF-Bos-TICAM1 WT plasmid. Alignment of pEF-Bos-TICAM1 and mutant plasmid, visualised using SnapGene (by Dotmatics). Nucleotide changes are marked with a red box. All constructs were created by me.

5.4 Optimisation of transient transfection in TICAM1-/- cells

To begin, transient transfection optimisation was completed in TICAM1-/- SV40-fibroblasts, previously shown to have undetectable TICAM1 protein expression (12), using Lipofectamine 3000 and pEF-Bos-TICAM1 WT plasmid, and cell viability and IFN-β induction was investigated (Figure 5.3). A plasmid concentration of 100ng was determined to be used with further experiments, as it achieved a ten-fold induction of IFN-β by luciferase assay while cell viability was at an acceptable level (Figure 5.3). This was investigated as it has been previously established that overexpression of *TICAM1* can trigger the TLR-3 downstream pathway, including rapid apoptosis which results in a high level of cell death (415, 416).

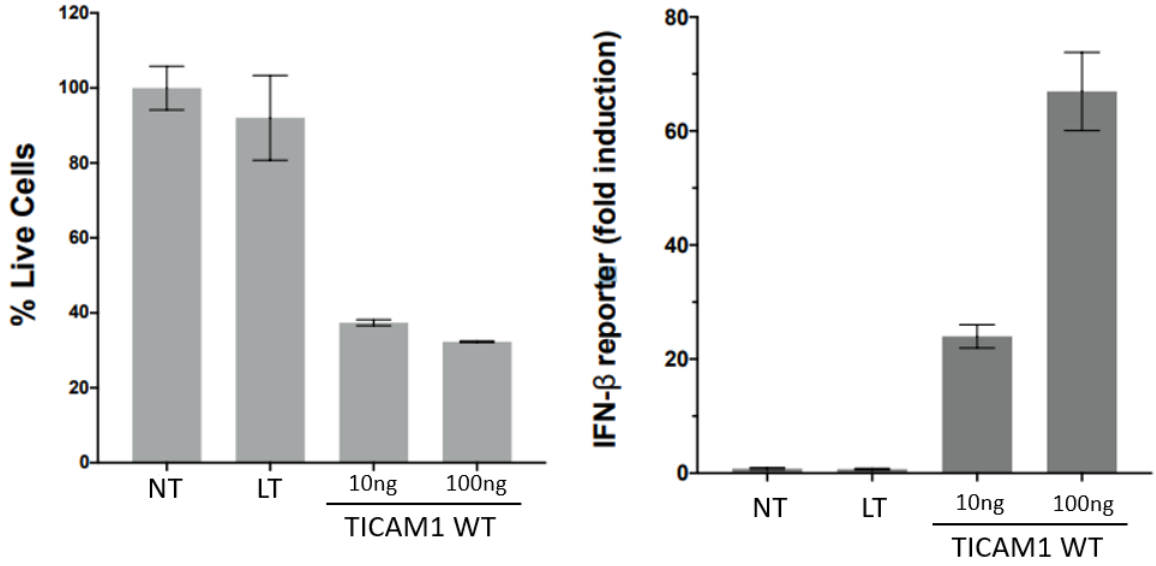


Figure 5.3 Transient transfection optimisation in TICAM1-/- cells

TICAM1-/- cells were seeded on 24-well plates and transiently transfected with TICAM1 WT plasmids at 10 ng and 100 ng using lipofectamine 3000 for 24 hours. A) MTS assay used to assess cell survival. B) IFN-β induction by luciferase assay. NT; non-transfected, LT; lipofectamine transfected control (no DNA), WT; wild type. Experiment performed by Ms Ira Sabli, n=3 independent experiments.

5.5 Loss of expression *TICAM1* variants identified using an overexpression system

5.5.1 Impact of *TICAM1* variants on protein expression by immunoblotting

TICAM1-/- SV40-fibroblast cells were transiently transfected pEF-Bos-TICAM1 constructs, in the presence of 20μM z-VAD-fmk to inhibit apoptosis. Post-transfection, protein was isolated from cell lysates and quantified as described in section 2.6. Western blots were then performed using an antibody recognising the HA antigen located on the C-terminal of TICAM1 and resolved via chemiluminescence (detailed protocol outlined in section 2.6). pEF-Bos-*TICAM1* constructs R141X and S186L were used as a loss-of-expression and hypomorphic controls, respectively (12). R141X was previously described to result in complete *TICAM1* deficiency, shown by undetectable protein expression by immunoblotting and absence of IFN-β or NF-κB promoter induction (12) S186L was previously described to be hypomorphic specifically in a doxycycline-inducible retroviral transduction system in which TICAM1 protein levels are more similar to endogenous expression, no hypomorphic traits were described when transiently transfected (12). Three variants showed a clear loss-of-expression with no detectable proteins showed by immunoblotting, and densitometry analysis shows a significant absence of detectable protein in T4I, S60C, and Q392K (Figure 5.4).

5.5.2 Visualisation of the impact of *TICAM1* variants with confocal microscopy

To confirm the loss-of-expression variants shown by the immunoblots, TICAM1-/- cells were transiently transfected with TICAM1 constructs C-terminally tagged with HA. (see detailed protocol in section 2.13). Immunofluorescence was visualised with confocal microscopy and cells transfected with T4I, S60C, and Q392K showed a lack of fluorescence of TICAM1-HA (Figure 5.5), therefore confirming a loss-of-expression phenotype for these three variants, corroborating the immunoblot data (Figure 5.4).

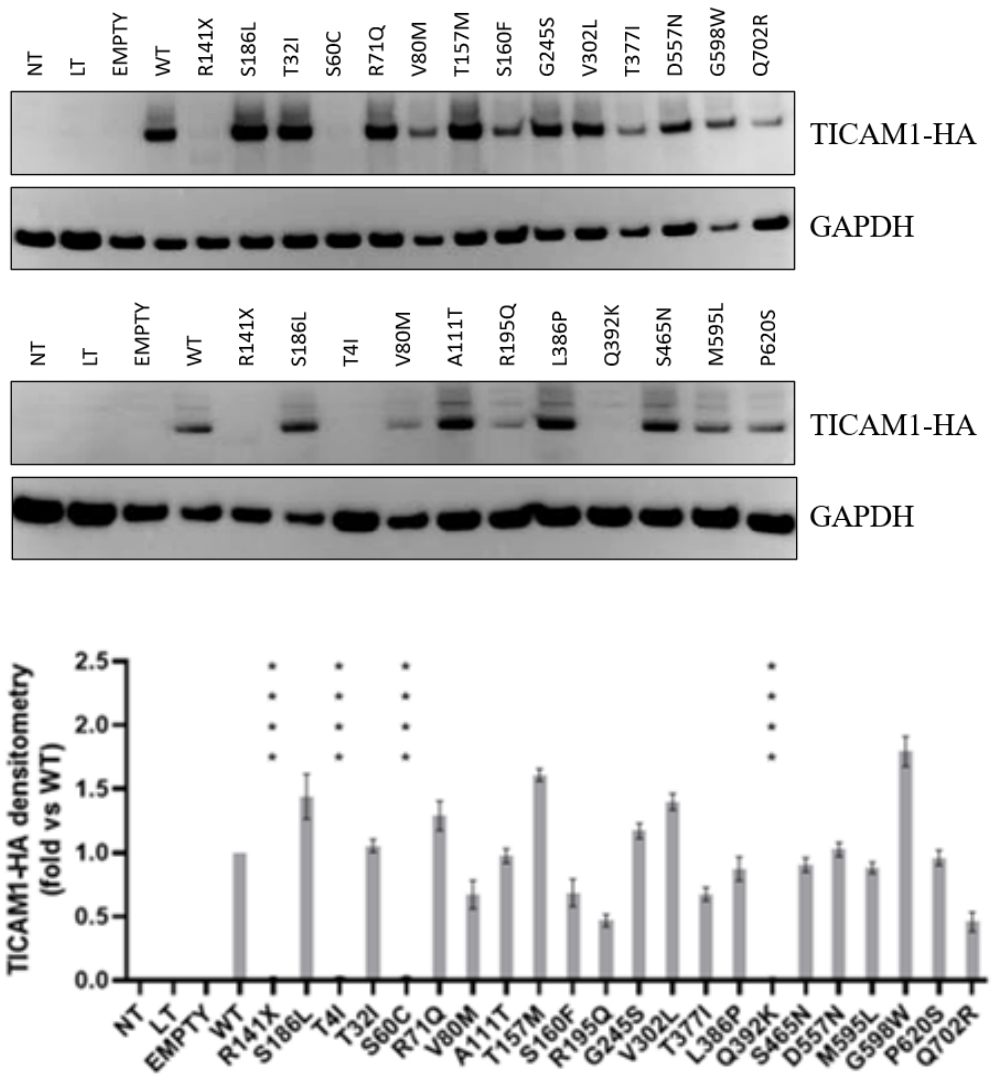


Figure 5.4 Impact of TICAM1 variants on protein expression by immunoblotting

TICAM1-/- cells transiently transfected with TICAM1 constructs. R141X and S186L were used as a loss-of-expression and hypomorphic controls, respectively. Immunoblots representative of n=3 and shows resolved TICAM1-HA proteins, with GAPDH used as a loading control. Densitometry analysed with ImageJ, data shown as mean ± SEM and the differences between variants and wild type were tested in one-way ANOVA, significance shown with ****p<0.0001. NT; non-transfected, LT; lipofectamine transfected, WT; wild type. Immunoblots were performed by both me and Ms Ira Sabli.

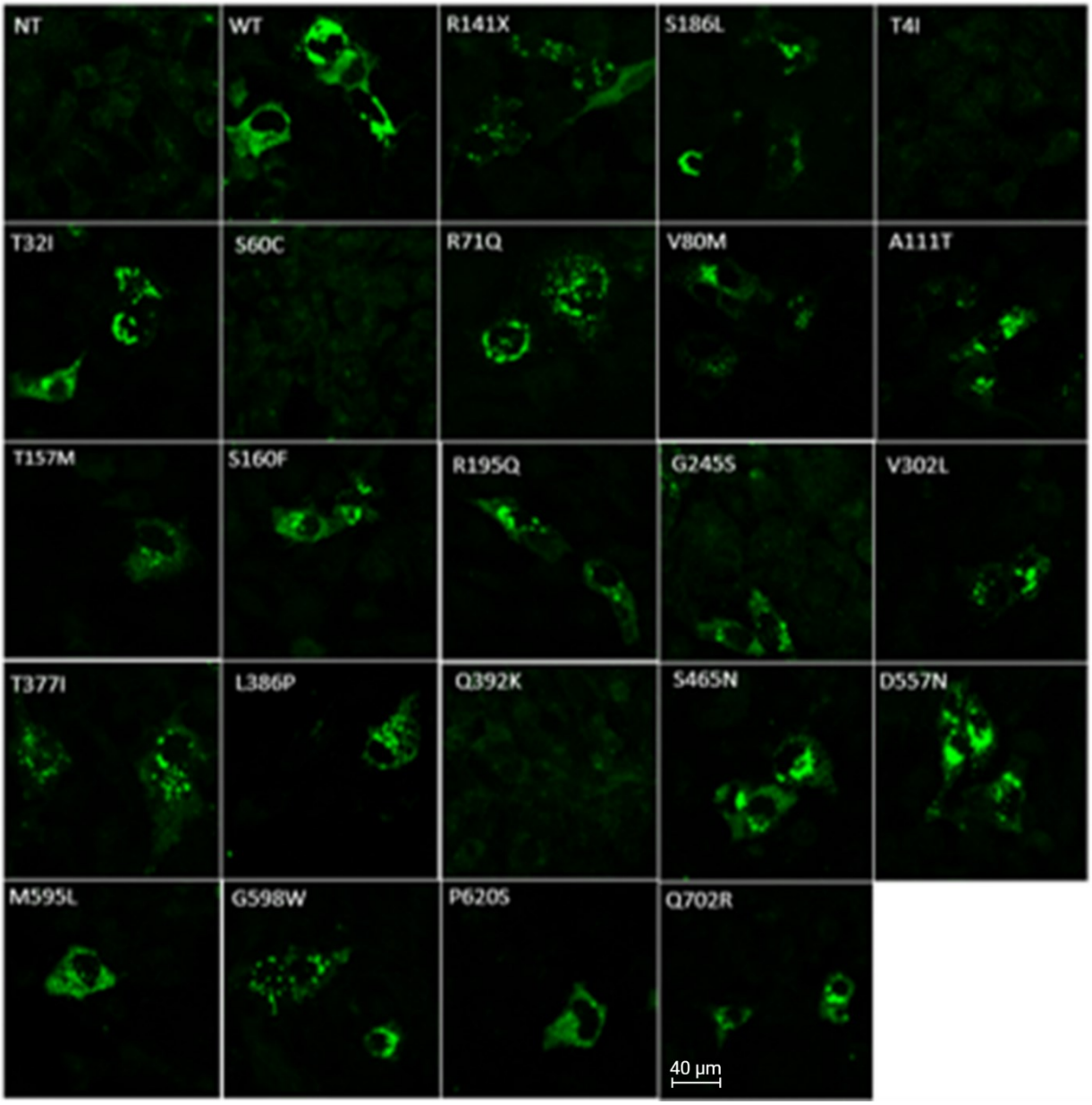


Figure 5.5 Impact of TICAM1 variants on protein expression by immunofluorescence

TICAM1-/- cells transiently transfected with TICAM1 constructs. Visualisation of immunofluorescence was achieved with confocal microscopy. R141X and S186L were used as a loss-of-expression and hypomorphic controls, respectively. Images are representative of n=3 and shows TICAM1-HA proteins. Scale bare represents 40 µM. NT; non-transfected, WT; wild type. Immunofluorescence was performed by Ms Ira Sabli.

5.6 Functional characterisation of *TICAM1* loss of expression variants shows impaired type I IFN and NF-κB signalling

All 20 *TICAM1* variants were tested in luciferase reporter assays to investigate their impact on signalling function. The three loss-of-expression variants identified by immunoblotting and immunofluorescence, T4I, S60C, and Q392K were also seen to be loss-of-function (LoF) with respect to induction of the IFN-β and NF-κB promoters, when compared to WT, and at a similar levels to the known LoF deleterious variant, R141X (12) (Figure 5.6). All other *TICAM1* variants did not show a loss of expression (Figure 5.6).

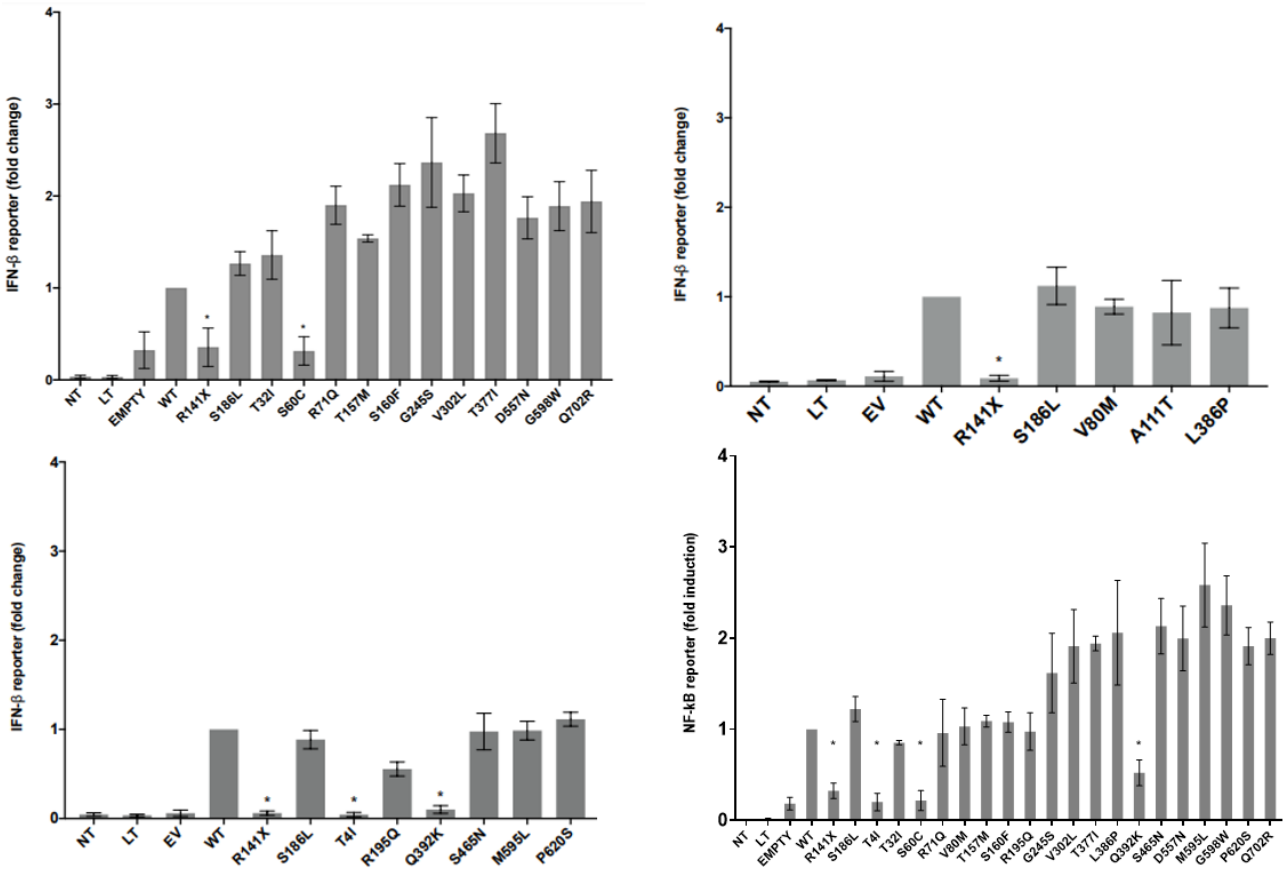


Figure 5.6 Impact of TICAM1 variants on type I IFN and NF-κB signalling

TICAM1-/- cells were transfected as described previously. R141X and S186L were used as a loss-of-expression and hypomorphic controls, respectively. IFN-β (top left, top right, bottom left) or NF-κB (bottom right) Firefly Luciferase was normalized against *Renilla* Luciferase activity and expressed values were analysed as fold change with respect to TICAM1 WT. Each graph shows n=3 from three independent biological repeats, values are graphically represented as mean ± SEM and the differences between variants and wild type were tested in one-way ANOVA, significance shown with *p<0.05. NT; non-transfected, LT; lipofectamine transfected, WT; wild type, EV; empty plasmid. Luciferase assays were performed by Ms Ira Sabli.

5.7 Discussion

Inborn errors of immunity resulting in deficiencies in the TLR3-IFN signalling pathway have been previously associated with other viral infections, most typically herpes simplex virus 1 (HSV-1) infection. This includes autosomal recessive (AR) *UNC93B* deficiency, AR, and autosomal dominant (AD) *TLR3* deficiencies, AD *TRAF3* deficiency, X-linked *NEMO* deficiency, and AR *STAT1* deficiency (see Table 1.6) (273, 275, 278, 279, 417, 418). Additionally, AD TANK-binding kinase 1 (*TBK1*) variant p.G159A, acting in a dominant negative manner, has shown failure in inducing both autophagy and an appropriate anti-viral IFN response upon both HSV-1 infection and stimulation with polyinosinic-polycytidylic acid (poly I:C), a synthetic molecule mimicking dsRNA structure (276, 419, 420). Also, both AD and AR *TICAM1* deficiencies, shown to impair TLR3 signalling, have also been implicated in HSV-1 infection, underlying herpes simplex encephalitis (HSE) in paediatric patients, with incomplete clinical penetrance (12). These published AD pS186L and AR p.R141X *TICAM1* variants, were used as hypomorphic and LoF controls during the work described in this chapter, and LoF resulting from the premature stop codon in p.R141X was co-corroborated internally with no detectable protein seen in immunoblots probed with antibody directly recognising TICAM1 (data not shown), and an absence of IFN-β and NF-κB promoter inductions, seen by reporter assay (Figure 5.6).

Due to inborn errors of immunity implicated with the development of critical pneumonia and ARDS in influenza infections because of deficiencies in TLR-3, IRF-7 and IRF-9, it was decided to investigate these genes along with 10 other loci directly related to these three genes that confer influenza susceptibility. This included genes in the TLR3-dependent type I IFN pathway *TICAM1, UNC93B1, TRAF3, TBK1, IRF-3*, and *NEMO/IKBKG*, and those located in the IRF7- and IRF9-dependent type I IFN amplification pathway including *IFNAR1, IFNAR2, STAT1,* and *STAT2* (289). Genetic aetiologies in all these genes have been linked with other viral diseases (summarised in Table 1.6). As mentioned previously, due to the time-sensitive nature of the topic at the time, this massive work was delegated throughout the COVIDHGE network, and our lab was tasked to investigate the impact of rare *TICAM1* variants.

In this study, 20 variants of *TICAM1* found in severe and/or critical COVID-19 pneumonia patients (Table 5.1) were successfully constructed into a mammalian expression vector pEF-Bos-TICAM1, described previously (12). The integration of the target SNP into the plasmid was confirmed by sanger sequencing followed by confirmation of no other second-site errors

introduced during the mutagenesis process that could impact observed variant-specific findings (Figure 5.2). It has been previously established that overexpression of *TICAM1,* which occurs during transient transfections, can trigger the TLR-3 downstream pathway, including rapid apoptosis which results in a high level of cell death (415, 416). Additionally, overexpression of TICAM1 can induce its own degradation through NDP52-mediated selective autophagy and, potentially, lysosomal degradation (421, 422). This is because, cellular TICAM1 levels are typically kept in tight regulation as it's a key adaptor molecule in the TLR pathway and increased expression can trigger the signalling system. To overcome apoptosis taking place in transient transfections and allow for investigations using applications downstream of transient transfections, z-VAD-fmk was supplemented into transfections. This compound is a cell-permeant pan-caspase inhibitor that prevents apoptosis by permanently binding to the catalytic site of caspase proteases. Use of this compound allowed for ~40% cell death rate in *TICAM1* transfections (Figure 5.3), which, while high, was at an acceptable level to conduct further experimentations, and showed good induction of IFN-β promoter representing an intact signalling pathway (Figure 5.3).

TICAM1-/- fibroblast cells were used as they provided a TICAM1 deficient background on which to assess expression or biological function of the TICAM1 variants. Primarily, protein expression was assessed using immunoblot (Figure 5.4) and immunofluorescence (Figure 5.5) to visualise the TICAM1-HA expressing proteins, and these showed mirrored results with three TICAM1 variants, T4I, S60C, and Q392K appearing to be loss-of-expression with no detectable protein. All other TICAM1 variants showed protein levels similar to or reduced compared to WT but densitometry showed that they did not reach the significance threshold (Figure 5.4). Protein lysates were also probed with anti-TICAM1 antibody (data not shown) and showed similar results to the anti-HA blots shown here, with a loss of function seen in T4I, S60C, Q392K, and R141X (Figure 5.4). To further characterise the impact of the variants on TICAM1 function, two luciferase reporter assays were performed to investigate induction of NF-κB and IFN-β promoter, key signalling molecules activated and downstream of TICAM1. These showed that the three loss-of-expression variants, T4I, S60C, and Q392K, were also loss-of-function with respect to NF-κB and IFN-β promoter induction, with induction levels significantly reduced compared to WT and in line with a previously established loss-of-function variant, R141X (Figure 5.6). Though three *TICAM1* variants were discovered to be LoF, the potential impact of the other variants should not be discarded as it has been shown that AD *TICAM1* variants can be hypomorphic and result in severe clinical phenotypes (12).

S186L was used as a hypomorphic control in these investigations, however hypomorphic phenotypes tend to be subtle and are typically not made apparent in an overexpression system, as was used in this research, as cellular levels can become saturated and subsequent analyses are not sensitive enough to reveal any impacts. However, S186L was found to be hypomorphic and showed significantly less cytokine production and impaired IFN-β, IFN-λ, and IL-6 induction in a doxycycline-inducible retroviral transduction system in which protein levels are closer to endogenous levels, also resulting in less cell death as protein levels are intracellularly tolerated (12). This method can therefore be further explored with the *TICAM1* variants identified in this cohort, to identify any more subtle hypomorphic effects on TICAM1 function. Furthermore, potential future work should also entail investigating dominant negative effects as these were discovered as monoallelic variants in patient samples and current work, discussed in this chapter, was limited to investigations in an overexpression system. The COVID-19 severe pneumonia patients possessing the three investigated LoF *TICAM1* variants, T4I, S60C, and Q392K did not carry any known variants at other candidate loci known to underlie inborn born errors of immunity (289).

In summary, three *TICAM1* variants were discovered to be LoF, T4I, S60C, and Q392K and this data together with other IEIs identified in this project that impact of type I IFN immunity have be shown to play a role in susceptibility to and severity of COVID-19. With autosomal-recessive deficiencies in *IRF-7* and *IFNAR1*, and autosomal-dominant deficiencies in *TLR-3, UNC93B1, TICAM1* (described in this chapter), *TBK1, IRF-3, IRF-7, IFNAR1,* and *IFNAR2* found to result in a stunting or absence of type I IFN production and amplification. All together, these deficiencies were shown to contribute to ~3.5% of patients developing life-threatening COVID-19 pneumonia and highlights the importance of type 1 IFN response in limiting SARS-CoV-2 infection (289).

Chapter 6

Chapter 6 BRC Paediatric Clinic (INGENIOS)

6.1 Background

One of the leading causes of childhood mortality in those under 5 years of age worldwide is infectious diseases along with other noncommunicable diseases such as pre-term birth complications, birth asphyxia, accidental trauma, and congenital anomalies, according to the WHO (423). There is typically an unproportional burden of infectious disease on the youngest and oldest members of society most likely due to an incomplete or deteriorating immune system. Children under 5 years of age are particularly vulnerable to infectious disease as they have immature immune systems, with maternal antibodies contributing to the infant's passive immunity waning within months of birth (424).

Analysis of death registration data between 2013-2015 in England and Wales showed that 3.3 deaths per 100,000 children were caused by infectious disease, with 1 in 5 of all childhood deaths resulting from infectious disease, though, notably, 55 % of these deaths had an underlying condition (425). From this subset, the average age of death was ~12 months and 63 % of deaths were from bacterial infections, 34 % from viral infections, and 2.5 % from fungal infections, with respiratory tract infections the most prevalent reported clinical presentation (425). Because of the still prevalent burden of infectious disease in some of the most vulnerable members of society, it is important to undertake research that improves diagnosis, understanding of disease mechanisms and informs therapeutic options. The development of NGS has presented a powerful tool to medical professionals in assisting with these endeavours.

Over the past two decades the paediatric infectious team at St Mary's Hospital (London, UK) has continually led NIHR and European Union funded projects investigating the genetics aetiologies and functional immunology of children with severe or unusual infection. The EUCLIDs FP7 project was one such project that was described in detail in Chapter 3 of this thesis. The ongoing work on these projects suggests that an estimated ~10% of children with serious infection phenotypes, and no known predisposing conditions, have identifiable genetic aetiologies to explain their disease, many of which are novel (426) (unpublished). Due to this, Imperial's Biomedical Research Centre (BRC), which is a partnership between Imperial College Healthcare NHS Trust and Imperial College London, awarded funding to create a new specialist clinic using non-NHS, research-based diagnostic approaches. This included functional immunological testing, NGS and subsequent variant analysis and research to

predominantly investigate severe and/or unusual infectious phenotypes in children. Previous elucidations of mechanisms in infectious diseases, for example complement deficiencies in invasive meningococcal disease or genetic variability in the IL12-IFN-γ pathway predisposing to atypical mycobacterial infections have resulted in improved diagnostics and better-informed therapy decision-making. Whole exome sequencing (WES) was the main NGS of choice in this clinic due to its cost-effectiveness. Although protein coding regions of the human genome constitutes less than 2% of the total DNA, more than 85% of all known disease causing gene variants are located in the exome (18). Other benefits in choosing to use WES was the ease of variant processing and analysis that does not require complicated analytical tools, compared to WGS or RNA-sequencing data, and the accompanying lower operational costs, including data storage (18). However, other NGS methods such as WGS or RNA-sequencing data can be used in specific cases where it is deemed necessary and appropriate.

This specialist clinic is part of the Investigating the GENetics of InfectiOn Study, INGENIOS. A key benefit of this clinic is the multi-disciplinary team (MDT) behind it, which includes clinicians, scientific researchers, research nurses, genomic medicine councillors, and a wide net of collaborators to call upon. The collective knowledge, skill, and expertise in this MDT enables a broad discussion of cases. The main goal of this project was to establish this specialist clinic, use it to recruit appropriate cases and investigate the possible mechanisms behind severe disease in children.

6.1.1 Aims

My aimed contributions to INGENIOS were to

- Design and implement a universal pipeline to investigate genetic variability in WES.
- Conduct genetic investigations using the designed pipeline in all INGENIOS recruits where WES data is available.
- Assist in wet lab work where appropriate, including DNA extraction, and relevant immunological testing.
- Participate and take a leading role in ongoing monthly MDT meetings, held to discuss new recruits, patient results and research, and present updated information on patients if/when they are seen again in clinic.

6.2 Summary of INGENIOS recruits

Appropriate patients who fit the study criteria of INGENIOS were recruited by the clinical team and admitted or referred to the Imperial College Healthcare NHS Trust (ICHNT) nationally, with a base at St Mary's Hospital (London, England). Recruitment criteria include presentation with a severe or unusual infectious disease in children who were otherwise healthy and have no known underlying risk factors for infection. Where possible, recruitment of a complete trio was preferred, a pedigree which represents an index case, the biological mother, and the biological father. In some cases, siblings and other family members were also recruited to assist in analysis relevant to disease segregation. At the time of writing, a total of 72 index cases have been recruited to INGENIOS, with accompanying family members.

Figure 6.1 summarises key demographic traits of the index cases, with 33/72 being female and 39/72 male, with a wide range of ethnicities. The most prominent ethnicities are those of north/mid/east European descent (35/72), which is expected as the clinic is geographically located in north-west Europe, followed by those of South Asian descent (12/72), likely owing to the multi-cultural diversity found within London. The recorded age of onset represents the age when the significant clinical episode occurred that prompted recruitment, this shows a wide distribution from 0-18 years with most illnesses presenting in those under 12 months of age (Figure 6.1). Figure 6.2 summarises the breadth of severe/unusual phenotypes present in INGENIOS, with most cases having infection with an identified causative pathogen (52/72) (Figure 6.2). A breakdown into broad categories reveals 17/72 bacterial infections, 32/72 viral infections, 22 of which identified HSV as the causative pathogen, 3/72 fungal infections, 13/72 which presented with an autoimmune/inflammation phenotype, typically of unknown cause, and 7/72 which were not categorised (Figure 6.2). There is a clear over-representation of HSV infections, reflecting the research interest of the project and referral of these cases to INGENIOS. A full comprehensive table of patient details is available in the appendix (Table 8.1).

After patients were recruited a biological sample was taken from them and from relevant family members, typically either blood or saliva, and DNA was extracted (detailed protocol found in section 2.4.7). The DNA quality was then checked and sent for WES to Novogene (UK, Cambridge centre). In some cases, where appropriate, a separate blood sample was obtained from case and control to perform additional immune functional testing. This typically occurred

in patients presenting with an HSV-1 infection as the lab was equipped to perform inhouse immunoassays, primarily testing patient cytokine response to stimulants and HSV-1 infection.

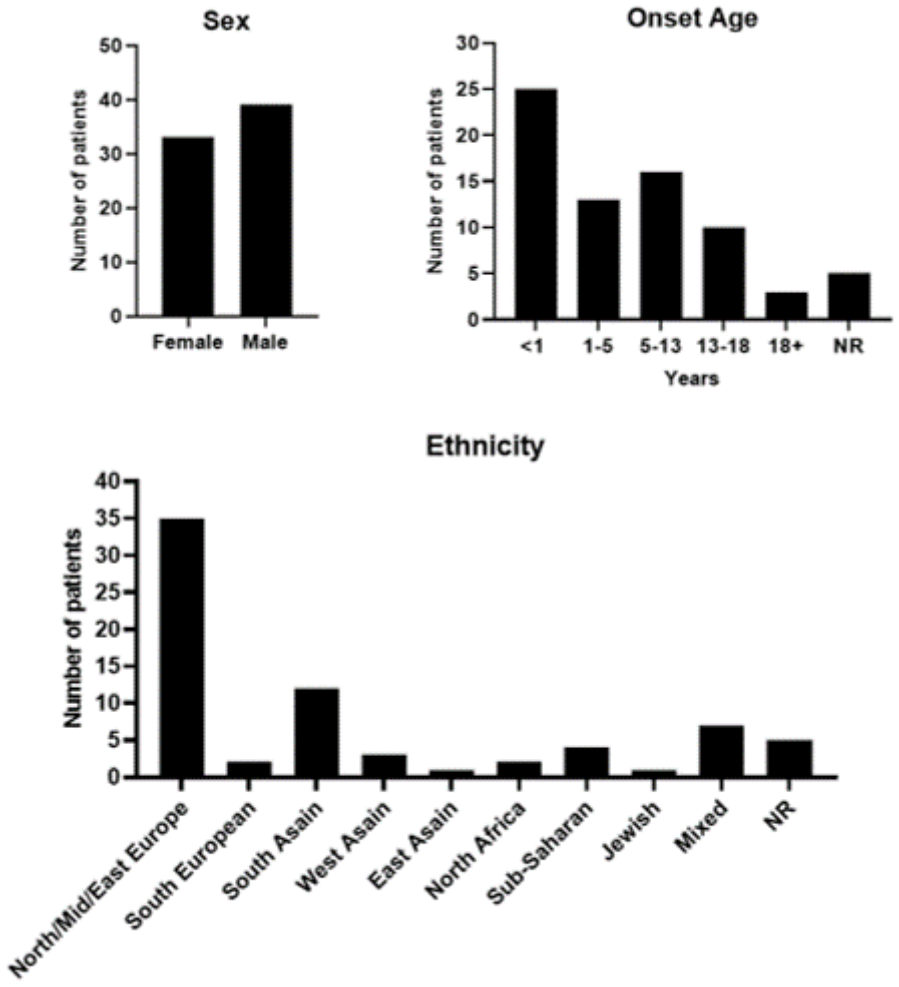


Figure 6.1 Demographic summary of INGENIOS patients

Distribution of major demographic information of the cohort recruited to the INGENIOS clinic, including sex (top left), age of onset of first significant clinical episode (top right), and ethnicities (bottom). Total number of patients (at the time of writing) = 72. NR; Not recorded.

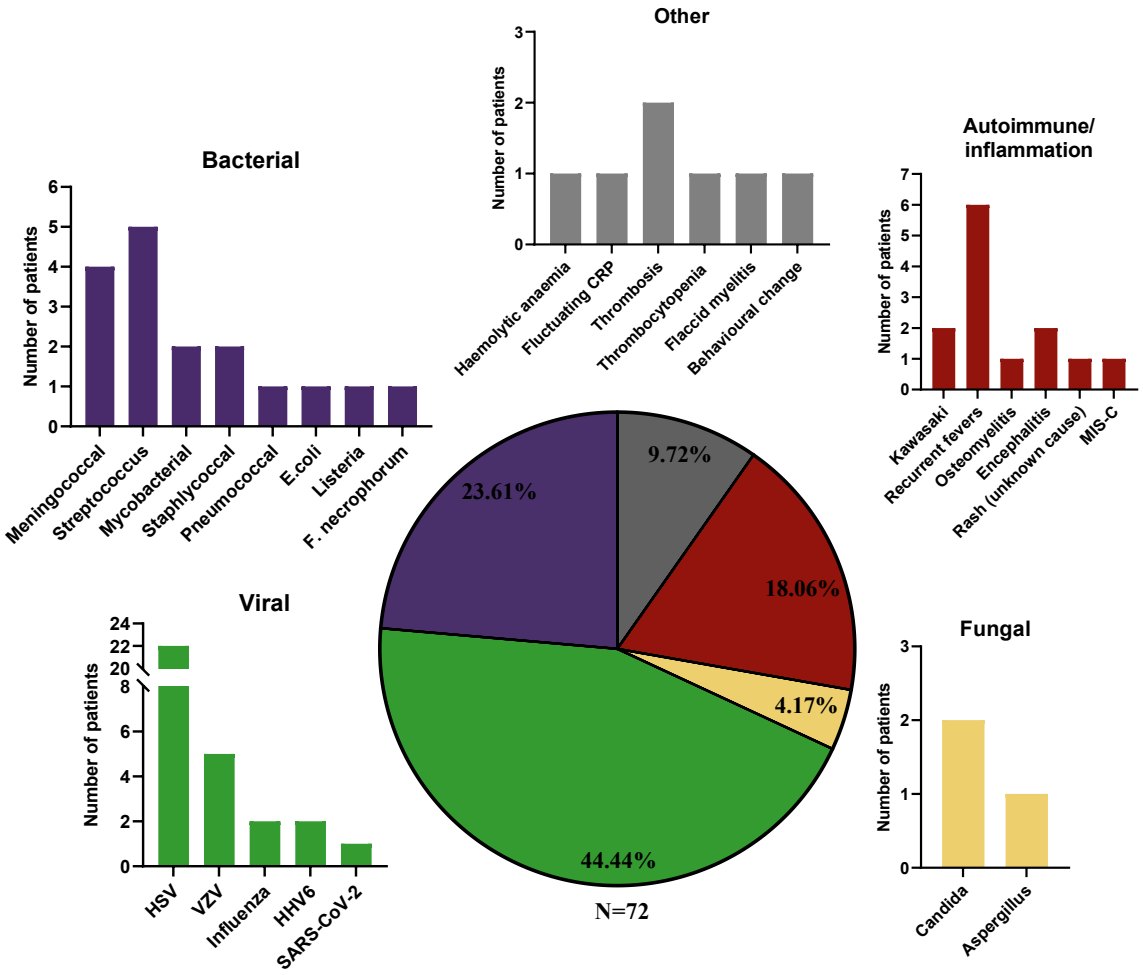


Figure 6.2 Clinical diagnoses summary of INGENIOS patients

Summary of the prominent phenotypes of the cohort recruited to the INGENIOS clinic, total n=72 represented by the pie chart in the middle. Surrounding bar charts represent the breakdown of clinical diagnosis or causative pathogen within each section. Total bacterial = 17, total viral = 32, total fungal = 3, total autoimmune/inflammation = 13, Total other = 7. CRP; C-reactive protein, HSV; herpes simplex virus, HHV6; Human herpesvirus 6, MIS-C; Multisystem inflammatory syndrome in children, SARS-CoV-2; severe acute respiratory syndrome coronavirus 2, VZV; varicella zoster virus.

6.3 Genetic analysis of WES

6.3.1 Initial WES processing

Whole exome sequencing was performed on an Illumina NovaSeq 6000 platform which generated paired end reads of 150 base pairs in length. This generates exomes with an effective mean coverage of 50x. With an average turnaround time of 3-4 weeks, the raw reads in FASTQ format were then returned to us from Novogene, and initially processed by Dr Evangelos Bellos for de-duplication and sequencing quality assessment using the Picard tools (http://broadinstitute.github.io/picard) and then aligned to the current human reference genome (build UCSC hg19) using bwa v0.7. Post-alignment, the exomes were jointly genotyped using GATK v3.8 (https://github.com/broadinstitute/gatk) (292). Initial variant annotation was achieved using Ensembl's Variant Effect Predictor (VEP) database and then the Variant Call Format (VCF) files were uploaded to an online software platform, QIAGEN Clinical Insight (QCI) Interpret (Figure 6.3). From here, variants observed in individuals can be viewed and handled in a user-friendly interactive format. The next step was to perform an initial genetic analysis that could be replicated in all patients consistently, and this pipeline was designed and implemented by me.

6.3.2 Development of a universal analysis pipeline

To begin, non-synonymous variants with a minor allele frequency of <1% was applied to filter out common, likely benign, polymorphisms. Next, it was important to discover if the patient had any known pathogenic variants in known genes associated with primary immunodeficiencies (PID) which could assist in a diagnosis, and limit further investigative resources and time. Therefore, a PID gene list was curated based on the most updated knowledge provided by The International Union of Immunological Societies (IUIS) expert committee on Inborn Errors of Immunity. This list currently consists of 485 genes underlying IEI phenotypes (7). Additionally, more comprehensive gene lists were independently curated for patients presenting with specific infectious diseases, from causative pathogens including HSV, *Neisseria meningitidis*, mycobacteria, and candidiasis, drawing from both publications and unpublished data shared by collaborators and colleagues (Figure 6.3) (full gene lists can be found in the appendix, section 8.1 Gene lists used to filter WES data). These gene lists were applied to WES data on QCI, to enable filtering of the sequence data at the gene level (not the variant level). Once a list of variants located in known disease-causing genes was produced using the patient WES it was then necessary to investigate on a variant level to

determine pathogenicity. This was accomplished in a multitude of ways. Firstly, I applied the accepted standards and guideline from the American College of Medical Genetics and Genomics (ACMG) via QCI software, to generate a list of variants categorised as pathogenic/ likely pathogenic/ uncertain significance/ likely benign/ benign/ conflicting interpretations (20). The relevance of each variant was considered according to the predicted mode of inheritance in each case, such that a heterozygous variant discovered in a gene that is known to cause disease in an autosomal recessive manner was deemed unlikely to result in a severe phenotype, except in very rare circumstances. Successfully filtered variants that are predicted pathogenic/likely pathogenic, fit the expected mode of inheritance of the patient in question, and have a published disease phenotype that fits the clinical history of the patient were then very closely scrutinised to determine the likelihood of their being a causative variant. This includes a comprehensive literature search on the variant itself to determine if it has been previously reported and in what context, the novelty of the variant, and assessing the type and location of variant itself and the impact it is likely to have on the protein. I accomplished this using multiple public resources, including Genomenon Mastermind (https://www.genomenon.com/mastermind/), the Human Gene Mutation Database (HGMD®) (https://www.hgmd.cf.ac.uk), NCBI ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), and the Genome Aggregation Database (https://gnomad.broadinstitute.org/). For example, a nonsense mutation resulting in an early termination, or an indel resulting in a frameshift, or a variant in a splice site, is likely to cause more damage to the protein than a more common missense variant. Protein impact can be scored *in silico* using predictive software such as Combined Annotation Dependent Depletion (CADD), Polymorphism Phenotyping v2 (PolyPhen-2), and Sorting Intolerant From Tolerant (SIFT). However, it is worth noting that in silico predictive software tends to favour identification of loss-of-function, damaging, variants and not potential gain-of-function variants, which are harder to predict. Therefore, predictive software was used with this limitation taken into consideration.

All information on the variants was then collated and presented at an MDT meeting, where a joint decision was made on the next steps to take, after consideration of the likelihood of a variant being causal in the patient's clinical presentation. MDT decisions included next steps in clinical management, and decisions about downstream functional work undertaken to investigate the biological impact of detected variants .

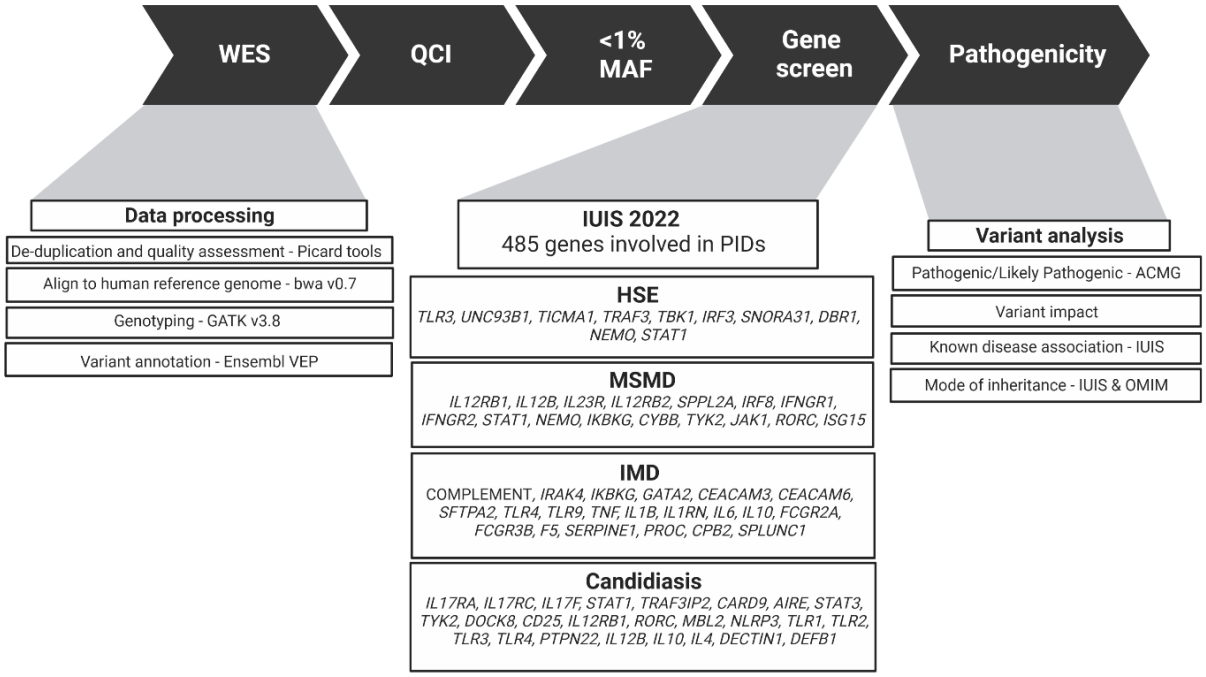


Figure 6.3 Analysis pipeline used for INGENIOS patients

Summary of the analysis pipeline designed to explore genetic variability using WES data of patients recruited to INGENIOS. Processed WES results were uploaded to QIAGEN Clinical Insight, where a cascade of filters were applied to investigate the presence of potential disease-causing variants. These steps included filtering out synonymous variants, and then filtering on low variant frequency, selection of known disease-causing genes, and prediction of pathogenicity. HSE; herpes simplex encephalitis, IMD; invasive meningococcal disease, MAF; minor allele frequency, MSMD; Mendelian susceptibility to mycobacterial disease, PID; primary immunodeficiency, QCI; QIAGEN Clinical Insight, WES; whole exome sequencing.

6.4 Discovery of genetic variants in paediatric patients

In most cases, implementing the designed pipeline described above achieves 1 of 6 outcomes that are discussed at the MDT meetings.

1. An established known pathogenic variant in a known disease-causing gene is found that matches the mode of inheritance and the patient's diagnosis, allowing for accurate informed care and where appropriate, referral.
2. A variant of unknown significance is discovered in a known disease-causing gene that matches the mode of inheritance and the patient's diagnosis, which typically results in functional testing to characterise the genotype.
3. An established known pathogenic variant in a known disease-causing gene is found that matches the patient's diagnosis, but the mode of inheritance is incorrect. This prompts discussion with expert collaborators and colleagues to determine the likely impact and decide on next steps.
4. An established known pathogenic variant in a known disease-causing gene is found that matches the mode of inheritance but does not match the patient's phenotype. In these cases, ethical relevance of reporting is discussed and may result in referral.
5. Either a novel variant, or a variant predicted to have high impact is discovered in a gene in a cellular pathway relevant to the patients' phenotype, that matches the mode of inheritance. In these cases, the research burden is discussed and if agreed further functional work is performed to characterise the variant.
6. No variants of interest are found in genes that match, or could lead to, the disease phenotype, and the patient is assigned a 'negative genetic report'.

Some examples of cases that fall into some of these categories will be described in the following sections.

6.4.1 De novo heterozygous *STAT3* in pneumococcal infection (1018)

Patient 1018 is a male of East Asian descent presenting with a history of severe pneumococcal pneumonia, multiple skin, ear, and upper respiratory tract infections and eczema. The first pneumococcal pneumonia episode occurred at around five years of age with right sided pneumonia and pleural effusion caused by *S. penumoniae* which was isolated from the pleural fluid and detected by PCR. This first episode required invasive treatment by drainage of the right sided empyema by video-assisted thoracoscopic surgery (VATS), and insertion of a chest drain. He recovered from this until around 10 years of age where he was admitted to St Mary's

Paediatric Intensive Care Unit (PICU) with right lobe pneumococcal pneumonia showing extensive right sided consolidation and pleural effusion, confirmed by urine antigen testing. He was conventionally ventilated for 5 days and treated with antibiotics. Due to his medical history, he was recruited into INGENIOS to determine if a genetic cause could be identified for his predisposition to severe pneumococcal disease. Notable immunological findings from NHS tests included very high serum IgE levels, 3598 IU/mL (normal range <200 IU/mL), indicative of potential Hyper-immunoglobulin E syndrome (HIES), and live vaccine failure for measles as no measles antibody was detected after vaccination with both doses. AD-HIES is characterized by chronic eczema, elevated IgE levels and recurrent infections of the skin, mucosa, and lungs (427, 428). There is no significant family medical history and DNA was obtained from the trio, index case, mother, and father, sent for WES, and analysed.

Genetic analysis of the trio WES was performed, using the pipeline described above (Figure 6.3), with all inheritance models interrogated. This analysis resulted in the discovery a heterozygous *de novo* pathogenic variant in Signal transducer and activator of transcription 1 (*STAT3*), p.R382Q (Table 6.1). This variant is classed as pathogenic by ACMG by fulfilling strong pathogenic (PS) criteria, PS2; De novo (both maternity and paternity confirmed) in a patient with the disease and no family history, and PS3; Well-established in vitro or in vivo functional studies supportive of a damaging effect on the gene or gene product, along with other moderate and supporting pathogenic criteria. *STAT3* p.R382Q is located in the STAT3 DNA-binding domain and is an established disease-causing variant in HIES, which aligns with 1018's immunological phenotype (429, 430). Published functional characterisation describes dominant negativity, and defective responses to cytokines, diminished STAT3 DNA-binding ability, and reduced secretion of anti-inflammatory molecules (429-431). Bacterial infections in STAT3-HIES are common and recurrent bacterial pneumonia is seen in 80% of STAT3-HIES patients, often complicated by pleural effusion (427, 432, 433). This clinic was therefore able to confidently report *STAT3* R382Q as a disease-causing variant in patient 1018.

Patient 1018 had a full recovery with good overall performance at hospital discharge, is currently on azithromycin prophylaxis and care has been handed over to the immunology team at Great Ormand Street Hospital (GOSH), with ongoing assessment to determine if a bone marrow transplant is necessary.

Table 6.1 *STAT3* variant summary discovered in INGENIOS patient 1018

Gene	*STAT3*	**Zygosity**	Het
dbSNP ID	*rs113994136*	**Inheritance**	De novo
HGVS description	NM_139276.3 c.1145G>A, p.R382Q	**Consequence**	Missense
		Read depth	64x
		ACMG	Pathogenic
Location GRCh38 (hg19)	Chromosome 17: g.42329642C>T	**CADD**	33
		gnomAD MAF	-

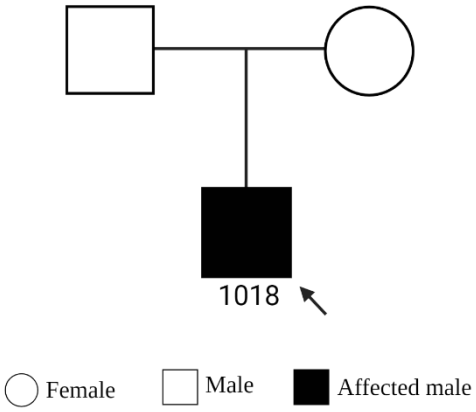


Figure 6.4 Pedigree of INGENIOS patient 1018

This pedigree shows INGENIOS patient 1018, indicated with the arrow, affected with recurrent pneumococcal pneumonia, the rest of the family remain well. A known pathogenic *STAT3* heterozygous variant p.R382Q was discovered de novo in the index case.

6.4.2 Known AD *STAT1* in oral thrush (1028)

Patient 1028 is a female of north/mid/east European descent with hypothyroidism who presented with recurrent oral thrush with the first episode occurring at 6 years of age. Familial history included recurrent mouth lesions with no known diagnosis in both the mother and older sister. *Candida albicans* was isolated from a mouth swab of the index case during an episode. DNA was obtained from both the index case and mother, sent for WES, and analysed. DNA was not available from the father or sister.

Genetic analysis of index case and mother WES was performed, using the pipeline described above (Figure 6.3), with pedigree segregation informing on a likely autosomal dominant inheritance, though all inheritance models were interrogated for completeness (Figure 6.5). This analysis resulted in the discovery of a known published pathogenic heterozygous missense variant in Signal transducer and activator of transcription 1 (*STAT1*), p.R274W (rs387906758) in patient 1028 that was inherited from the mother (Table 6.2). ACMG computed a pathogenic classification as the variant aligns with strong pathogenic (PS) criteria, PS3; Well-established

in vitro or in vivo functional studies supportive of a damaging effect on the gene or gene product, and PS4; The prevalence of the variant in affected individuals is significantly increased compared with the prevalence in controls, along with other moderate and supporting pathogenic criteria fulfilment. *STAT1* p.R274W has been described as a gain-of-function variant associated with autosomal dominant Chronic Mucocutaneous Candidiasis (CMC) by independent sources (434, 435). Published functional characterisation of p.R274W determined hyperphosphorylation of STAT1 leading to defective responses in type 1 and type 17 helper T cells (Th1 and Th17) necessary for clearing CMC with IL-17A & IL-17F cytokines, and impaired STAT3-dependent transcriptional activity, with the IFN-γ receptor pathway maintained (434, 435). The original published case studies on gain-of-function *STAT3* in CMC describes a family pedigree with an almost identical phenotype to patient 1028 and affected family members, in which both female index case and mother are affected with CMC with candida and hypothyroidism and a younger sister is also affected with CMC with candida (434).

This strong evidence led us to determine *STAT3* p.R274W to be a causative disease-causing variant in patient 1028 and care was taken over by the Great Ormand Street Hospital (GOSH) immunology team.

Table 6.2 *STAT1* variant summary discovered in INGENIOS patient 1028

Gene	*STAT1*	**Zygosity**	Het
dbSNP ID	rs387906758	**Inheritance**	AD
HGVS description	NM_007315.4 c.820G>A, p.R274W	**Consequence**	Missense
		Read depth	25x
		ACMG	Pathogenic
Location GRCh38 (hg19)	Chromosome 2: g. 191859911 G>A	**CADD**	32
		gnomAD MAF	-

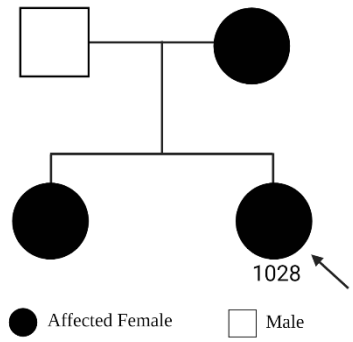


Figure 6.5 Pedigree of INGENIOS patient 1028

This pedigree shows INGENIOS patient 1028, indicated with the arrow, affected with recurrent oral thrush, a phenotype shared with her older sister and mother. A known pathogenic *STAT1* variant p.R274W was discovered in the index case, segregated autosomal dominantly with the mother. Sister and father were not genetically investigated.

6.4.3 Novel X-linked *CFP* in IMD (1034)

Patient 1034 is a male of north/mid/east European descent and presented with a severe case meningococcal sepsis, with blood culture detecting *Neisseria meningitidis* serogroup Y, at aged 7 that resulted in a four-limb amputation. At the time of illness, he had not been immunised but has since been immunised with the MenACWY vaccine and he is currently well after the disease incident and has been advised by the clinical team to re-start on regular long-term penicillin prophylaxis. Family history at the time of recruitment showed both parents and older brother to be well with no significant clinical episodes. DNA was obtained from both the index case and mother, sent for WES, and analysed. A DNA sample was obtained from the father, but quality control revealed the DNA to be of too low a quality to continue with WES, and a new DNA sample has since been requested.

Genetic analysis of index case and mother WES was performed, using the pipeline described above (Figure 6.3), with all inheritance models interrogated. This analysis resulted in the discovery a novel high impact hemizygous variant in complement factor properdin (*CFP*) (Figure 6.6). Chromosome X-Linked (XL) recessive damaging variants in *CFP*, hemizygous in males and homozygous in females, are well established to cause invasive meningococcal disease as they result in *CFP* deficiency. CFP is a positive regulator of the alternative complement pathway as it can bind to and stabilise a central element, C3b, thereby prolonging its functional activity within the pathway (120). *CFP* deficiencies have been associated with non-recurrent and rapidly progressive fatal meningococcaemia, with around 50% of *CFP* deficient IMD patients infected by Nm serogroups W and Y (121). Notably, the causative Nm serogroup in patient 1034 meningococcal sepsis was serogroup Y. Details of previously published *CFP* deficiencies associated with IMD can be found in Table 1.2.

The predicted deleterious *CFP* variant discovered in patient 1034, c.940+1G>C, results in a splice site loss located immediately after exon 6. ACMG pathogenic criteria were fulfilled with a very strong pathogenic score, PVS1, awarded due to the presence of a null variant (in a gene where LoF is a known mechanism of disease (Table 6.3). As this is a novel variant, not previously described but predicted to have high impact and located in a known disease-causing gene that matches the mode of inheritance and patients' diagnoses, further research to characterise the impact of the novel *CFP* variant will be undertaken.

Table 6.3 *CFP* variant summary discovered in INGENIOS patient 1034

Gene	*CFP*	**Zygosity**	Hemi
dbSNP ID	-	**Inheritance**	XL
HGVS description	NM_002621.2	**Consequence**	Splice site loss
	c.940+1G>C	**Read depth**	52x
		ACMG	Likely pathogenic
Location GRCh38	Chromosome X:	**CADD**	33
(hg19)	g. 47486171C>G	**gnomAD MAF**	-

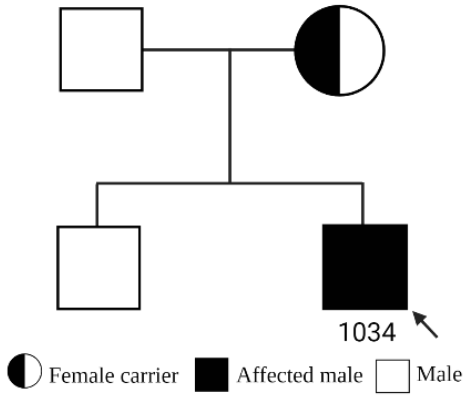


Figure 6.6 Pedigree of INGENIOS patient 1034

This pedigree shows INGENIOS patient 1034, indicated with the arrow, affected with severe meningococcal sepsis, a healthy older brother, and healthy parents. A novel CFP variant, predicted to result in a splice site loss is currently being investigated as the disease-causing variant and was XL-inherited from the mother. Brother and father have not been genetically investigated.

6.4.4 AD *CTLA4* in HSE (1002)

Patient 1002 is a female of north/mid/east European descent who presented with HSV-1 encephalitis (HSE) at 9 years of age. She subsequently presented with a rapid onset of new behavioural changes, 4 months after the initial HSE episode, and was diagnosed with - N-methyl-D-aspartate receptor (NMDAR) encephalitis. There was no known parental history of HSV infections, and the younger brother remained healthy with no significant clinical episodes. The patient's mother had a history of vitiligo and of epilepsy. Anti-NMDAR antibody testing returned a negative result during the initial HSV-1 encephalitis episode but was positive at the subsequent encephalitis episode. Detection of HSV-1 in her cerebrospinal fluid was confirmed with polymerase chain reaction (PCR) in the HSE episode but was negative in the NMDAR encephalitis episode. A magnetic resonance imaging (MRI) brain scan during the initial HSE episode revealed a parenchymal bleed, an area of infarction, and inflamed anteromedial temporal lobes, sub frontal regions, and limbic region. She was treated with intravenous acyclovir to clear the infection, followed by oral valaciclovir continuing prophylaxis. There were no significant new MRI changes when she re-presented with NMDAR encephalitis.

Patient 1002 was discharged from hospital care a month after the NMDAR encephalitis onset with new persisting attention deficit hyperactivity disorder and behaviour management issues.

DNA samples from the full trio ofindex, mother, and father, were obtained and sent for WES and analysis using the designed pipeline was carried out (Figure 6.3). There were no variants of interest revealed by the HSE gene screen and no clear potentially causative variants in genes typically associated with the presenting phenotype found in the PID gene screen. However, an autosomal dominant variant in Cytotoxic T-Lymphocyte Associated Protein 4 (*CTLA4),* was uncovered in the PID gene screen, a missense variant resulting in amino acid change p.G109E (Table 6.4). CTLA-4 is a key negative immune regulator involved in the suppression of proliferation and differentiation of T cells, via regulatory T cells (Tregs) (436, 437). The main mechanism used by CTLA4 to perform this regulation is by preferentially binding CD80 and CD86 with higher affinity, thereby successfully competing with costimulatory receptor CD28 on antigen-presenting cells (APCs) (438, 439). Once bound CTLA-4 performs transendocytosis, whereby it removes CD80 and CD86 from the surface of the APC therefore reducing APC-mediated activation of T cells (440, 441). Autosomal dominant variants in *CTLA4* have been shown to result in haploinsufficiency, impaired ligand binding, or impaired CTLA-4 dimerization, causing immune dysregulation and immunodeficiency in humans (442-445). Clinical presentation of *CTLA4* haploinsufficiency is highly variable, with incomplete penetrance. Findings include a decrease in circulating T and B cells, autoimmune cytopenia, enteropathy, interstitial lung disease, extra-lymphoid lymphocytic infiltration, and recurrent infections (7).

The specific *CTLA4* p.G109E variant uncovered in INGENIOS patient 1002 had been previously published in association with *CTLA4* haploinsufficiency and first described in a female patient of European descent who presented at age 3 with immune dysregulation syndrome including, lymphoproliferation, splenomegaly, hepatomegaly, lymphadenopathy, and both upper and lower respiratory tract infections with *Staphylococcus aureus* and human papillomavirus isolated (444). The INGENIOS MDT consulted with an expert in *CTLA4* haploinsufficiency, Dr Bodo Grimbacher (Freiburg, Germany). After consulting with Dr Grimbacher, whole blood samples of patient 1002 and her mother, who also shares the *CTLA4* p.G109E variant (Figure 6.7), were sent to his team to perform a transendocytosis assay to determine if CTLA-4 function had been affected.

Results of the transendocytosis assay confirmed CTLA4 insufficiency in patient 1002, showing a decrease percentage of intracellular CTLA-4 expression in activated Treg cells of 1002 (24.0%) and her mother (16.6 %) in comparison to the travel control (37.5 %) (Figure 6.8A). A reduced percentage of CTLA4-dependent transendocytosis in Treg cells of 1002 (67.3 %) and her mother (66.4 %) in comparison to the travel control (81.4 %) was also confirmed (Figure 6.8B). This experiment concluded that 1002 and her mother had impaired CTLA-4 biology but not to the full extent usually seen in CTLA-4 insufficient individuals, typically characterized by a complex immune dysregulation syndrome (444) (personal communication with Dr Bodo Grimbacher). This allowed the MDT to consider that there was no clear evidence that *CTLA4* p.G109E was disease-causing in relevance to HSE, but it was noted that there was a confirmed functional impairment that was known to predispose to autoimmune complications, which might include the subsequent NMDAR encephalitis in the index patient, and the presence of vitiligo in the mother. This case represents a potential complex genotype involved in NMDAR encephalitis, and if proven with further research, would represent the first association of *CTLA4* insufficiency in this illness.

Table 6.4 *CTLA4* variant summary discovered in INGENIOS patient 1002

Gene	*CTLA4*	**Zygosity**	Het
dbSNP ID	rs144988077	**Inheritance**	AD
HGVS description	NM_005214.5 c.326G>A, p.G109E	**Consequence**	Missense
		Read depth	246x
		ACMG	Uncertain
Location GRCh38 (hg19)	Chromosome 2 g.203870802G>A	**CADD**	17.58
		gnomAD MAF	0.025

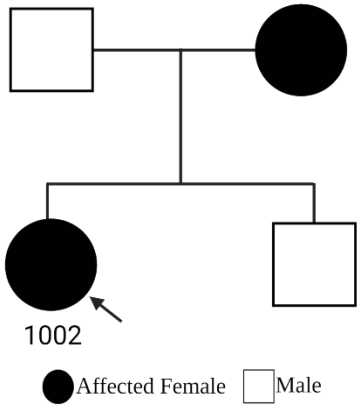


Figure 6.7 Pedigree of INGENIOS patient 1002

This pedigree shows INGENIOS patient 1002, indicated with the arrow, affected with herpes simplex encephalitis (HSE), with all other family members unaffected. A heterozygous *CTLA4* variant of uncertain significance, inherited from the mother as autosomal dominant is currently being functionally investigated as a potential disease-causing or disease-exacerbating variant 1002, mother, and father have been genetically investigated, the brother has not.

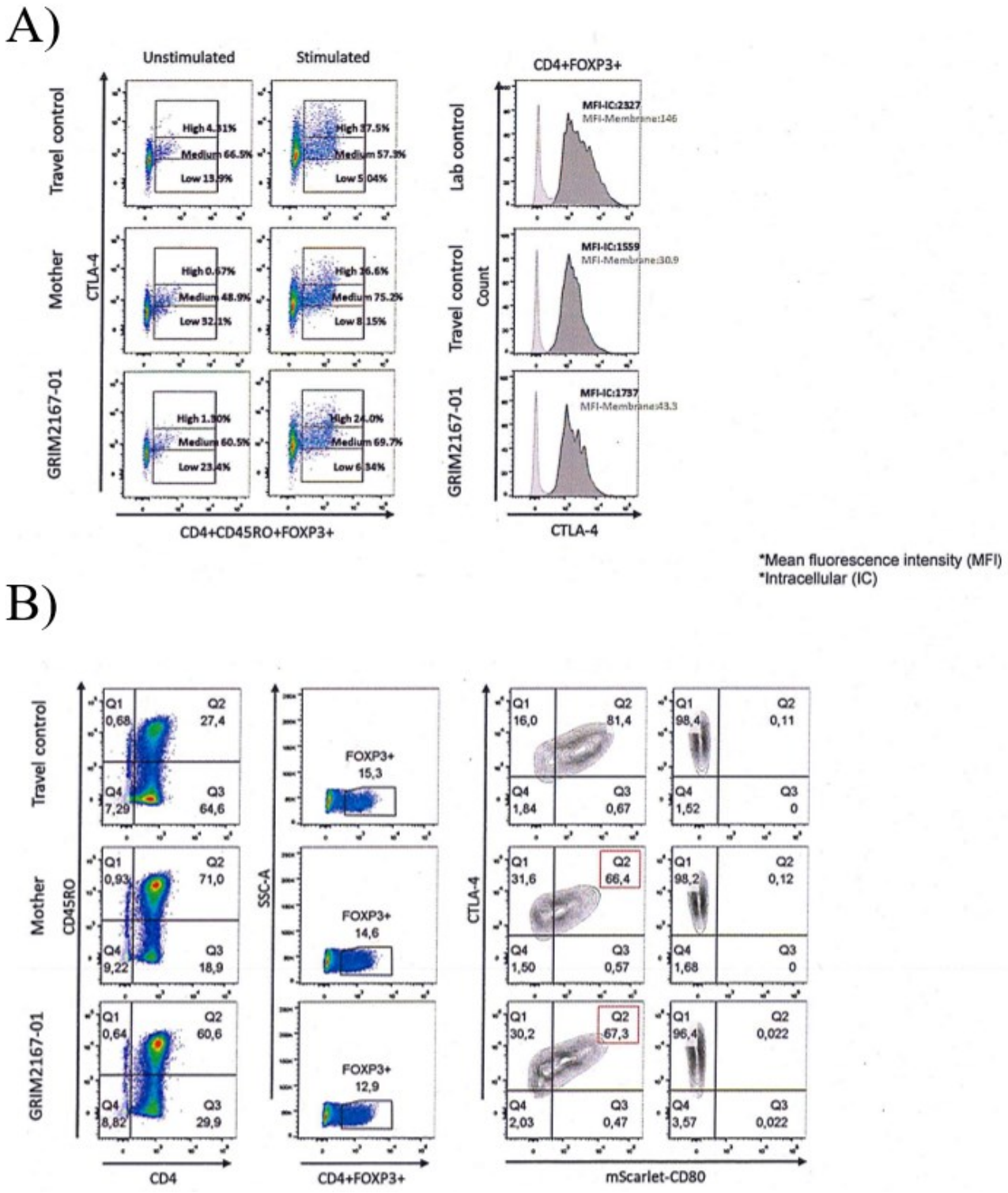


Figure 6.8 ***CTLA4*** **transendocytosis assay**

Performed by Dr Bodo Grimbacher and their team (Freiburg, Germany). 1002, mother of 1002, and a travel control were investigated for A) CTLA4 expression in activated Treg cells and B) CTLA4 function by examining transendocytosis of CD80-mScarlet (a CTLA-4 ligand) by flow cytometry in isolated primary human CD4+ T cells activated with CD3/CD28 beads for 16 hours in the presence of CD80-mScarlet. Cells were gated for using flow cytometry on CD4+CD45RO+FOXP3+ and analysed for mScarlet uptake.

6.3 Discussion

The premise of INGENIOS is to provide a route for personalised genetic and functional analysis to enable delivery of the best possible, well-informed, clinical care. For recruitment, the clinic capitalised on the existing tertiary referral pathway of children with severe infectious conditions to the ICHNT paediatric infectious diseases and paediatric intensive care teams. Children admitted or referred to ICHNT with mainly severe, atypical, recurrent, or persistent infections are considered for recruitment into INGENIOS by which the following investigations can be applied, next generation sequencing of DNA (predominantly WES), thorough genetic investigations, and where appropriate, targeted functional immunological investigations not currently available through the National Health Service (NHS). These functional investigations can include, but are not limited to, investigating cellular responses to immune/inflammatory mediators, or pattern associated molecular patterns (PAMPs), including measurement of cytokine production, proliferation, and cell mortality. Additionally, gene-specific assays can be used, where available, to investigate the impact of variants detected by WES. For instance, the transendocytosis *CTLA4* experiments performed for patient 1002 were able to directly test the impact of the detected *CTLA4* variants (Figure 6.8).

The main strength of INGENIOS is the in-depth genetic investigation. I implemented an analysis pipeline (Figure 6.3) that allows for WES data from a clinically variable cohort to be processed cohesively. Multiple steps make up the pipeline and each was carefully considered and optimised before integrating into the final form. WES processing, initially performed by Dr Bellos, is important to accurately call and align variants, and remove obvious sequencing artefacts that could obscure real data. The current tools used to perform this can be updated as variant calling software improves and develops.

A cut-off MAF of <1% was used to enrich for rare variants that are more likely to cause severe and/or unusual disease. A more stringent cut off of 0.1 % was considered suitable for highly homogeneous and fully penetrant diseases, but the more relaxed MAF of <1% was considered better where there might be disease with incomplete clinical penetrance (19, 20). Consideration of the background patient ethnicity was important for interpretation of filtering on the basis of MAF, as variant frequency within different ethnicities can vary, likely due to genetic drift during population expansion (446).

Next, WES data were screened for known pathogenic variants found in genes associated with PIDs. This step enabled rapid and accurate reporting of known causative variants back to the patient and their clinical care team. However, in cases where known causative variants were not revealed further research was undertaken. The next step was to interrogate the variants found in all known PID genes, a classification list which is updated periodically and reported by the IUIS expert committee on Inborn Errors of Immunity (7). This allowed for the discovery of novel causative variants in genes already associated with the presenting phenotype, as with patient 1034 (Table 6.3 and Figure 6.6). I also curated specialised and exhaustive candidate gene lists for specific clinical presentations – including infections by specified pathogens for which there is available evidence for IEIs that are known to predispose to or influence disease development. These lists were generated by intense and comprehensive literature searches.

Variants that passed through this filter cascade were then evaluated based on their predicted impact on protein coding, using the ACMG classification of interpreted pathogenicity (20), and taking account of the inheritance pattern, and clinical phenotype. The inheritance pattern analysis considered whether the variants detected were expected to be disease-causing in either a monoallelic or biallelic manner and made use of trio data to determine whether the pattern of disease within the family was compatible with the variant inheritance pattern. It is particularly useful to have WES from a trio when disease patterns are multi-generational, as it allows for familial segregation analysis to be explored, for example in the case of patient 1028 (Figure 6.5). It is also important to determine the likelihood of a sibling possessing the same disease-causing variant, for example in the case of patient 1018 his investigations revealed a known pathogenic variant in *STAT3* resulting in severe recurrent pneumococcal pneumonia (Table 6.1 and Figure 6.5). However, it was classified as de novo as neither parent was a carrier, therefore it is unlikely for another child born from the same parents to acquire this specific variant.

The results from the genetic, and functional, investigations are reviewed by clinicians, researchers and clinical scientists in a monthly MDT which lays the foundations for diagnostic reporting in this specialist area. Thus far, INGENIOS has been able to assist in positively identifying, or accurately informing care, for around 10% of recruits (personal communication) and has provided opportunities to investigate potentially new pathways in disease, discover novel variants in known disease associated genes, and contribute to knowledge in the IEI field.

Chapter 7

Chapter 7 Discussion

7.1 Overview

A common feature of the infectious diseases discussed in this thesis is that only a small proportion of the exposed population will develop significant and severe clinical symptoms while the majority will have self-limiting or mild disease. In 2019, prior to the SARS-CoV-2 viral pandemic, the World Health Organisation (WHO) reported over 7.7 million deaths caused by infectious disease worldwide (1). Differences in host susceptibility and predisposition to development of severe clinical manifestations of infectious disease has been investigated for decades and attribution of host genetic factors in infectious disease originated with adoption and twin studies in the twentieth century (447, 448). More modern genetic approaches including WES, linkage analysis, and gene enrichment studies using large case vs control cohorts, have since been able to identify single gene inborn errors of immunity (IEI) in infectious disease (14, 25, 412). IEI's are responsible for more than 485 known clinical syndromes, including many primary immunodeficiencies (7).

Invasive meningococcal disease is rare but has significant mortality with a case fatality rate of 10%, with up to 19% of survivors suffering from life-long severe sequelae with a reduced quality of life (49, 50). Seminal studies by Goldschneider and colleagues in the 1960s demonstrated that establishment of serum bactericidal antibodies acquired against circulating invasive strains and nasopharyngeal carriage was the most important host defence mechanism that correlated to protection against Nm infection (44, 45). Both rare monogenic variants and common polymorphisms have been associated with susceptibility to and severity of IMD (described in Chapter 1), and their discoveries have assisted in elucidating the host susceptibility, pathogenesis of Nm, and aetiology of IMD.

The infectious diseases in this thesis all have significant mortality and/or morbidity. IMD has significant mortality at an overall rate of over 50% in untreated individuals and 10 % in those treated (47, 48). COVID-19 has a case fatality rate of ~1 %, with 20% of patients requiring hospitalisation, (267). *S. aureus* has a case fatality rate of ~25 %(207), *S. pneumoniae* causes invasive disease in ~35 per 100,000 in the infant and over 65 age groups (228, 229), and invasive GAS is responsible for around 500 000 deaths annually worldwide (246-248). Individuals with complement deficiencies, typically autosomal recessive disorders, account for less than 3 % of IMD and are known to be at increased risk of recurrent IMD, as they have a

diminished ability to limit meningococci dissemination (11, 121, 122). This indicates that there are more genetic aetiologies to be discovered that can assist in elucidating disease pathogenesis.

This thesis resulted from the data acquired during EUCLIDS and was positioned to identify novel undescribed genetic aetiologies that underlie invasive bacterial infections, using IMD as the primary model. Whole exome sequencing (WES) was employed as the primary genetic methodology during this study as more than 85 % of all known disease-causing gene variants are located in the protein coding regions of the genome, making it a cost-effective method, in both actual sequencing costs and operational costs, for investigating genetic disorders associated with infectious disease (16). This study also aimed to functionally validate and mechanistically characterise these genetic findings in a relevant disease context. Throughout this project I was a part of a new specialist paediatric clinic that investigated WES, and immunological experiments where necessary, of children with severe or unusual infections, in real time. During this thesis, I also collaborated with the COVIDHGE consortium in a project that came about suddenly with the advent of the SARS-CoV-2 pandemic. The scope of that project was to investigate how genetic variability of loci involved in type I IFN response could contribute to the wide clinical variability in phenotypes observed during widespread SARS-CoV-2 infection. Similarly to other aspects of my thesis, WES and functional characterisation was utilised to investigate this.

7.2 Main findings

Initial gene burden analyses of WES data from EUCLIDS recruits uncovered *JAK2* as a top hit enriched specifically in IMD, and *KRT10* enriched in the overall invasive bacterial disease cohort. The biological functions of *JAK2* are aplenty and its role in cytokine signalling and innate immunity made it a promising candidate to investigate in the context of bacterial infection (299, 300). This thesis has identified at least three gain-of-function *JAK2* variants that constitutively activate the JAK/STAT signalling pathway, p.F556V, p.E846D, and p.L871V, determined by immunoblotting of phosphorylated JAK2 and phosphorylated STAT1, and verified by flow cytometry of increased activated STAT1 expression (Chapter 3). Both F556V, and E846D have been previously described as GoF *JAK2* variants, but only in the specific context of myeloproliferative neoplasms and not in association with infectious disease (308-310). *JAK2* L871V is a novel variant, that has not been previously described. *JAK2* variants discovered in invasive bacterial disease cases, exclusive of IMD, G127A, M187I, V392M, S523T, G571S, and I724T have not yet differed from JAK2 WT under any tested

circumstances. This does not necessarily prove that GoF *JAK2* variants have an exclusive impact only in IMD cases as they may impact *JAK2* function in a way that I have not tested for yet or there may be *JAK2* variants in bacterial disease patients still yet undiscovered that show similar GoF traits to those described here. So far, I have optimised the experiments to be used with new patient samples and, was able to perform these experiments on whole blood *JAK2* V617F samples, V617F is a known GoF variant. Utilising an *in vitro* static adhesion assay to mock imitate an endothelial environment has shown that GoF *JAK2* V617F neutrophils have increased attachment to their endothelial ligands, VCAM1 and ICAM1, at baseline and that both control and V617F neutrophils show a marked increase in attachment upon stimulation with formalin-fixed Nm, *S. aureus, S. pneumoniae,* and GAS. I have also shown that this is a Rap-1 dependent process in V617F neutrophils but not in control neutrophils as they potentially use other pathways independent of Rap1, such as Rho PAR-2 mediated signalling (368) (Chapter 3). At least one *JAK2* IMD patient identified in this cohort was able to be recalled, fortuitously this patient has the E846D variant, shown to be one of the clear GoF variants. New blood samples will be donated shortly to test with in these experiments and I anticipate seeing similar GoF trends in the E846D cells as seen in the V617F cells. The characterisation of this mechanism as potentially being the causative pathway for the development of severe disease in IMD cases would corroborate with previously published data that shows GoF *JAK2* variants to promote venous thrombosis through enhanced adhesion of granulocytes to the endothelia (294). This study shows, for the first time, association of GoF *JAK2* variants in IMD cases and begins to characterise constitutive activation of JAK2, via a neutrophil-mediated pathway, as contributing to poor clinical outcome in IMD. This could result in use of already licensed *JAK2* inhibitors, such as ruxolitinib, to treat patients with meningococcal infection who test positive for *JAK2* GoF variants to prevent the development of severe sepsis.

Previously described genetic aetiologies reported to be associated with IMD, have been implicated in influencing coagulation and thrombosis (Table 1.5). Severe IMD is often characterised by coagulopathy, which can result in the most severe clinical thrombotic complication, purpura fulminans, developing in up to 20 % of cases and can progress into skin necrosis requiring grafting or amputations (51, 187, 188). Polymorphisms in genes involved in regulation of coagulation by the fibrinolytic system have been associated with development of these severe IMD outcomes (51). These include the factor V Leiden mutation (FV^{L}) in *F5*, polymorphisms in the *PROC* promoter region, p.T325I in *CPB2*, and indels in the *SERPINE1* promoter region (Table 1.5) (11). Results from the static adhesion assays suggest that *JAK2*

GoF variants may contribute to increased severity of IMD by causing increased attachment of neutrophils to the endothelium, which lead to tissue damage, constitutive activation of the inflammatory response and can result in sepsis and thrombosis (323, 362, 363).

Further gene enrichment analysis of the EUCLIDS cohort revealed *KRT10* to be enriched in overall invasive bacterial infections, caused by meningococcus, pneumococcus, streptococcus, staphylococcus. Variants in *KRT10* associated with epidermolytic hyperkeratosis and ichthyosis with confetti can result in aberrant KRT10 intermediate filament function, hereby affecting KRT10's role in providing structural integrity in the cytoskeleton (380, 394-398). Epidermolytic hyperkeratosis involves lysis of the epidermal granular layer forming painful and unpleasant blisters, and secondary focal infections with bacteria, typically *Staphylococcus aureus* is common (449). However, these secondary *S. aureus* infections are normally superficial, mostly malodorous, and tend not to result in invasive disease (449). To note, none of the reported variants known to be causative for epidermolytic hyperkeratosis or ichthyosis with confetti are present in our invasive bacterial disease cohort.

KRT10 genetic variability has not been previously associated with invasive infectious disease but wild type KRT10 is known to bind to both *S. aureus* ClfB and *S. pneumonia* PsrP and encourages bacterial colonisation and biofilm formation (225, 382-384). This interaction between KRT10 and bacterial surface proteins encouraged further investigations and supplied insight to form a hypothesis that AD *KRT10* variants could modulate bacterial adhesion and invasion and contribute to pathogenesis. To investigate this, I created 9 *KRT10* missense variants identified in the EUCLIDS cohort and all showed detectable protein via immunoblotting, though variants S42F, R86H, and K334Q showed reduced KRT10 expression in HEK293 cells and S53N, and G575V showed reduced KRT10 expression in HaCaT cells. However, with the exception of G575V expression in HaCaT cells, all other variants showed KRT10-FLAG expression similar to wild-type in both HEK293 and HaCaT cells. Immunofluorescence investigations supports reduced expression of G575V in HaCaT cells but no other KRT10 variants show substantially reduced expression compared to WT, this leads me to believe that the antibody used to directly probe KRT10 in immunoblotting needs to be further investigated. This can be accomplished by redoing an antibody titration experiment to achieve optimal antibody dilution and by probing with other available KRT10 antibodies. Immunofluorescence investigations of subcellular WT KRT10-FLAG localisation also showed a well-defined intermediate filament network spread uniformly throughout the cell that is typical of WT KRT10. Whereas KRT10 S53N, R86H, F102S, K334Q, and G575V appear to

aggregate in the cytosol with an appearance of intracellular punctate appearance (Chapter 4). In future experiments it will be necessary to co-stain with other subcellular markers to determine the exact location of KRT10 expression and localisation.

Principal bacterial assays investigating the adhesion and invasion of staphylococci in cells expressing *KRT10* variants showed significant increased staphylococci-cell association with S42F, R86H, F102S, G124V, and G575V expression in HEK293 cells and S42F, R86H, G124V, and G575V expression in undifferentiated HaCaT cells. Increased number of intracellular staphylococci was also measured in HEK293 cells expressing G124V and G575V, and HaCaT cells expressing S42F and G575V. Notably, KRT10 G575V is the only variant to be significant in both adhesion and invasion in both cell types. As *S. aureus* ClfB is known to bind to the KRT10 C-terminal glycine and serine motifs where G575 is located this may suggest the hypothesis that G575V improves *S. aureus* adhesion and subsequent invasion (383). However, as discussed in Chapter 4, there has been difficulties replicating the significance of these results that will need to be addressed before definitive concluding remarks can be made. Additionally, though patients with known PIDs were excluded from WES selection of the EUCLIDS cohort, I will be undertaking a more thorough investigation of *KRT10* patients to definitively exclude the presence of known IEIs, using the rigorous methodologies developed in Chapter 6. Thus far, genetic variability underlying pneumococcal and staphylococcal diseases are the result of deficiencies arising from IEIs (Chapter 1, section 1.5), the data presented in this thesis shows the first instance where genetic variability in a nonimmune related gene could influence susceptibility to invasive disease. If proven, this could reveal novel therapeutics for the treatment of invasive bacterial disease in *KRT10* patients. For example, it has been shown in patients with epidermolytic hyperkeratosis and ichthyosis caused by *KRT10* mutations that retinoid treatment reduces the formation of KRT10 aggregates and can rescue the filament structure in keratinocytes, though the exact mechanism remains elusive (450, 451). Currently my data suggests that *KRT10* variants discovered in invasive bacterial disease patients form KRT10 aggregates and may play a role in modulating bacterial adhesion and invasion. This could align with the impact of *KRT10* variants in epidermolytic hyperkeratosis and ichthyosis where retinoid treatment were successfully used to rescue mutant KRT10 function, revealing a potential treatment avenue for invasive bacterial disease patients with *KRT10* variants.

IEIs resulting in deficiencies in the TLR3-IFN signalling pathway have been previously associated with viral infections, most notably herpes simplex virus 1 and influenza. The

COVID Human Genetic Effort (COVIDHGE, https://www.covidhge.com) proposed that causative IEIs, in pathways associated with other viral infections, may be involved in development of COVID-19 pneumonia with complete or incomplete penetrance in young and healthy patients. Here we provide evidence that rare *TICAM1* variants discovered in COVID-19 pneumonia patients result in TICAM1 deficiencies within the context of the TLR3-IFN signalling pathway. There was a clear loss of protein expression shown by immunoblotting and immunofluorescence and a marked decrease in IFN-β and NF-κB promoter induction in at least three rare *TICAM1* variants tested, T4I, S60C, and Q392K (Chapter 5). The COVID-19 severe pneumonia patients possessing these three variants did not carry any known variants at other candidate loci known to underlie inborn born errors of immunity (289). As the patients were heterozygous for the mutation, further functional investigation is required to decipher whether this effect is due to haploinsufficiency or dominant negativity.

This data proposes that type I IFN is important in defence against life-threatening COVID-19 pneumonia, as defects in the TLR3-IFN signalling pathway can impair induction of key anti-viral signalling molecules (289). Some key risk factors have been associated with development of severe COVID-19, chief of which are increased age, male sex, obesity, and other pre-existing co-morbidities (267, 408, 409). Additionally, two classes of inborn errors have been associated with severe COVID-19, firstly inborn errors underlying the high production of autoantibodies neutralising type I IFNs which has been reported to account for ~20% of fatal COVID-19 cases in the over-80s population (452). Secondly, monogenic inborn errors of type I IFN immunity, specifically TLR3 and interferon regulatory factor 7 (IRF7)–dependent type I IFN immunity have been described to play a role in susceptibility to and severity of COVID-19 in ~3.5% of patients with life-threatening COVID-19 pneumonia (289). In these patients, autosomal-recessive deficiencies in *IRF-7* and *IFNAR1*, and autosomal-dominant deficiencies in *TLR-3, UNC93B1, TICAM1, TBK1, IRF-3, IRF-7, IFNAR1,* and *IFNAR2* were found to result in a stunting or absence of type I IFN production and amplification. This major collaborative study assisted in elucidating the critical role of TLR3 and type I IFN immunity in limiting SARS-CoV-2 infection (289).

Applying developed genetic investigations to the specialist paediatric clinic which recruited children with severe or unusual infections into INGENIOS contributed to identifying, or accurately informing care, for at least 10% of recruits (Chapter 6). Some case studies presented in this thesis include the discovery of a de novo heterozygous *STAT3* variant, R382Q, in pneumococcal infection which is a known dominant negative variant that results in hyper-

immunoglobulin E syndrome (HIES). Further investigations revealed that bacterial infections in STAT3-HIES are common and recurrent bacterial pneumonia is seen in 80% of STAT3-HIES patients, often complicated by pleural effusion (427, 432, 433). This aligned exactly with the patient phenotype and allowed accurate reporting of a disease-causing phenotype which led to referral to an appropriate immunology team and a possibility of a bone marrow transplant to cure his disease. INGENIOS has also facilitated research of a potentially complex genotype involved in NMDAR encephalitis, and if proven with further functional studies, would represent the first ever association of *CTLA4* insufficiency in this illness.

The main strength of INGENIOS is the in-depth genetic investigation. I implemented an analysis pipeline (Figure 6.3) that allows for WES data from a clinically variable cohort to be processed cohesively. Genetic research is only as good as the technology available and current tools can present some technical limitations, including WES difficulty in sequencing GC-rich repetitive regions and the necessity of using a noncomprehensive capture library (453). However, the current tools used to perform this can be updated as variant calling software improves and develops, while the robust logical concept of variant filtering designed can remain as a solid foundation. INGENIOS has provided opportunities to investigate potentially new pathways in disease, discover novel variants in known disease associated genes, and contribute to knowledge in the IEI field (Chapter 6).

7.3 Study limitations and future recommendations

This study has utilised functional validation of promising gene candidates directed by WES collected from large well phenotype defined patient cohorts as a primary method to discover novel aetiologies in severe and unusual infectious disease. However, while this method provides statistical genetic significance validated by functional data, there are some limitations to consider. Firstly, recruitment of rare disease samples such as invasive meningococcal disease, presents a significant task for the clinical team responsible. Secondly, in relation to recruitment, the need for rigorous phenotyping of clinical outcomes, disease, and demographic data is essential to properly categorise disease groups and perform subsequent analysis. Thirdly, the route, timing, and geographical location of enrolment may present difficulties for future subsequent investigations with the cohort aging and contact details changing. WES analysis revealed 16 *JAK2* variants in a total of 35 patient recruited into EUCLIDs, including 18 IMD cases. However, thus far only one patient, recruited with IMD, has been successfully

recontacted and accommodating in supplying new relevant samples, such as whole blood samples, for further functional characterisation studies.

Technical limitations are also inherent in WES methodologies, for example current WES relies on public databases for known target exome regions and undiscovered regions are not considered, resulting in a noncomprehensive capture library (453). WES also has difficulties in successfully sequencing GC rich regions and repetitive regions with adequate coverage, and it can not accurately detect copy number variants, of which have been previously associated with rare disease (454). The biggest limitation of WES is that only captures ~2% of the entire genome, the protein coding regions, whereas other techniques such as whole genome sequencing captures most if not all of it. However the argument for using WES is that within the 2% of the human genome it covers, over 85% of disease-causing variants are found there (23) and a comparative study showed that WGS in clinical diagnostics is only 0.81 % improved over WES (26), a modest increase that may not offset the raised cost of WGS for some. The high skill level, operational cost, and computational power necessary for WGS and data analysis allows WES to remain a viable choice for genetic studies.

Another restriction in this project was the difficulty in recalling previously recruited patients and precuring new relevant patient samples. To attempt to overcome this plasmids construct were created to continue to investigate genetic variability with functional characterisation assays. However, the transient transfection methods employed during this project can create an exogenous expression system and as the efficiency of transient transfection is not at 100% this can introduce variability in measured results. To overcome this, I have made sure to use cell lines with no detectable endogenous expression of target proteins, determined by immunoblotting. Furthermore, I have generated stable cell lines expressing *JAK2* variants of interest, that, upon sequence confirmation, will be used to continue functional analyses. Other genome editing techniques such as base editing (455) and prime editing (456) using the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) systems can also be utilised to introduce genetic variants of interest in appropriate cell lines.

Research in meningococcal disease lacks any suitable animal models for investigations, as Nm is a human restrictive pathogen, and use of *in vitro* systems can limit the significance of results once transferred to an *in vivo* setting. Neutrophil chemotaxis and recruitment are particularly tricky to investigate as these events occur naturally in the free-flowing system of the human endothelial bloodstream. Investigates of neutrophil β-integrins binding to their ligands,

VCAM1 and ICAM1 have been well established (294, 361, 457, 458), and modifications on these protocols have allowed us to begin to measure the impact of *JAK2* variants on neutrophil adherence, which could promote sepsis as excessive neutrophil activity can induce endothelial cells to shift to a pro-inflammatory and pro-coagulant phenotype (459). Sepsis was one of the major clinical manifestations seen in the EUCLIDS IMD *JAK2* patient cohort and this can be difficult to characterise in an in vitro model. To overcome this restriction, I shall be collaborating with Dr Milka Sarris (University of Cambridge) to attempt to characterise the impact of *JAK2* mutants in vivo using an established zebrafish inflammation model (371). The premise of this pilot work is to prepare *JAK2* zebrafish ortholog (jak2a) RNA that will be used for transgenesis into zebrafish embryos whose lineage have undergone transgenic modification to express fluorescently tagged neutrophils (371, 372). The larvae expressing the jak2a ortholog variants can then be challenged with molecules such as LPS and/or bacteria and real-time imaging of microbe-leukocyte interactions can be performed, including measurement of neutrophil recruitment to the challenge site. I aligned human JAK2 and zebrafish jak2a and identified six of the *JAK2* variants in EUCLIDs patients to have orthologs in zebrafish jak2a, JAK2 L113V, L393V, F556V, R564L, F844L, and L871V to jak2a L90V, L359V, F520V, R528L, F807L, and L834V respectively. I created these jak2a orthologs using site directed mutagenesis (data not shown) and my colleague is currently subcloning the jak2a ortholog inserts into vectors to be used for transgenesis. This is an exciting opportunity to investigate the impact of *JAK2* ortholog variants *in vivo.*

7.4 Concluding remarks

Whole exome sequencing has provided an avenue of research that is patient driven and pushes towards significant findings of rare variants associated with infectious disease. The findings of the work presented in this project highlight the usefulness of WES in directing subsequent functional characterisation in novel genes likely to be causative in disease. This includes the discovery of gain-of-function *JAK2* variants that maybe associated in driving severe disease in IMD. Loss-of-function *TICAM1* variants associated with impaired type I IFN signalling in SARA-CoV-2 viral infection. Finally, potential *KRT10* gain-of-function variants encouraging increased bacterial adhesion and colonisation, that may be associated with a number of invasive bacterial infections. The technologies and methodologies developed during this work has also been applied to improving diagnostic genetic reporting and informed patient care of children with severe and/or unusual infections recruited into the Imperial College Biomedical Research Council funded specialist paediatric clinic and accompanying study, INGENIOS. In this work, efforts were made to understand the impact of novel genetic aetiologies in susceptibility to and severity of infectious disease, discovered by whole exome sequencing and functionally validated.

References

1. Organisation WH. The top 10 causes of death. . 2019.
2. Lederberg J. J. B. S. Haldane (1949) on infectious disease and evolution. Genetics. 1999;153(1):1-3.
3. Liston A, Humblet-Baron S, Duffy D, Goris A. Human immune diversity: from evolution to modernity. Nat Immunol. 2021;22(12):1479-89.
4. Li Y, Oosting M, Smeekens SP, Jaeger M, Aguirre-Gamboa R, Le KTT, et al. A Functional Genomics Approach to Understand Variation in Cytokine Production in Humans. Cell. 2016;167(4):1099-110 e14.
5. Meyer D, VR CA, Bitarello BD, DY CB, Nunes K. A genomic perspective on HLA evolution. Immunogenetics. 2018;70(1):5-27.
6. Carrington M, Nelson GW, Martin MP, Kissner T, Vlahov D, Goedert JJ, et al. HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. Science. 1999;283(5408):1748-52.
7. Bousfiha A, Moundir A, Tangye SG, Picard C, Jeddane L, Al-Herz W, et al. The 2022 Update of IUIS Phenotypical Classification for Human Inborn Errors of Immunity. J Clin Immunol. 2022;42(7):1508-20.
8. Picard C, Casanova JL, Puel A. Infectious diseases in patients with IRAK-4, MyD88, NEMO, or IkappaBalpha deficiency. Clin Microbiol Rev. 2011;24(3):490-7.
9. Picard C, von Bernuth H, Ghandil P, Chrabieh M, Levy O, Arkwright PD, et al. Clinical features and outcome of patients with IRAK-4 and MyD88 deficiency. Medicine (Baltimore). 2010;89(6):403-25.
10. Bustamante J. Mendelian susceptibility to mycobacterial disease: recent discoveries. Hum Genet. 2020;139(6-7):993-1000.
11. Hodeib S, Herberg JA, Levin M, Sancho-Shimizu V. Human genetics of meningococcal infections. Hum Genet. 2020;139(6-7):961-80.
12. Sancho-Shimizu V, Perez de Diego R, Lorenzo L, Halwani R, Alangari A, Israelsson E, et al. Herpes simplex encephalitis in children with autosomal recessive and dominant TRIF deficiency. J Clin Invest. 2011;121(12):4889-902.
13. Rabbee N, Betensky RA. Power calculations for familial aggregation studies. Genet Epidemiol. 2004;26(4):316-27.
14. Amos W, Driscoll E, Hoffman JI. Candidate genes versus genome-wide associations: which are better for detecting genetic susceptibility to infectious disease? Proc Biol Sci. 2011;278(1709):1183-8.
15. Arias TD, Jorge LF, Barrantes R. Uses and misuses of definitions of genetic polymorphism. A perspective from population pharmacogenetics. Br J Clin Pharmacol. 1991;31(1):117-9.
16. Belkadi A, Bolze A, Itan Y, Cobat A, Vincent QB, Antipenko A, et al. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. Proc Natl Acad Sci U S A. 2015;112(17):5473-8.
17. Krumm N, Sudmant PH, Ko A, O'Roak BJ, Malig M, Coe BP, et al. Copy number variation detection and genotyping from exome sequence data. Genome Res. 2012;22(8):1525-32.
18. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, et al. Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet. 2011;12(11):745-55.
19. Whiffin N, Minikel E, Walsh R, O'Donnell-Luria AH, Karczewski K, Ing AY, et al. Using high-resolution variant frequencies to empower clinical genome interpretation. Genet Med. 2017;19(10):1151-8.
20. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405-24.
21. Hamamy H, Antonarakis SE, Cavalli-Sforza LL, Temtamy S, Romeo G, Kate LP, et al. Consanguineous marriages, pearls and perils: Geneva International Consanguinity Workshop Report. Genet Med. 2011;13(9):841-7.

22. Meyts I, Bosch B, Bolze A, Boisson B, Itan Y, Belkadi A, et al. Exome and genome sequencing for inborn errors of immunity. J Allergy Clin Immunol. 2016;138(4):957-69.
23. Petersen BS, Fredrich B, Hoeppner MP, Ellinghaus D, Franke A. Opportunities and challenges of whole-genome and -exome sequencing. BMC Genet. 2017;18(1):14.
24. Starokadomskyy P, Gemelli T, Rios JJ, Xing C, Wang RC, Li H, et al. DNA polymerase-alpha regulates the activation of type I interferons through cytosolic RNA:DNA synthesis. Nat Immunol. 2016;17(5):495-504.
25. Davila S, Wright VJ, Khor CC, Sim KS, Binder A, Breunis WB, et al. Genome-wide association study identifies variants in the CFH region associated with host susceptibility to meningococcal disease. Nat Genet. 2010;42(9):772-6.
26. Meienberg J, Bruggmann R, Oexle K, Matyas G. Clinical sequencing: is WGS the better WES? Hum Genet. 2016;135(3):359-62.
27. Tønjum T, Putten Jv. Infectious Diseases (Fourth Edition). Cohen J, Powderly WG, Opal SM, editors: Elsevier; 2017.
28. Weichselbaum A. Ueber die Aetiologie der akuten Meningitis cerebro-spinalis. Fortschr Med. 1887.
29. Vieusseux M. Mémoire sur la maladie qui a regné a Genêve au printemps de 1805. J Med Chir Pharmacol. 1805.
30. McGahey K. Report on an outbreak of epidemic cerebrospinal meningitis in Zungeru during February and March 1905. Journal of Tropical Medicine. 1905;8:210–6.
31. McGahey K. Report of the outbreak of cerebrospinal meningitis at Yola. Journal of Tropical Medicine. 1905;8:216-7.
32. Horn A. Report on an investigation of cerebro-spinal fever in the northern territories of the Gold Coast in 1908. Journal of Tropical Medicine and Hygiene. 1908;11:358–65.
33. Stephens DS, Hoffman LH, McGee ZA. Interaction of Neisseria meningitidis with human nasopharyngeal mucosa: attachment and entry into columnar epithelial cells. J Infect Dis. 1983;148(3):369-76.
34. Hayakawa K, Itoda I, Shimuta K, Takahashi H, Ohnishi M. Urethritis caused by novel Neisseria meningitidis serogroup W in man who has sex with men, Japan. Emerg Infect Dis. 2014;20(9):1585-7.
35. Kanemitsu N, Hayashi I, Satoh N, Hayakawa T, Mitsuya H, Hayase Y, et al. Acute urethritis caused by Neisseria meningitidis. Int J Urol. 2003;10(6):346-7.
36. Rouphael NG, Stephens DS. Neisseria meningitidis: biology, microbiology, and epidemiology. Methods Mol Biol. 2012;799:1-20.
37. Serruto D, Bottomley MJ, Ram S, Giuliani MM, Rappuoli R. The new multicomponent vaccine against meningococcal serogroup B, 4CMenB: immunological, functional and structural characterization of the antigens. Vaccine. 2012;30 Suppl 2(0 2):B87-97.
38. Virji M. Pathogenic neisseriae: surface modulation, pathogenesis and infection control. Nat Rev Microbiol. 2009;7(4):274-86.
39. Virji M, Makepeace K, Ferguson DJ, Watt SM. Carcinoembryonic antigens (CD66) on epithelial cells and neutrophils are receptors for Opa proteins of pathogenic neisseriae. Mol Microbiol. 1996;22(5):941-50.
40. Cartwright KAV, Stuart JM, Jones DM, Noah ND. The Stonehouse survey: nasopharyngeal carriage of meningococci and Neisseria lactamica. Epidemiology and Infection. 1987;99:591-601.
41. Caugant DA, Hoiby EA, Rosenqvist E, Froholm LO, Selander RK. Transmission of Neisseria meningitidis among asymptomatic military recruits and antibody analysis. Epidemiol Infect. 1992;109(2):241-53.
42. Kremastinou J, Tzanakaki G, Levidiotou S, Markou F, Themeli E, Voyiatzi A, et al. Carriage of Neisseria meningitidis and Neisseria lactamica in northern Greece. FEMS Immunol Med Microbiol. 2003;39(1):23-9.
43. Tyski S, Grzybowska W, Dulny G, Berthelsen L, Lind I. Phenotypical and genotypical characterization of Neisseria meningitidis carrier strains isolated from Polish recruits in 1998. Eur J Clin Microbiol Infect Dis. 2001;20(5):350-3.
44. Goldschneider I, Gotschlich EC, Artenstein MS. Human immunity to the meningococcus. II. Development of natural immunity. J Exp Med. 1969;129(6):1327-48.

45. Goldschneider I, Gotschlich EC, Artenstein MS. Human Immunity To The Meningococcus. I. The Role of Humoral Antibodies J Exp Med. 1969;129(6):1307-26.
46. Jones GR, Christodoulides M, Brooks JL, Miller AR, Cartwright KA, Heckels JE. Dynamics of carriage of Neisseria meningitidis in a group of military recruits: subtype stability and specificity of the immune response following colonization. J Infect Dis. 1998;178(2):451-9.
47. Edwards KM, Jones LM, Stephens DS. Clinical features of mild systemic meningococcal disease with characterization of bacterial isolates. Clin Pediatr (Phila). 1985;24(11):617-20.
48. Caugant DA, Maiden MC. Meningococcal carriage and disease--population biology and evolution. Vaccine. 2009;27 Suppl 2:B64-70.
49. Cohn AC, MacNeil JR, Clark TA, Ortega-Sanchez IR, Briere EZ, Meissner HC, et al. Prevention and control of meningococcal disease: recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Recomm Rep. 2013;62(RR-2):1-28.
50. Pace D, Pollard AJ. Meningococcal disease: clinical presentation and sequelae. Vaccine. 2012;30 Suppl 2:B3-9.
51. Lecuyer H, Borgel D, Nassif X, Coureuil M. Pathogenesis of meningococcal purpura fulminans. Pathog Dis. 2017;75(3).
52. Nadel S. Treatment of Meningococcal Disease. J Adolesc Health. 2016;59(2 Suppl):S21-8.
53. Coureuil M, Join-Lambert O, Lecuyer H, Bourdoulous S, Marullo S, Nassif X. Mechanism of meningeal invasion by Neisseria meningitidis. Virulence. 2012;3(2):164-72.
54. Yazdankhah SP, Caugant DA. Neisseria meningitidis: an overview of the carriage state. J Med Microbiol. 2004;53(Pt 9):821-32.
55. Harrison LH, Trotter CL, Ramsay ME. Global epidemiology of meningococcal disease. Vaccine. 2009;27 Suppl 2:B51-63.
56. Organization WH. Meningococcal meningitis. World Health Organization 2018.
57. Jafri RZ, Ali A, Messonnier NE, Tevi-Benissan C, Durrheim D, Eskola J, et al. Global epidemiology of invasive meningococcal disease. Popul Health Metr. 2013;11(1):17.
58. Stephens DS, Greenwood B, Brandtzaeg P. Epidemic meningitis, meningococcaemia, and Neisseria meningitidis. The Lancet. 2007;369(9580):2196-210.
59. Parikh SR, Campbell H, Gray SJ, Beebeejaun K, Ribeiro S, Borrow R, et al. Epidemiology, clinical presentation, risk factors, intensive care admission and outcomes of invasive meningococcal disease in England, 2010-2015. Vaccine. 2018;36(26):3876-81.
60. Wang JF, Caugant DA, Li X, Hu X, Poolman JT, Crowe BA, et al. Clonal and antigenic analysis of serogroup A Neisseria meningitidis with particular reference to epidemiological features of epidemic meningitis in the People's Republic of China. Infect Immun. 1992;60(12):5267-82.
61. Greenwood BM, Bradley AK, Smith AW, Wall RA. Mortality from meningococcal disease during an epidemic in The Gambia, West Africa. Trans R Soc Trop Med Hyg. 1987;81(4):536-8.
62. Mazamay S, Guegan JF, Diallo N, Bompangue D, Bokabo E, Muyembe JJ, et al. An overview of bacterial meningitis epidemics in Africa from 1928 to 2018 with a focus on epidemics "outside-the-belt". BMC Infect Dis. 2021;21(1):1027.
63. Lingappa JR, Al-Rabeah AM, Hajjeh R, Mustafa T, Fatani A, Al-Bassam T, et al. Serogroup W-135 meningococcal disease during the Hajj, 2000. Emerg Infect Dis. 2003;9(6):665-71.
64. Booy R, Gentile A, Nissen M, Whelan J, Abitbol V. Recent changes in the epidemiology of Neisseria meningitidis serogroup W across the world, current vaccination policy choices and possible future strategies. Hum Vaccin Immunother. 2019;15(2):470-80.
65. Toros B, Hedberg ST, Unemo M, Jacobsson S, Hill DM, Olcen P, et al. Genome-Based Characterization of Emergent Invasive Neisseria meningitidis Serogroup Y Isolates in Sweden from 1995 to 2012. J Clin Microbiol. 2015;53(7):2154-62.
66. Oldfield NJ, Harrison OB, Bayliss CD, Maiden MC, Ala'Aldeen DA, Turner DP. Genomic Analysis of Serogroup Y Neisseria meningitidis Isolates Reveals Extensive Similarities Between Carriage-Associated and Disease-Associated Organisms. J Infect Dis. 2016;213(11):1777-85.
67. Whitney AM, Coulson GB, von Gottberg A, Block C, Keller N, Mayer LW, et al. Genotypic comparison of invasive Neisseria meningitidis serogroup Y isolates from the United States, South Africa, and Israel, isolated from 1999 through 2002. J Clin Microbiol. 2009;47(9):2787-93.
68. Djibo S, Nicolas P, Alonso JM, Djibo A, Couret D, Riou JY, et al. Outbreaks of serogroup X meningococcal meningitis in Niger 1995-2000. Trop Med Int Health. 2003;8(12):1118-23.

69. Chen C, Zhang TG, He JG, Wu J, Chen LJ, Liu JF, et al. A first meningococcal meningitis case caused by serogroup X Neisseria meningitidis strains in China. Chin Med J (Engl). 2008;121(7):664-6.
70. Fazio C, Starnino S, Solda MD, Sofia T, Neri A, Mastrantonio P, et al. Neisseria meningitidis serogroup X sequence type 2888, Italy. Emerg Infect Dis. 2010;16(2):359-60.
71. Parikh SR, Campbell H, Bettinger JA, Harrison LH, Marshall HS, Martinon-Torres F, et al. The everchanging epidemiology of meningococcal disease worldwide and the potential for prevention through vaccination. J Infect. 2020;81(4):483-98.
72. Beebeejaun K, Parikh SR, Campbell H, Gray S, Borrow R, Ramsay ME, et al. Invasive meningococcal disease: Timing and cause of death in England, 2008-2015. J Infect. 2020;80(3):286-90.
73. Subbarao S, Campbell H, Ribeiro S, Clark SA, Lucidarme J, Ramsay M, et al. Invasive Meningococcal Disease, 2011-2020, and Impact of the COVID-19 Pandemic, England. Emerg Infect Dis. 2021;27(9):2495-7.
74. Moore AE, MacNeil JR, Wang X, Joseph SJ, Lorentzson L, Thomas S, et al. Emergence of Localized Serogroup W Meningococcal Disease in the United States - Georgia, 2006-2016. MMWR Morb Mortal Wkly Rep. 2018;67(32):894-7.
75. Barroso DE, Castineiras TM, Cabral AC, Vicente AC, Rebelo MC, Cerqueira EO, et al. Neisseria meningitidis ST-11 clonal complex bearing capsule serogroups B, C, or W in Brazil. J Infect. 2013;66(6):547-50.
76. Ambrosio L, Neri A, Fazio C, Rossolini GM, Vacca P, Riccobono E, et al. Genomic analysis of Neisseria meningitidis carriage isolates during an outbreak of serogroup C clonal complex 11, Tuscany, Italy. PLoS One. 2019;14(5):e0217500.
77. Feagins AR, Sadji AY, Topaz N, Itsko M, Halatoko JWA, Dzoka A, et al. Neisseria meningitidis Serogroup C Clonal Complex 10217 Outbreak in West Kpendjal Prefecture, Togo 2019. Microbiol Spectr. 2022;10(2):e0192321.
78. Rosenstein NE, Perkins BA, Stephens DS, Popovic T, Hughes JM. Meningococcal disease. N Engl J Med. 2001;344(18):1378-88.
79. Brandtzaeg P, Kierulf P, Gaustad P, Skulberg A, Bruun JN, Halvorsen S, et al. Plasma endotoxin as a predictor of multiple organ failure and death in systemic meningococcal disease. J Infect Dis. 1989;159(2):195-204.
80. Kimberlin DW, Jackson MA, Long SS, Brady MT. Menigococcal infections: American Academy of Pediatrics Committee on Infectious Diseases; 2018.
81. Wyle FA, Artenstein MS, Brandt BL, Tramont EC, Kasper DL, Altieri PL, et al. Immunologic response of man to group B meningococcal polysaccharide vaccines. J Infect Dis. 1972;126(5):514-21.
82. Pizza M, Scarlato V, Masignani V, Giuliani MM, Arico B, Comanducci M, et al. Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. Science. 2000;287(5459):1816-20.
83. Martinon-Torres F, Banzhoff A, Azzari C, De Wals P, Marlow R, Marshall H, et al. Recent advances in meningococcal B disease prevention: real-world evidence from 4CMenB vaccination. J Infect. 2021;83(1):17-26.
84. Ladhani SN, Andrews N, Parikh SR, Campbell H, White J, Edelstein M, et al. Vaccination of Infants with Meningococcal Group B Vaccine (4CMenB) in England. N Engl J Med. 2020;382(4):309-17.
85. Marshall HS, McMillan M, Koehler AP, Lawrence A, Sullivan TR, MacLennan JM, et al. Meningococcal B Vaccine and Meningococcal Carriage in Adolescents in Australia. N Engl J Med. 2020;382(4):318-27.
86. Pathan N, Faust SN, Levin M. Pathophysiology of meningococcal meningitis and septicaemia. Arch Dis Child. 2003;88(7):601-7.
87. Chow JC, Young DW, Golenbock DT, Christ WJ, Gusovsky F. Toll-like receptor-4 mediates lipopolysaccharide-induced signal transduction. J Biol Chem. 1999;274(16):10689-92.
88. Ingalls RR, Lien E, Golenbock DT. Membrane-associated proteins of a lipopolysaccharide-deficient mutant of Neisseria meningitidis activate the inflammatory response through toll-like receptor 2. Infect Immun. 2001;69(4):2230-6.

89. Mogensen TH, Paludan SR, Kilian M, Ostergaard L. Live Streptococcus pneumoniae, Haemophilus influenzae, and Neisseria meningitidis activate the inflammatory response through Toll-like receptors 2, 4, and 9 in species-specific patterns. J Leukoc Biol. 2006;80(2):267-77.
90. Waage A, Brandtzaeg P, Halstensen A, Kierulf P, Espevik T. The complex pattern of cytokines in serum from patients with meningococcal septic shock. Association between interleukin 6, interleukin 1, and fatal outcome. J Exp Med. 1989;169(1):333-8.
91. Waage A, Halstensen A, Espevik T. Association between tumour necrosis factor in serum and fatal outcome in patients with meningococcal disease. Lancet. 1987;1(8529):355-7.
92. Waage A, Halstensen A, Shalaby R, Brandtzaeg P, Kierulf P, Espevik T. Local production of tumor necrosis factor alpha, interleukin 1, and interleukin 6 in meningococcal meningitis. Relation to the inflammatory response. J Exp Med. 1989;170(6):1859-67.
93. Hack CE, De Groot ER, Felt-Bersma RJ, Nuijens JH, Strack Van Schijndel RJ, Eerenberg-Belmer AJ, et al. Increased plasma levels of interleukin-6 in sepsis. Blood. 1989;74(5):1704-10.
94. Hu X, Li J, Fu M, Zhao X, Wang W. The JAK/STAT signaling pathway: from bench to clinic. Signal Transduct Target Ther. 2021;6(1):402.
95. Mantovani A, Cassatella MA, Costantini C, Jaillon S. Neutrophils in the activation and regulation of innate and adaptive immunity. Nat Rev Immunol. 2011;11(8):519-31.
96. Stephens DS. Biology and pathogenesis of the evolutionarily successful, obligate human bacterium Neisseria meningitidis. Vaccine. 2009;27 Suppl 2(Suppl 2):B71-7.
97. Morgan BP, Marchbank KJ, Longhi MP, Harris CL, Gallimore AM. Complement: central to innate immunity and bridging to adaptive responses. Immunol Lett. 2005;97(2):171-9.
98. Papayannopoulos V. Neutrophil extracellular traps in immunity and disease. Nat Rev Immunol. 2018;18(2):134-47.
99. Lappann M, Danhof S, Guenther F, Olivares-Florez S, Mordhorst IL, Vogel U. In vitro resistance mechanisms of Neisseria meningitidis against neutrophil extracellular traps. Mol Microbiol. 2013;89(3):433-49.
100. Flexner S. THE RESULTS OF THE SERUM TREATMENT IN THIRTEEN HUNDRED CASES OF EPIDEMIC MENINGITIS. J Exp Med. 1913;17(5):553-76.
101. Flexner S, Jobling JW. An Analysis of Four Hundred Cases of Epidemic Meningitis Treated with the Anti-Meningitis Serum. J Exp Med. 1908;10(5):690-733.
102. Lewis LA, Ram S, Prasad A, Gulati S, Getzlaff S, Blom AM, et al. Defining targets for complement components C4b and C3b on the pathogenic neisseriae. Infect Immun. 2008;76(1):339-50.
103. Heesterbeek DA, Bardoel BW, Parsons ES, Bennett I, Ruyken M, Doorduijn DJ, et al. Bacterial killing by complement requires membrane attack complex formation via surface-bound C5 convertases. EMBO J. 2019;38(4).
104. Janeway CAJ, Travers P, Walport M, al. e. The complement system and innate immunity. Immunobiology: The Immune System in Health and Disease New York, USA: Garland Science; 2001.
105. Ross SC, Rosenthal PJ, Berberich HM, Densen P. Killing of Neisseria meningitidis by human neutrophils: implications for normal and complement-deficient individuals. J Infect Dis. 1987;155(6):1266-75.
106. Figueroa J, Andreoni J, Densen P. Complement deficiency states and meningococcal disease. Immunol Res. 1993;12(3):295-311.
107. Kim JJ, Mandrell RE, Griffiss JM. Neisseria lactamica and Neisseria meningitidis share lipooligosaccharide epitopes but lack common capsular and class 1, 2, and 3 protein epitopes. Infect Immun. 1989;57(2):602-8.
108. Robinson K, Neal KR, Howard C, Stockton J, Atkinson K, Scarth E, et al. Characterization of humoral and cellular immune responses elicited by meningococcal carriage. Infect Immun. 2002;70(3):1301-9.
109. Palmeira P, Quinello C, Silveira-Lessa AL, Zago CA, Carneiro-Sampaio M. IgG placental transfer in healthy and pathological pregnancies. Clin Dev Immunol. 2012;2012:985646.
110. Pollard AJ, Galassini R, van der Voort EM, Booy R, Langford P, Nadel S, et al. Humoral immune responses to Neisseria meningitidis in children. Infect Immun. 1999;67(5):2441-51.
111. Janeway CA, Jr. How the immune system protects the host from infection. Microbes Infect. 2001;3(13):1167-71.

112. Mantis NJ, Rol N, Corthesy B. Secretory IgA's complex roles in immunity and mucosal homeostasis in the gut. Mucosal Immunol. 2011;4(6):603-11.
113. Woof JM, Russell MW. Structure and function relationships in IgA. Mucosal Immunol. 2011;4(6):590-7.
114. Lim D, Gewurz A, Lint TF, Ghaze M, Sepheri B, Gewurz H. Absence of the sixth component of complement in a patient with repeated episodes of meningococcal meningitis. J Pediatr. 1976;89(1):42-7.
115. Petersen BH, Graham JA, Brooks GF. Human deficiency of the eighth component of complement. The requirement of C8 for serum Neisseria gonorrhoeae bactericidal activity. J Clin Invest. 1976;57(2):283-90.
116. Mashbat B, Bellos E, Hodeib S, Bidmos F, Thwaites RS, Lu Y, et al. A Rare Mutation in SPLUNC1 Affects Bacterial Adherence and Invasion in Meningococcal Disease. Clin Infect Dis. 2020;70(10):2045-53.
117. Callaghan MJ, Rockett K, Banner C, Haralambous E, Betts H, Faust S, et al. Haplotypic diversity in human CEACAM genes: effects on susceptibility to meningococcal disease. Genes Immun. 2008;9(1):30-7.
118. Jack DL, Cole J, Naylor SC, Borrow R, Kaczmarski EB, Klein NJ, et al. Genetic Polymorphism of the Binding Domain of Surfactant Protein–A2 Increases Susceptibility to Meningococcal Disease. Clinical Infectious Diseases. 2006;43:1426-33.
119. Mashbat B, Bellos E, Hodeib S, Bidmos F, Thwaites RS, Lu Y, et al. A rare mutation in SPLUNC1 underlies meningococcal disease affecting bacterial adherence and invasion. Clin Infect Dis. 2019.
120. Lewis LA, Ram S. Meningococcal disease and the complement system. Virulence. 2014;5(1):98-126.
121. Figueroa JE, Densen P. Infectious diseases associated with complement deficiencies. Clin Microbiol Rev. 1991;4(3):359-95.
122. Fijen CA, Kuijper EJ, te Bulte MT, Daha MR, Dankert J. Assessment of complement deficiency in patients with meningococcal disease in The Netherlands. Clin Infect Dis. 1999;28(1):98-105.
123. Martinon-Torres F, Png E, Khor CC, Davila S, Wright VJ, Sim KS, et al. Natural resistance to Meningococcal Disease related to CFH loci: Meta-analysis of genome-wide association studies. Sci Rep. 2016;6:35842.
124. Schneider MC, Prosser BE, Caesar JJ, Kugelberg E, Li S, Zhang Q, et al. Neisseria meningitidis recruits factor H using protein mimicry of host carbohydrates. Nature. 2009;458(7240):890-3.
125. Westberg J, Fredrikson GN, Truedsson L, Sjoholm AG, Uhlen M. Sequence-based analysis of properdin deficiency: identification of point mutations in two phenotypic forms of an X-linked immunodeficiency. Genomics. 1995;29(1):1-8.
126. Fredrikson GN, Westberg J, Kuijper EJ, Tijssen CC, Sjoholm AG, Uhlen M, et al. Molecular characterization of properdin deficiency type III: dysfunction produced by a single point mutation in exon 9 of the structural gene causing a tyrosine to aspartic acid interchange. J Immunol. 1996;157(8):3666-71.
127. van den Bogaard R, Fijen CA, Schipper MG, de Galan L, Kuijper EJ, Mannens MM. Molecular characterisation of 10 Dutch properdin type I deficient families: mutation analysis and X-inactivation studies. Eur J Hum Genet. 2000;8(7):513-8.
128. Helminen M, Seitsonen S, Jarva H, Meri S, Jarvela IE. A novel mutation W388X underlying properdin deficiency in a Finnish family. Scand J Immunol. 2012;75(4):445-8.
129. Marujo F, Costa LC, Duarte R, Brito MJ, Cordeiro A, Neves C, et al. Invasive Meningococcal Disease Unraveling a Novel Mutation in the C5 Gene in a Portuguese Family. Pediatr Infect Dis J. 2019;38(4):416-8.
130. Owen EP, Wurzner R, Leisegang F, Rizkallah P, Whitelaw A, Simpson J, et al. A complement C5 gene mutation, c.754G>A:p.A252T, is common in the Western Cape, South Africa and found to be homozygous in seven percent of Black African meningococcal disease cases. Mol Immunol. 2015;64(1):170-6.
131. Wang X, Fleischer DT, Whitehead WT, Haviland DL, Rosenfeld SI, Leddy JP, et al. Inherited human complement C5 deficiency. Nonsense mutations in exons 1 (Gln1 to Stop) and 36 (Arg1458 to

Stop) and compound heterozygosity in three African-American families. J Immunol. 1995;154(10):5464-71.
132. Pfarr N, Prawitt D, Kirschfink M, Schroff C, Knuf M, Habermehl P, et al. Linking C5 deficiency to an exonic splicing enhancer mutation. J Immunol. 2005;174(7):4172-7.
133. Delgado-Cervino E, Fontan G, Lopez-Trascasa M. C5 complement deficiency in a Spanish family. Molecular characterization of the double mutation responsible for the defect. Mol Immunol. 2005;42(1):105-11.
134. Parham KL, Roberts A, Thomas A, Wurzner R, Henderson HE, Potter PC, et al. Prevalence of mutations leading to complete C6 deficiency (C6Q0) in the Western Cape, South Africa and detection of novel mutations leading to C6Q0 in an Irish family. Mol Immunol. 2007;44(10):2756-60.
135. Nishizaka H, Horiuchi T, Zhu ZB, Fukumori Y, Nagasawa K, Hayashi K, et al. Molecular bases for inherited human complement component C6 deficiency in two unrelated individuals. J Immunol. 1996;156(6):2309-15.
136. Hobart MJ, Fernie BA, Fijen KA, Orren A. The molecular basis of C6 deficiency in the western Cape, South Africa. Hum Genet. 1998;103(4):506-12.
137. Zhu ZB, Totemchokchyakarn K, Atkinson TP, Volanakis JE. Molecular defects leading to human complement component C6 deficiency in an African-American family. Clin Exp Immunol. 1998;111(1):91-6.
138. Barroso S, Lopez-Trascasa M, Merino D, Alvarez AJ, Nunez-Roldan A, Sanchez B. C7 deficiency and meningococcal infection susceptibility in two spanish families. Scand J Immunol. 2010;72(1):38-43.
139. Nishizaka H, Horiuchi T, Zhu ZB, Fukumori Y, Volanakis JE. Genetic bases of human complement C7 deficiency. J Immunol. 1996;157(9):4239-43.
140. Ki CS, Kim JW, Kim HJ, Choi SM, Ha GY, Kang HJ, et al. Two novel mutations in the C7 gene in a Korean patient with complement C7 deficiency. J Korean Med Sci. 2005;20(2):220-4.
141. Fernie BA, Orren A, Sheehan G, Schlesinger M, Hobart MJ. Molecular bases of C7 deficiency: three different defects. J Immunol. 1997;159(2):1019-26.
142. Barroso S, Sanchez B, Alvarez AJ, Lopez-Trascasa M, Lanuza A, Luque R, et al. Complement component C7 deficiency in two Spanish families. Immunology. 2004;113(4):518-23.
143. Barroso S, Rieubland C, Jose alvarez A, Lopez-Trascasa M, Bart PA, Nunez-Roldan A, et al. Molecular defects of the C7 gene in two patients with complement C7 deficiency. Immunology. 2006;118(2):257-60.
144. Dellepiane RM, Dell'Era L, Pavesi P, Macor P, Giordano M, De Maso L, et al. Invasive meningococcal disease in three siblings with hereditary deficiency of the 8(th) component of complement: evidence for the importance of an early diagnosis. Orphanet J Rare Dis. 2016;11(1):64.
145. Saucedo L, Ackermann L, Platonov AE, Gewurz A, Rakita RM, Densen P. Delineation of additional genetic bases for C8 beta deficiency. Prevalence of null alleles and predominance of C-->T transition in their genesis. J Immunol. 1995;155(10):5022-8.
146. Arnold DF, Roberts AG, Thomas A, Ferry B, Morgan BP, Chapel H. A novel mutation in a patient with a deficiency of the eighth component of complement associated with recurrent meningococcal meningitis. J Clin Immunol. 2009;29(5):691-5.
147. Kira R, Ihara K, Takada H, Gondo K, Hara T. Nonsense mutation in exon 4 of human complement C9 gene is the major cause of Japanese complement C9 deficiency. Hum Genet. 1998;102(6):605-10.
148. Zoppi M, Weiss M, Nydegger UE, Hess T, Spath PJ. Recurrent meningitis in a patient with congenital deficiency of the C9 component of complement. First case of C9 deficiency in Europe. Arch Intern Med. 1990;150(11):2395-9.
149. Sprong T, Roos D, Weemaes C, Neeleman C, Geesing CL, Mollnes TE, et al. Deficient alternative complement pathway activation due to factor D deficiency by 2 novel mutations in the complement factor D gene in a family with meningococcal infections. Blood. 2006;107(12):4865-70.
150. Biesma DH, Hannema AJ, van Velzen-Blad H, Mulder L, van Zwieten R, Kluijt I, et al. A family with complement factor D deficiency. J Clin Invest. 2001;108(2):233-40.
151. Slade C, Bosco J, Unglik G, Bleasel K, Nagel M, Winship I. Deficiency in complement factor B. N Engl J Med. 2013;369(17):1667-9.

152. Vyse TJ, Morley BJ, Bartok I, Theodoridis EL, Davies KA, Webster AD, et al. The molecular basis of hereditary complement factor I deficiency. J Clin Invest. 1996;97(4):925-33.
153. Alba-Dominguez M, Lopez-Lera A, Garrido S, Nozal P, Gonzalez-Granado I, Melero J, et al. Complement factor I deficiency: a not so rare immune defect: characterization of new mutations and the first large gene deletion. Orphanet J Rare Dis. 2012;7:42.
154. Da Silva Reis E, Baracho GV, Sousa Lima A, Farah CS, Isaac L. Homozygous hereditary C3 deficiency due to a premature stop codon. J Clin Immunol. 2002;22(6):321-30.
155. Jonsson G, Truedsson L, Sturfelt G, Oxelius VA, Braconier JH, Sjoholm AG. Hereditary C2 deficiency in Sweden: frequent occurrence of invasive infection, atherosclerosis, and rheumatic disease. Medicine (Baltimore). 2005;84(1):23-34.
156. Haralambous E, Dolly SO, Hibberd ML, Litt DJ, Udalova IA, O'Dwyer C, et al. Factor H, a regulator of complement activity, is a major determinant of meningococcal disease susceptibility in UK Caucasian patients. Scand J Infect Dis. 2006;38(9):764-71.
157. Bradley DT, Bourke TW, Fairley DJ, Borrow R, Shields MD, Zipfel PF, et al. Susceptibility to invasive meningococcal disease: polymorphism of complement system genes and Neisseria meningitidis factor H binding protein. PLoS One. 2015;10(3):e0120757.
158. Faber J, Schuessler T, Finn A, Murdoch C, Zenz W, Habermehl P, et al. Age-dependent association of human mannose-binding lectin mutations with susceptibility to invasive meningococcal disease in childhood. Pediatr Infect Dis J. 2007;26(3):243-6.
159. Hibberd ML, Sumiya M, Summerfield JA, Booy R, Levin M. Association of variants of the gene for mannose-binding lectin with susceptibility to meningococcal disease. Meningococcal Research Group. Lancet. 1999;353(9158):1049-53.
160. Lundbo LF, Sorensen HT, Clausen LN, Hollegaard MV, Hougaard DM, Konradsen HB, et al. Mannose-Binding Lectin Gene, MBL2, Polymorphisms Do Not Increase Susceptibility to Invasive Meningococcal Disease in a Population of Danish Children. Open Forum Infect Dis. 2015;2(4):ofv127.
161. Garred P, Michaelsen TE, Bjune G, Thiel S, Svejgaard A. A low serum concentration of mannan-binding protein is not associated with serogroup B or C meningococcal disease. Scand J Immunol. 1993;37(4):468-70.
162. von Bernuth H, Picard C, Jin Z, Pankla R, Xiao H, Ku CL, et al. Pyogenic bacterial infections in humans with MyD88 deficiency. Science. 2008;321(5889):691-6.
163. Spinner MA, Sanchez LA, Hsu AP, Shaw PA, Zerbe CS, Calvo KR, et al. GATA2 deficiency: a protean disorder of hematopoiesis, lymphatics, and immunity. Blood. 2014;123(6):809-21.
164. Anim M, Sogkas G, Schmidt G, Dubrowinskaja N, Witte T, Schmidt RE, et al. Vulnerability to Meningococcal Disease in Immunodeficiency Due to a Novel Pathogenic Missense Variant in NFKB1. Front Immunol. 2021;12:767188.
165. Arbour NC, Lorenz E, Schutte BC, Zabner J, Kline JN, Jones M, et al. TLR4 mutations are associated with endotoxin hyporesponsiveness in humans. Nat Genet. 2000;25(2):187-91.
166. Sanders MS, van Well GT, Ouburg S, Lundberg PS, van Furth AM, Morre SA. Single nucleotide polymorphisms in TLR9 are highly associated with susceptibility to bacterial meningitis in children. Clin Infect Dis. 2011;52(4):475-80.
167. Nadel S, Newport MJ, Booy R, Levin M. Variation in the Tumor Necrosis Factor-a Gene Promoter Region May Be Associated with Death from Meningococcal Disease. The Journal of Infectious Diseases. 1996;174:878-80.
168. Read RC, Teare DM, Pridmore AC, Naylor SC, Timms JM, Kaczmarski EB, et al. The tumor necrosis factor polymorphism TNF (-308) is associated with susceptibility to meningococcal sepsis, but not with lethality. Crit Care Med. 2009;37(4):1237-43.
169. Titmarsh CJ, Moscovis SM, Hall S, Tzanakaki G, Kesanopoulos K, Xirogianni A, et al. Comparison of cytokine gene polymorphisms among Greek patients with invasive meningococcal disease or viral meningitis. J Med Microbiol. 2013;62(Pt 5):694-700.
170. Balding J, Healy CM, Livingstone WJ, White B, Mynett-Johnson L, Cafferkey M, et al. Genomic polymorphic profiles in an Irish population with meningococcaemia: is it possible to predict severity and outcome of disease? Genes Immun. 2003;4(8):533-40.
171. Domingo P, Muniz-Diaz E, Baraldes MA, Arilla M, Barquet N, Pericas R, et al. Relevance of genetically determined host factors to the prognosis of meningococcal disease. Eur J Clin Microbiol Infect Dis. 2004;23(8):634-7.

172. Read RC, Camp NJ, di Giovine FS, Borrow R, Kaczmarski EB, Chaudhary AGA, et al. An Interleukin-1 Genotype Is Associated with Fatal Outcome of Meningococcal Disease. The Journal of Infectious Diseases. 2000;182(5):1557-60.
173. Tanaka T, Narazaki M, Kishimoto T. IL-6 in inflammation, immunity, and disease. Cold Spring Harb Perspect Biol. 2014;6(10):a016295.
174. Platonov AE, Kuijper EJ, Vershinina IV, Shipulin GA, Westerdaal N, Fijen CA, et al. Meningococcal disease and polymorphism of FcgammaRIIa (CD32) in late complement component-deficient individuals. Clin Exp Immunol. 1998;111(1):97-101.
175. Platonov AE, Shipulin GA, Vershinina IV, Dankert J, van de Winkel JG, Kuijper EJ. Association of Human FcgRIIa (CD32) Polymorphism with Susceptibility to and Severity of Meningococcal Disease. The Journal of Infectious Diseases. 1997;27:746-50.
176. van der Pol W-L, Huizinga TWJ, Vidarsson G, van der Linden MW, Jansen MD, Keijsers V, et al. Relevance of Fcg Receptor and Interleukin-10 Polymorphisms for Meningococcal Disease. The Journal of Infectious Diseases. 2001;184:1548-55.
177. Fijen CA, Bredius RG, Kuijper EJ, Out TA, De Haas M, De Wit AP, et al. The role of Fcgamma receptor polymorphisms and C3 in the immune defence against Neisseria meningitidis in complement-deficient individuals. Clin Exp Immunol. 2000;120(2):338-45.
178. Smith I, Vedeler C, Halstensen A. FcgammaRIIa and FcgammaRIIIb polymorphisms were not associated with meningococcal disease in Western Norway. Epidemiol Infect. 2003;130(2):193-9.
179. Frans G, Moens L, Schrijvers R, Wuyts G, Bouckaert B, Schaballie H, et al. PID in Disguise: Molecular Diagnosis of IRAK-4 Deficiency in an Adult Previously Misdiagnosed With Autosomal Dominant Hyper IgE Syndrome. J Clin Immunol. 2015;35(8):739-44.
180. Huppmann AR, Leiding JW, Hsu AP, Raffeld M, Uzel G, Pittaluga S, et al. Pathologic Findings in NEMO Deficiency: A Surgical and Autopsy Survey. Pediatr Dev Pathol. 2015;18(5):387-400.
181. Faber J, Henninger N, Finn A, Zenz W, Zepp F, Knuf M. A toll-like receptor 4 variant is associated with fatal outcome in children with invasive meningococcal disease. Acta Paediatr. 2009;98(3):548-52.
182. Faber J, Meyer CU, Gemmer C, Russo A, Finn A, Murdoch C, et al. HUMAN TOLL-LIKE RECEPTOR 4 MUTATIONS ARE ASSOCIATED WITH SUSCEPTIBILITY TO INVASIVE MENINGOCOCCAL DISEASE IN INFANCY. The Pediatric Infectious Disease Journal. 2006;25(1):80-1.
183. Read RC, Cannings C, Naylor SC, Timms JM, Maheswaran R, Borrow R, et al. Variation within Genes Encoding Interleukin-1 and the Interleukin-1 Receptor Antagonist Influence the Severity of Meningococcal Disease. Annals of Internal Medicine. 2003;138(7):534-41.
184. Endler G, Marculescu R, Starkl P, Binder A, Geishofer G, Muller M, et al. Polymorphisms in the Interleukin-1 Gene Cluster in Children and Young Adults with Systemic Meningococcemia. Clinical Chemistry 2006;52(3):511-4.
185. Bredius RG, Derkx BH, Fijen CA, de Wit TP, de Haas M, Weening RS, et al. Fc gamma receptor IIa (CD32) polymorphism in fulminant meningococcal septic shock in children. J Infect Dis. 1994;170(4):848-53.
186. Domingo P, Muniz-Diaz E, Baraldes MA, Arilla M, Barquet N, Pericas R, et al. Associations between Fc gamma receptor IIA polymorphisms and the risk and prognosis of meningococcal disease. Am J Med. 2002;112(1):19-25.
187. Kondaveeti S, Hibberd ML, Booy R, Nadel S, Levin M. Effect of the Factor V Leiden mutation on the severity of meningococcal disease. Pediatr Infect Dis J. 1999;18(10):893-6.
188. Powars D, Larsen R, Johnson J, Hulbert T, Sun T, Patch MJ, et al. Epidemic meningococcemia and purpura fulminans with induced protein C deficiency. Clin Infect Dis. 1993;17(2):254-61.
189. Sackesen C, Secmeer G, Gurgey A, Kanra G, Ecevit Z, Ceyhan M, et al. Homozygous Factor V Leiden mutation in a child with meningococcal purpura fulminans. Pediatr Infect Dis J. 1998;17(1):87.
190. Binder A, Endler G, Rieger S, Geishofer G, Resch B, Mannhalter C, et al. Protein C promoter polymorphisms associate with sepsis in children with systemic meningococcemia. Hum Genet. 2007;122(2):183-90.

191. Spek CA, Greengard JS, Griffin JH, Bertina RM, Reitsma PH. Two mutations in the promoter region of the human protein C gene both cause type I protein C deficiency by disruption of two HNF-3 binding sites. J Biol Chem. 1995;270(41):24216-21.
192. Brandtzaeg P, Sandset PM, Joo GB, Ovstebo R, Abildgaard U, Kierulf P. The quantitative association of plasma endotoxin, antithrombin, protein C, extrinsic pathway inhibitor and fibrinopeptide A in systemic meningococcal disease. Thromb Res. 1989;55(4):459-70.
193. Emonts M, de Bruijne EL, Guimaraes AH, Declerck PJ, Leebeek FW, de Maat MP, et al. Thrombin-activatable fibrinolysis inhibitor is associated with severity and outcome of severe meningococcal infection in children. J Thromb Haemost. 2008;6(2):268-76.
194. Kremer Hovinga JA, Franco RF, Zago MA, Ten Cate H, Westendorp RG, Reitsma PH. A functional single nucleotide polymorphism in the thrombin-activatable fibrinolysis inhibitor (TAFI) gene associates with outcome of meningococcal disease. J Thromb Haemost. 2004;2(1):54-7.
195. Haralambous E, Hibberd ML, Hermans PW, Ninis N, Nadel S, Levin M. Role of functional plasminogen-activator-inhibitor-1 4G/5G promoter polymorphism in susceptibility, severity, and outcome of meningococcal disease in Caucasian children. Crit Care Med. 2003;31(12):2788-93.
196. Westendorp RG, Hottenga JJ, Slagboom PE. Variation in plasminogen-activator-inhibitor-1 gene and risk of meningococcal septic shock. Lancet. 1999;354(9178):561-3.
197. Geishofer G, Binder A, Muller M, Zohrer B, Resch B, Muller W, et al. 4G/5G promoter polymorphism in the plasminogen-activator-inhibitor-1 gene in children with systemic meningococcaemia. Eur J Pediatr. 2005;164(8):486-90.
198. Binder A, Endler G, Muller M, Mannhalter C, Zenz W. 4G4G genotype of the plasminogen activator inhibitor-1 promoter polymorphism associates with disseminated intravascular coagulation in children with systemic meningococcemia. J Thromb Haemost. 2007;5(10).
199. Shi C, Sui Z, Li L, Yang R. No Association of SERPINE1 -675 Polymorphism With Sepsis Susceptibility: A Meta-Analysis. Medicine (Baltimore). 2015;94(45):e1173.
200. Hermans PW, Hibberd ML, Booy R, Daramola O, Hazelzet JA, de Groot R, et al. 4G/5G promoter polymorphism in the plasminogen-activator-inhibitor-1 gene and outcome of meningococcal disease. Meningococcal Research Group. Lancet. 1999;354(9178):556-60.
201. Friedrich R, Panizzi P, Fuentes-Prior P, Richter K, Verhamme I, Anderson PJ, et al. Staphylocoagulase is a prototype for the mechanism of cofactor-induced zymogen activation. Nature. 2003;425(6957):535-9.
202. McAdow M, Missiakas DM, Schneewind O. Staphylococcus aureus secretes coagulase and von Willebrand factor binding protein to modify the coagulation cascade and establish host infections. J Innate Immun. 2012;4(2):141-8.
203. Foster TJ. The MSCRAMM Family of Cell-Wall-Anchored Surface Proteins of Gram-Positive Cocci. Trends Microbiol. 2019;27(11):927-41.
204. Cunnion KM, Hair PS, Buescher ES. Cleavage of complement C3b to iC3b on the surface of Staphylococcus aureus is mediated by serum complement factor I. Infect Immun. 2004;72(5):2858-63.
205. Sharp JA, Echague CG, Hair PS, Ward MD, Nyalwidhe JO, Geoghegan JA, et al. Staphylococcus aureus surface protein SdrE binds complement regulator factor H as an immune evasion tactic. PLoS One. 2012;7(5):e38407.
206. Wertheim HF, Melles DC, Vos MC, van Leeuwen W, van Belkum A, Verbrugh HA, et al. The role of nasal carriage in Staphylococcus aureus infections. Lancet Infect Dis. 2005;5(12):751-62.
207. Williams RE. Healthy carriage of Staphylococcus aureus: its prevalence and importance. Bacteriol Rev. 1963;27(1):56-71.
208. Lowy FD. Staphylococcus aureus infections. N Engl J Med. 1998;339(8):520-32.
209. Wozniak JM, Mills RH, Olson J, Caldera JR, Sepich-Poore GD, Carrillo-Terrazas M, et al. Mortality Risk Profiling of Staphylococcus aureus Bacteremia by Multi-omic Serum Analysis Reveals Early Predictive and Pathogenic Signatures. Cell. 2020;182(5):1311-27 e14.
210. Thammavongsa V, Kim HK, Missiakas D, Schneewind O. Staphylococcal manipulation of host immune responses. Nat Rev Microbiol. 2015;13(9):529-43.
211. Kubica M, Guzik K, Koziel J, Zarebski M, Richter W, Gajkowska B, et al. A potential new pathway for Staphylococcus aureus dissemination: the silent survival of S. aureus phagocytosed by human monocyte-derived macrophages. PLoS One. 2008;3(1):e1409.

212. Baron S. Staphylococcus. Baron S, editor. Galveston (TX): University of Texas Medical Branch at Galveston; 1996.
213. Genestier AL, Michallet MC, Prevost G, Bellot G, Chalabreysse L, Peyrol S, et al. Staphylococcus aureus Panton-Valentine leukocidin directly targets mitochondria and induces Bax-independent apoptosis of human neutrophils. J Clin Invest. 2005;115(11):3117-27.
214. Tong SY, Davis JS, Eichenberger E, Holland TL, Fowler VG, Jr. Staphylococcus aureus infections: epidemiology, pathophysiology, clinical manifestations, and management. Clin Microbiol Rev. 2015;28(3):603-61.
215. David MZ, Daum RS. Community-associated methicillin-resistant Staphylococcus aureus: epidemiology and clinical consequences of an emerging epidemic. Clin Microbiol Rev. 2010;23(3):616-87.
216. Laupland KB, Lyytikainen O, Sogaard M, Kennedy KJ, Knudsen JD, Ostergaard C, et al. The changing epidemiology of Staphylococcus aureus bloodstream infection: a multinational population-based surveillance study. Clin Microbiol Infect. 2013;19(5):465-71.
217. Klevens RM, Morrison MA, Nadle J, Petit S, Gershman K, Ray S, et al. Invasive methicillin-resistant Staphylococcus aureus infections in the United States. JAMA. 2007;298(15):1763-71.
218. Tong SY, Bishop EJ, Lilliebridge RA, Cheng AC, Spasova-Penkova Z, Holt DC, et al. Community-associated strains of methicillin-resistant Staphylococcus aureus and methicillin-susceptible S. aureus in indigenous Northern Australia: epidemiology and outcomes. J Infect Dis. 2009;199(10):1461-70.
219. Tong SY, van Hal SJ, Einsiedel L, Currie BJ, Turnidge JD, Australian New Zealand Cooperative on Outcomes in Staphylococcal S. Impact of ethnicity and socio-economic status on Staphylococcus aureus bacteremia incidence and mortality: a heavy burden in Indigenous Australians. BMC Infect Dis. 2012;12:249.
220. Boisson B. The genetic basis of pneumococcal and staphylococcal infections: inborn errors of human TLR and IL-1R immunity. Hum Genet. 2020;139(6-7):981-91.
221. Boisson B, Laplantine E, Prando C, Giliani S, Israelsson E, Xu Z, et al. Immunodeficiency, autoinflammation and amylopectinosis in humans with inherited HOIL-1 and LUBAC deficiency. Nat Immunol. 2012;13(12):1178-86.
222. Israel L, Wang Y, Bulek K, Della Mina E, Zhang Z, Pedergnana V, et al. Human Adaptive Immunity Rescues an Inborn Error of Innate Immunity. Cell. 2017;168(5):789-800 e10.
223. Geno KA, Gilbert GL, Song JY, Skovsted IC, Klugman KP, Jones C, et al. Pneumococcal Capsules and Their Types: Past, Present, and Future. Clin Microbiol Rev. 2015;28(3):871-99.
224. Henriques-Normark B, Tuomanen EI. The pneumococcus: epidemiology, microbiology, and pathogenesis. Cold Spring Harb Perspect Med. 2013;3(7).
225. Sanchez CJ, Shivshankar P, Stol K, Trakhtenbroit S, Sullam PM, Sauer K, et al. The pneumococcal serine-rich repeat protein is an intra-species bacterial adhesin that promotes bacterial aggregation in vivo and in biofilms. PLoS Pathog. 2010;6(8):e1001044.
226. Hammerschmidt S, Talay SR, Brandtzaeg P, Chhatwal GS. SpsA, a novel pneumococcal surface protein with specific binding to secretory immunoglobulin A and secretory component. Mol Microbiol. 1997;25(6):1113-24.
227. Luna CM, Pulido L, Niederman MS, Casey A, Burgos D, Leiva Aguero SD, et al. Decreased relative risk of pneumococcal pneumonia during the last decade, a nested case-control study. Pneumonia (Nathan). 2018;10:9.
228. Alqahtani AS, Tashani M, Ridda I, Gamil A, Booy R, Rashid H. Burden of clinical infections due to S. pneumoniae during Hajj: A systematic review. Vaccine. 2018;36(30):4440-6.
229. Wahl B, O'Brien KL, Greenbaum A, Majumder A, Liu L, Chu Y, et al. Burden of Streptococcus pneumoniae and Haemophilus influenzae type b disease in children in the era of conjugate vaccines: global, regional, and national estimates for 2000-15. Lancet Glob Health. 2018;6(7):e744-e57.
230. Canvin JR, Marvin AP, Sivakumaran M, Paton JC, Boulnois GJ, Andrew PW, et al. The role of pneumolysin and autolysin in the pathology of pneumonia and septicemia in mice infected with a type 2 pneumococcus. J Infect Dis. 1995;172(1):119-23.
231. Malley R, Henneke P, Morse SC, Cieslewicz MJ, Lipsitch M, Thompson CM, et al. Recognition of pneumolysin by Toll-like receptor 4 confers resistance to pneumococcal infection. Proc Natl Acad Sci U S A. 2003;100(4):1966-71.

232. Brown JS, Hussell T, Gilliland SM, Holden DW, Paton JC, Ehrenstein MR, et al. The classical pathway is the dominant complement pathway required for innate immunity to Streptococcus pneumoniae infection in mice. Proc Natl Acad Sci U S A. 2002;99(26):16969-74.
233. Chapman SJ, Khor CC, Vannberg FO, Frodsham A, Walley A, Maskell NA, et al. IkappaB genetic polymorphisms and invasive pneumococcal disease. Am J Respir Crit Care Med. 2007;176(2):181-7.
234. Croucher NJ, Finkelstein JA, Pelton SI, Mitchell PK, Lee GM, Parkhill J, et al. Population genomics of post-vaccine changes in pneumococcal epidemiology. Nat Genet. 2013;45(6):656-63.
235. Flannery B, Heffernan RT, Harrison LH, Ray SM, Reingold AL, Hadler J, et al. Changes in invasive Pneumococcal disease among HIV-infected adults living in the era of childhood pneumococcal immunization. Ann Intern Med. 2006;144(1):1-9.
236. Boisson B, Honda Y, Ajiro M, Bustamante J, Bendavid M, Gennery AR, et al. Rescue of recurrent deep intronic mutation underlying cell type-dependent quantitative NEMO deficiency. J Clin Invest. 2019;129(2):583-97.
237. Paciolla M, Pescatore A, Conte MI, Esposito E, Incoronato M, Lioi MB, et al. Rare mendelian primary immunodeficiency diseases associated with impaired NF-kappaB signaling. Genes Immun. 2015;16(4):239-46.
238. Hanson EP, Monaco-Shawver L, Solt LA, Madge LA, Banerjee PP, May MJ, et al. Hypomorphic nuclear factor-kappaB essential modulator mutation database and reconstitution system identifies phenotypic and immunologic diversity. J Allergy Clin Immunol. 2008;122(6):1169-77 e16.
239. Lundbo LF, Harboe ZB, Clausen LN, Hollegaard MV, Sorensen HT, Hougaard DM, et al. Genetic Variation in NFKBIE Is Associated With Increased Risk of Pneumococcal Meningitis in Children. EBioMedicine. 2016;3:93-9.
240. Cunningham MW. Pathogenesis of group A streptococcal infections. Clin Microbiol Rev. 2000;13(3):470-511.
241. Cunningham MW. Pathogenesis of group A streptococcal infections and their sequelae. Adv Exp Med Biol. 2008;609:29-42.
242. Horstmann RD, Sievertsen HJ, Knobloch J, Fischetti VA. Antiphagocytic activity of streptococcal M protein: selective binding of complement control protein factor H. Proc Natl Acad Sci U S A. 1988;85(5):1657-61.
243. Horstmann RD, Sievertsen HJ, Leippe M, Fischetti VA. Role of fibrinogen in complement inhibition by streptococcal M protein. Infect Immun. 1992;60(12):5036-41.
244. Walker MJ, Barnett TC, McArthur JD, Cole JN, Gillen CM, Henningham A, et al. Disease manifestations and pathogenic mechanisms of Group A Streptococcus. Clin Microbiol Rev. 2014;27(2):264-301.
245. Ralph AP, Carapetis JR. Group a streptococcal diseases and their global burden. Curr Top Microbiol Immunol. 2013;368:1-27.
246. Arif N, Yousfi S, Vinnard C. Deaths from necrotizing fasciitis in the United States, 2003-2013. Epidemiol Infect. 2016;144(6):1338-44.
247. Schmitz M, Roux X, Huttner B, Pugin J. Streptococcal toxic shock syndrome in the intensive care unit. Ann Intensive Care. 2018;8(1):88.
248. Organization WH. Disease Outbreak News; Increased incidence of scarlet fever and invasive Group A Streptococcus infection - multi-country. 2022 [Available from: .
249. Davies HD, McGeer A, Schwartz B, Green K, Cann D, Simor AE, et al. Invasive group A streptococcal infections in Ontario, Canada. Ontario Group A Streptococcal Study Group. N Engl J Med. 1996;335(8):547-54.
250. Parks T, Elliott K, Lamagni T, Auckland K, Mentzer AJ, Guy R, et al. Elevated risk of invasive group A streptococcal disease and host genetic variation in the human leucocyte antigen locus. Genes Immun. 2020;21(1):63-70.
251. Liu DX, Liang JQ, Fung TS. Human Coronavirus-229E, -OC43, -NL63, and -HKU1 (Coronaviridae).2021.
252. Lim YX, Ng YL, Tam JP, Liu DX. Human Coronaviruses: A Review of Virus-Host Interactions. Diseases. 2016;4(3).

253. Fung TS, Liu DX. Human Coronavirus: Host-Pathogen Interaction. Annu Rev Microbiol. 2019;73:529-57.
254. Anderson RM, Fraser C, Ghani AC, Donnelly CA, Riley S, Ferguson NM, et al. Epidemiology, transmission dynamics and control of SARS: the 2002-2003 epidemic. Philos Trans R Soc Lond B Biol Sci. 2004;359(1447):1091-105.
255. Ellis P, Somogyvari F, Virok DP, Noseda M, McLean GR. Decoding Covid-19 with the SARS-CoV-2 Genome. Curr Genet Med Rep. 2021;9(1):1-12.
256. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020;579(7798):270-3.
257. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature. 2003;426(6965):450-4.
258. Tai W, He L, Zhang X, Pu J, Voronin D, Jiang S, et al. Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. Cell Mol Immunol. 2020;17(6):613-20.
259. Sungnak W, Huang N, Becavin C, Berg M, Queen R, Litvinukova M, et al. SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. Nat Med. 2020;26(5):681-7.
260. Lukassen S, Chua RL, Trefzer T, Kahn NC, Schneider MA, Muley T, et al. SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells. EMBO J. 2020;39(10):e105114.
261. V'Kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. Coronavirus biology and replication: implications for SARS-CoV-2. Nat Rev Microbiol. 2021;19(3):155-70.
262. Diamond MS, Kanneganti TD. Innate immunity: the first line of defense against SARS-CoV-2. Nat Immunol. 2022;23(2):165-76.
263. Takeuchi O, Akira S. Pattern recognition receptors and inflammation. Cell. 2010;140(6):805-20.
264. Lester SN, Li K. Toll-like receptors in antiviral innate immunity. J Mol Biol. 2014;426(6):1246-64.
265. Samuel CE. Antiviral actions of interferon. Interferon-regulated cellular proteins and their surprisingly selective antiviral activities. Virology. 1991;183(1):1-11.
266. Clemens MJ. PKR--a protein kinase regulated by double-stranded RNA. Int J Biochem Cell Biol. 1997;29(7):945-9.
267. Lamers MM, Haagmans BL. SARS-CoV-2 pathogenesis. Nat Rev Microbiol. 2022;20(5):270-84.
268. Gao C, Zhu L, Jin CC, Tong YX, Xiao AT, Zhang S. Proinflammatory cytokines are associated with prolonged viral RNA shedding in COVID-19 patients. Clin Immunol. 2020;221:108611.
269. Mizumoto K, Kagaya K, Zarebski A, Chowell G. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. Euro Surveill. 2020;25(10).
270. Nishiura H, Kobayashi T, Miyama T, Suzuki A, Jung SM, Hayashi K, et al. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). Int J Infect Dis. 2020;94:154-5.
271. Berlin DA, Gulick RM, Martinez FJ. Severe Covid-19. N Engl J Med. 2020;383(25):2451-60.
272. Cascella M, Rajnik M, Aleem A, Dulebohn SC, Di Napoli R. Features, Evaluation, and Treatment of Coronavirus (COVID-19). StatPearls. Treasure Island (FL)2022.
273. Casrouge A, Zhang SY, Eidenschenk C, Jouanguy E, Puel A, Yang K, et al. Herpes simplex virus encephalitis in human UNC-93B deficiency. Science. 2006;314(5797):308-12.
274. Andersen LL, Mork N, Reinert LS, Kofod-Olsen E, Narita R, Jorgensen SE, et al. Functional IRF3 deficiency in a patient with herpes simplex encephalitis. J Exp Med. 2015;212(9):1371-9.
275. Guo Y, Audry M, Ciancanelli M, Alsina L, Azevedo J, Herman M, et al. Herpes simplex virus encephalitis in a patient with complete TLR3 deficiency: TLR3 is otherwise redundant in protective immunity. J Exp Med. 2011;208(10):2083-98.
276. Herman M, Ciancanelli M, Ou YH, Lorenzo L, Klaudel-Dreszler M, Pauwels E, et al. Heterozygous TBK1 mutations impair TLR3 immunity and underlie herpes simplex encephalitis of childhood. J Exp Med. 2012;209(9):1567-82.

277. Lim HK, Seppanen M, Hautala T, Ciancanelli MJ, Itan Y, Lafaille FG, et al. TLR3 deficiency in herpes simplex encephalitis: high allelic heterogeneity and recurrence risk. Neurology. 2014;83(21):1888-97.
278. Perez de Diego R, Sancho-Shimizu V, Lorenzo L, Puel A, Plancoulaine S, Picard C, et al. Human TRAF3 adaptor molecule deficiency leads to impaired Toll-like receptor 3 response and susceptibility to herpes simplex encephalitis. Immunity. 2010;33(3):400-11.
279. Zhang SY, Jouanguy E, Ugolini S, Smahi A, Elain G, Romero P, et al. TLR3 deficiency in patients with herpes simplex encephalitis. Science. 2007;317(5844):1522-7.
280. Lim HK, Huang SXL, Chen J, Kerner G, Gilliaux O, Bastard P, et al. Severe influenza pneumonitis in children with inherited TLR3 deficiency. J Exp Med. 2019;216(9):2038-56.
281. Casanova JL, Anderson MS. Unlocking life-threatening COVID-19 through two types of inborn errors of type I IFNs. J Clin Invest. 2023;133(3).
282. Ullah MO, Sweet MJ, Mansell A, Kellie S, Kobe B. TRIF-dependent TLR signaling, its functions in host defense and inflammation, and its potential as a therapeutic target. J Leukoc Biol. 2016;100(1):27-45.
283. Sharma S, tenOever BR, Grandvaux N, Zhou GP, Lin R, Hiscott J. Triggering the interferon antiviral response through an IKK-related pathway. Science. 2003;300(5622):1148-51.
284. Fitzgerald KA, McWhirter SM, Faia KL, Rowe DC, Latz E, Golenbock DT, et al. IKKepsilon and TBK1 are essential components of the IRF3 signaling pathway. Nat Immunol. 2003;4(5):491-6.
285. Sato S, Sugiyama M, Yamamoto M, Watanabe Y, Kawai T, Takeda K, et al. Toll/IL-1 receptor domain-containing adaptor inducing IFN-beta (TRIF) associates with TNF receptor-associated factor 6 and TANK-binding kinase 1, and activates two distinct transcription factors, NF-kappa B and IFN-regulatory factor-3, in the Toll-like receptor signaling. J Immunol. 2003;171(8):4304-10.
286. Sasai M, Tatematsu M, Oshiumi H, Funami K, Matsumoto M, Hatakeyama S, et al. Direct binding of TRAF2 and TRAF6 to TICAM-1/TRIF adaptor participates in activation of the Toll-like receptor 3/4 pathway. Mol Immunol. 2010;47(6):1283-91.
287. Tatematsu M, Ishii A, Oshiumi H, Horiuchi M, Inagaki F, Seya T, et al. A molecular mechanism for Toll-IL-1 receptor domain-containing adaptor molecule-1-mediated IRF-3 activation. J Biol Chem. 2010;285(26):20128-36.
288. Mork N, Kofod-Olsen E, Sorensen KB, Bach E, Orntoft TF, Ostergaard L, et al. Mutations in the TLR3 signaling pathway and beyond in adult patients with herpes simplex encephalitis. Genes Immun. 2015;16(8):552-66.
289. Zhang Q, Bastard P, Liu Z, Le Pen J, Moncada-Velez M, Chen J, et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. Science. 2020;370(6515).
290. Watling D, Guschin D, Muller M, Silvennoinen O, Witthuhn BA, Quelle FW, et al. Complementation by the protein tyrosine kinase JAK2 of a mutant cell line defective in the interferon-gamma signal transduction pathway. Nature. 1993;366(6451):166-70.
291. Schoggins JW, Wilson SJ, Panis M, Murphy MY, Jones CT, Bieniasz P, et al. A diverse range of gene products are effectors of the type I interferon antiviral response. Nature. 2011;472(7344):481-5.
292. Auwera Gvd, O'Connor BD. Genomics in the cloud : using Docker, GATK, and WDL in Terra. First edition. ed. Sebastopol, CA: O'Reilly Media; 2020. xxiv, 467 pages p.
293. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020;581(7809):434-43.
294. Edelmann B, Gupta N, Schnoeder TM, Oelschlegel AM, Shahzad K, Goldschmidt J, et al. JAK2-V617F promotes venous thrombosis through beta1/beta2 integrin activation. J Clin Invest. 2018;128(10):4359-71.
295. Byun M, Abhyankar A, Lelarge V, Plancoulaine S, Palanduz A, Telhan L, et al. Whole-exome sequencing-based discovery of STIM1 deficiency in a child with fatal classic Kaposi sarcoma. J Exp Med. 2010;207(11):2307-12.
296. Hoischen A, van Bon BW, Gilissen C, Arts P, van Lier B, Steehouwer M, et al. De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. Nat Genet. 2010;42(6):483-5.
297. Mensah AA, Campbell H, Clark SA, Ribeiro S, Lucidarme J, Bai X, et al. Outcomes of meningococcal serogroup B disease in children after implementation of routine infant 4CMenB

vaccination in England: an active, prospective, national surveillance study. Lancet Child Adolesc Health. 2023;7(3):190-8.
298. Flexner S. The Results of the Serum Treatment in Thirteen Hundred Cases of Epidemic Meningitis. J Exp Med. 1913;17(5):553-76.
299. Imada K, Leonard WJ. The Jak-STAT pathway. Mol Immunol. 2000;37(1-2):1-11.
300. Yamaoka K, Saharinen P, Pesu M, Holt VE, 3rd, Silvennoinen O, O'Shea JJ. The Janus kinases (Jaks). Genome Biol. 2004;5(12):253.
301. Chen E, Mullally A. How does JAK2V617F contribute to the pathogenesis of myeloproliferative neoplasms? Hematology Am Soc Hematol Educ Program. 2014;2014(1):268-76.
302. McLornan D, Percy M, McMullin MF. JAK2 V617F: a single mutation in the myeloproliferative group of disorders. Ulster Med J. 2006;75(2):112-9.
303. Baxter EJ, Scott LM, Campbell PJ, East C, Fourouclas N, Swanton S, et al. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. Lancet. 2005;365(9464):1054-61.
304. Braunstein EM, Moliterno AR. Back to biology: new insights on inheritance in myeloproliferative disorders. Curr Hematol Malig Rep. 2014;9(4):311-8.
305. Rolles B, Mullally A. Molecular Pathogenesis of Myeloproliferative Neoplasms. Curr Hematol Malig Rep. 2022;17(6):319-29.
306. Cerquozzi S, Tefferi A. Blast transformation and fibrotic progression in polycythemia vera and essential thrombocythemia: a literature review of incidence and risk factors. Blood Cancer J. 2015;5(11):e366.
307. Iurlo A, Cattaneo D, Gianelli U. Blast Transformation in Myeloproliferative Neoplasms: Risk Factors, Biological Findings, and Targeted Therapeutic Options. Int J Mol Sci. 2019;20(8).
308. Kapralova K, Horvathova M, Pecquet C, Fialova Kucerova J, Pospisilova D, Leroy E, et al. Cooperation of germ line JAK2 mutations E846D and R1063H in hereditary erythrocytosis with megakaryocytic atypia. Blood. 2016;128(10):1418-23.
309. Milosevic Feenstra JD, Nivarthi H, Gisslinger H, Leroy E, Rumi E, Chachoua I, et al. Whole-exome sequencing identifies novel MPL and JAK2 mutations in triple-negative myeloproliferative neoplasms. Blood. 2016;127(3):325-32.
310. Wu QY, Ma MM, Tong YX, Zhu YY, Liu Y, Cao J, et al. Effects of JAK2 V556F mutation on the JAK2's activity, structural stability and the transformation of Ba/F3 cells. Int J Biol Macromol. 2018;117:271-9.
311. Mambet C, Babosova O, Defour JP, Leroy E, Necula L, Stanca O, et al. Cooccurring JAK2 V617F and R1063H mutations increase JAK2 signaling and neutrophilia in myeloproliferative neoplasms. Blood. 2018;132(25):2695-9.
312. Lanikova L, Babosova O, Swierczek S, Wang L, Wheeler DA, Divoky V, et al. Coexistence of gain-of-function JAK2 germ line mutations with JAK2V617F in polycythemia vera. Blood. 2016;128(18):2266-70.
313. Mazurkiewicz-Munoz AM, Argetsinger LS, Kouadio JL, Stensballe A, Jensen ON, Cline JM, et al. Phosphorylation of JAK2 at serine 523: a negative regulator of JAK2 that is stimulated by growth hormone and epidermal growth factor. Mol Cell Biol. 2006;26(11):4052-62.
314. Benton CB, Boddu PC, DiNardo CD, Bose P, Wang F, Assi R, et al. Janus kinase 2 variants associated with the transformation of myeloproliferative neoplasms into acute myeloid leukemia. Cancer. 2019;125(11):1855-66.
315. Shahin OA, Chifotides HT, Bose P, Masarova L, Verstovsek S. Accelerated Phase of Myeloproliferative Neoplasms. Acta Haematol. 2021;144(5):484-99.
316. Ma W, Kantarjian H, Zhang X, Yeh CH, Zhang ZJ, Verstovsek S, et al. Mutation profile of JAK2 transcripts in patients with chronic myeloproliferative neoplasias. J Mol Diagn. 2009;11(1):49-53.
317. Siemiatkowska A, Bieniaszewska M, Hellmann A, Limon J. JAK2 and MPL gene mutations in V617F-negative myeloproliferative neoplasms. Leuk Res. 2010;34(3):387-9.
318. Gill H, Ip HW, Yim R, Tang WF, Pang HH, Lee P, et al. Next-generation sequencing with a 54-gene panel identified unique mutational profile and prognostic markers in Chinese patients with myelofibrosis. Ann Hematol. 2019;98(4):869-79.
319. Bahar B, Barton K, Kini AR. The role of the Exon 13 G571S JAK2 mutation in myeloproliferative neoplasms. Leuk Res Rep. 2016;6:27-8.

320. Hammaren HM, Virtanen AT, Raivola J, Silvennoinen O. The regulation of JAKs in cytokine signaling and its breakdown in disease. Cytokine. 2019;118:48-63.
321. Kishimoto T. IL-6: from its discovery to clinical applications. Int Immunol. 2010;22(5):347-52.
322. Guschin D, Rogers N, Briscoe J, Witthuhn B, Watling D, Horn F, et al. A major role for the protein tyrosine kinase JAK1 in the JAK/STAT signal transduction pathway in response to interleukin-6. EMBO J. 1995;14(7):1421-9.
323. Chen Z, Zhang H, Qu M, Nan K, Cao H, Cata JP, et al. Review: The Emerging Role of Neutrophil Extracellular Traps in Sepsis and Sepsis-Associated Thrombosis. Front Cell Infect Microbiol. 2021;11:653228.
324. Hubbard AK, Rothlein R. Intercellular adhesion molecule-1 (ICAM-1) expression and cell signaling cascades. Free Radic Biol Med. 2000;28(9):1379-86.
325. Oh ST. Neutralize the neutrophils! Neutrophil beta1/beta2 integrin activation contributes to JAK2-V617F-driven thrombosis. J Clin Invest. 2018;128(10):4248-50.
326. Elbim C, Chollet-Martin S, Bailly S, Hakim J, Gougerot-Pocidalo MA. Priming of polymorphonuclear neutrophils by tumor necrosis factor alpha in whole blood: identification of two polymorphonuclear neutrophil subpopulations in response to formyl-peptides. Blood. 1993;82(2):633-40.
327. Lauterbach M, O'Donnell P, Asano K, Mayadas TN. Role of TNF priming and adhesion molecules in neutrophil recruitment to intravascular immune complexes. J Leukoc Biol. 2008;83(6):1423-30.
328. Montecucco F, Steffens S, Burger F, Da Costa A, Bianchi G, Bertolotto M, et al. Tumor necrosis factor-alpha (TNF-alpha) induces integrin CD11b/CD18 (Mac-1) up-regulation and migration to the CC chemokine CCL3 (MIP-1alpha) on human neutrophils through defined signalling pathways. Cell Signal. 2008;20(3):557-68.
329. Silveira AA, Dominical VM, Lazarini M, Costa FF, Conran N. Simvastatin abrogates inflamed neutrophil adhesive properties, in association with the inhibition of Mac-1 integrin expression and modulation of Rho kinase activity. Inflamm Res. 2013;62(2):127-32.
330. O'Shea JJ, Holland SM, Staudt LM. JAKs and STATs in immunity, immunodeficiency, and cancer. N Engl J Med. 2013;368(2):161-70.
331. Matsuda T, Feng J, Witthuhn BA, Sekine Y, Ihle JN. Determination of the transphosphorylation sites of Jak2 kinase. Biochem Biophys Res Commun. 2004;325(2):586-94.
332. Argetsinger LS, Kouadio JL, Steen H, Stensballe A, Jensen ON, Carter-Su C. Autophosphorylation of JAK2 on tyrosines 221 and 570 regulates its activity. Mol Cell Biol. 2004;24(11):4955-67.
333. Argetsinger LS, Stuckey JA, Robertson SA, Koleva RI, Cline JM, Marto JA, et al. Tyrosines 868, 966, and 972 in the kinase domain of JAK2 are autophosphorylated and required for maximal JAK2 kinase activity. Mol Endocrinol. 2010;24(5):1062-76.
334. Feener EP, Rosario F, Dunn SL, Stancheva Z, Myers MG, Jr. Tyrosine phosphorylation of Jak2 in the JH2 domain inhibits cytokine signaling. Mol Cell Biol. 2004;24(11):4968-78.
335. Robertson SA, Koleva RI, Argetsinger LS, Carter-Su C, Marto JA, Feener EP, et al. Regulation of Jak2 function by phosphorylation of Tyr317 and Tyr637 during cytokine signaling. Mol Cell Biol. 2009;29(12):3367-78.
336. Scott LM, Beer PA, Bench AJ, Erber WN, Green AR. Prevalance of JAK2 V617F and exon 12 mutations in polycythaemia vera. Br J Haematol. 2007;139(3):511-2.
337. Scott LM, Tong W, Levine RL, Scott MA, Beer PA, Stratton MR, et al. JAK2 exon 12 mutations in polycythemia vera and idiopathic erythrocytosis. N Engl J Med. 2007;356(5):459-68.
338. Guglielmelli P, Loscocco GG, Mannarelli C, Rossi E, Mannelli F, Ramundo F, et al. JAK2V617F variant allele frequency >50% identifies patients with polycythemia vera at high risk for venous thrombosis. Blood Cancer J. 2021;11(12):199.
339. Kralovics R, Passamonti F, Buser AS, Teo SS, Tiedt R, Passweg JR, et al. A gain-of-function mutation of JAK2 in myeloproliferative disorders. N Engl J Med. 2005;352(17):1779-90.
340. Rumi E, Pietra D, Ferretti V, Klampfl T, Harutyunyan AS, Milosevic JD, et al. JAK2 or CALR mutation status defines subtypes of essential thrombocythemia with substantially different clinical course and outcomes. Blood. 2014;123(10):1544-51.

341. Aral B, Courtois M, Ragot S, Bourgeois V, Bottolier-Lemallaz E, Briandet C, et al. Germline JAK2 L611S mutation in a child with thrombocytosis. Haematologica. 2018;103(8):e372-e3.
342. Etheridge SL, Cosgrove ME, Sangkhae V, Corbo LM, Roh ME, Seeliger MA, et al. A novel activating, germline JAK2 mutation, JAK2R564Q, causes familial essential thrombocytosis. Blood. 2014;123(7):1059-68.
343. Mead AJ, Rugless MJ, Jacobsen SE, Schuh A. Germline JAK2 mutation in a family with hereditary thrombocytosis. N Engl J Med. 2012;366(10):967-9.
344. Rumi E, Harutyunyan AS, Casetti I, Pietra D, Nivarthi H, Moriggl R, et al. A novel germline JAK2 mutation in familial myeloproliferative neoplasms. Am J Hematol. 2014;89(1):117-8.
345. Marty C, Saint-Martin C, Pecquet C, Grosjean S, Saliba J, Mouton C, et al. Germ-line JAK2 mutations in the kinase domain are responsible for hereditary thrombocytosis and are resistant to JAK2 and HSP90 inhibitors. Blood. 2014;123(9):1372-83.
346. Ungureanu D, Wu J, Pekkala T, Niranjan Y, Young C, Jensen ON, et al. The pseudokinase domain of JAK2 is a dual-specificity protein kinase that negatively regulates cytokine signaling. Nat Struct Mol Biol. 2011;18(9):971-6.
347. Saharinen P, Takaluoma K, Silvennoinen O. Regulation of the Jak2 tyrosine kinase by its pseudokinase domain. Mol Cell Biol. 2000;20(10):3387-95.
348. Panovska-Stavridis I, Eftimov A, Ivanovski M, Pivkova-Veljanovska A, Cevreska L, Hermouet S, et al. Essential Thrombocythemia Associated With Germline JAK2 G571S Variant and Somatic CALR Type 1 Mutation. Clin Lymphoma Myeloma Leuk. 2016;16(5):e55-7.
349. Tun PWW, Buka RJ, Graham J, Dyer P. Heterozygous, germline JAK2 E846D substitution as the cause of familial erythrocytosis. Br J Haematol. 2022;198(5):923-6.
350. Park HS, Son BR, Shin KS, Kim HK, Yang Y, Jeong Y, et al. Germline JAK2 V617F mutation as a susceptibility gene causing myeloproliferative neoplasm in first-degree relatives. Leuk Lymphoma. 2020;61(13):3251-4.
351. Mead AJ, Chowdhury O, Pecquet C, Dusa A, Woll P, Atkinson D, et al. Impact of isolated germline JAK2V617I mutation on human hematopoiesis. Blood. 2013;121(20):4156-65.
352. Yoshimitsu M, Hachiman M, Uchida Y, Arima N, Arai A, Kamada Y, et al. Essential thrombocytosis attributed to JAK2-T875N germline mutation. International Journal of Hematology. 2019;110(5):584-90.
353. Jones AV, Chase A, Silver RT, Oscier D, Zoi K, Wang YL, et al. JAK2 haplotype is a major risk factor for the development of myeloproliferative neoplasms. Nat Genet. 2009;41(4):446-9.
354. Kilpivaara O, Mukherjee S, Schram AM, Wadleigh M, Mullally A, Ebert BL, et al. A germline JAK2 SNP is associated with predisposition to the development of JAK2(V617F)-positive myeloproliferative neoplasms. Nat Genet. 2009;41(4):455-9.
355. Van Loo P, Campbell PJ. ABSOLUTE cancer genomics. Nat Biotechnol. 2012;30(7):620-1.
356. Coffee B, Cox HC, Bernhisel R, Manley S, Bowles K, Roa BB, et al. A substantial proportion of apparently heterozygous TP53 pathogenic variants detected with a next-generation sequencing hereditary pan-cancer panel are acquired somatically. Hum Mutat. 2020;41(1):203-11.
357. Judkins T, Leclair B, Bowles K, Gutin N, Trost J, McCulloch J, et al. Development and analytical validation of a 25-gene next generation sequencing panel that includes the BRCA1 and BRCA2 genes to assess hereditary cancer risk. BMC Cancer. 2015;15:215.
358. Raymond VM, Gray SW, Roychowdhury S, Joffe S, Chinnaiyan AM, Parsons DW, et al. Germline Findings in Tumor-Only Sequencing: Points to Consider for Clinicians and Laboratories. J Natl Cancer Inst. 2016;108(4).
359. Depner M, Fuchs S, Raabe J, Frede N, Glocker C, Doffinger R, et al. The Extended Clinical Phenotype of 26 Patients with Chronic Mucocutaneous Candidiasis due to Gain-of-Function Mutations in STAT1. J Clin Immunol. 2016;36(1):73-84.
360. Tefferi A, Lasho TL, Schwager SM, Steensma DP, Mesa RA, Li CY, et al. The JAK2(V617F) tyrosine kinase mutation in myelofibrosis with myeloid metaplasia: lineage specificity and clinical correlates. Br J Haematol. 2005;131(3):320-8.
361. Gupta N, Edelmann B, Schnoeder TM, Saalfeld FC, Wolleschak D, Kliche S, et al. JAK2-V617F activates beta1-integrin-mediated adhesion of granulocytes to vascular cell adhesion molecule 1. Leukemia. 2017;31(5):1223-6.

362. Martin-Fernandez M, Tamayo-Velasco A, Aller R, Gonzalo-Benito H, Martinez-Paz P, Tamayo E. Endothelial Dysfunction and Neutrophil Degranulation as Central Events in Sepsis Physiopathology. Int J Mol Sci. 2021;22(12).
363. Shen XF, Cao K, Jiang JP, Guan WX, Du JF. Neutrophil dysregulation during sepsis: an overview and update. J Cell Mol Med. 2017;21(9):1687-97.
364. Criss AK, Seifert HS. A bacterial siren song: intimate interactions between Neisseria and neutrophils. Nat Rev Microbiol. 2012;10(3):178-90.
365. Brown KA, Treacher DF. Neutrophils as potential therapeutic targets in sepsis. Discov Med. 2006;6(33):118-22.
366. Palmer CS, Kimmey JM. Neutrophil Recruitment in Pneumococcal Pneumonia. Front Cell Infect Microbiol. 2022;12:894644.
367. Wang X, Qin W, Zhang Y, Zhang H, Sun B. Endotoxin promotes neutrophil hierarchical chemotaxis via the p38-membrane receptor pathway. Oncotarget. 2016;7(45):74247-58.
368. Yagi Y, Otani H, Ando S, Oshiro A, Kawai K, Nishikawa H, et al. Involvement of Rho signaling in PAR2-mediated regulation of neutrophil adhesion to lung epithelial cells. Eur J Pharmacol. 2006;536(1-2):19-27.
369. Saunders CA, Majumdar R, Molina Y, Subramanian BC, Parent CA. Genetic manipulation of PLB-985 cells and quantification of chemotaxis using the underagarose assay. Methods Cell Biol. 2019;149:31-56.
370. Pedruzzi E, Fay M, Elbim C, Gaudry M, Gougerot-Pocidalo MA. Differentiation of PLB-985 myeloid cells into mature neutrophils, shown by degranulation of terminally differentiated compartments in response to N-formyl peptide and priming of superoxide anion production by granulocyte-macrophage colony-stimulating factor. Br J Haematol. 2002;117(3):719-26.
371. Coombs C, Georgantzoglou A, Walker HA, Patt J, Merten N, Poplimont H, et al. Chemokine receptor trafficking coordinates neutrophil clustering and dispersal at wounds in zebrafish. Nat Commun. 2019;10(1):5166.
372. Poplimont H, Georgantzoglou A, Boulch M, Walker HA, Coombs C, Papaleonidopoulou F, et al. Neutrophil Swarming in Damaged Tissue Is Orchestrated by Connexins and Cooperative Calcium Alarm Signals. Curr Biol. 2020;30(14):2761-76 e7.
373. Liu L, Oza S, Hogan D, Perin J, Rudan I, Lawn JE, et al. Global, regional, and national causes of child mortality in 2000-13, with projections to inform post-2015 priorities: an updated systematic analysis. Lancet. 2015;385(9966):430-40.
374. Martin NG, Sadarangani M, Pollard AJ, Goldacre MJ. Hospital admission rates for meningitis and septicaemia caused by Haemophilus influenzae, Neisseria meningitidis, and Streptococcus pneumoniae in children in England over five decades: a population-based observational study. Lancet Infect Dis. 2014;14(5):397-405.
375. Ho M, Thompson B, Fisk JN, Nebert DW, Bruford EA, Vasiliou V, et al. Update of the keratin gene family: evolution, tissue-specific expression patterns, and relevance to clinical disorders. Hum Genomics. 2022;16(1):1.
376. Kim J, Villadsen R. The Expression Pattern of Epidermal Differentiation Marker Keratin 10 in the Normal Human Breast and Breast Cancer Cells. J Histochem Cytochem. 2020;68(8):561-70.
377. Ali D, Alhattab D, Jafar H, Alzubide M, Sharar N, Bdour S, et al. Differential Marker Expression between Keratinocyte Stem Cells and Their Progeny Generated from a Single Colony. Int J Mol Sci. 2021;22(19).
378. Kintarak S, Whawell SA, Speight PM, Packer S, Nair SP. Internalization of Staphylococcus aureus by human keratinocytes. Infect Immun. 2004;72(10):5668-75.
379. Moran MC, Pandya RP, Leffler KA, Yoshida T, Beck LA, Brewer MG. Characterization of Human Keratinocyte Cell Lines for Barrier Studies. JID Innov. 2021;1(2):100018.
380. Pan Y, Feng C, Wang H, Lee M, Tang Z, Lin Z. Ichthyosis with confetti caused by new and recurrent mutations in KRT10 associated with varying degrees of keratin 10 mis-localization. J Dermatol Sci. 2020;98(1):35-40.
381. Foster TJ, Geoghegan JA, Ganesh VK, Hook M. Adhesion, invasion and evasion: the many functions of the surface proteins of Staphylococcus aureus. Nat Rev Microbiol. 2014;12(1):49-62.

382. O'Brien LM, Walsh EJ, Massey RC, Peacock SJ, Foster TJ. Staphylococcus aureus clumping factor B (ClfB) promotes adherence to human type I cytokeratin 10: implications for nasal colonization. Cell Microbiol. 2002;4(11):759-70.
383. Walsh EJ, O'Brien LM, Liang X, Hook M, Foster TJ. Clumping factor B, a fibrinogen-binding MSCRAMM (microbial surface components recognizing adhesive matrix molecules) adhesin of Staphylococcus aureus, also binds to the tail region of type I cytokeratin 10. J Biol Chem. 2004;279(49):50691-9.
384. Shivshankar P, Sanchez C, Rose LF, Orihuela CJ. The Streptococcus pneumoniae adhesin PsrP binds to Keratin 10 on lung cells. Mol Microbiol. 2009;73(4):663-79.
385. Rothnagel JA, Dominey AM, Dempsey LD, Longley MA, Greenhalgh DA, Gagne TA, et al. Mutations in the rod domains of keratins 1 and 10 in epidermolytic hyperkeratosis. Science. 1992;257(5073):1128-30.
386. Rothnagel JA, Fisher MP, Axtell SM, Pittelkow MR, Anton-Lamprecht I, Huber M, et al. A mutational hot spot in keratin 10 (KRT 10) in patients with epidermolytic hyperkeratosis. Hum Mol Genet. 1993;2(12):2147-50.
387. Cheraghlou S, Atzmony L, Roy SF, McNiff JM, Choate KA. Mutations in KRT10 in epidermolytic acanthoma. J Cutan Pathol. 2020;47(6):524-9.
388. Chipev CC, Yang JM, DiGiovanna JJ, Steinert PM, Marekov L, Compton JG, et al. Preferential sites in keratin 10 that are mutated in epidermolytic hyperkeratosis. Am J Hum Genet. 1994;54(2):179-90.
389. Cheng J, Syder AJ, Yu QC, Letai A, Paller AS, Fuchs E. The genetic basis of epidermolytic hyperkeratosis: a disorder of differentiation-specific epidermal keratin genes. Cell. 1992;70(5):811-9.
390. Paller AS, Syder AJ, Chan YM, Yu QC, Hutton E, Tadini G, et al. Genetic and clinical mosaicism in a type of epidermal nevus. N Engl J Med. 1994;331(21):1408-15.
391. Syder AJ, Yu QC, Paller AS, Giudice G, Pearson R, Fuchs E. Genetic mutations in the K1 and K10 genes of patients with epidermolytic hyperkeratosis. Correlation between location and disease severity. J Clin Invest. 1994;93(4):1533-42.
392. Muller FB, Huber M, Kinaciyan T, Hausser I, Schaffrath C, Krieg T, et al. A human keratin 10 knockout causes recessive epidermolytic hyperkeratosis. Hum Mol Genet. 2006;15(7):1133-41.
393. Tsubota A, Akiyama M, Kanitakis J, Sakai K, Nomura T, Claudy A, et al. Mild recessive bullous congenital ichthyosiform erythroderma due to a previously unidentified homozygous keratin 10 nonsense mutation. J Invest Dermatol. 2008;128(7):1648-52.
394. Burger B, Ghosh A, Ng CKY, Piscuoglio S, Spoerri I, Itin PH, et al. Discovery of heterozygous KRT10 alterations in MAUIE cases underlines the importance of regular skin cancer screening in ichthyosis with confetti. Br J Dermatol. 2020;183(5):954-5.
395. Lim YH, Qiu J, Saraceni C, Burrall BA, Choate KA. Genetic Reversion via Mitotic Recombination in Ichthyosis with Confetti due to a KRT10 Polyalanine Frameshift Mutation. J Invest Dermatol. 2016;136(8):1725-8.
396. Renz P, Imahorn E, Spoerri I, Aushev M, March OP, Wariwoda H, et al. Arginine- but not alanine-rich carboxy-termini trigger nuclear translocation of mutant keratin 10 in ichthyosis with confetti. J Cell Mol Med. 2019;23(12):8442-52.
397. Saito R, Boyce A, Hsu CK, Rashidghamat E, Hide M, Wedgeworth EK, et al. Predictive phenotyping of inherited ichthyosis by next-generation DNA sequencing. Br J Dermatol. 2017;176(1):249-51.
398. Spoerri I, Brena M, De Mesmaeker J, Schlipf N, Fischer J, Tadini G, et al. The phenotypic and genotypic spectra of ichthyosis with confetti plus novel genetic variation in the 3' end of KRT10: from disease to a syndrome. JAMA Dermatol. 2015;151(1):64-9.
399. Sajjan U, Ackerley C, Forstner J. Interaction of cblA/adhesin-positive Burkholderia cepacia with squamous epithelium. Cell Microbiol. 2002;4(2):73-86.
400. Budak G, Eren Ozsoy O, Aydin Son Y, Can T, Tuncbag N. Reconstruction of the temporal signaling network in Salmonella-infected human cells. Front Microbiol. 2015;6:730.
401. Schulte T, Lofling J, Mikaelsson C, Kikhney A, Hentrich K, Diamante A, et al. The basic keratin 10-binding domain of the virulence-associated pneumococcal serine-rich protein PsrP adopts a novel MSCRAMM fold. Open Biol. 2014;4(1):130090.

402. McAleese FM, Walsh EJ, Sieprawska M, Potempa J, Foster TJ. Loss of clumping factor B fibrinogen binding activity by Staphylococcus aureus involves cessation of transcription, shedding and cleavage by metalloprotease. J Biol Chem. 2001;276(32):29969-78.
403. Hernandez N, Melki I, Jing H, Habib T, Huang SSY, Danielson J, et al. Life-threatening influenza pneumonitis in a child with inherited IRF9 deficiency. J Exp Med. 2018;215(10):2567-85.
404. Ciancanelli MJ, Huang SX, Luthra P, Garner H, Itan Y, Volpi S, et al. Infectious disease. Life-threatening influenza and impaired interferon amplification in human IRF7 deficiency. Science. 2015;348(6233):448-53.
405. Ning S, Pagano JS, Barber GN. IRF7: activation, regulation, modification and function. Genes Immun. 2011;12(6):399-414.
406. Sole-Violan J, Sologuren I, Betancor E, Zhang S, Pérez C, Herrera-Ramos E, et al., editors. Lethal influenza virus A H1N1 infection in two relatives with autosomal dominant GATA-2 deficiency. 33rd International Symposium on Intensive Care and Emergency Medicine; 2013; Brussels, Belgium: Crit Care
407. Sologuren I, Martinez-Saavedra MT, Sole-Violan J, de Borges de Oliveira E, Jr., Betancor E, Casas I, et al. Lethal Influenza in Two Related Adults with Inherited GATA2 Deficiency. J Clin Immunol. 2018;38(4):513-26.
408. Grasselli G, Zangrillo A, Zanella A, Antonelli M, Cabrini L, Castelli A, et al. Baseline Characteristics and Outcomes of 1591 Patients Infected With SARS-CoV-2 Admitted to ICUs of the Lombardy Region, Italy. JAMA. 2020;323(16):1574-81.
409. O'Driscoll M, Ribeiro Dos Santos G, Wang L, Cummings DAT, Azman AS, Paireau J, et al. Age-specific mortality and immunity patterns of SARS-CoV-2. Nature. 2021;590(7844):140-5.
410. Zhang SY, Zhang Q, Casanova JL, Su HC, Team C. Severe COVID-19 in the young and healthy: monogenic inborn errors of immunity? Nat Rev Immunol. 2020;20(8):455-6.
411. Zietz M, Zucker J, Tatonetti NP. Testing the association between blood type and COVID-19 infection, intubation, and death. medRxiv. 2020.
412. Zeberg H, Paabo S. The major genetic risk factor for severe COVID-19 is inherited from Neanderthals. Nature. 2020;587(7835):610-2.
413. Casanova JL, Su HC, Effort CHG. A Global Effort to Define the Human Genetics of Protective Immunity to SARS-CoV-2 Infection. Cell. 2020;181(6):1194-9.
414. Zhang Q. Human genetics of life-threatening influenza pneumonitis. Hum Genet. 2020;139(6-7):941-8.
415. Han KJ, Su X, Xu LG, Bin LH, Zhang J, Shu HB. Mechanisms of the TRIF-induced interferon-stimulated response element and NF-kappaB activation and apoptosis pathways. J Biol Chem. 2004;279(15):15652-61.
416. Kaiser WJ, Offermann MK. Apoptosis induced by the toll-like receptor adaptor TRIF is dependent on its receptor interacting protein homotypic interaction motif. J Immunol. 2005;174(8):4942-52.
417. Audry M, Ciancanelli M, Yang K, Cobat A, Chang HH, Sancho-Shimizu V, et al. NEMO is a key component of NF-kappaB- and IRF-3-dependent TLR3-mediated immunity to herpes simplex virus. J Allergy Clin Immunol. 2011;128(3):610-7 e1-4.
418. Dupuis S, Jouanguy E, Al-Hajjar S, Fieschi C, Al-Mohsen IZ, Al-Jumaah S, et al. Impaired response to interferon-alpha/beta and lethal viral disease in human STAT1 deficiency. Nat Genet. 2003;33(3):388-91.
419. Ahmad L, Mashbat B, Leung C, Brookes C, Hamad S, Krokowski S, et al. Human TANK-binding kinase 1 is required for early autophagy induction upon herpes simplex virus 1 infection. J Allergy Clin Immunol. 2019;143(2):765-9 e7.
420. Sabli IK, Sancho-Shimizu V. Inborn errors of autophagy and infectious diseases. Curr Opin Immunol. 2021;72:272-6.
421. Tao S, Zhu L, Lee P, Lee WM, Knox K, Chen J, et al. Negative control of TLR3 signaling by TICAM1 down-regulation. Am J Respir Cell Mol Biol. 2012;46(5):660-7.
422. Inomata M, Niida S, Shibata K, Into T. Regulation of Toll-like receptor signaling by NDP52-mediated selective autophagy is normally inactivated by A20. Cell Mol Life Sci. 2012;69(6):963-79.

423. Organisation WH. Child mortality (under 5 years) 2022 [Available from:
.
424. Niewiesk S. Maternal antibodies: clinical significance, mechanism of interference with immune responses, and possible vaccination strategies. Front Immunol. 2014;5:446.
425. Ferreras-Antolin L, Oligbu G, Okike IO, Ladhani S. Infection is associated with one in five childhood deaths in England and Wales: analysis of national death registrations data, 2013-15. Arch Dis Child. 2020;105(9):857-63.
426. Martinon-Torres F, Salas A, Rivero-Calle I, Cebey-Lopez M, Pardo-Seco J, Herberg JA, et al. Life-threatening infections in children in Europe (the EUCLIDS Project): a prospective cohort study. Lancet Child Adolesc Health. 2018;2(6):404-14.
427. Chandesris MO, Melki I, Natividad A, Puel A, Fieschi C, Yun L, et al. Autosomal dominant STAT3 deficiency and hyper-IgE syndrome: molecular, cellular, and clinical features from a French national survey. Medicine (Baltimore). 2012;91(4):e1-e19.
428. Grimbacher B, Holland SM, Puck JM. Hyper-IgE syndromes. Immunol Rev. 2005;203:244-50.
429. Minegishi Y, Saito M, Tsuchiya S, Tsuge I, Takada H, Hara T, et al. Dominant-negative mutations in the DNA-binding domain of STAT3 cause hyper-IgE syndrome. Nature. 2007;448(7157):1058-62.
430. Giacomelli M, Tamassia N, Moratto D, Bertolini P, Ricci G, Bertulli C, et al. SH2-domain mutations in STAT3 in hyper-IgE syndrome patients result in impairment of IL-10 function. Eur J Immunol. 2011;41(10):3075-84.
431. He J, Shi J, Xu X, Zhang W, Wang Y, Chen X, et al. STAT3 mutations correlated with hyper-IgE syndrome lead to blockage of IL-6/STAT3 signalling pathway. J Biosci. 2012;37(2):243-57.
432. Gernez Y, Freeman AF, Holland SM, Garabedian E, Patel NC, Puck JM, et al. Autosomal Dominant Hyper-IgE Syndrome in the USIDNET Registry. J Allergy Clin Immunol Pract. 2018;6(3):996-1001.
433. Kroner C, Neumann J, Ley-Zaporozhan J, Hagl B, Meixner I, Spielberger BD, et al. Lung disease in STAT3 hyper-IgE syndrome requires intense therapy. Allergy. 2019;74(9):1691-702.
434. van de Veerdonk FL, Plantinga TS, Hoischen A, Smeekens SP, Joosten LA, Gilissen C, et al. STAT1 mutations in autosomal dominant chronic mucocutaneous candidiasis. N Engl J Med. 2011;365(1):54-61.
435. Zheng J, van de Veerdonk FL, Crossland KL, Smeekens SP, Chan CM, Al Shehri T, et al. Gain-of-function STAT1 mutations impair STAT3 activity in patients with chronic mucocutaneous candidiasis (CMC). Eur J Immunol. 2015;45(10):2834-46.
436. Walker LS. Treg and CTLA-4: two intertwining pathways to immune tolerance. J Autoimmun. 2013;45(100):49-57.
437. Read S, Malmstrom V, Powrie F. Cytotoxic T lymphocyte-associated antigen 4 plays an essential role in the function of CD25(+)CD4(+) regulatory cells that control intestinal inflammation. J Exp Med. 2000;192(2):295-302.
438. Bour-Jordan H, Bluestone JA. Regulating the regulators: costimulatory signals control the homeostasis and function of regulatory T cells. Immunol Rev. 2009;229(1):41-66.
439. Collins AV, Brodie DW, Gilbert RJ, Iaboni A, Manso-Sancho R, Walse B, et al. The interaction properties of costimulatory molecules revisited. Immunity. 2002;17(2):201-10.
440. Qureshi OS, Zheng Y, Nakamura K, Attridge K, Manzotti C, Schmidt EM, et al. Trans-endocytosis of CD80 and CD86: a molecular basis for the cell-extrinsic function of CTLA-4. Science. 2011;332(6029):600-3.
441. Walker LS, Sansom DM. The emerging role of CTLA4 as a cell-extrinsic regulator of T cell responses. Nat Rev Immunol. 2011;11(12):852-63.
442. Kuehn HS, Ouyang W, Lo B, Deenick EK, Niemela JE, Avery DT, et al. Immune dysregulation in human subjects with heterozygous germline mutations in CTLA4. Science. 2014;345(6204):1623-7.
443. Schubert D, Bode C, Kenefeck R, Hou TZ, Wing JB, Kennedy A, et al. Autosomal dominant immune dysregulation syndrome in humans with CTLA4 mutations. Nat Med. 2014;20(12):1410-6.
444. Schwab C, Gabrysch A, Olbrich P, Patino V, Warnatz K, Wolff D, et al. Phenotype, penetrance, and treatment of 133 cytotoxic T-lymphocyte antigen 4-insufficient subjects. J Allergy Clin Immunol. 2018;142(6):1932-46.

445. Zeissig S, Petersen BS, Tomczak M, Melum E, Huc-Claustre E, Dougan SK, et al. Early-onset Crohn's disease and autoimmunity associated with a variant in CTLA-4. Gut. 2015;64(12):1889-97.
446. Hofer T, Ray N, Wegmann D, Excoffier L. Large allele frequency differences between human continental groups are more likely to have occurred by drift during range expansions than by selection. Ann Hum Genet. 2009;73(1):95-108.
447. Cooke GS, Hill AV. Genetics of susceptibility to human infectious disease. Nat Rev Genet. 2001;2(12):967-77.
448. Petersen L, Nielsen GG, Andersen PK, Sorensen TI. Case-control study of genetic and environmental influences on premature death of adult adoptees. Genet Epidemiol. 2002;23(2):123-32.
449. Peter Rout D, Nair A, Gupta A, Kumar P. Epidermolytic hyperkeratosis: clinical update. Clin Cosmet Investig Dermatol. 2019;12:333-44.
450. Li H, Torma H. Retinoids reduce formation of keratin aggregates in heat-stressed immortalized keratinocytes from an epidermolytic ichthyosis patient with a KRT10 mutation*. Acta Derm Venereol. 2013;93(1):44-9.
451. Virtanen M, Gedde-Dahl T, Jr., Mork NJ, Leigh I, Bowden PE, Vahlquist A. Phenotypic/genotypic correlations in patients with epidermolytic hyperkeratosis and the effects of retinoid therapy on keratin expression. Acta Derm Venereol. 2001;81(3):163-70.
452. Bastard P, Rosen LB, Zhang Q, Michailidis E, Hoffmann HH, Zhang Y, et al. Autoantibodies against type I IFNs in patients with life-threatening COVID-19. Science. 2020;370(6515).
453. Ku CS, Cooper DN, Polychronakos C, Naidoo N, Wu M, Soong R. Exome sequencing: dual role as a discovery and diagnostic tool. Ann Neurol. 2012;71(1):5-14.
454. Weischenfeldt J, Symmons O, Spitz F, Korbel JO. Phenotypic impact of genomic structural variation: insights from and for human disease. Nat Rev Genet. 2013;14(2):125-38.
455. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature. 2016;533(7603):420-4.
456. Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. Nature. 2019;576(7785):149-57.
457. Togni M, Swanson KD, Reimann S, Kliche S, Pearce AC, Simeoni L, et al. Regulation of in vitro and in vivo immune functions by the cytosolic adaptor protein SKAP-HOM. Mol Cell Biol. 2005;25(18):8052-63.
458. Nagel W, Zeitlmann L, Schilcher P, Geiger C, Kolanus J, Kolanus W. Phosphoinositide 3-OH kinase activates the beta2 integrin adhesion pathway and induces membrane recruitment of cytohesin-1. J Biol Chem. 1998;273(24):14853-61.
459. Zhang H, Wang Y, Qu M, Li W, Wu D, Cata JP, et al. Neutrophil, neutrophil extracellular traps and endothelial cell dysfunction in sepsis. Clin Transl Med. 2023;13(1):e1170.

Chapter 8 Appendix

8.1 Gene lists used to filter WES data

8.1.1 Mendelian susceptibility to mycobacterial diseases (MSMD)

IL12RB1, IL12B, IL23R, IL12RB2, SPPL2A, IRF8, IFNGR1, IFNGR2, STAT1, NEMO, IKBKG, CYBB, TYK2, JAK1, RORC, ISG15

8.1.2 Herpes simplex virus infection

TLR3, UNC93B1, TICAM1, TRAF3, TBK1, IRF3, SNORA31, DBR1, NEMO, STAT1

8.1.3 Neisseria meningitidis infection

SPLUNC1, CFP, C2, C3, C5, C6, C7, C8B, C9, CFD, CFB, CFI, IRAK4, IKBKG, GATA2, CEACAM3, CEACAM6, SFTPA2, CFHR3, CFH, MBL2, TLR4, TLR9, TNF, IL1B, IL1RN, IL-6 , IL10, FCGR2A, FCGR3B, F5, SERPINE1, PROC, CPB2

8.1.4 Candidiasis

IL17RA, IL17RC, IL17F, STAT1, TRAF3IP2, CARD9, AIRE, STAT3, TYK2, DOCK8, CD25, IL12RB1, RORC, MBL2, NLRP3, TLR1, TLR2, TLR3, TLR4, PTPN22, IL12B, IL10, IL4, DECTIN1, DEFB1

8.2 INGENIOS patient details

Table 8.1 Recruited INGENIOS patient details

Patient ID	Sex	Age at Onset (months)	Ethnicity	Clinical Diagnosis	Pathogen detected	WES	Trio recruited
1001	F	188	North/Mid/East Europe	Meningococcal B Septicaemia - right lower leg necrosis - compartment syndrome	*N. meningitidis* B via blood culture	Y	N
1002	F	113	North/Mid/East Europe	HSV Encephalitis - subsequent NMDAR encephalitis 20/02/2018	HSV-1 via CSF PCR	Y	Y
1003	M	84	North/Mid/East Europe	VZV meningitis	VZV via CSF PCR	Y	N
1004	M	2	North/Mid/East Europe	E. coli bacteraemia, E. coli meningitis with subdural collection	E.coli vis blood and CSF culture	Y	Y
1005	F	2	Polish and Middle East	HHV-6 severe illness 2012, bone marrow suppression, DAT-negative haemolysis (ongoing)	HHV-6 via blood PCR and bone marrow	Y	Y
1006	F	1	Polish and Middle East	Unexplained haemolytic anaemia	HHV-6 via blood PCR	Y	Y

1007	M	135	South Asian	Fungal chest infection candida parasitosis on BAL	Candida via blood PCR	Y	N
1008	F	181	South Asian	Left lilac vein thrombosis		N	N
1009	M	5	North/Mid/East Europe	HSV-1 encephalitis in childhood, subsequent recurrent superficial herpetic lesions on face without significant progression	HSV-1 via CSF PCR	Y	Y
1010	M		North African	Immune thrombocytopenia and autoimmune neutropenia		N	N
1011	M	82	South Asian	Fever R50		Y	Y
1012	F	85	North African	Recurrent fever		Y	Y
1013	F	39	South Asian	HSV-1 encephalitis	HSV-1 via CSF PCR	Y	Y
1014	M		North/Mid/East Europe	Persistent MAI lung infection, low interferon gamma production	Nasopharyngeal aspirate PCR - detected influenza A H1, RSV, Rhinovirus/enterovirus, parainfluenza 4	N	N
1015	M	63	North/Mid/East Europe; Native American	Kawasaki disease		Y	Y

1016	M	199	North/Mid/East Europe	Subdural empyema with intracerebral abscess with frontal sinus extension	Fusobacterium necrophorum via blood culture	Y	Y
1017	F	202	North/Mid/East Europe	HSV Encephalitis	HSV-1 via CSF PCR	Y	Y
1018	M			Pneumococcal disease		Y	Y
1019	M	8	North/Mid/East Europe	Type B Neisseria meningitidis - sepsis	*N. meningitidis* B via blood culture	N	N
1020	F	2	Sub-Saharan	Listeria meningitis, raised ICP, VP shunt	Listeria monocytogenes via CSF PCR	Y	N
1021	F	149		HSE at birth and again in 2019		Y	Y
1022	M		South Asian	Disseminated rifampicin-resistant BCG (meningitis, miliary disease, colitis), Severe colitis - defunctioning ileostomy - reversed Sep. 2019 after 6 months TPN	Mycobacterium Bovis (BCG strain) via sputum culture	Y	Y
1023	F	2	North/Mid/East Europe	HSV infection of skin and eyes	HSV-1 detected via eye swab PCR NOT via CSF	Y	Y
1024	M			Neonatal HSE (fatal) and mum with HSV1 colitis		Y	Y

1025	M	195	North/Mid/East Europe	Cerebral aspergilloma with left frontal lobe abscess extending into left orbit - left orbital exenteration	Aspergillus detected in pus from left frontal burrhole aspirations	Y	N
1026	M	107	South Asian	Autoimmune encephalitis		Y	Y
1027	M	14	North/Mid/East Europe	Autoimmune encephalitis (influenza A)	Influenza A via nasopharyngeal aspirate and CSF PCR	Y	Y
1028	F	72	North/Mid/East Europe	Recurrent oral thrush	Candida albicans via mouth swab	Y	N
1029	M	93	Sub-Saharan	Meningococcal septic arthritis of right knee, SARS-CoV-2 co-infection	N. meningitidis (group W135 type 2a subtype p1.5 p1.2) from right knee joint fluid	Y	N
1031	M	145	South East Asian	Sydenham Chorea	ASOT positive	Y	Y
1031S	M	112	South East Asian	Recurrent fever	ASOT positive	Y	Y
1032	M		South European	Recurrent HSV lesions around left eye since 2012	HSV-1 via eye swab	Y	Y
1033	F	2	North/Mid/East Europe	Recurrent fever		Y	N

1034	M		North/Mid/East Europe	Severe meningococcal sepsis with 4-limb amputation		Y	Y
1035	F		North/Mid/East Europe	Sepsis - 2 episodes of meningococcal disease, with no immune deficit found		Y	Y
1036	M		North/Mid/East Europe	HSV-1 infection (skin/eye/mouth) in neonatal period, and again at age 2 with encephalitis		Y	Y
1037	M	5	North/Mid/East Europe	Recurrent fever from August 2019, associated with shivering and rigors. Fever lasts 8-10 days and recurs in 3 weeks.		Y	Y
1038	M		South Asian	Ongoing fluctuating CRP (nil focus identified)		Y	Y
1039	M	157	Sub-Saharan	Musculoskeletal - osteomyelitis in the right distal femoral metaphysis		Y	N
1040	M	93	North/Mid/East Europe	16/11/2020 - unwell with sore throat; since then behaviour change, inc. episodes of screaming, shouting and hitting face and head		Y	N

1042	M	2	North/Mid/East Europe	Meningitis; bacteraemia Streptococcus pneumoniae serotype 35F	beta-haemolytic streptococcus group C/G via nose/throat swab culture	Y	Y
1043	M		North/Mid/East Europe	Recurrent shingles		Y	N
1044	F	73	South European	Recurrent right peri-orbital HSV infections	HSV-1 via vesicle and eye swab	Y	Y
1045	F	2	North/Mid/East Europe	Neonatal HSV-2 Encephalitis	HSV-2 via CSF PCR	Y	Y
1046	F		Jewish	Atypical mycobacterial infection with Mycobacterium avium, with large abscess formation in the neck. Completed a year of 4-drug treatment (Clarithromycin, Moxifloxacin, Ethambutol, Rifampicin)	Mycobacterium avium with Achromobacter from lymph node biopsy	Y	Y
1047	F	27	Mixed: Turkish and Scottish	HHV-6 in blood; chronic ITP	HHV6 via blood PCR	Y	Y
1048	M			x2 episodes HSV-1: after birth (severe, disseminated), and age 7 (HSE)		Y	Y

1049	M	13	South Asian (Indian)	Rash of unknown cause		Y	Y
1050	M			Kawasaki disease with strong family history		Y	Y
1051	M	141	North/Mid/East Europe	VZV Meningitis	VZV via PCR	Y	Y
1052	F	5	Mother of English and Syrian descent. Father Syrian.	HSV-1 encephalitis.	HSV-1 via CSF PCR	Y	N
1053	F	165	North/Mid/East Europe	COVID pneumonitis	SARS-CoV-2 PCR positive	Y	Y
1054	M	14	West Asian	Streptococcus pneumoniae; haemolytic uraemic syndrome.	streptococcus pneumonia from nasopharyngeal aspirate and blood culture	Y	Y
1055	F	2	Mother of Iranian and Singaporean heritage and father Brazilian	Neonatal HSV-2 infection. Vesicular lesions	HSV-2 via blood and eye swab PCR	Y	N

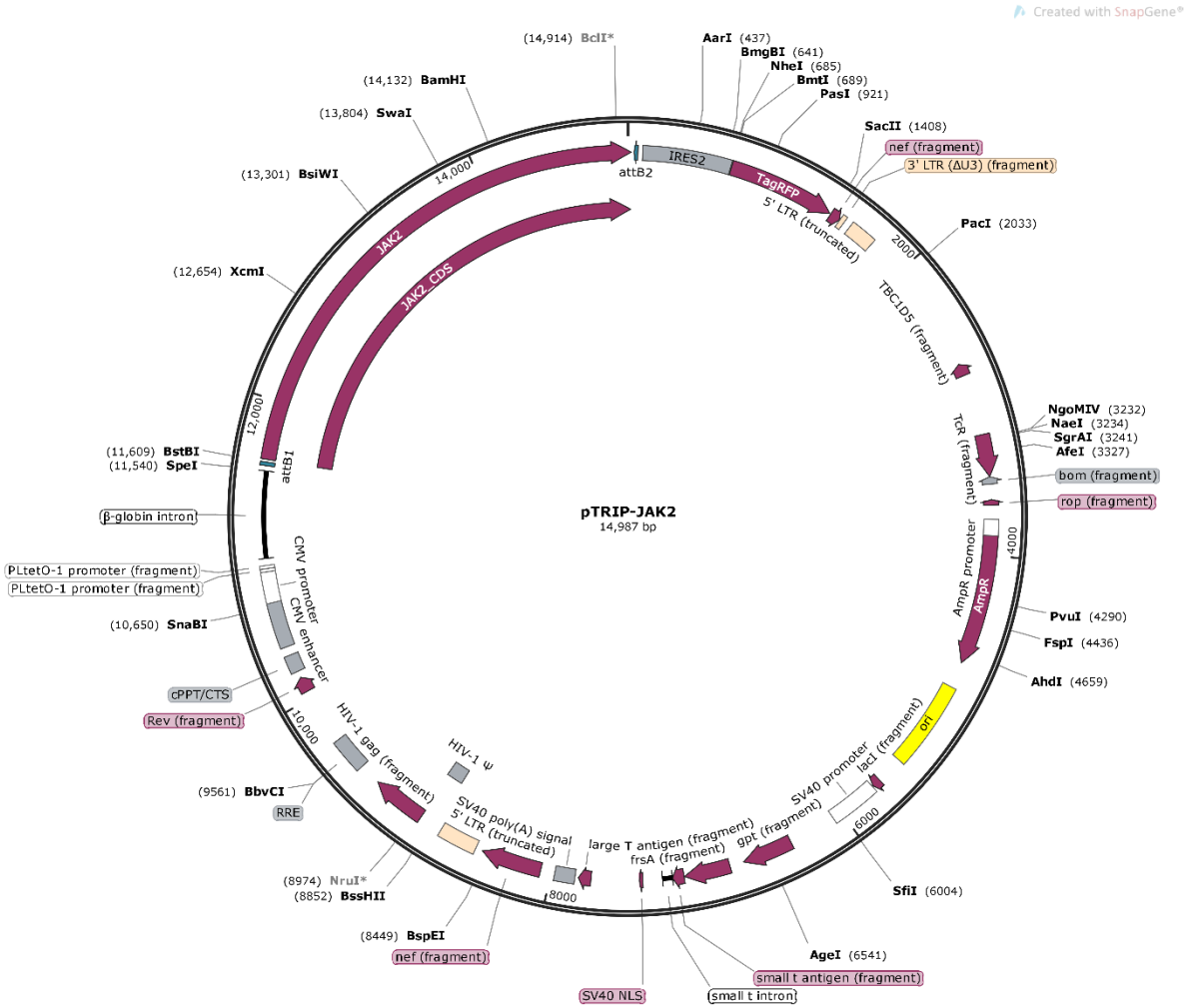
1056	F	3	North/Mid/East Europe	HSV encephalitis	HSV-1 via CSF PCR	Y	Y
1057	F	168	Sub Saharan	Meningoencephalitis		Y	Y
1058	F	33	North/Mid/East Europe	Acute flaccid myelitis		Y	Y
1059	F	14	North/Mid/East Europe	Staphylococcus aureus meningitis. HHV6 in blood and bone marrow, completed 14 days of foscarnet. Bicytopena- neutropenia and anaemia (requiring blood transfusions)	Data missing	Y	Y
1060	M	283	West Asian	Disseminated VZV infection	Blood and forehead swab PCR	Y	N
1061	F	0	North/Mid/East Europe	Recurrent HSV-2 meningitis (prophylaxis treatment since first episode of HSV-2)	HSV-2 via CSF PCR	Y	N
1062	M	6	West Asian	MRSA pneumonia with right loculated empyemas	Pleural fluid MRSA positive	Y	N
1063	F	159	North/Mid/East Europe	2x episodes of PIMS-TS		N	N
1065	M	1	North/Mid/East Europe	HSV-2 encephalitis; neonatal HSV-2 infection skin, eye and mouth.	HSV-2 via blood and skin PCR	Y	Y

1066	F		North/Mid/East Europe	Recurrent Shingles		Y	N
1067	F		South Asian	VZV meningitis, full recovery. Measles despite vaccination, full recovery 2012. COVID 2 episodes 16/12/21, 02/22.		Y	N
1070	F	250	North/Mid/East Europe	Right sided Mastoiditis and cerebral Venous Sinus Thrombosis post head injury.		Y	Y
1071	F	19	North/Mid/East Europe	Recurrent cutaneous HSV infection	HSV IgG		
1072	M	51	South Asian	Recurrent monthly fevers without cause		Y	Y
1073	F	24	Mixed: Italian and Egyptian	Recurrent Herpes	HSV IgG	Y	Y
1075	M	96	North/Mid/East Europe	Recurrent labial, hand and arm herpes simplex lesions	HSV IgG, Measles IgG	Y	N
1076	M	36	North/Mid/East Europe	Recurrent orofacial herpes	HSV IgG	Y	Y
1077	F	169	East Asian	Streptococcus Bacteraemia	Streptococcus Intermedius via Blood culture:	Y	Y

					Streptococcus Milleri via Pus from Liver Abscess		

8.3 Plasmid Maps

8.3.1 pTRIP-JAK2 wildtype plasmid map



8.3.2 pEF-bos-TICAM1 wildtype plasmid map



8.3.3 pCMV-KRT10 wildtype plasmid map

