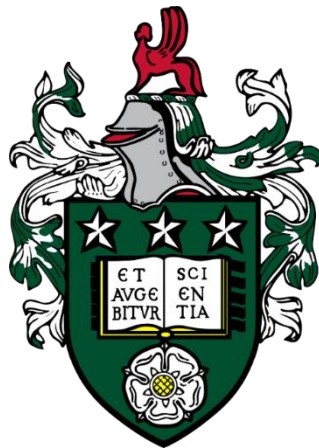


Development of Gene Expression Assay to Stratify Autoimmune Connective Tissue Diseases

Zoe Frances Wigston

Submitted in accordance with the requirements for the degree of
Master of Science by Research



The University of Leeds
School of Medicine

September 2022

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Zoe Frances Wigston to be identified as Author of this work has been asserted by her in accordance with the Copyright, Designs and Patents Act 1988.

Acknowledgements

The work within this report could not be possible without the guidance and support from my supervisors; Dr Edward Vital, Dr Miriam Wittmann and Miss Michelle Wilson. Most importantly, I could not have undertaken this research without the support from Miss Michelle Wilson who conducted advanced factor analysis on R packages, and alongside Dr Elizabeth Hensor provided me with continuous statistical analysis support. I would like to extend my gratitude to Dr Agata Burska for supporting me on my journey as an early career researcher, for guiding and coaching me through qPCR techniques and for her continuous support throughout this research. I am also thankful to Dr Lucy Carter for writing support.

I am indebted to the patients who have volunteered in the participating studies. Also, this research would not have been possible without clinical professionals in the Leeds Lupus group for consenting and caring for these patients, and for that I would like to thank Dr Edward Vital, Dr Katherine Dutton, Dr Sabih-UI Hassan, Dr Khaled Mahmoud, Dr Lucy Carter, Dr Jack Arnold, and Dr Md Yuzaiful Md Yusof who were supported by Huma Cassamoali and Sabina Khan.

A special thanks to our collaborators at the University of Manchester, Dr Darren Plant, for providing me with samples from BRIGGSS study. Additionally, to our collaborators in the EULAR taskforce for granting me early access to the manuscript: *Points to consider for the measurement and reporting of interferon activity in clinical research and future clinical application, for granting me early access of the publication*, to which a large portion of my work is based on.

I am grateful to have been surrounded by my colleagues, now lifelong friends, at Leeds Institute of Rheumatic and Musculoskeletal Medicine who have supported me through my research, and I would like to acknowledge Rekha Parmar and Jo Howorth for mentoring me through my personal development. Finally, a big thanks to my partner Tom and surrounding friends for their continuous support during this time.

Abstract

SLE is a complex multisystem autoimmune disease, causing difficulties when diagnosing and treating. IFN dysregulation is a hallmark of SLE pathogenesis which leads to the activation of self-antigens in autoimmunity. Subtypes of IFN signal through specific receptors, however, downstream activation produces pleiotropic downstream effects, with distinct overlap, making the IFN pathway a challenge to study. Quantification of IFN stimulated gene expression acts as an indirect measurement of IFN protein levels and is known as IFN signature. Previously, the Leeds Lupus group have established a clinically relevant two-scoring system to measure IFN activity in PBMCs and were named IFN-Score-A and IFN-Score-B, however, IFN signatures have not yet reached routine clinical practice. In 2019 a EULAR task force highlighted points to consider and terminology consensus when reporting IFN assays, including reasonable justification of assay choices. ISG measurement in a whole blood (WB) sample type is more feasible than PBMC so development of a WB assay was an aim of this report. Firstly, quantified IFN-Score-A and -B determined by TaqMan array were analysed in 45 matched PBMC and WB sample types which showed the most clinically useful PBMC derived IFN-Score-B cannot be interchanged between sample types. Secondly, factor analysis reinterrogation of 31 ISGs in WB derived two-IFN-Scores: IFN-Score-C (CCL8, CXCL10, IFI27, ISG15 and LAMP3) and IFN-Score-D (CASP1, CEACAM1, SOCS1 and TRIM38) and explained 97% of the variation. IFN-Score-A best defined SLE, RA and healthy patient groups compared to other IFN-Scores and IFN-Score-B and -D expressed strong correlations within WB suggesting the potential for future clinical utility of IFN-Score-D. Finally, RefFinder identified YWHAZ, PGK1 and GUSB to be the most stable reference genes from a group of 16 candidate genes in a WB SLE and healthy cohort. In conclusion, novel whole blood IFN scores warrant further evaluation in clinical validation studies.

Table of Contents

| | |
|--|-----------|
| Acknowledgements | 2 |
| Abstract | 3 |
| List of Abbreviations | 7 |
| Chapter 1. Literature Review | 10 |
| 1.1 Systemic Lupus Erythematosus | 10 |
| 1.1.1 Clinical classification of SLE | 10 |
| 1.1.2 Treatment of SLE | 12 |
| 1.2 SLE Pathogenesis | 13 |
| 1.2.1 Genetic susceptibility | 13 |
| 1.2.2 Autoreactive immune response | 13 |
| 1.3 IFN biology in SLE | 14 |
| 1.3.1 Types of IFN | 15 |
| 1.3.2 IFN-I signalling | 16 |
| 1.3.3 IFN-I producing cells | 17 |
| 1.3.4 IFN targeted therapy | 18 |
| 1.4 Biomarkers in SLE | 19 |
| 1.4.1 Overview | 19 |
| 1.4.2 IFN signature in SLE | 20 |
| 1.4.3 Development of IFN-Scores: Leeds Lupus group | 20 |
| 1.4.4 EULAR task force | 22 |
| 1.5 Hypothesis | 26 |
| 1.6 Aims | 26 |
| 1.7 Objectives | 26 |
| Chapter 2: General Materials and Methods | 27 |
| 2.1 Regulatory approval | 27 |
| 2.2 Blood collection and sample processing | 28 |
| 2.3 RNA isolation | 29 |
| 2.4 ISG transcript quantification | 29 |

| | | |
|--|--|-----------|
| 2.5 | Calculation of IFN-Scores | 33 |
| 2.6 | Statistical analysis..... | 33 |
| Chapter 3: Comparison of IFN-Scores Between PBMC and Whole Blood Sample | | |
| Types..... | | 34 |
| 3.1 | Background..... | 34 |
| 3.2 | Objective..... | 34 |
| 3.3 | Methods..... | 35 |
| 3.3.1 | Sample inclusion and sample processing | 35 |
| 3.3.2 | Gene transcript quantification | 35 |
| 3.3.3 | Statistical analysis | 35 |
| 3.4 | Results..... | 36 |
| 3.5 | Discussion | 39 |
| Chapter 4: IFN-Scores in Whole Blood: Factor Analysis..... | | |
| | | 42 |
| 4.1 | Background..... | 42 |
| 4.2 | Objective..... | 42 |
| 4.3 | Methods..... | 43 |
| 4.3.1 | Sample inclusion and sample processing | 43 |
| 4.3.2 | Gene transcript quantification | 43 |
| 4.3.3 | Factor analysis..... | 43 |
| 4.3.4 | Statistical analysis | 44 |
| 4.4 | Results..... | 46 |
| 4.4.1 | Whole blood factor analysis reveals distinct two-scores for whole blood ISG expression | 46 |
| 4.4.2 | Comparison of PBMC and novel whole blood IFN-Scores | 47 |
| 4.4.3 | Comparison of PBMC and whole blood sample types in newly developed whole blood IFN-Scores..... | 50 |
| 4.4.4 | Comparison of original IFN-Scores to newly developed whole blood IFN-Scores..... | 51 |
| 4.4.5 | Comparison of novel and existing whole blood IFN-Scores across disease groups..... | 55 |

| | | |
|--|--|-----------|
| 4.5 | Discussion | 57 |
| Chapter 5: Exploring Appropriate Reference Genes in SLE Patient Whole Blood | | |
| When Measuring ISGs | | |
| | | 61 |
| 5.1 | Background..... | 61 |
| 5.2 | Objectives..... | 62 |
| 5.3 | Methods..... | 63 |
| 5.3.1 | Sample inclusion and sample processing | 63 |
| 5.3.2 | Gene transcript quantification | 63 |
| 5.3.3 | Exclusion of 18s..... | 63 |
| 5.3.4 | Exploring reference genes | 63 |
| 5.3.5 | Calculation of IFN-Scores | 64 |
| 5.3.6 | Statistical analysis | 64 |
| 5.4 | Results..... | 66 |
| 5.4.1 | Algorithm based reference genes stability..... | 66 |
| 5.4.2 | Comparing expression across candidate reference genes..... | 67 |
| 5.4.3 | Evaluating the effects of normalising IFN-Scores to different reference genes..... | 70 |
| 5.5 | Discussion | 72 |
| Chapter 6: Discussion | | 75 |
| References..... | | 82 |

List of Abbreviations

| | |
|------------|--|
| ACR | American College of Rheumatology |
| AGS | Aicardi Goutières Syndrome |
| ANA | Anti-Nuclear Antibody |
| BAFF | B-Cell Activating Factor |
| BILAG | British Isles Lupus Assessment Group |
| BRAGGSS | Biologics In Rheumatoid Arthritis Genetics and Genomics Study Syndicate |
| CI | Confidence Interval |
| CLE | Cutaneous Lupus Erythematosus |
| CONVAS | Connective Tissue Disease and Vasculitis Cohort Cross-Sectional and Longitudinal Clinical And Basic Science Evaluation Study |
| Ct | Cycle Threshold |
| CTD | Connective Tissue Disease |
| DEFINITION | Defining Interferon Mediated Autoimmune Conditions |
| DNA | Deoxyribonucleic Acid |
| EGFR | Epidermal Growth Factor Receptor |
| ELISA | Enzyme Linked Immunosorbent Assay |
| EULAR | European League Against Rheumatism |
| FRET | Fluorescence Resonance Energy Transfer |
| GAPDH | Glyceraldehyde 3 Phosphatase Dehydrogenase |
| GAS | γ -Activated Sites |
| HC | Healthy Control |
| HLA | Human Leucocyte Antigen |
| IC | Immune Complex |
| IFC | Integrated Fluidic Circuit |
| IFN | Interferon |
| IFNAR | Interferon Alpha Receptor |

| | |
|-------------|---|
| IFN-I | Type 1 Interferon |
| IFN-II | Type 2 Interferon |
| IFN-III | Type 3 Interferon |
| IFNLR1 | Interferon Lambda Receptor 1 |
| IgM | Immunoglobulin M |
| IL-6 | Interleukin 6 |
| IL-10R2 | Interleukin 10 Receptor 2 |
| IRF7 | Interferon Regulatory Factor 7 |
| ISG | Interferon Stimulated Genes |
| ISGF3 | Interferon Stimulated Gene Factor 3 |
| ISRE | Interferon Response Element |
| JAK | Janus Tyrosine Kinase |
| KMO | Kaiser-Meyer-Olkin |
| LOA | Limits of Agreement |
| M1.2 | Module 1.2 |
| M3.4 | Module 3.4 |
| M5.12 | Module 5.12 |
| MASTERPLANS | Prediction of Lupus Treatment Response Study |
| MHC | Major Histocompatibility Complex |
| MIQE | Minimum Information for Publication of Quantitative Real-Time PCR Experiments |
| NET | Neutrophil Extracellular Trap |
| NK | Natural Killer Cell |
| NKT | Natural Killer T-Cell |
| PBMC | Peripheral Blood Mononuclear Cell |
| PMN | Polymorphonuclear |
| RA | Rheumatoid Arthritis |
| RG | Reference Gene |

| | |
|--------|--|
| RNA | Ribonuclear Acid |
| RT | Room Temperature |
| RT-PCR | Real Time Polymerase Chain Reaction |
| SD | Standard Deviation |
| SiMoA | Single Molecule Assay |
| SLE | Systemic Lupus Erythematosus |
| SLICC | Systemic Lupus International Collaborating Clinics |
| SLR | Systematic Literature Review |
| STAT1 | Signal Transducer and Activator of Transcription |
| TE | Tris-Hydrochloric Acid Ethylenediaminetetraacetic Acid |
| TLR | Toll Like Receptor |
| TNF | Tumour Necrosis Factor |
| TULIP | Treatment Of Uncontrolled Lupus Via the Interferon Pathway |
| TYK2 | Tyrosine Kinase 2 |
| WB | Whole Blood |

Chapter 1. Literature Review

1.1 Systemic Lupus Erythematosus

Systemic Lupus Erythematosus (SLE) is a complex multisystem autoimmune disease that presents with multiple clinical manifestations, which may cause challenges to diagnose and treat. Breakdown of immune self-tolerance results in immune-mediated tissue injury in a multitude of organ systems including the skin, musculoskeletal, renal, and hematologic systems. SLE is preceded by a phase of asymptomatic autoantibody positivity (1) of autoreactive antinuclear antibodies (ANA) and can be detected in serum up to 10 years before clinical feature presentation. However, ANA are present in up to 25% of the general population, which less than 1% develop autoimmune disease, of which are classified as “At-Risk” (2).

A combination of genetic, environmental, and hormonal factors are known to be susceptibility features of the disease (1). Age, sex and ethnicity are impacting factors of SLE prevalence, with increase affect in female to males at an average ratio of 9:1 in the UK and highest incidence is amongst women at the reproductive ages (1). Ethnic ancestry also influences disease and organ involvement (3).

1.1.1 Clinical classification of SLE

Classification systems for SLE have been established to assist in differentiating the disease from other symptomatically and clinically similar connective tissue diseases (CTDs). The systemic lupus international collaborating clinics (SLICC) classification criteria for SLE, originally developed in 2012, were revised in 2019 by an expert task force from the European League Against Rheumatism (EULAR) and American College of Rheumatology (ACR) and was driven by data in the literature and expert experience (4, 5).

In the revised system the first classification stage requires positive ANA immunofluorescence at a threshold of $1 \geq 80$ titre on Hep-2 cells. The next stage comprises of scores based on additive weighted criteria, including clinical and immunological domains. The criteria are cumulative across the course of disease and do not have to be simultaneously present. A positive ANA plus 10 or more points (Table 1), satisfies classification for SLE in a research context. Finally, SLICC criteria were developed for research classification rather than diagnostic purposes, therefore deviations from criteria based on clinical judgements may be applied (6).

Table 1. European League Against Rheumatism (EULAR) and American College of Rheumatology (ACR) SLE criteria classification (6)

| | |
|--|---------------|
| A positive ANA test of $\geq 1:80$ on Hep-2 cells | |
| ↓ | |
| If positive, continue to additive criteria | |
| ↓ | |
| Additive criteria Must have at least one criterion at least 10 points The highest scoring criteria is counted towards final score for each domain | |
| Clinical domains and criteria | Weight |
| Constitutional | |
| <i>Fever</i> | 2 |
| Hematologic | |
| <i>Leukopenia</i> | 3 |
| <i>Thrombocytopenia</i> | 4 |
| <i>Autoimmune haemolysis</i> | 4 |
| Neuropsychiatric | |
| <i>Delirium</i> | 2 |
| <i>Psychosis</i> | 3 |
| <i>Seizure</i> | 5 |
| Mucocutaneous | |
| <i>Non-scarring alopecia</i> | 2 |
| <i>Oral ulcers</i> | 2 |
| <i>Subacute cutaneous OR discoid lupus</i> | 4 |
| <i>Acute cutaneous lupus</i> | 6 |
| Serosa | |
| <i>Pleural or pericardial effusion</i> | 5 |
| <i>Acute pericarditis</i> | 6 |
| Musculoskeletal | |
| <i>Joint involvement</i> | 6 |
| Renal | |
| <i>Proteinuria $>0.5g/24h$</i> | 4 |
| <i>Renal biopsy Class II or V lupus nephritis</i> | 8 |
| <i>Renal biopsy Class III or IV lupus nephritis</i> | 10 |
| Immunology domains criteria | |
| Antiphospholipid antibodies | |
| <i>Anti-cardiolipin antibodies OR</i> <i>Anti-$\beta 2GP1$ antibodies OR</i> <i>Lupus anticoagulant</i> | 2 |
| Complement proteins | |
| <i>Low C3 OR low C4</i> | 3 |
| <i>Low C3 AND low C4</i> | 4 |
| SLE-specific antibodies | |
| <i>Anti-dsDNA antibody OR</i> <i>Anti-smith antibody</i> | 6 |

1.1.2 Treatment of SLE

SLE is a chronic relapsing and remitting disease, and management aims to control symptoms, improving quality of life, reduce the severity of organ involvement caused by inflammation and to prevent disease flares. Due to the heterogenetic nature of SLE, response to therapy varies widely and treatment approaches must be tailored for individuals. Biomarkers to guide treatment course are however, still limited (1).

The major classes of SLE therapies are glucocorticoids, antimalarial drugs including hydroxychloroquine, immunosuppressants such as mycophenolate mofetil and methotrexate, and more recent targeted biologic therapies. In 2011, Belimumab was the first drug to be licenced for the treatment of SLE (7) however, other therapeutic agents such as immunosuppressants and glucocorticoids are used to manage the symptoms.

Treatment intensity depends on the severity of the disease. For example, current guidelines suggest the use of antimalarial drugs in mild skin or joint disease and a more aggressive treatment including combination of immunosuppressants and glucocorticoids may be necessary for moderate-severe multi system disease and lupus nephritis. Extended exposure to immunosuppressants are associated with risks to infections, diabetes and osteoporosis (4, 8).

The established biologic agents used in the treatment of SLE, Rituximab and Belimumab, both target B cells and were developed based on overwhelming evidence for the role of B cell in the pathogenesis of SLE. They are both monoclonal antibodies that deplete B cells by targeting cell surface CD20 (9) and soluble B cell activating factor (BAFF) protein respectively (7). However, a significant subset of patients fail to respond to B cell targeted therapy and further research into the pathogenesis of SLE has identified new therapeutic targets including the development of biologic therapy targeting interferon (IFN) pathway, discussed later in the report (8).

1.2 SLE Pathogenesis

1.2.1 Genetic susceptibility

There is increasing evidence that predisposing hereditary genetic abnormalities contribute to the risk of SLE. For example, disease concordance is 25% in monozygotic twins compared to 2% in dizygotic twins (1). Moreover, a large nationwide study conducted in Denmark from 1977-2013 revealed SLE risk was 10.3-fold higher among first degree relatives of affected individuals and 3.6-fold increased with second or third degree relatives compared to the general population (10).

One of the first identified genetic risk loci was within human leucocyte antigen (HLA) class proteins also known as major histocompatibility complex (MHC) (11, 12). The other major group of SLE risk variants are associated with type 1 IFN (IFN-I) signalling, production and response and this is discussed further below (13).

In combination with genetic susceptibility, environmental triggers such as UV radiation, viral infection, smoking, and certain drugs are recognised to initiate autoimmunity (14).

1.2.2 Autoreactive immune response

SLE is regarded as the breakdown of self-tolerance with the development of tissue injury and inflammation as a result of auto-reactive immune cells and autoantibodies.

Autoantibodies in SLE overwhelmingly target nuclear self-antigens, believed to reflect the insufficient clearance of cellular death material, a hallmark of SLE pathogenesis. There are over 180 autoantibodies associated with SLE that are mostly complementary to single-stranded and double-stranded deoxyribonucleic acid (DNA), Ro/La antigens, and ribonuclear proteins (15).

Apoptosis, secondary necrosis, release of extracellular traps upon neutrophil death: NETosis and autophagy are among the many cellular death pathways that release cellular nuclear material, which then form immune complexes (IC) upon binding with autoantibodies. The deposition of ICs in SLE propagates inflammation and tissue injury through complement fixation, and activation of diverse immune cells, including dendritic cells. UV exposure (16) and recurring infections are the main triggers for SLE cell death pathogenesis and can occur in different ways depending on the cell type signal.

B cells are central to the adaptive immune system and are the main antibody producing cells in the humoral response to bacterial or viral pathogens. B cell development begins in the bone marrow where pluripotent hematopoietic stem cells mature and differentiate into immature B cells and present a unique membrane bound B cell receptor (immunoglobulin M

(IgM)) corresponding to specific antigens. Furthermore, B cell maturation takes place in the secondary lymphoid organs. Following antigen encounter, B cells are activated through T-cell dependant or independent mechanisms and a subset of activated cells will ultimately differentiate to become antibody producing plasmablasts or long lived plasma cells or memory B cells (17).

The breach of self-tolerance is initiated when autoreactive B cells are generated to recognise self-antigen and nuclear components. In normal B cell development, autoreactive B cells bind to self-antigens with high affinity and therefore undergo clonal deletion or are rendered anergic through a process of negative selection. This process is ignored in SLE (18) and ANAs are formed that are complementary to uncleared nuclear material and form ICs in the periphery. Of course, self-antigens can never be removed, and ICs are deposited into organs such as the kidneys or skin where they promote unwanted immune responses (19).

1.3 IFN biology in SLE

Interferons (IFNs) are so called, as they have the ability to interfere with viral replication (20) and produce an inflammatory response via the innate or adaptive immune system. IFNs also act as a signal to neighbouring cells to inform them of an infection, preparing and protecting them from further damage. In normal responses, IFNs production reduces when viral infections are no longer a threat, however, in SLE, a constant production of ICs uncontrollable IFN production.

Since the 1970's, IFN-I pathway activation has been implicated in SLE disease activity (21) and prompted efforts towards developing IFN-I blockade therapy (22, 23). Several bodies of evidence to support the role of IFN-I in SLE pathogenesis:

Firstly, over half of currently identified genetic susceptibility loci in SLE are related to IFN production or pathway (24). For example, variants of IFN-I pathway activation genes have been reported in but not limited to *IRF5*, *IRF7*, *IRF8*, *STAT4*, *OPN*, *IFIH1* and *TYK2* (13). Additionally, there is a hereditary association with abnormal IFN-I pathway activation genes where first degree relatives present with 20% more serum IFN- α than healthy unrelated individuals (25). Interestingly, there are no genetic abnormalities with IFN- α itself however, there are with the IFN- κ (*IFNK*) transcript. A genome-wide study revealed associations between increased serum IFN-I activity and genetic abnormalities in *IFNK* identifying other types of IFN-I have importance in SLE pathogenesis (26).

Secondly, in the early 2000's the term 'interferonopathy' was coined, which described diseases related to the increased regulation of IFN-I such as Aicardi-Goutières syndrome

(AGS) (27). AGS has been associated with a large spectrum of mutations across the *TREX1* gene, a 3'-5' exonuclease, leading to failure of clearing apoptotic nuclear material thus promoting an IFN-I immune response. While these interferonopathies are not identical to SLE, they do have common features such as chilblain lupus lesions. In combination to these findings, rare monogenic forms of SLE have been associated with single polymorphisms to the *TREX1* gene, and presents overlapping clinical features of AGS further validating the involvement of IFN-I dysregulation in SLE (28, 29).

Thirdly, high levels of circulating IFN-I in the blood of SLE patients has led to the discovery of the IFN signature: the quantification of a group of IFN stimulated genes (ISGs). A high IFN-I signature has been reported in the blood and tissues of SLE patients and has been proposed as a biomarker to distinguish SLE from other autoimmune diseases, prognosis and response to treatment (30).

Finally, observational studies of IFN- α treated malignancies induced an SLE like disease but discontinued when IFN- α therapy stopped. It is therefore no surprise that research has been directed into IFN blockade for the treatment of SLE and reagents such as Anifrolumab: a monoclonal antibody that blocks the IFN-I receptor that has now been licensed for treatment of SLE (31).

Collectively these provide convincing evidence for the involvement of IFN in SLE pathogenesis reveals further targets for therapeutic methods.

1.3.1 Types of IFN

The three families of IFN: type 1 (IFN-I), type 2 IFN (IFN-II) and type 3 IFN (IFN-III) are subdivided due to their unique signalling receptor, protein structure and what cells they are secreted by (Table 2).

IFN-I are the largest family of IFN divided into 5 subtypes (α , $-\beta$, $-\omega$, $-\epsilon$ and $-\kappa$) and are secreted by all nucleated cells (32, 33). All nucleated cells express the IFN-I receptor IFNAR1/IFNAR2 and transcribe ISGs with the IFN-stimulated response element (ISRE) as their promoter (34) following activation of JAK1/TYK2 and/or STAT1/STAT2 intracellular signalling cascades. IFN- γ is the only subclass of IFN-II (34, 35), and is produced by NK, NKT and T cells. Like IFN-I receptor, the IFN-II receptor (IFNGR1/IFNGR2) is ubiquitously expresses on all nucleated cells and activates intracellular cascades involving STAT1 homodimers (36). Finally, the 4 subtypes of IFN-III lambda are recognised by the IFNLR1/IL-10R2 receptor complex and signal via interferon-stimulated gene factor 3 (ISGF3) with more restricted expression, predominately on epithelial cells (37, 38).

To summarise, all IFN subtypes induce ISG transcription, leading to the activation of innate and adaptive immune responses to virus and tumours, or in an autoimmune case, self-antigen. IFN receptor activation produces pleiotropic downstream effects, with distinct overlapping signalling, making the IFN pathway a challenge to study (39).

Table 2. IFN subtypes: cell secretion and receptor targets

| IFN type | Subtype | Secreted by | Receptor | Receptor expression | Reference |
|----------|---|---------------------|------------------|------------------------|-------------------|
| IFN-I | <ul style="list-style-type: none"> • IFN-α (further divided into 13 subtypes) • IFN-β • IFN-ω • IFN-ϵ • IFN-κ | All nucleated cells | IFNAR1 / IFNAR2 | All nucleated cells | (34) (40) |
| IFN-II | IFN- γ | NK, NKT, T cells | IFNGR1 / IFNGR2 | All nucleated cells | (35) (34) (36) |
| IFN-III | <ul style="list-style-type: none"> • IFN-λ1 / IL-29 • IFN-λ2 / IL-28A • IFN-λ3 / IL-28B • IFN-λ4 | All nucleated cells | IFNLR1 / IL-10R2 | Epithelial cells, pDCs | (37) (38) |

IFN: Interferon, IFNAR: Interferon alpha receptor, IFNGR: Interferon gamma receptor, IFNL: Interferon lambda receptor, IL: interleukin, NK: natural killer, NKT: Natural killer T cell, pDC: plasma dendritic cell.

1.3.2 IFN-I signalling

IFNAR1 and IFNAR2 form the heterodimeric transmembrane structure of IFN-I receptor and is ubiquitously expressed on nucleated cells (41). IFN-I ligand activation of IFNAR subunit associated proteins; tyrosine kinase-2 (TYK2) and Janus tyrosine kinase (JAK1) respectively become phosphorylated and, in turn, activate STAT2 and STAT1. Two signalling cascades can be activated by i) homodimer formation of STAT1-STAT1 transcription factor binds to the IFN- γ activated sites (GAS) (42) or ii) activated STAT1-STAT2 binds to IRF9 to form complex interferon-stimulated gene factor 3 (ISGF3) to initiate ISG transcription. Like STAT1 homodimer, ISGF3 initiates ISG expression but binds to the IFN-stimulated response element (ISRE) promoter region instead of GAS (43) and downstream transcription of ISGs rely on which promoter region is activated and or type of IFN stimulation. IFN-II also initiates STAT1 homodimer signalling pathways through IFNGR1/IFNGR2 stimulation and IFN-III initiates ISGF3 activation through IFNLR1/IL-10R2 stimulation creating an overlap of ISG expression (44).

The production of IFN-I acts as an effective tool to fight viral infections by induction of an antiviral state of infected and neighbouring cells by upregulation of certain enzymatic ISGs,

thus preventing further viral infection. However, dysregulated IFN-I promotes unwanted immune responses to autoantibodies in SLE.

For example, high serum IFN-I initiates IFN priming effects by increasing expression of STAT1 and IRF9 in undifferentiated macrophages, which can increase ISG production without IFN stimulation (45). Also, there is an increase in B cell survival including induction of plasma cell differentiation and class switching of Ig antibodies to favour production of high affinity autoantibodies. These autoantibodies create anti-DNA or anti-RNP ICs further activating IFN-I production through TLR pathways. Since healthy pDCs produce IFN in response to ANA containing ICs, a model in SLE was suggested wherein IFN-I activated antigen presenting cells, pDCs, rapidly translate further IFN-I, creating a vicious cycle of IFN-I production in autoimmunity (46-48). Recent data suggests this may not be the case, discussed below.

1.3.3 IFN-I producing cells

In non-dysfunctional state, IFNs are produced in response to pathogens upon detection of viral nucleic acids by pattern recognition receptors on innate immune cells, particularly pDCs: bone marrow derived, antigen presenting cells, that have the professional ability to produce IFN-I via TLR-7 and -9 signalling that recognise ribonucleic acid (RNA) or DNA components of ICs (49, 50).

In SLE however, recent evidence suggests that IFN production does not originate from pDCs as previously believed. For example, experimental autoimmune murine models have shown a decrease in circulating pDCs and their ability to produce IFN-I with no migration of pDCs into target organs such as kidneys (51). Other SLE human studies have detected no IFN- α in any circulating cell subset (52) and that these cells have lost their ability to produce IFN- α upon TLR-7 and -9 activation (53).

Moreover, this hypothesis suggests that in the absence of inflammation, pDCs levels and function would expect to be normal. However, studies have shown that the dysfunctional pDCs that are present in active SLE are the same as those found in ANA positive At-Risk patients that never develop organ inflammation or require treatment (53) and therefore suggestion dysfunctional pDCs are independent of inflammation or disease.

SLE IFN-I production is chronically increased in other tissues, especially the skin compared to blood, which may explain the persistent high IFN-I activity. New theories suggest that sustained IFN-I production, in particular IFN- κ , seems to be due to a continual response of keratinocytes to UV light (54). *IFNK* expression is significantly increased in lesional skin of

patients with cutaneous lupus erythematosus (CLE) and it increases keratinocyte sensitivity to UV irradiation (55, 56). *In vitro* studies of SLE and At-Risk non-lesional skin biopsies without signs of cutaneous inflammation, showed a higher *IFNK* expression with TLR3 and RIG-I stimulation (53). Therefore, there is enough data to indicate that skin itself, and not infiltrating cells, drives a high IFN-I response is delivered by keratinocytes even at At-Risk patients (ANA positive).

1.3.4 IFN targeted therapy

Given the pleiotropic effects of IFN-I signalling, blocking IFN-I protein or IFN-I receptors has versatile therapeutic potential to target autoimmune diseases related to IFN-I over production. Clinical trial design has proven difficult, due to the heterogeneity of SLE however, there is promise of new therapy, Anifrolumab, an anti-IFNAR1 human monoclonal antibody that inhibits all IFN-I signalling (57). Anifrolumab is now licenced in Europe and United States of America to treat adults with moderate to severe SLE and is the first drug for the treatment of SLE since the approval of Belimumab in 2011 (7). Phase III (Treatment of Uncontrolled Lupus via the Interferon Pathway: TULIP-2) clinical trials of Anifrolumab met the primary end points by reducing disease activity by 52 weeks compared to placebo (58).

With the licensing of Anifrolumab, in addition to existing B cell targeted therapy, clinicians may now be faced with a choice between different classes of therapy for SLE. It also remains to be understood at what stage of the disease and in what sequence these therapies will yield maximum clinical benefits for example some patients with strongly IFN-I driven disease may benefit from Anifrolumab at an earlier stage before trying conventional immunosuppressants. A recent study demonstrated that Anifrolumab is more effective in patients with a high interferon gene signature (IFNGS) (59). In order to inform treatment decisions, clinicians need robust biomarkers for IFN-I driven disease.

1.4 Biomarkers in SLE

1.4.1 Overview

A biomarker is defined as “a biological molecule found in blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or of a condition or disease” (60) and are utilised for clinical diagnosis and therapeutic research (61). There are many established biomarkers in cancer biology, used in screening or cancer stage determination (60). For example, mutation screening of epidermal growth factor receptor (*EGFR*) in non-small cell lung cancer, determines response to treatment of *EGFR* inhibitors such as Erlotinib which is superior to conventional chemotherapy (62).

Effective biomarker development could help to resolve several important clinical challenges in SLE, for example distinguishing SLE from symptomatically similar disease such as RA and Sjogren’s syndrome, assisting early identification of poor prognosis disease from milder forms, predicting major organ involvement and disease flares, and importantly, guiding early personalised treatment strategies (63). Excluding serum ANA detection and complement deficiency, clinical validation of SLE has not yet reached biomarker level. Furthermore, as ANAs are present in many of the general population, (2, 64) and negative ANA results of some SLE patients (6), improved biomarkers are necessary.

A comprehensive literature review by Arriens et al recognised numerous pro-inflammatory cytokines; TNF- α , IL-1, IL-6, IL-12, IL-17, IL-21 and IL-23 that were found to discriminate between SLE and healthy individuals, and although these contribute to understanding disease pathogenesis, they have not progressed to clinical application or the focus of new drug developments (65).

Clinically traditional biomarkers i.e., anti-dsDNA and low-level complement are indicative of SLE disease activity, but sensitivity and specificity for predicting treatment outcomes and organ involvement have been inconsistent. More recently, increased serum BAFF shows some potential as novel B cell biomarkers as it positively correlates with serum anti-dsDNA, and rising BAFF levels, following rituximab, were closely associated with subsequent flare (66). Moreover, discovery of BAFF has allowed drug development of Belimumab, a monoclonal antibody targeting and inhibiting BAFF mechanism of action (7).

Beyond B-cell development focus, there is increasing need to establish biomarkers for evaluating IFN-mediated pathways in SLE. There is now promising data in relation to measuring ISG transcriptomes reviewed by Fujio et al (67) and many other biomarker that utilise transcriptomics with IFN gaining increased prominence. However, given the pleiotropic effect of IFN signalling there are considerable challenges to implementing IFN

biomarkers in clinical practice, and an ongoing EULAR task force is initiating points to consider when reporting IFN activity in clinical research (68) discussed later in this report.

1.4.2 IFN signature in SLE

Since determining serum IFN on a protein level was an insensitive detection method, research has moved towards an indirect way of measuring IFN pathway activation: the IFN signature. The IFN signature is described as a measurement of a group ISGs and was first reported in 2003 by Baechler et al (69). It is either described in a categorical manner as IFN 'high' or 'low' or as continuous variable known as an IFN-Score (69, 70). Measuring the IFN signature currently has little consensus, and the literature reports various ways of IFN assay methodologies (69, 71).

Unbiased clustering from blood genome wide studies have organised clinically relevant transcripts into modules that relate to a particular cell subset or immune function (72). Expression levels of certain modules were found to overlap between different diseases, and global analysis of all modules were disease specific giving insight into disease pathogenesis. In SLE, the IFN signature clustered into 3 different modules: M1.2, M3.2 and M5.12 which appear to have varying stability in SLE patients. Module 1.2 is stable over time and induced by IFN-I whereas modules 3.4 and 5.12 respond to both IFN-I and -II and are more responsive to disease activity levels, reiterating the diversity and complexity of the IFN pathway (73). This innovative analysis approach builds on the body of evidence that IFN signature is positively correlated to SLE disease activity and progression.

1.4.3 Development of IFN-Scores: Leeds Lupus group

As microarray analysis used in the initial discovery of the SLE modular transcriptome is expensive and time consuming, there is a need for refining and simplifying the quantitative measurement of IFN signature for development towards clinical application. Based upon the comprehensive description of IFN Modules (M1.2, M3.4 and 5.12) in the SLE blood transcriptome by Chiche et al (73), the Leeds Lupus group utilised these findings to build a continuous IFN-Scoring system (70).

A selection of 29 highly expressed genes in SLE from the three IFN modules, with the addition of 2 commonly reported ISGs (*IFI27*, *IFI6*) were quantified using TaqMan on the Fluidigm® platform: a high-throughput fluidics gene expression technology. Statistical factor analysis can help reveal unobserved variables that are impacting upon the behaviour of IFN annotated modules, and this method was used in this study to determine ISGs score

selection. Two factors were shown to account for 84% of variability in IFN status. 26 out of 31 measured ISGs were clustered into two factors and described as: IFN-Score-A (*IFI27, IFI44, ISG15, GBP1, IFI44L, IRF7, IFIT1, RSAD2, CXCL10, CCL8, CEACAM1, XAF1*); a group of ISGs, previously reported in existing literature as typical of the IFN-I signature, and IFN-Score-B (*IFIH1, LAMP3, NT5C3B, UBE2L6, TAP1, STAT1, SERPING1, SOCS1, SP100, TRIM38, IFI16, UNC93B1, BST2, PHF11*); a groups of ISGs not so commonly used in the literature which may be more dynamically regulated and responsive to other IFN-subtypes. The remaining genes (*CASP1, IFI6, HERC5, EIF2AK2 & MX1*) cross loaded into both factors and were excluded from IFN-Scores (70).

Exploration of continuous IFN-Scores proved to be more informative than bimodal (high or low) scoring system for stratifying disease. For example, a variation of diagnosis was found within the IFN-high and -low groups, with many healthy individuals to be categorised as IFN-high, when a continuous measurement was able to decipher between diagnostic groups (Table 3).

IFN-Scores were firstly developed in unsorted peripheral blood mononuclear cells (PBMCs) and since then, they have been proven useful in whole blood (WB) Tempus™ samples. The novel IFN-Score-B has found to be a better predictor of clinical outcomes than classical ISGs in IFN-Score-A and details of the findings in PBMCs and WB can be found in the table below (Table 3).

Table 3. Reported findings of IFN-Scores

| IFN-Score-A significant findings | IFN-Score B significant findings |
|--|--|
| <ul style="list-style-type: none"> - Differentiates SLE from Rheumatoid Arthritis (RA) and HC PBMCs (70) - Associated with cutaneous and haematological SLE PBMCs (70) - Higher score value in At-Risk non-lesioned skin than blood PBMCs (74) | <ul style="list-style-type: none"> - Differentiates SLE and RA from HC PBMCs (70) - Associated with cutaneous and haematological SLE PBMCs (70) - Predicts progression to SLE in ANA positive, At-Risk individuals PBMCs (74) - Predicts progression to RA in CCP positive individuals PBMCs (75) - Associated with response to Rituximab Whole blood (76) - Associated with imaging-proven synovitis Whole blood (77) |

1.4.4 EULAR task force

There is a large body of evidence to suggest that using an IFN signature, albeit type-I and or -II, as a biomarker might ultimately assist clinical care, but progress towards adoption in clinical settings faces several challenges. Particularly around terminology and unifying approaches to ISG selection, measurement and establishing clinical associations. The literature reports numerous IFN assay detection methods, however, consensus is needed for future research for IFN assays to reach clinical practice. In 2019, a EULAR task force was established of a multi-disciplinary team, consisting of 17 members, in order to review published data around IFN-I assays and to enable the facilitation of these into clinical practice and are listed below.

A systematic literature review studied published data (up to Oct 2019) that reported IFN-I assays in basic and clinical research of RMDs and described the population studied, type of assay, material analysed, pathway element, detailed description of method and calculation of reported result, validity reliability and feasibility of each method (68).

It is well known that IFN pathway activation is easily measured via qPCR, so there is no surprise that this was the most reported method, although, reporting of IFN-Scores varied. For instance, 82/122 studies reported a continuous IFN-Score from a set of ISGs, 30/122 reported individual gene transcript expression, 4/122 measured IFN protein transcript, 4/122 reported transcript expression of IFN induced chemokines and 2/122 studies reported groups of ISGs in a categorical signature. Within these studies, different combinations of ISGs were measured but the rationale of ISG choice was not always given. IFN activity was verified with secondary methods including comparison of previously published scores, validation against IFN- α protein levels using ELISA or SiMoA, and against the expression of IFN-stimulated proteins or individual ISGs. Immunoassays, microarray, reported cell assay, DNA methylation, flow cytometry, cytopathic effect assay, RNA sequencing, Plaque reduction assay, Nanostring and bisulfite sequencing were amongst the other methods reported in the literature.

The collection of evidence led the panel to develop the following points to consider when reporting IFN-I assays (78) [*Press release*] :

1. Task force consensus terminology should be considered for reporting IFN assays measurement (Table 4).
2. **Existing assays measure different aspects of the IFN pathway; they do not reflect the entirety of the pathway, and some are not specific for IFN-I. The most appropriate assay will depend on the research or clinical question and should be justified.**

3. Publications on novel IFN-I pathway assays should report whether they specifically reflect IFN-I, and to the extent possible, which IFN-I is measured.
4. For assays that evaluate pathways downstream of the IFN-I receptor (e.g. IFN stimulated gene expression or protein scores) the choice of components needs to be justified. For gene expression scores, the known subsets of IFN-stimulated genes should be described separately.
5. IFN-I pathway is consistently activated in several RMDs, but assays measuring IFN-I pathway activation cannot be currently recommended for diagnostic purposes.
6. IFN-I pathway assays define more severe subgroups within many RMDs so should be considered for stratification studies.
7. IFN-I pathway activation is associated with disease activity in some RMDs, especially SLE and myositis, but its added value in clinical decision-making is uncertain.
8. IFN-I pathway assays can predict disease exacerbations, in particular flare occurrence in SLE patients, but further work should be performed to determine to what extent they outperform current instruments.
9. IFN-I pathway assays might predict progression from pre-clinical autoimmunity to clinical disease.
10. In SLE, IFN-I pathway assays may be useful in predicting response to IFN-I targeting therapies.
11. IFN-I pathway assay results may be affected by some treatments (e.g. IFN-targeted therapies and high-dose glucocorticoids), and timing of sample collection should take this into account and be reported.

The continuous two score system (IFN-Score-A and -B) for quantification of IFN pathway activation is emerging as a leading candidate biomarker framework for SLE (70, 74-77). The work presented, aims to advance the development of these biomarkers, in line with EULAR task force recommendations by addressing point 2. The key point to consider is that the IFN-I pathway is a complex overlapping signalling system, and a IFN assays only capture sections of the pathway. ISGs may also be influenced by other non-IFN immune mediators, therefore, not all assays will be specific to IFN-I. There is no gold standard for IFN assays, however, the EULAR task force recommends reporting the appropriateness of IFN assays based on justification of theoretical, experimental, feasibility and clinical evidence requirements.

Table 4. Task force consensus terminology when reporting IFN-I assay.

| Term | Abbreviation | Definition |
|---|---------------------|--|
| Interferon | IFN | Proteins with anti-viral activity; IFNs are mediators of an anti-viral response. They belong to the Type I, Type II and Type III IFN families. |
| Type I interferon | IFN-I | The IFNs alpha, beta, omega, kappa, epsilon, secreted by any nucleated cell, and binding to the IFNAR, which is expressed on any nucleated cell. |
| Type II interferon | IFN-II | IFN gamma, mostly secreted by T cells, binding to the IFNGR, which is expressed on most leucocytes. |
| Type III interferon | IFN-III | IFN lambda, which are structurally more similar to IL-10 but share downstream signalling and gene expression with IFN-I. |
| Interferon-stimulated genes | ISGs | Genes whose expression is known to be upregulated by any kind of IFN. Individual ISGs may not exclusively represent Type I IFN pathway activation. |
| Type I Interferon pathway activation | | Any evidence for function of the components of the Type I IFN pathway. This includes secretion of a Type I IFN protein, binding to the IFNAR, initiation of JAK/STAT signalling pathways, expression of IFN-stimulated genes, expression of IFN-stimulated proteins. |
| Type I interferon pathway assay | | An assay measuring one or more components of the Type I IFN pathway at a molecular or functional level. |
| Interferon stimulated gene expression signature | | A qualitative description of coordinated expression of a set of ISGs that is indicative of Type I IFN pathway activation. |
| Interferon stimulated gene expression score | | A quantitative variable derived from expression of a defined set of ISGs that is indicative of Type I IFN pathway activation. |

| Term | Abbreviation | Definition |
|-------------------------------------|---------------------|---|
| Interferon stimulated protein score | | A variable derived from expression of a defined set of soluble biomarkers known to be upregulated by IFN, although not specific for Type I IFN. |
| Interferonopathy | | Monogenic diseases in which there is constitutive Type I IFN pathway activation with a causal role in pathology. The clinical picture may resemble rheumatic musculoskeletal diseases. However, most diseases with IFN pathway activation are not Interferonopathies. |

1.5 Hypothesis

IFN gene expression assays can be refined to improve their ability to stratify autoimmune diseases.

1.6 Aims

The overall aim of this project is to define the optimal IFN gene expression assay for adoption into routine clinical laboratories.

This project aims to address some of the EULAR points to consider with our IFN-Scores. Although the IFN-Scores have previously been validated by *in vitro* stimulation, there are several unanswered questions concerning their content and construct, validity and feasibility. This report mainly focuses on moving towards simpler sample collection from PBMCs to WB Tempus™.

1.7 Objectives

1. Assess the effect of sample type on existing IFN-Scores-A and -B by comparing paired samples from the DEFINITION study.
2. Define the optimal selection of target ISGs for WB samples using factor analysis.
3. To determine the most stable WB reference gene panel, robust to influence of IFN-I.

Chapter 2: General Materials and Methods

2.1 Regulatory approval

Samples were collected under four ethically approved studies outlined below.

DEFINITION: Defining Interferon Mediated autoimmune conditions, and healthy controls were under the ethical approval of CONVAS: Connective Tissue Disease and Vasculitis Cohort Cross-sectional and Longitudinal Clinical and Basic Science Evaluation study.

CONVAS is an observational study based on routine clinical care that allows for collection of clinical data imaging and biomarkers on patients with a broad range of connective tissue diseases. Ethical approval was granted by Yorkshire and The Humber, Leeds East Research Ethics Committee (10/H1306/88). The University of Leeds were contracted with the administrative sponsorship.

USEFUL: Ultrasound Evaluation for Musculoskeletal Lupus, was a prospective observational study on patients with lupus arthritis receiving glucocorticoid therapy. USEFUL was sponsored by The University of Leeds and, ethical and health research authority approval was granted by the Northwest - Greater Manchester Central Research Ethics Committee (16-NW-0060). The University of Leeds was contracted with the administrative sponsorship.

MASTERPLANS: Prediction of Lupus Treatment Response Study, consortium aiming to stratify therapy in SLE for which University of Leeds led the workstream for analysis of IFN-Scores. Samples evaluated were obtained from BILAG-BR: British Isles Lupus Assessment Group Biologics Register, a prospective UK-wide registry evaluating the safety and efficacy of biologics in SLE, sponsored by the University of Manchester and ethical approval was granted by Northwest–Greater Manchester West NRES Committee (REC: 09/H1014/64) and UK Health Research Authority (IRAS ref. 24407).

BRAGGSS: Biologics in Rheumatoid Arthritis Genetics and Genomics Study Syndicate study-ethics was approved by the Northwest 6 Central Manchester South Research Ethics Committee (COREC 04/Q1403/37) and all patients provided written consent.

All patients provided written informed consent and these studies were conducted in accordance with the principles of the Declaration of Helsinki.

Table 5 includes number of participants used from each study, sample type used, recruitment criteria and where these samples were used in this report.

Table 5. Outline of participants in each study

| Study name | Number of participants | Sample type use in research | Recruitment criteria | Relevant chapters in thesis |
|-------------------|---|------------------------------------|--|------------------------------------|
| DEFINITION | 36 matched PBMCs and WB (1 failed PBMC) | PBMC & Tempus™ | ANA positive At-Risk individuals, SLE active, inactive | All |
| USEFUL | 96 | Tempus™ | SLE patients with musculoskeletal symptoms | 0 |
| MASTERPLANS | 223 | Tempus™ | SLE patient due to start a biologic | 0 |
| BRAGGSS | 30 | Tempus™ | RA patients | 0 |
| Healthy controls | 10 | PBMCs & Tempus™ | Healthy individuals with no autoimmune disease | All |

2.2 Blood collection and sample processing

PBMCs: 24mls of blood were drawn from participants directly into EDTA vacutainers and processed no longer than 4 hours after collection. PBMCs were separated using LeucoSep tubes (Greiner Bio-One): a density gradient centrifugation method. Blood was poured directly onto the porous barrier of the LeucoSep tube and centrifuged at 800 x g for 15 minutes at room temperature (RT) in a swinging bucket rotor. The brake was set to 0 on the centrifuge to avoid the blood layers from mixing after separation. The top layer of plasma was discarded by an auto pipette and the enriched PBMCs, seen by eye as a white buffy coat layer, was transferred into a new centrifuge tube and washed with 50ml of PBS and centrifuged at 300 x g for 10 minutes. The supernatant was discarded, and the pelleted PBMCs were treated with red cell lysis buffer (0.89% NH₄Cl, 0.1% KHCO₃, 0.02% EDTA, in ddH₂O) at RT for 10 minutes. After incubation, cells were centrifuged at 300 x g for 10 minutes and washed again with PBS. 2X10⁶ PBMCs were then pelleted down and resuspended in 300ul of RNA lysis buffer (Cambridge bioscience; Zymo research) and stored at -80°C until RNA isolation was performed.

Whole blood: 3ml of blood were drawn from participants directly into Tempus™ Blood RNA Tubes (Applied Biosystems™), inverted to thoroughly to mix blood with RNA stabilisers and stored at -80°C until RNA isolation was performed.

2.3 RNA isolation

PBMCs: Total RNA was isolated using the Quick-RNA™ Miniprep Kit (Cambridge bioscience; Zymo research): a Clean Spin™ column technology. Firstly, PBMCs in RNA lysis buffer were completely thawed and brought to RT before proceeding with the protocol. Genomic DNA and cellular debris were removed using the Spin-Away™ Filter. The flow through was mixed with 95-100% ethanol in preparation to transfer the mixture to the Zymo-Spin™ IIICG Column. After the RNA was centrifuged through the column membrane, DNase I was pipetted directly onto the membrane and incubated for 15 minutes at RT to remove any additional contaminating genomic DNA. The column membrane was washed with RNA prep buffer and twice more with RNA wash buffer (containing ethanol) before eluting RNA in 50ul of DNase/RNase-free water.

Whole blood: Total RNA was isolated directly from thawed Tempus™ blood via column separation using Preserved Blood RNA Purification Kit I (Norgen Biotek, Canada): a column-based extraction kit. Tempus™ blood was vortexed vigorously with a diluent provided with the kit for 30 seconds then centrifuged at 4000 x g for 30 minutes at 4°C in swing bucket centrifuge for 30 minutes to pellet the RNA. Supernatant was discarded and the pellet was resuspended with lysis solution then mixed with 95-100% ethanol. The mixture was added to the column and centrifuged to bind the RNA to the membrane and treated with DNase I (Norgen's RNase-Free DNase I Kit) for 15 minutes at RT. The column was washed with wash solution 3 times before eluting the RNA with 100ul elution solution.

For both sample types, RNA quantity and quality was measured using Nanodrop 1000 (ThermoFisher Scientific) and stored at -80°C immediately.

2.4 ISG transcript quantification

Reverse transcription & preamplification of target genes: RNA of <100ng was reverse transcribed with a mixture of random and oligo deoxythymidine primers (Fluidigm® Reverse Transcription Master Mix) using the thermocycler settings in Table 6. Next, the copied DNA (cDNA) was pre-amplified with ISGs and housekeeping genes, see Table 9 for TaqMan assay ID information. A pooled assay mix was created with addition of 1X Tris-Hydrochloric acid Ethylenediaminetetraacetic acid (TE) buffer (Promega) to be a final concentration of

0.2X which was then combined with cDNA and Preamp Master Mix (Fluidigm®). Pre-amplification was performed using 14 cycles (Table 7) and the product was diluted 1:5 with TE buffer.

Table 6. Thermal cycler settings for reverse transcription

| Temperature | Time | Condition |
|-------------|------------|-----------------------|
| 25 | 5 minutes | Pre-anneal of primers |
| 42 | 30 minutes | Reverse transcription |
| 85 | 5 minutes | Denature enzyme |
| 4 | ∞ | |

Table 7. Thermal cycler settings for preamplification of target genes

| Temperature | Time | Cycles | Condition |
|-------------|------------|--------|----------------------|
| 95 | 2 minutes | 1 | Initial denaturation |
| 95 | 15 seconds | 14 | Denature cDNA |
| 60 | 4 minutes | | Anneal & extension |
| 4 | ∞ | | |

Set up of integrated fluidic circuit (IFC) gene expression array: Pre-amplified target cDNA was prepared with gene expression sample loading reagent (Fluidigm®) and TaqMan universal master mix (Applied Biosystems). In addition, TaqMan assays were prepared with assay loading reagent (Fluidigm®). PBMCs stimulated with high concentrations of IFNs were used as a positive control along with 'no template control' and 'no reverse transcriptase' as negative controls. Samples and TaqMans were loaded on a Fluidigm® 96.96 Dynamic Array™ integrated fluidic circuit IFCs chip and gene expression assays were performed using the BioMark™ HD System gene expression Standard v1 programme (Table 8).

Real time qPCR analysis: IFC runs were analysed on the Fluidigm Real-Time PCR Analysis software and cycle threshold (Ct) values were automatically determined using user detector thresholds.

Table 8. Real Time PCR BioMark™ HD System gene expression Standard v1 programme

| Temperature | Time | Cycles | Condition |
|-------------|------------|--------|--------------|
| 50 | 2 minutes | 1 | Thermal mix |
| 70 | 30 minutes | 1 | |
| 25 | 10 minutes | 1 | |
| 50 | 2 minutes | 1 | UNG |
| 95 | 10 minutes | 1 | Hot start |
| 95 | 15 seconds | 40 | Denaturation |
| 60 | 1 minute | | Annealing |

Table 9. ISG TaqMan assay IDs (Applied Biosystems)

| IFN-Score- | Abbreviation | Gene Name | Function | TaqMan Assay ID |
|------------|--------------|---|-----------------------------|-----------------|
| A | CCL8 | C-C Motif Chemokine Ligand 8 | Chemotactic factor | Hs04187715_m1 |
| A | CEACAM1 | Carcinoembryonic Antigen-Related Cell Adhesion Molecule 1 | Cell-cell adhesion molecule | Hs00989786_m1 |
| A | CXCL10 | C-X-C Motif Chemokine Ligand 10 | Th1 proinflammatory | Hs01124251_g1 |
| A | GBP1 | Guanylate Binding Protein 1 | Anti-viral | Hs00977005_m1 |
| A | IFI27 | Interferon Alpha Inducible Protein 27 | Apoptosis signally pathway | Hs01086373_g1 |
| A | IFI44 | Interferon Induced Protein 44 | Cytoskeleton | Hs00951349_m1 |
| A | IFI44L | Interferon Induced Protein 44 Like | Anti-viral | Hs00915292_m1 |
| A | IFIT1 | Interferon Induced Protein With Tetratricopeptide Repeats 1 | Anti-viral | Hs01911452_s1 |
| A | IRF7 | Interferon Regulatory Factor 7 | Binds to ISRE | Hs01014809_g1 |
| A | ISG15 | ISG15 Ubiquitin Like Modifier | Chemotactic factor | Hs00192713_m1 |
| A | RSAD2 | Radical S-Adenosyl Methionine Domain Containing 2 | Anti-viral | Hs00369813_m1 |
| A | XAF1 | X inhibitor of apoptosis Associated Factor 1 | Inhibitor of apoptosis | Hs01550142_m1 |
| B | BST2 | Bone Marrow Stromal Cell Antigen 2 | Development of B Cells | Hs01561315_m1 |

| IFN-Score- | Abbreviation | Gene Name | Function | TaqMan Assay ID |
|-------------------|---------------------|--|---|------------------------|
| B | IFI16 | Interferon Gamma Inducible Protein 16 | Inhibits cell growth | Hs00194261_m1 |
| B | IFIH1 | Interferon Induced with Helicase C Domain 1 | Anti-viral | Hs01070332_m1 |
| B | LAMP3 | Lysosomal Associated Membrane Protein 3 | Plays a role in dendritic cells (adaptive immunity) | Hs00180880_m1 |
| B | NT5C3B | 5'-Nucleotidase, Cytosolic IIB | 5'-nucleotidase activity | Hs00369454_m1 |
| B | PHF11 | PHD Finger Protein 11 | Th1 proinflammatory | Hs00211573_m1 |
| B | SERPING1 | Serpin Family G Member 1 | Complement cascade regulator | Hs00163781_m1 |
| B | SOCS1 | Suppressor Of Cytokine Signalling 1 | Negative feedback of cytokine signalling | Hs00705164_s1 |
| B | SP100 | SP100 Nuclear Antigen | Tumour suppressor | Hs00162109_m1 |
| B | STAT1 | Signal Transducer and Activator Of Transcription 1 | Induces ISG transcription | Hs01013996_m1 |
| B | TAP1 | Transporter 1, ATP Binding Cassette Subfamily B Member | Antigen transport for MHC class I signalling | Hs00388675_m1 |
| B | TRIM38 | Tripartite Motif Containing 38 | inhibiting TLR3-mediated type I interferon signalling | Hs00197164_m1 |
| B | UBE2L6 | Ubiquitin Conjugating Enzyme E2 L6 | Degrades abnormal proteins | Hs01125548_m1 |
| B | UNC93B1 | Unc-93 Homolog B1, TLR Signalling Regulator | Regulates TLR signalling | Hs00276771_m1 |
| Cross Loading | HERC5 | HECT And RLD Domain Containing E3 Ubiquitin Protein Ligase 5 | Anti-viral | Hs00180943_m1 |
| Cross Loading | IFI6 | Interferon Alpha Inducible Protein 6 | Regulates apoptosis | Hs00242571_m1 |
| Cross Loading | MX1 | MX Dynamin Like GTPase 1 | Anti-viral | Hs00895608_m1 |

| IFN-Score- | Abbreviation | Gene Name | Function | TaqMan Assay ID |
|---------------|--------------|---|---------------------------------------|-----------------|
| Cross Loading | CASP1 | Caspase 1 | Cleaves and activates other cytokines | Hs00354836_m1 |
| Cross Loading | EIF2AK2 | Eukaryotic Translation Initiation Factor 2 Alpha Kinase 2 | Anti-viral | Hs00169345_m1 |

2.5 Calculation of IFN-Scores

Ct values were exported from the real-time PCR software and undetected Ct values were singly imputed using the R package 'nondetects'.

For Chapters 3 and 4, delta Ct gene expression was normalised to *PPIA* (TaqMan ID: Hs99999904_m1), and IFN-Scores were determined by the median value of all ISGs in the score (Table 9). Calculations of IFN-Scores are described separately in Chapter 5. Finally, to better visualise IFN-Scores, delta Ct values were reflected by multiplying the value by -1. This means that the higher the value the higher the expression and vice versa.

2.6 Statistical analysis

All graphs were created in GraphPad Prism software 9.3.1. Statistical analyses of Pearson's correlation and Bland-Altman plots were also analysed in GraphPad Prism. Independent T tests were calculated using SPSS and R software and packages were used for factor analysis. Details of each tests are described in each Chapter methods section.

Chapter 3: Comparison of IFN-Scores Between PBMC and Whole Blood Sample Types

3.1 Background

Continuous IFN-Score-A and -B were first published in 2018 (70) using unsorted PBMCs and, as previously discussed, they show promise as IFN biomarkers with strong associations with several important clinical outcomes in SLE (70, 74). Compared to isolated PBMCs, WB samples obtained in RNA stabilising Tempus™ collection system is much simpler and quicker processing technique and thus may be more accessible for development as a clinical platform.

For example, benefits of using a WB sample type over PBMCs includes clear advantages for blood processing and storage: freshly drawn blood into Tempus™ or PAXGENE tubes can be stored directly at -80°C for up to 6 years (79). The manufacturer also states that RNA will be stable at room temperature for up to five days, allowing option to transport samples to laboratories for processing using regular mail services, or even international shipping.

In contrast, PBMCs must be processed within 4 hours, or an excess of cellular apoptosis can lead to degradation in RNA integrity. Moreover, on site laboratory facilities and specialised expertise are required to isolated PBMCs and this process can take up to two hours. PBMCs require long term storage in a liquid nitrogen facility ($\leq -150^{\circ}\text{C}$) (80). Alternately, isolated PBMCs can be stored in RNA lysis buffer at -80°C before RNA extraction. These additional considerations will ultimately impact upon operating costs and wider availability as a clinical assay.

3.2 Objective

To assess whether PBMC and WB sample types can be used interchangeably to calculate existing IFN-Scores-A and -B.

3.3 Methods

3.3.1 Sample inclusion and sample processing

For this series of experiments, paired PBMCs and WB obtained at the same blood draw were analysed from 45 subjects enrolled in the DEFINITION study. To evaluate IFN-Scores across the range of disease, subjects with baseline active SLE (n=12), inactive SLE (n=10), At-Risk progressors (n=6), At-Risk non-progressors (n=7) and healthy donors (n=10) were included (Table 5). PBMC and WB Tempus™ sample collection and RNA extraction was carried out as per section 2.3.

3.3.2 Gene transcript quantification

Thirty-one ISGs (Table 9) gene transcripts plus reference gene *PPIA* were quantified by TaqMan Fluidigm as described in section 2.4 and IFN-Score-A and -B were calculated as described in section 0. IFN-Score are represented as reflected scores (score multiplied by -1) so that numerically higher scores represent higher expression, while preserving the original distribution.

3.3.3 Statistical analysis

Statistical analysis and graphs were performed and created using GraphPad Prism version 9.3.1. Box and whisker plots show the median, interquartile ranges, and minimum and maximum values with paired t-test assessing IFN-Score-A and -B differences between PBMC and WB sample types.

The relationship of IFN-Score-A and -B from 45 matched PBMC and WB samples were compared, and Pearson's correlation calculated the correlation coefficient (R) to determine the strength of the relationship comparing the two sample types, where $r=1$ / -1 is perfect positive / negative correlation and $r=0$ is no correlation between sample types.

Bland-Altman were used to visualise and assess the agreement and differences between IFN-Score-A and -B in PBMC and WB sample types. Values were plotted by the average of the two measurements (x-axis) and the difference of the two measurements (y-axis). Perfect agreement is when there is no difference between the two sample types and the bias=0. Points scatters with a consistent spread across values represent a systematic difference. Details of samples and comparisons are noted in respective figure legends.

3.4 Results

Comparison of IFN-Score-A and -B in PBMC and WB sample types revealed that WB has a significantly higher mean expression compared to PBMCs ($p < 0.0001$). Mean IFN-Score-A expression in PBMCs was -5.14 (1.95) Vs. -3.70 (2.08) in WB. Mean IFN-Score-B expression in PBMCs was -4.68 (0.72) Vs. -3.45 (0.73) in WB (Figure 1).

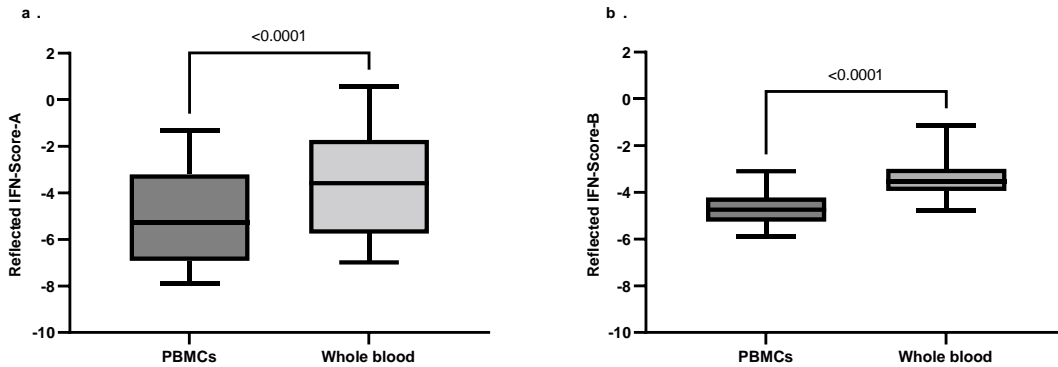


Figure 1. IFN-Score-A and -B display an overall higher expression in Tempus™ whole blood compared to PBMCs. 45 matched PBMCs and WB samples calculated a) Reflected IFN-Score-A and b) Reflected IFN-Score-B. Box and whisker plots display median, interquartile ranges and minimum and maximum values. Paired t-test calculated the significance between the two sample types, p values displayed on the graph.

Pearson's correlation was calculated to see whether a standardised difference could be identified and used to convert between the sample types (Figure 2). IFN-Score-A displays a strong correlation between sample types ($r=0.93$, $p < 0.0001$, Figure 2a) however a weaker correlation was found with IFN-Score-B ($r=0.45$, $p=0.0019$, Figure 2b). Although the correlation was significant in both IFN-Scores, this alone is not enough to calculate agreement between sample types as they do not measure the differences between each point, and there was still substantial disagreement. The difference between sample types was more consistent for IFN-Score-A than -B.

Bland-Altman is a method that is used to assess the agreement between two clinical measurement methods (81) and therefore has been used in determining the bias between the mean difference in this data (Figure 3). Perfect agreements (solid line) between sample types were not expected.

The bias of IFN-Score-A in PBMCs vs. WB was -1.44 (upper LOA=0.08, lower LOA= -2.96, Figure 3a) reflected delta Ct and -1.23 (upper LOA=0.26, lower LOA =-2.73, Figure 3b) in IFN-Score-B in PBMCs vs. WB. The variation in LOA is relative to the range of the values and the size of the bias for both scores and for IFN-Score-A, the bias is more systematic.

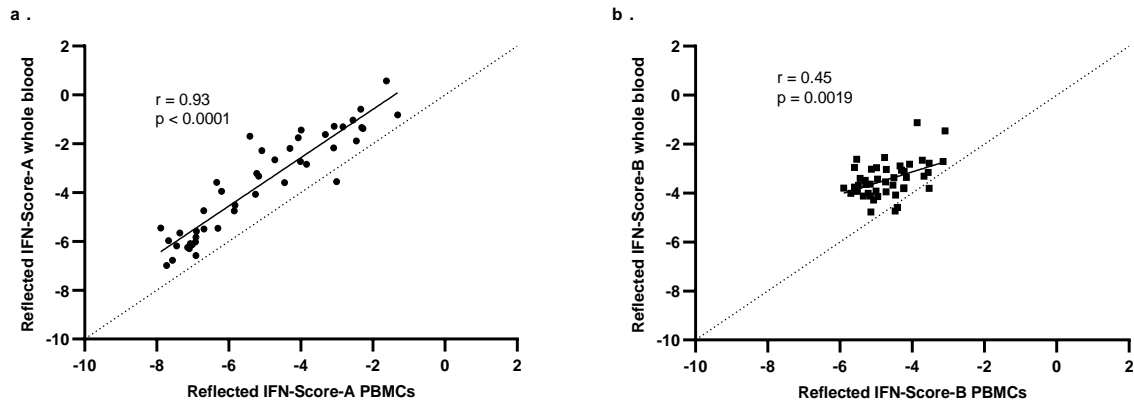


Figure 2. PBMC and Tempus™ whole blood have a stronger relationship in IFN-Score-A than IFN-Score-B. Comparison of 45 match PBMCs vs. WB samples in a) Reflected IFN-Score-A and b) Reflected IFN-Score-B. The dotted line represents perfect correlation, and the solid line represents the regression line of the points on the graph. Pearson's correlation calculated the r and p values displayed on each data set.

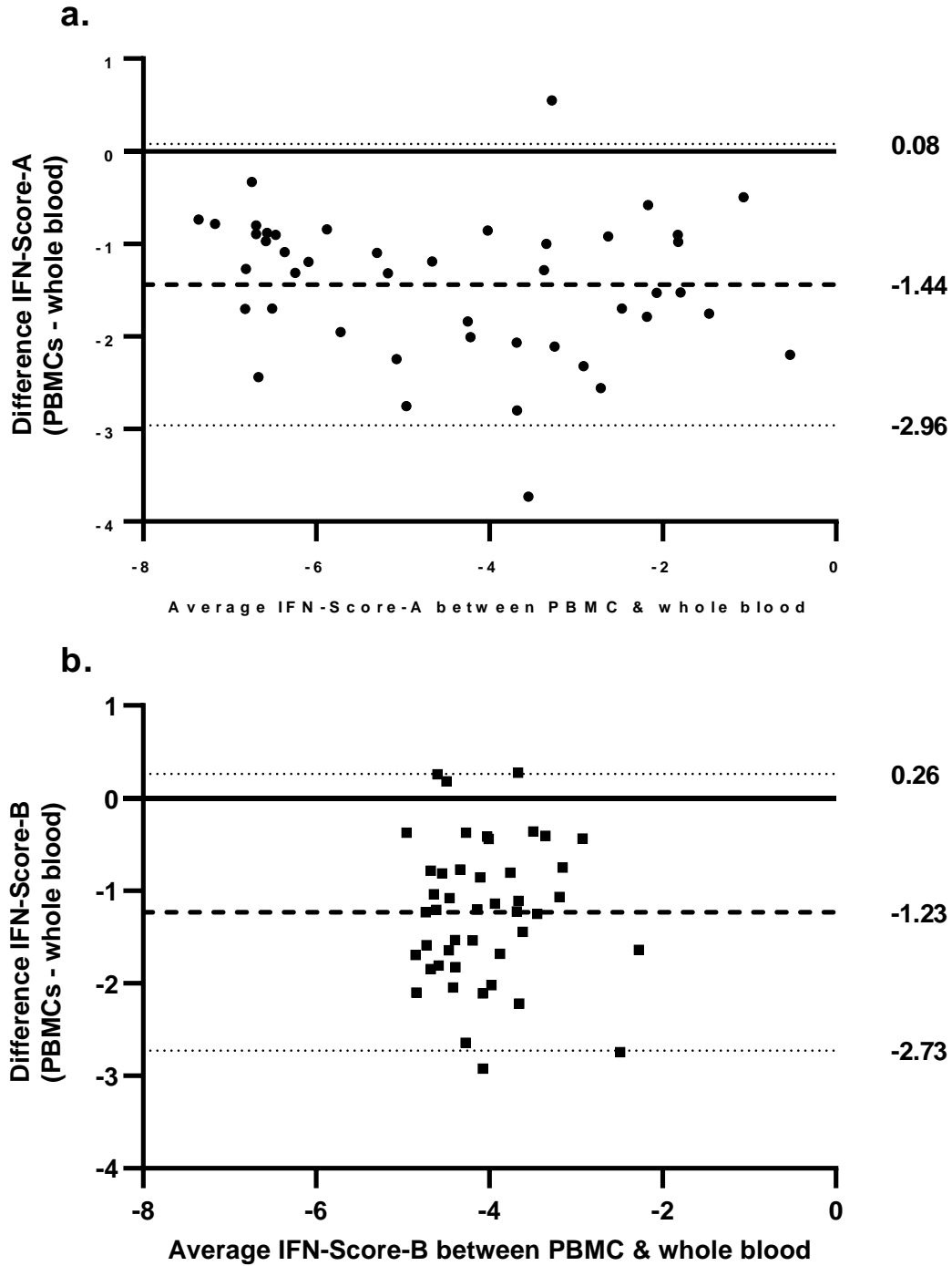


Figure 3. Bland-Altman plots determine limits of agreement of IFN-Score between PBMC and WB sample types. Agreement of IFN-Scores in 45 match PBMCs and WB: a) IFN-Score-A and b) IFN-Score-B. The average value was plotted against the difference between both sample types for each sample ($n=45$). The solid line ($y=0$) represents perfect agreement, the thick dotted black line represents the bias, and the smaller dotted lines represent the upper and lower limits of agreement, indicated on the graph.

3.5 Discussion

Key results

WB sample type is a more technically feasible way of processing IFN-I assays in clinical settings. Since clinical validation of IFN-Score-A and -B to date has been performed using a PBMC sample type (70, 74) this work set out to compare IFN-Scores between sample types to test how readily measurement using WB samples might be used interchangeably with PBMC data. This analysis revealed the following:

1. WB sample type produced significantly higher IFN-Score-A and -B expression compared to a PBMC sample type.
2. Comparisons between different sample types showed strong positive correlation in IFN-Score-A with apparent systematic agreement. However, there was only a weak positive correlation in IFN-Score-B.

Collectively these results indicate that IFN-Score-A could potentially be used in WB samples with further investigation. However, agreement between PBMC and WB sample types in IFN-Score-B was limited.

Context with other data

Firstly, there are key differences between PBMCs and WB. In a study of 8 mild asthmatics, He et al found that 704/730 transcripts had higher expression levels in matched WB samples compared to PBMCs, and WB transcript measurements were less variable (84). Additionally, PBMCs lack the presence of polymorphonuclear leukocytes (PMN) such as neutrophils, eosinophils, platelets, reticulocytes, and red blood cells. In particular, neutrophils are the most abundant PMNs in non-disease, and levels increase in active SLE compared to healthy (82). Whitney et al, discovered that ~2000 gene transcripts to have at least a 2-fold change between PBMC and WB sample types that were associated with various cell types (83). Furthermore, since all nucleated cells respond to IFN-I, greater quantification of transcripts in WB is expected with their greater repertoire of cells (34).

The disagreement of IFN-Score-B between sample type was the most distinctive result. One explanation to this could be that neutrophil ISG expression is influenced by alternative IFN signalling cascades to other immune cells in PBMCs. For example, IFN-I, -II and -III signalling cascades activate pleiotropic downstream effects, with distinct overlap in gene transcription (39). IFN-Score-B uniquely contains ISGs from module M5.12 which were found to be induced by IFN-II (IFN- γ) as well as IFN-I (73). To support this, Ellis et al,

reported that PMNs signalled via STAT1 and STAT3 transcription factors upon IFN- γ activation which was not witness in other cell types (84).

In line with neutrophils dominating WB cellular components, Coit et al discovered there were significant demethylation of ISG CG sites (upregulated due to increased access to the gene) in SLE neutrophils in contrast to healthy control neutrophils (85), which could potentially explain the dissimilarities between PBMC and WB in IFN-Score-B. Additionally, M1.2 ISGs, unique to IFN-Score-A, are longitudinally stable due to epigenetic changes, and the strong correlation between sample types suggest that this is true across difference cell types (73, 86).

IFN-Score-A and -B were previously analysed in sorted PBMC cell subsets. The balance of IFN-Score-A and IFN-Score-B didn't seem to differ between subsets, although there was a higher expression in some subsets, especially monocytes in both SLE and healthy (70). Studies of isolated neutrophils reveal that ISGs were upregulated in patients with active Tuberculosis compared to healthy control (87), suggesting that neutrophils influence a higher IFN-Score that is seen in WB compared to PBMCs SLE.

Moreover, there is plenty of evidence that demonstrates the importance of neutrophils in SLE pathogenesis especially through their cellular death process, NETosis which is heightened in SLE compared to healthy. During this process, extracellular traps are released that contain nuclear material, initiating further immune response and IFN production (82). There has been a continuous discussion in the literature regarding low density neutrophils to play a part in SLE pathogenesis. Although, low density neutrophils are isolated along with PBMCs as it suggests in their name and therefore does not support arguments towards differences between PBMC and WB ISG expression. However, studies comparing low and normal density neutrophils (present in WB only) show higher expression of IFN-I and -II ISGs compared to immature, low density neutrophils, which are present in PBMCs (88). This suggests that activation of ISGs are upregulated upon terminal maturation stages of mature neutrophils (only present in WB).

Theories of increased neutrophil ISG expression by influences of epigenetics and varying IFN signalling pathways have not been evaluated in this study. However, reinterrogation of ISGs could account for unexplained variants, discussed in the next chapter.

Limitations

A limitation of this study is that variables such as age, sex, ethnicity, different diseases, measure of full blood count haven't been accounted for in the way they might influence ISG expression.

Future Work

A true explanation of why sample types do not correlate in IFN-Score-B may be complex to investigate, requiring, further studies into measurement of IFN-Scores for each cell subset.

Furthermore, it is disappointing to see that the most clinically relevant IFN-Score-B cannot simply be interchanged between sample types. Therefore, the recognised approach, factor analysis as used initially for IFN-Score-A and -B will be applied to a WB cohort to detect other unmeasured variables described in the next chapter.

Conclusions:

When reviewing the literature, researchers need to be aware of sample type description, as this data suggests assay measurements cannot always be interchangeable. These results raise questions about the contribution of neutrophils to the IFN-Score, and whether that has special pathogenic significance. The data suggests there are unmeasured variables which account for differences in PBMCs and WB ISGs and these appear most marked for the IFN-Score-B genes than IFN-Score-A.

In clinical practice and multicentre trials, it is more feasible to collect WB samples. IFN-Scores must therefore be developed and validated for that specific sample type. Next, the measurement of ISGs in WB is explored in the same way that PBMC IFN-Scores were previously discovered.

Chapter 4: IFN-Scores in Whole Blood: Factor Analysis

4.1 Background

Previous work has shown that two-scores (IFN-Score-A and-B) derived from genes selected from previously defined IFN modules (73) were able to differentiate between various clinical subsets (Table 3) (70). IFN-Score-A consisted of 12 classically reported ISGs whereas IFN-Score-B included 14 less frequently reported ISGs (Table 9). Moreover, IFN-Score-B has sparked particular interest due to predicting disease progression and response to therapy (74, 76).

Analysis of IFN-Scores using a WB platform is desirable for development of clinical assays due to a range of technical considerations discussed earlier in this report. The development of IFN-Score-A and-B was carried out in PBMCs however, this report indicates insufficient alignment between IFN-Scores, particularly IFN-Score-B, measured in WB and PBMCs (70).

This suggests a new arrangement of ISGs may be required in a WB sample type. Therefore, this chapter uses factor analysis, as applied in the original development of IFN-Score-A and B (70) to establish whether a two factor system can sufficiently describe ISG activity in WB.

It is essential that any clinically applicable potential SLE biomarker can reliably differentiate SLE from healthy individuals but also from other related autoimmune diseases. Rheumatoid arthritis (RA) is a chronic inflammatory autoimmune disease, with symptoms mainly involving painful and swollen joints but it can also involve constitutional upset, fatigue and lung involvement which may overlap symptomatically with SLE. Not only do RA and SLE demonstrate similar clinical features but both display an upregulated IFN signature (89, 90). When IFN-Score-A and -B were created, factor analysis included SLE, RA and healthy controls which can influence factor clustering of underlying factors and they displayed a difference between SLE and RA (70), therefore this chapter will extend to evaluate subjects with SLE and RA as well as healthy subjects.

4.2 Objective

To use factor analysis to discover new IFN-Scores in WB and explore how these relate to known PBMC IFN-Scores and patient groups.

4.3 Methods

4.3.1 Sample inclusion and sample processing

For this section of the report, 342 established SLE patients from the MASTERPLANS, USEFUL and DEFINITION, 6 baseline At-Risk progressors, 7 baseline At-Risk non-progressors, 30 RA patients and 10 healthy donor WB samples were utilised in factor analysis and 45 PBMCs from DEFINITION study were used in sample comparison analysis (Figure 4). Studies are described in more detail in in section 2.1. WB Tempus™ and PBMC sample collection and RNA extraction was carried out as per section 2.3.

4.3.2 Gene transcript quantification

The same 31 ISGs (Table 9) gene transcripts, as evaluated in the initial development of IFN-Score-A and -B in PBMCs were re-quantified here in WB, in addition to reference gene *PPIA*. Gene expression was quantified by TaqMan Fluidigm, as described in section 2.4 and IFN-Score-A and -B were calculated as described in section 0. IFN-Score are represented as reflected scores (score multiplied by -1) so that numerically higher scores represent higher expression, while preserving the original distribution.

4.3.3 Factor analysis

Factor analysis is a statistical approach for exploring the underlying structure of the data and to determine whether a reduced number of latent (not directly observed) variables could describe the variation in the data (91) and was used in pool samples of SLE, ANA positive At-Risk, RA and healthy individuals as described in section 4.3.1. Undetected delta Ct values were singly imputed using the R 'nondetects' package (92) and adequacy of the sample size was established using the Kaiser-Meyer-Olkin measure ('KMO' R function in the 'psych' R package), where a score greater than 0.5 is sufficient in sample sizes above 300 (93). Correlation within the data was explored using the 'ggcorrplot' R function in the 'ggcorrplot' package and Barlett's test of sphericity using the R function 'cortest.bartlett' in the 'psych' package (94, 95). The 'det' R function in base R was used to calculate the determinant of the correlation matrix and ISGs were removed based on minimising the squared multiple correlation until the determinant was above 0.00001.

Horn's Parallel Analysis was carried out using the 'paran' R function in the 'paran' R package to determine the optimal number of factors to retain, however if fewer factors than recommended explained a high percentage (~80%) of the variation the simpler solution was selected (96). A promax rotation was applied to the final solution using the 'fa' R function in

the 'psych' package and ISGs with a factor loading less than 0.4 on all factors were removed. The process was repeated until the final solution contained ISGs with a loading of greater than 0.4 on at least one factor.

A score for each of the final factors was then constructed by taking the ISGs that were loaded above 0.4 on each respective factor and summing the median ΔC_t values of those ISGs. Any ISGs that loaded greater than 0.4 on more than one factor were considered to be cross loaded and was not included in the construction of the scores.

4.3.4 Statistical analysis

The relationship between newly developed IFN-Score(s), in 45 PBMC and WB sample types, were explored using Pearson's correlation coefficient (r), correlation plots and Bland-Altman plots.

Old and new IFN-Scores were calculated in 395 WB samples and relationships between each IFN-Score were analysed using Pearson's correlation coefficient (r) and correlation plots.

To analyse the clinical relevance of the new WB IFN-Score(s), differences between Healthy, SLE and RA diagnostic groups, were determined by Hedges' g^* effect size which is recommended when equal variances are not assumed. Results were obtained from an online calculation platform which was created by the authors of this method for those who are not familiar with using statistical software, R, and can be found here:

<https://effectsize.shinyapps.io/deffsize/> (97). Effect sizes ≤ 0.2 were considered small $0.5 > 0.8$ were considered medium and ≥ 0.8 were considered large.

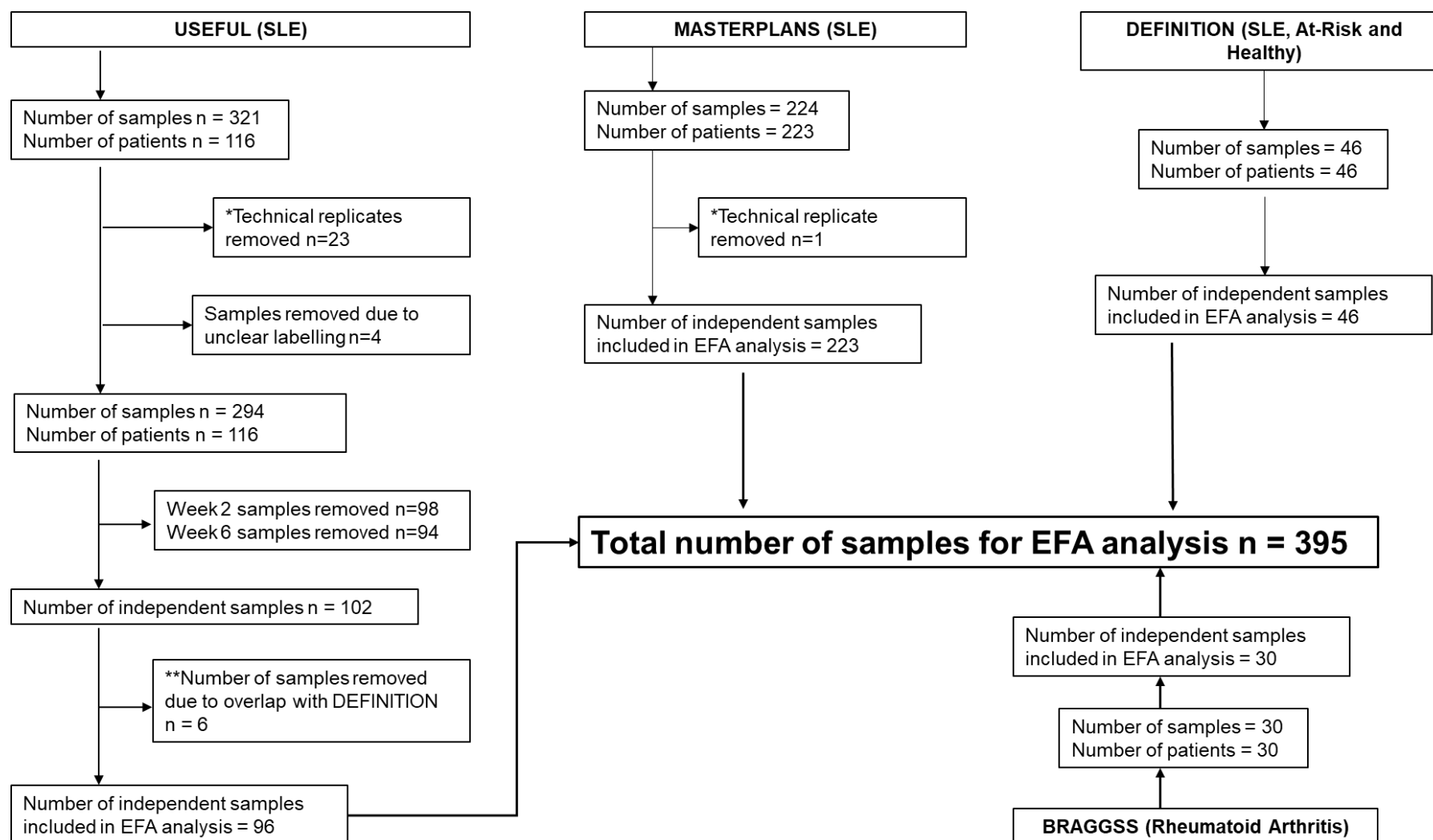


Figure 4. Consolidated Standards of Reporting Trials (CONSORT) diagram of samples included in exploratory factor analysis. *Technical replicate and **overlapping patient samples removed according to run date (earlier run retained).

4.4 Results

4.4.1 Whole blood factor analysis reveals distinct two-scores for whole blood ISG expression

Factor analysis in PBMCs initially described a two-factor system for describing ISG expression in the form of IFN-Score-A and -B. In order to determine whether an IFN-Score(s) could be developed using WB, factor analysis was therefore performed on the same TaqMan quantified 31 gene set in WB to understand underlying variables that might be affecting IFN pathway activity.

TaqMan quantified 31 ISGs (Table 9) to determine whether the gene expression values of multiple genes were driven by unobserved (latent) continuous variables by factor analysis. Kaiser–Meyer–Olkin (KMO) values of 0.96 confirmed sample size adequacy of factor analysis and Bartlett's test of sphericity confirmed correlation between genes ($p < 0.001$). To achieve correlation matrix determinant greater than 0.00001, 18 genes were removed based on minimising the squared multiple correlation (*HERC5*, *IFI44*, *RSAD2*, *UBE2L6*, *XAF1*, *IFI16*, *STAT1*, *MX1*, *EIFAK2*, *IRF7*, *TAP1*, *SP100*, *IFI44L*, *IFIH1*, *SERPING1*, *UNC93B1*, *PHF11*, *IFIT1*). Horn's Parallel Analysis suggested that 6 factors be retained however 91% variation was explained by two factors and so the simpler solution was retained. An oblique rotation was applied and resulted in 2 genes being removed for loading less than 0.04.

The process was then repeated on the reduced set of genes; the determinant was above 0.00001 and KMO = 0.89. Horn's Parallel Analysis suggested that 4 factors be retained however 97% variation was explained by two factors and so the simpler solution was again retained. Oblique rotation was again applied, and all genes had loadings greater than 0.04 on one or both of the retained factors.

From the 31 ISGs quantified, the two-factor solution derived in this WB analysis comprised a reduced gene set to those initially included in PBMC IFN-Score-A and IFN-Score-B. The resulting WB two factor system was therefore denoted as IFN-Score-C and IFN-Score-D (Table 10).

CCL8 loaded most strongly onto IFN-Score-C and *CASP1* loaded the most onto IFN-Score-D. *BST2* and *GBP1* were cross loaded and excluded from both factors. Table 10 displays the ISGs loaded into each factors along with the eigenvalues.

4.4.2 Comparison of PBMC and novel whole blood IFN-Scores

Compared to IFN-Score-A and -B developed and validated in PBMC sample type, the IFN-Score-C and -D system identified here in WB, utilises fewer ISGs in each score: five ISGs in IFN-Score-C and four ISGs in IFN-Score-D (Table 10).

IFN-Score-C closely aligns with IFN-Score-A in that IFN-Score-C shares 4/5 ISGs with IFN-Score-A: *CCL8*, *CXCL10*, *IFI27* and *ISG15*, and *LAMP3* previously clustered to IFN-Score-B. IFN-Score-D comprises 2/4 ISGs shared with IFN-Score-B: *SOCS1* and *TRIM38*, one ISG with IFN-Score-A *CEACAM1* and *CASP1* did not correspond with either of the original IFN-Scores.

The 31 genes analysed in the development of IFN-Scores were selected from three IFN-annotated modules described in a modular analysis of the blood SLE transcriptome by microarray (73). To further understand how closely the novel IFN-Score-C and -D system for WB corresponds to the original IFN-Score-A and -B derived in PBMCs, the modular origin of the composite genes were compared between IFN-Scores-A and-B, and IFN-Scores-C and -D. Similarities in the parent modules of IFN-Score genes supports the resemblance between IFN-Score-A to IFN-Score-C and between IFN-Score-B and IFN-Score-D (Table 11). IFN-Score-A and -C contain ISGs in modules M1.2 and M3.4. As suggested by the lower number of overlapping genes comprising IFN-Score-D and IFN-Score-B, the resemblance between these appears less than that of IFN-Score-A and IFN-Score-C. Although, IFN-Score-B does not fully align with IFN-Score-D, both contain ISGs mainly from M5.12 and M3.4 and only IFN-Score-B consists of ISGs from M1.2.

Table 10. Whole blood factor analysis: ISG Eigenvalues values. Values highlighted in green indicate the ISGs included in IFN-Score-C and values highlighted in orange indicate ISG included in IFN-Score-D. *BST2* and *GBP1* were cross loaded and excluded from further analysis.

| Gene | Factor 1: IFN-Score-C Eigenvalues | Factor 2: IFN-Score-D Eigenvalues |
|----------------|--|--|
| BST2 | 0.441 | 0.527 |
| CASP1 | -0.084 | 0.934 |
| CCL8 | 1.008 | -0.296 |
| CEACAM1 | 0.006 | 0.887 |
| CXCL10 | 0.797 | -0.078 |
| GBP1 | 0.419 | 0.508 |
| IFI27 | 0.836 | 0.031 |
| ISG15 | 0.655 | 0.237 |
| LAMP3 | 0.729 | 0.202 |
| SOCS1 | -0.154 | 0.58 |
| TRIM38 | 0.226 | 0.726 |

Table 11. Modular comparisons of ISGs in each IFN-Score

| | Module | IFN-Score-A | IFN-Score-B | IFN-Score-C | IFN-Score-D |
|----------|--------|-------------|-------------|-------------|-------------|
| BST2 | 5.12 | | * | | |
| CASP1 | 5.12 | | | | * |
| CCL8 | 3.4 | * | | * | |
| CEACAM1 | 3.4 | * | | | * |
| CXCL10 | 1.2 | * | | * | |
| EIF2AK2 | 3.4 | | | | |
| GBP1 | 3.4 | * | | | |
| HERC5 | 1.2 | | | | |
| IFI16 | 5.12 | | * | | |
| IFI27 | - | * | | * | |
| IFI44 | 1.2 | * | | | |
| IFI44L | 1.2 | * | | | |
| IFI6 | - | | | | |
| IFIH1 | 1.2 | | * | | |
| IFIT1 | 1.2 | * | | | |
| IRF7 | 3.4 | * | | | |
| ISG15 | 1.2 | * | | * | |
| LAMP3 | 1.2 | | * | * | |
| MX1 | 1.2 | | | | |
| NT5C3B | 5.12 | | * | | |
| PHF11 | 5.12 | | * | | |
| RSAD2 | 1.2 | * | | | |
| SERPING1 | 1.2 | | * | | |
| SOCS1 | 3.4 | | * | | * |
| SP100 | 5.12 | | * | | |
| STAT1 | 3.4 | | * | | |
| TAP1 | 5.12 | | * | | |
| TRIM38 | 5.12 | | * | | * |
| UBE2L6 | 3.4 | | * | | |
| UNC93B1 | 5.12 | | * | | |
| XAF1 | 1.2 | * | | | |

4.4.3 Comparison of PBMC and whole blood sample types in newly developed whole blood IFN-Scores

As sample type comparisons in IFN-Score-A and -B, revealed that IFN-Score-B correlated poorly between PBMC and WB sample types, the new WB derived IFN-Score-C and -D were also compared between PBMC and WB sample types of 45 matching samples from the DEFINITION study.

As presented in the previous chapter, there was a positive correlation between PBMCs and WB in IFN-Score-A, but a much less convincing relationship was seen in IFN-Score-B across sample types. Overall, the correlation between different sample types in IFN-Score-C and -D appeared slightly less well preserved across sample types. IFN-Score-C showed a positive correlation between sample types ($r=0.77$, $p<0.001$, Figure 5a) and there was a weaker positive relationship with IFN-Score-D ($r=0.30$, $p=0.0419$, Figure 5b).

Analysis of agreement using Bland-Altman agreement plots, showed a weak degree of agreement for IFN-Score -C (Bias=-1.74, upper LOA=1.22, lower LOA=-4.70, Figure 5c). When there is an overall lower average between sample types, WB expression is higher and when there is an overall higher average between sample types, PBMC expression is higher. IFN-Score-D had a bias of -1.62, upper LOA of 0.20 and a lower LOA of -3.44, Figure 5d).

Collectively, these findings indicate that, like IFN-Score-A and -B, the novel WB derived IFN-Score-C and -D cannot be interchanged between sample types and further suggests that cell types within WB, for example, are highly influential to the pattern of ISG expression.

4.4.4 Comparison of original IFN-Scores to newly developed whole blood IFN-Scores

When both the original and newly developed score were calculated in WB samples, the weakest positive correlation was found between IFN-Score-C and -D ($r=0.54$, $p<0.0001$, Figure 6f) which was strikingly different to the relationship between IFN-Score-A and -B ($r=0.83$, $p<0.0001$, Figure 6a). IFN-Score-C shared a moderate positive correlation with IFN-Score-A ($r=0.77$, $p<0.0001$, Figure 6b) compared to a weaker correlation with IFN-Score-B ($r=0.66$, $p<0.0001$ Figure 6d), which is not surprising when IFN-Score-C shares 4/5 common ISGs with IFN-Score-A, compared to 1/5 with IFN-Score-B (Table 11). There was a positive correlation between IFN-Score-D and -A ($r=0.71$, $p<0.0001$, Figure 6c) and an impressively strong positive correlation with IFN-Score-B ($r=0.92$, $p<0.0001$, Figure 6e).

IFN-Score-A and -B were developed in PBMCs and IFN-Score-C and -D were developed in WB, which showed similarities between ISGs included in the scores. Therefore, it was hypothesised that WB IFN-Score-C and -D, might be surrogates of PBMC IFN-Score-A and -B respectively. To test this, samples were analysed comparing IFN-Score-A in PBMCs with -C in WB and IFN-Score-B in PBMCs with -D in WB (Figure 7). In 45 matched PBMC and WB samples from the DEFINITION study, Pearson's correlation was calculated to test whether the similar scores (IFN-Score-A with -C and IFN-Score-B with -D) produce comparable results (Figure 7 a & b).

There is a strong positive relationship between IFN-Score-A and -C when calculated in their respective sample type ($r=0.79$, $p<0.0001$, Figure 7a) compared to a weaker positive relationship between IFN-Score-B and -D ($r=0.47$, $p=0.001$, Figure 7b). The LOA is larger between IFN-Score-A and -C (Bias=1.73, upper LOA=4.08, lower LOA=-0.59, Figure 7c) compared to the LOA IFN-Score-B and -D (Bias=0.25, upper LOA=1.74, lower LOA=-1.24, Figure 7d). Notably, when there is an overall lower average between IFN-Score-A in PBMC and IFN-Score-C in WB, IFN-Score-C in WB is higher and when there is an overall higher expression, IFN-Score-A in PBMCs is higher.

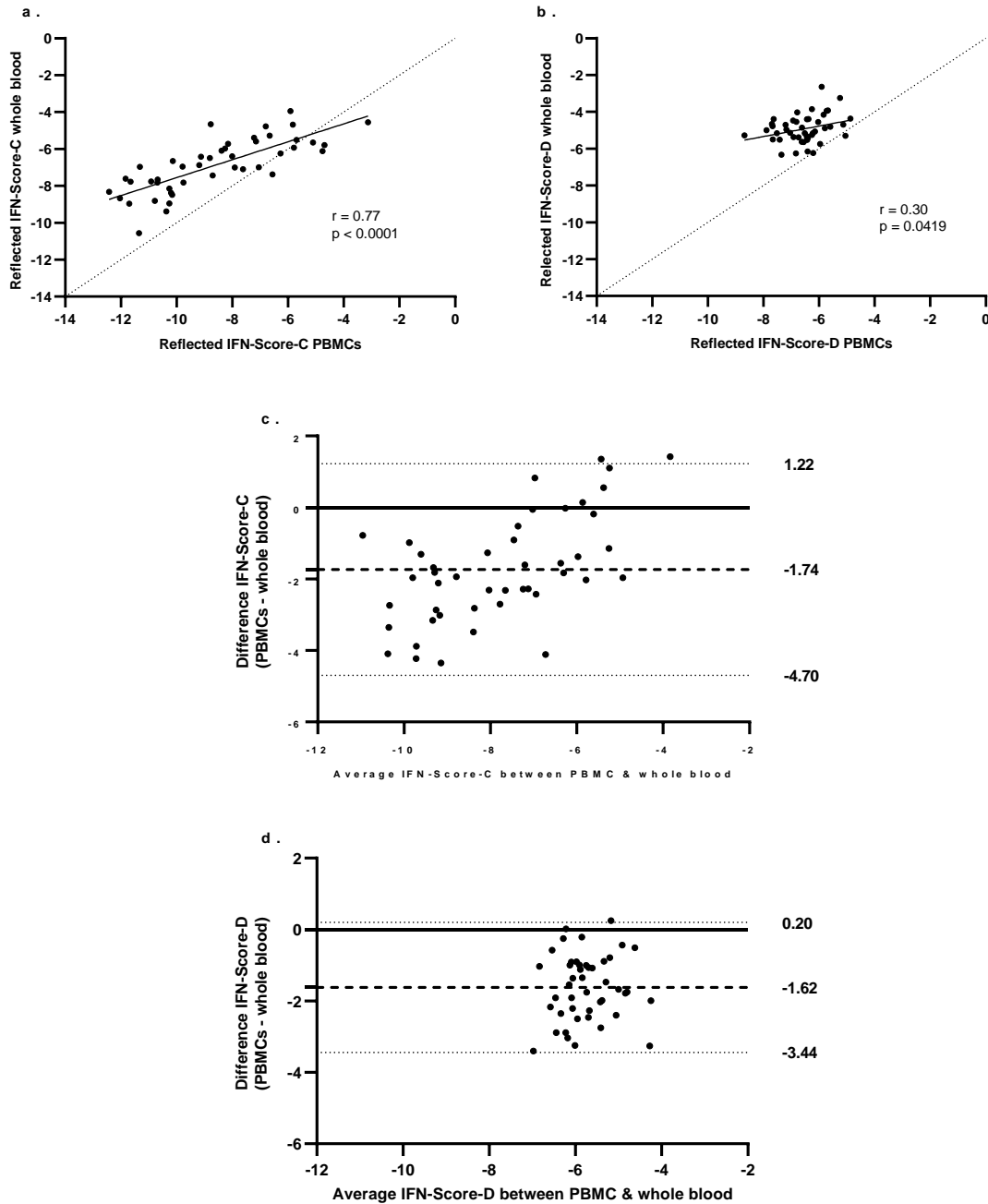


Figure 5. Comparison of IFN-Score-C and -D between PBMC and WB sample types. IFN-Scores were analysed in 45 matched PBMC and WB samples from active SLE ($n=12$), inactive SLE ($n=10$), At-Risk progressor ($n=6$), At-Risk non-progressor ($n=7$) and healthy donors ($n=10$). Dotted line on graphs a & b represent perfect agreement. Pearson's correlation coefficient r and p value represented on each plot. Bland-Altman plots c & d: the solid line ($y=0$) represents perfect agreement, the thick dotted black line represents the bias, and the smaller dotted lines represent the upper and lower limits of agreement, represented on the plots.

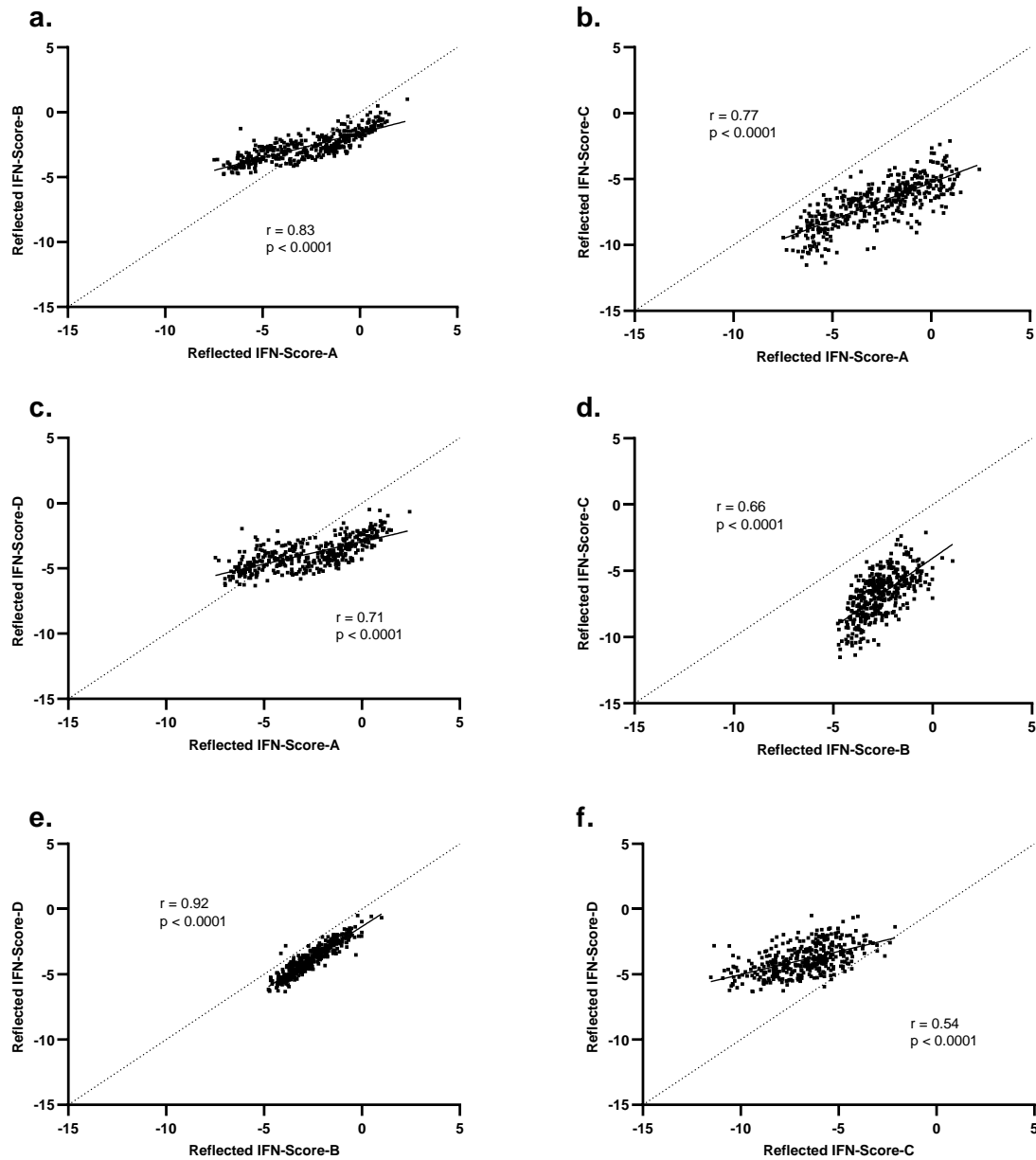


Figure 6. Correlation comparison of each IFN-Score in whole blood. Tempus™ whole blood data including SLE $n=347$, RA $n=30$ and healthy $n=10$, and At-Risk non progressor $n=7$. Reflected IFN-Scores displayed to identify larger values as higher gene expression and smaller values as lower gene expression. The dotted line represents perfect correlation. Pearson's correlation coefficient r and p values are represented for each plot.

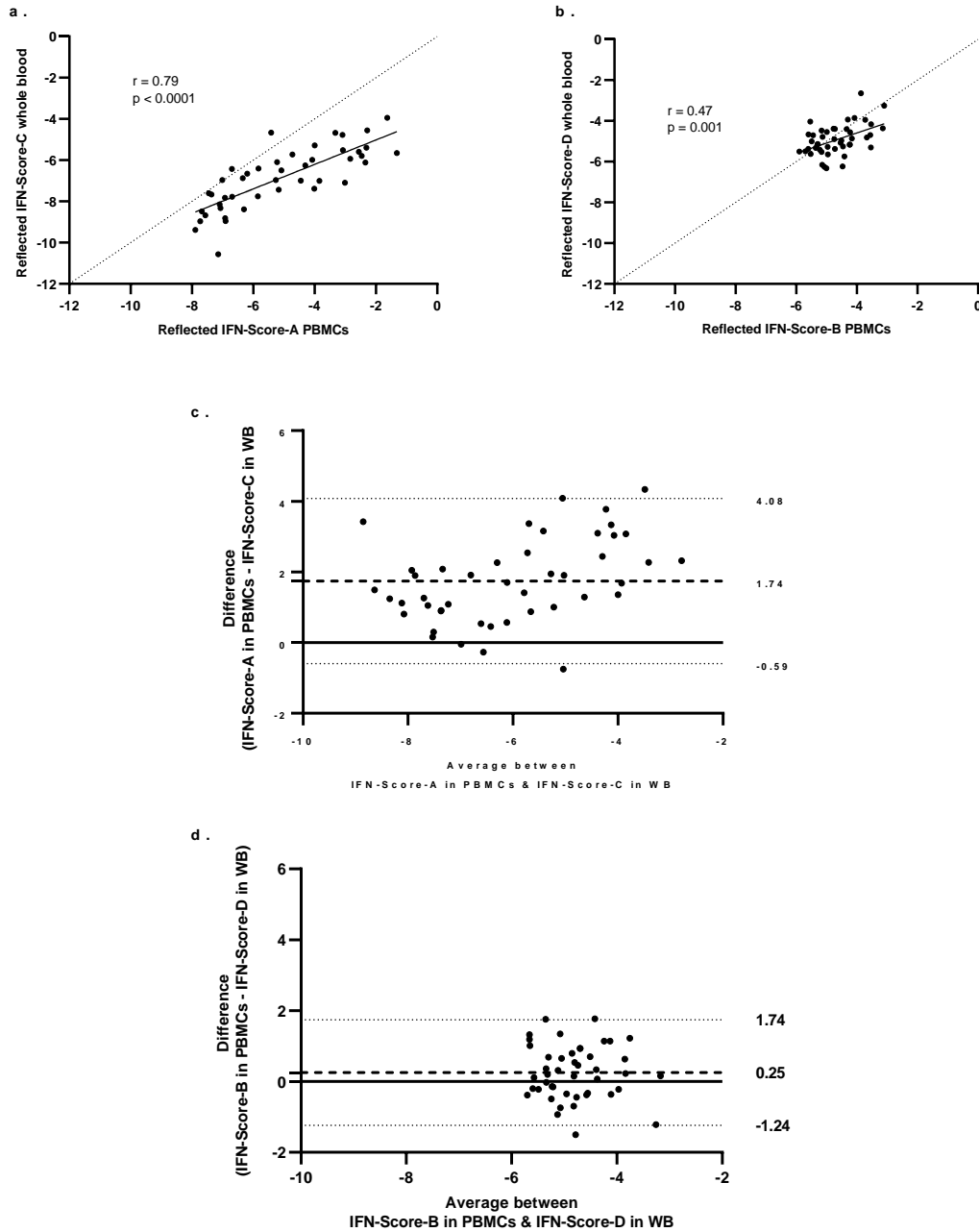


Figure 7. Correlation and Bland-Altman agreement plots between IFN-Score-A in PBMCs with IFN-Score-C in WB and IFN-Score-B in PBMCs with IFN-Score-D in WB. Since IFN-Score-A and -B were developed in PBMCs and IFN-Score-C and -D were developed in WB, comparisons were made between IFN-Score-A Vs -C and IFN-Score-B Vs -D as they showed to share similar properties. IFN-Scores were analysed in 45 matched PBMC and WB samples, active SLE ($n=12$), inactive SLE ($n=10$), At-Risk progressor ($n=6$), At-Risk non-progressor ($n=7$) and healthy donors ($n=10$). Dotted line on graphs a & b represent perfect agreement. Pearson's correlation coefficient r and p values are represented for each plot. Bland-Altman plots c & d: the solid line ($y=0$) represents perfect agreement, the thick dotted black line represents the bias, and the smaller dotted lines represent the upper and lower limits of agreement, represented on the graphs.

4.4.5 Comparison of novel and existing whole blood IFN-Scores across disease groups

Next, the ability of the existing and new scores measured in WB to discriminate between healthy subjects, SLE and RA was evaluated (Figure 8). Although IFN-Score-A was developed in PBMCs, WB calculated IFN-Score-A displayed the most definition between patient groups, with Hedges' g^* calculating large effect sizes between the averages all groups in (Healthy Vs SLE =1.59, 95% CI [1.10, 1.99], Healthy Vs RA =1.19, [0.36, 1.99], SLE Vs RA =0.88, [0.63, 1.13]).

IFN-Score-B and -D measured in WB both showed 'good' ability to discriminate between disease groups, with similar large effect sizes between healthy Vs. SLE (IFN-Score-B=1.12, [0.66, 1.58], IFN-Score-D=1.45, [0.96, 1.94]) and healthy Vs RA (IFN-Score-B=1.5, [0.69, 2.28], IFN-Score-D=2, [1.11, 2.85]). However, both IFN-Scores failed to display clear classification between SLE and RA (IFN-Score-B=0.07, [-0.21, 0.35], IFN-Score-D=0, [-0.29, 0.29]).

These findings indicate that despite apparent similarities in gene and modular composition of IFN-Scores-A and -C there were marked differences in their ability to distinguish between autoimmune disease groups and this could have implications for further development as biomarkers. For example, IFN-Score-A produced large effect sizes between all diagnostic groups whereas IFN-Score-C only has small effect sizes between all diagnostic groups (Healthy Vs SLE =0.453, [-0.03, 0.94], Healthy Vs RA =0.18, [-0.51, 0.86], SLE Vs RA =0.30, [-0.01, 0.61]).

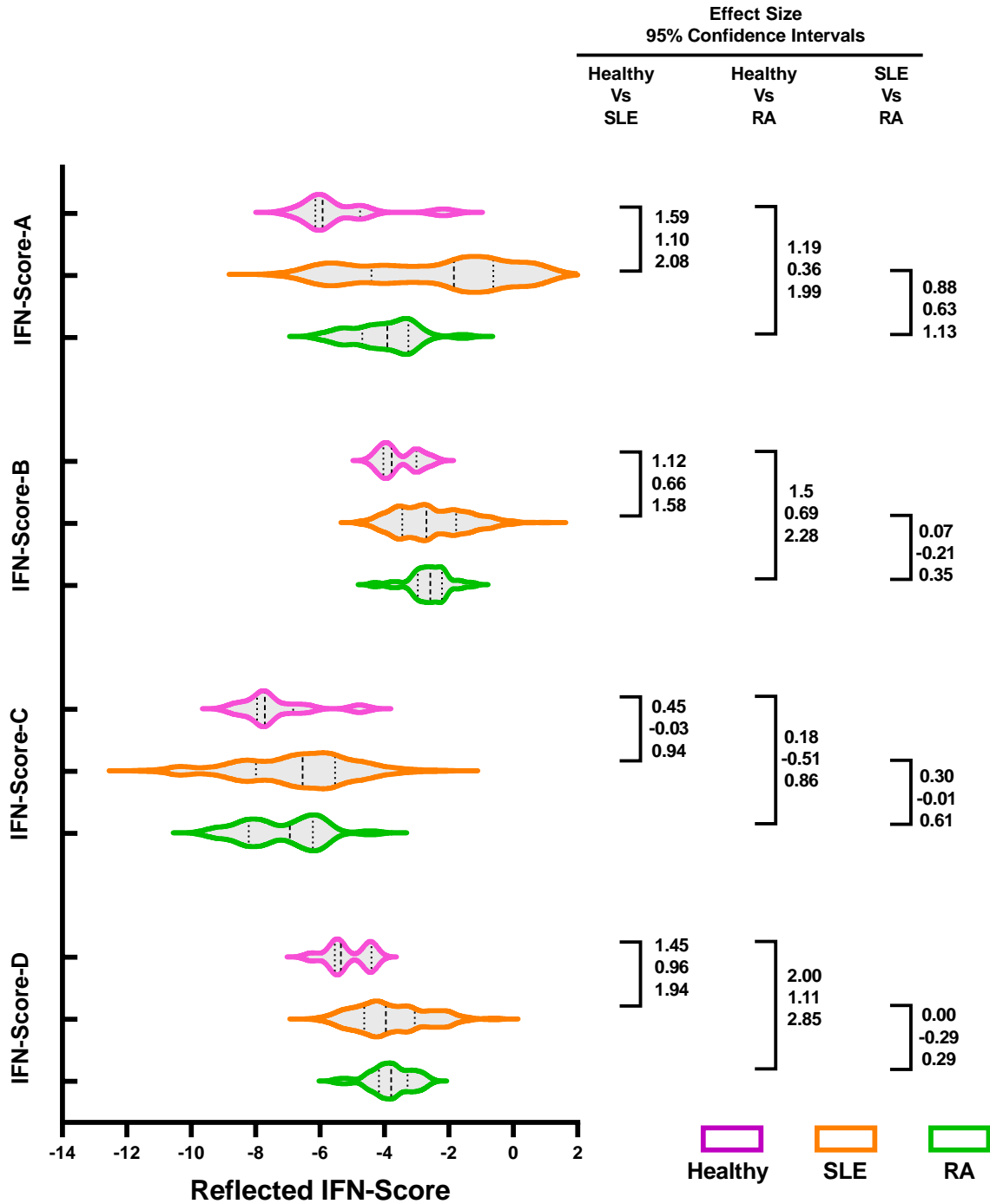


Figure 8. Reflected delta Ct Scores of IFN-Score-A, -B, -C and -D, comparison between diagnostic groups SLE, HC and RA. Violin plots represent the median value as bold dashed line and upper and lower quartiles as dotted lines. Healthy in pink $n=10$, SLE in orange $n=342$ and RA in green $n=30$ are displayed in each IFN-Score and effect values calculated between each diagnostic group within each score. Effect sizes calculated using Hedges' g^* presented as top value and lower and upper 95% confidence intervals presented underneath (97). Effect sizes ≤ 0.2 were considered small $0.5 > 0.8$ were considered medium and ≥ 0.8 were considered large.

4.5 Discussion

Key results

Initially, it is disappointing to see sample types cannot be interchanged reliably when measuring IFN-Score-A and -B however, this has led to new opportunities to analyse IFN pathogenesis in a WB sample type and investigates the potential for a more robust biomarker.

To refine and develop on IFN-Score signatures in SLE, a WB sample type is a more efficient way to process samples compared to PBMCs. Factor analysis of 31 ISGs (Table 9) in PBMCs, previously created a two-score system of ISG clustering named IFN-Score-A and IFN-Score-B. The current work demonstrated in chapter 3 indicated that IFN-Scores-A and -B show a poor degree of correlation and agreement when quantified in different sample types. This was particularly the case for IFN-Score-B, which has previously shown the most clinically relevant associations (70, 74-77).

To address this, factor analysis was repeated in WB samples of the same 31 ISGs drawn from modular SLE blood transcriptome (70), to test whether a different underlying factors explain the structure and variability of ISG expression in WB. This work has revealed the following:

1. A two-factor system (IFN-Score-C and -D) of ISGs established in WB accounts for 96% of the data compared to previous PBMC (IFN-Score-A and -B) analysis which accounts for 84% variation in ISG expression.
2. IFN-Scores-C and -D comprised a lower number of genes, with partially overlapping parent IFN modules (M1.2, M3.4, M5.12) to IFN-Score-A and -B which benefits the aim of refining IFN assays for clinical practice.
3. Correlation between PBMC and WB sample types showed weak positive correlation of IFN-Score-C however, weak correlation of IFN-Score-D.
4. WB comparisons identified weaker positive between IFN-Score-C and -D which was strikingly different to the positive correlation between IFN-Score-A and -B. Notably, IFN-Score-B and -D shared an impressively strong positive correlation.

5. Comparison of IFN-Scores dependent of their sample type origin, revealed IFN-Score-C and -D were not simply the WB equivalent of IFN-Scores-A and -D in PBMCs. IFN-Score-A in PBMC relation with -C in WB showed bias in agreement plots and IFN-Score-B in PBMCs and -D in WB had weak positive correlations.
6. IFN-Score-A better stratified between SLE, RA and healthy patient groups compared to IFN-Score- B, -C and -D in WB. IFN-Score-B and -D showed significant difference between healthy and SLE, and healthy and RA but failed to distinguish between SLE and RA. IFN-Score-C was the weakest among all IFN-Scores to stratify patient groups, with no significant differences.

Context with other data

These results suggest that the underlying factors influencing IFN-Score clustering are not equivalent within PBMC and WB autoimmunity. One suggestion could be the type of cytokine influence driving the IFN-Scores. For example, Chiche et al, described M1.2 and M3.4 to be induced by IFN- β more than IFN- α and both M3.4 and M5.12 were also induced by IFN-II (IFN- γ) as well as IFN-I (73). In this study, factor analysis clustering of ISGs found IFN-Score-C to include ISGs in M1.2 and M3.4 and IFN-Score-D equally consisted of ISGs in M3.4 and M5.12, suggesting other subtypes of IFN could explain the differences in clustering. Also, M1.2 is reported to be the most longitudinally stable IFN module whereas M5.12 is more responsive to disease activity, constituting these ISGs to be more robust, clinically relevant biomarker.

The degree of correlations and agreement between IFN-Scores was not simply a function of how similar they were in gene composition. Despite the differences in IFN-Score-B and -D ISG composition, the strong correlation between scores suggest they entail similar underlying factors which could be a contribution of IFN signalling cascade. For example, giving the broad range of transcription factor initiation upon IFN signalling, alternative types of IFN could induce similar transcription factors and or transcription promoter regions (39).

As highlighted previously, there are distinct differences in cellular composition between PBMC and WB sample types alluding to the granulocyte compartment of WB being the underlying influences of WB IFN-Scores, particularly IFN-Score-C, which showed weak correlation between other IFN-Scores. An RNA sequencing study of 2 female and 2 male healthy individuals revealed that ISGs from IFN-Score-C such as *CCL8*, *CXCL10* and *IFI27* were mostly upregulated in monocyte cell subsets, and *ISG15* and *LAMP3* were mostly

upregulated in granulocyte cells subsets such as neutrophils (98). Moreover, all IFN-Score-D transcripts *CASP1*, *CEACAM1*, *SOCS1* and *TRIM38* were mostly upregulated in neutrophils and basophils (98). This highlights the differences between ISG regulation in cell types however, results could be uniquely different in autoimmunity.

This gives possible explanation to the difference in factor analysis results compared to previous PBMC IFN-Scores-A and -B, in which cell types such as neutrophils, influence ISG transcripts. Especially when neutrophils play a large role in IFN dysregulation in SLE pathogenesis, by releasing NETs and inducing IFN production and high neutrophil transcript scores in juvenile SLE patients, correlate with high IFN signature (99-102).

Limitations

One limitation of this study is that a limited number of Healthy samples, potentially reducing the variety of ISG transcription level for factor analysis. Previously there were 14.9% of healthy patients for factor analysis that determined SLE from RA, whereas this study contained 2.5% of healthy patients within the total sample inclusion.

A reliable and robust biomarker must distinguish between similar diseases and patients whose classification may not yet be clear, in a scenario that matches a clinical question. SLE and RA are clinically and pathogenically quite different and can usually be distinguished easily without any blood tests. Other relevant comparisons may be SLE with Sjogren's syndrome, which is more similar to SLE than RA with similar joint involvement. However, this project was not designed to be a diagnostic study.

Also, RA PBMC sample type was unable to be retrieved as RA WB were collected from an external site. This would have been useful to test whether new IFN-Scores could distinguish between disease also in PBMCs.

Factor analysis results depend on the diversity of individuals who contribute to the sample. The greater diversity of immune states among the participants increases the ability of the analysis to find factors. This was addressed by including SLE, RA and HC. However, it is possible that results would differ if other autoimmune diseases were included, or greater variety of clinical states for each disease.

Future work

Firstly, to test the relevance and reliability of IFN-Score-C and -D biomarker potential, IFN-Scores will need to be measured in an independent cohort against other clinical factors such as pre/post treatment, pre/post diagnosis, and comparison of differential clinical feature.

Future work will explore WB separated cell type *in vitro* stimulations to see what exactly is driving the scores, this could include different IFN subtypes and cytokines that have been noted to drive IFN signature.

Conclusions:

Although in these results, IFN-Score-A differentiated the patient groups most clearly, in previous work IFN-Score-B has performed better for specific questions. For example, to predict development of SLE in At-Risk individuals, results were stronger for IFN-Score-B than -A (74). Since these studies analysed PBMCs, a validation study using WB would require an IFN-Score that captures similar information to IFN-Score-B using a WB sample. IFN-Score-D appears to meet this need and could act as a surrogate WB IFN-Score for IFN-Score-B in PBMCs, moreover, it uses a smaller number of genes so it simpler and cheaper. However, an additional study would be required to validate IFN-Score-D for the same clinical questions.

Chapter 5: Exploring Appropriate Reference Genes in SLE Patient Whole Blood When Measuring ISGs

5.1 Background

Real time polymerase chain reaction (RT-PCR) is a commonly used, highly sensitive method to quantify targeted RNA. RNA is reverse transcribed into cDNA and amplified according to target gene sequence (TaqMan) where the Ct value is determined (103). Reference genes (RG), also known as a house keeping gene or endogenous control are genes that are consistently and uniformly expressed and are essential to normalise quantity of target gene expression (104). Early qPCR studies commonly used *β-actin (ACTB)*, *GAPDH* and *18s* as RGs (105) however, recent studies have shown these genes respond to diverse biological factors and were not selected as having the most stable expression when screened in panels of 10-20 RGs (106). An optimised RG should be uniformly expressed in all experimental conditions, in order to account for technical differences such as, RNA quantity, enzyme efficiency, sample collection and preparation (107, 108). According to the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE), unless a single RG has been fully validated, the use of at least 3 RGs is advised to avoid bias when interpreting results (109, 110).

Moreover, RGs should be validated by researchers using their experimental set up and determined by using a validated algorithm such as publicly available software packages: NormFinder (111), geNorm (112), BestKeeper (113), and comparative delta Ct method (114). In brief, NormFinder ranks candidate RGs according to their stability value, with anything below 0.15 considered to be stable (111). By calculating the average pairwise variation to other RGs, geNorm outputs a gene expression stability value of 'M' also with a cut-off point of 0.15 (112). BestKeeper, ranks RGs depending on their SD values within each gene (113). Finally, the comparative delta Ct method calculates the SD of each gene between each sample and ranks the genes with the lowest SD to be the most stable RG (114). As each algorithm is different, it is recommended that more than one software packages be used, which is provided by RefFinder: a free online tool which creates comprehensive stability rankings by combining all 4 algorithms (www.heartcure.com.au/reffinder/) (115).

The previously described IFN-Scores-A and -B and novel WB derived scores IFN-Score-C and -D have thus far been normalised to a single RG, PPIA (70). This was previously demonstrated as stably expressed and non-responsive to IFN- γ stimulation in a study of human keratinocytes cell lines (116). As the IFN pathway activation has diverse effects multiple cell types and gene expression can be sensitive to multiple subtypes of IFN specific

validation of RGs for use in ISG assays should be tailored to experimental conditions, sample type and disease status. Some of the most established and widely used RGs (*ACTB*, *GAPDH*, *HPRT*, *18s* and *B2M*) have shown inconsistent expression in different cell types or disease such as upregulation of *ACTB* in leukaemia tumour samples (103, 110, 117). As outlined in previous chapters, IFN-Scores-A and -B, as well as IFN-Scores-C and -D derived here in WB show promising potential for further development towards clinical application. It is therefore critical to now verify the appropriate RG for these assays and this approach aligns with the conclusions of EULAR task force on IFN assays introduced earlier in this report. In view of the various benefits of a WB analysis workstream as a clinical assay, this work focusses exclusively on WB.

5.2 Objectives

To determine the most stable RG to measure IFN-Scores in SLE and healthy WB samples using RefFinder.

5.3 Methods

5.3.1 Sample inclusion and sample processing

For this section of the report, 6 healthy control bloods were used and patient SLE samples were derived from the DEFINITION study (SLE inactive (n=7), SLE active (n=9) (total n=22). WB Tempus™ sample collection and RNA extraction was carried out as per section 2.3.

5.3.2 Gene transcript quantification

16 candidate RGs were selected from existing ISG related literature summarised by the EULAR task force PICO1 systematic literature review (68). Table 12 describes the candidate RGs and their functions with the references to the literature. As well as the 16 RGs, the 31 ISGs used in the derivation of IFN-Scores (Table 9) gene transcripts were determined by TaqMan Fluidigm, as described in section 2.4. IFN-Score are represented as reflected scores (score multiplied by -1) so that numerically higher scores represent higher expression, while preserving the original distribution.

5.3.3 Exclusion of 18s

From the 16 candidate RG evaluated, 18s was eliminated from further analysis due to excessive abundance of the gene in all samples with an average Ct value of 2.5. 18s is an rRNA which can make up the bulk of extracted RNA and leading to under representation of overall RNA (118). Moreover, RGs should be of similar Ct values to gene of interest which ensures validation of the same kinetic interactions during qPCR, therefore (118)18s was excluded in evaluation for stable RG software.

5.3.4 Exploring reference genes

Exploration of appropriate RGs were explored using the publicly available, free to access online software; RefFinder (115). Raw Ct values of SLE and HCs were imported together into the software, and rankings of gene stability were outputted. The RefFinder software calculated a comprehensive gene stability weight using a combination of 4 algorithms from other RG software: geNorm, NormFinder, BestKeeper, and the comparative Delta-Ct method. Outputs of all algorithm rankings are shown in the results and the top three most stable genes from RefFinder's comprehensive gene stability rankings were used in further validation.

5.3.5 Calculation of IFN-Scores

Delta Ct values of ISG transcripts were determined by normalising to i) the geometric mean of the top three most stable reference gene, ii) previously used *PPIA* and iii) the least stable reference gene. IFN-Scores -A, -B, -C and -D were then calculated according to the delta Ct values to each normalisation method. Table 11 represents the ISG belonging to each score.

5.3.6 Statistical analysis

To compare the effect of normalising ISGs to different RGs, independent t-tests were calculated between SLE and healthy patient groups. Prior to this, Levene's test of homogeneity was violated and therefore equal variances were not assumed.

Pearson's correlation calculated the correlation coefficient (R) to determine the strength of the relationships comparing IFN-Scores normalised to different RGs, where $r=1$ / -1 is perfect positive / negative correlation and $r=0$ is no correlation between sample types. Details of samples and comparisons are noted in respective figure legends.

Table 12. List of Reference Gene TaqMan Assays: Abbreviations, full name, function, TaqMan assay ID and a literature reference of where the gene was used as a reference gene. Information sourced from Genecards.org (119). * HMBS and IPO8 were not found as reference genes in the SLR but are commonly used reference genes in other studies and were therefore included.

| Abbreviation | Name | Function | TaqMan ID | Reference |
|---------------|--|--|---------------|---|
| 18S | 18S ribosomal RNA | Small ribosomal subunit, translation | Hs03003631_g1 | (120) (121) (122) (123) (124) (125) |
| ACTB | β -actin | Cytoskeleton | Hs99999903_m1 | (126) |
| B2M | β -2-microglobulin | Subunit of MHC class I, antigen presentation | Hs00187842_m1 | (124) (126) |
| GAPDH | Glyceraldehyde-3-phosphate dehydrogenase | Glycolysis | Hs99999905_m1 | (127) (128) (129) (122) (130) (131) (124) (126) |
| GUSB | β -glucuronidase | Breakdown of mucopolysaccharides | Hs99999908_m1 | (132) (133) (134) |
| HMBS* | Hydroxymethylbilane synthase (Porphobilinogen deaminase) | Porphyrin metabolism | Hs00609296_g1 | (135) |
| HPRT1 | Hypoxanthine guanine phosphoribosyl transferase | Generation of purine nucleotides | Hs99999909_m1 | (125) (136) (137) (138) (52) |
| IPO8* | Importin-8 | Nuclear protein import | Hs00183533_m1 | (139) |
| PGK1 | Phosphoglycerate kinase-1 | Glycolysis | Hs00943178_g1 | (140) |
| POLR2A | RNA polymerase II, subunit A | Transcription | Hs00172187_m1 | (141) |
| PPIA | Peptidyl proline isomerase A (Cyclophilin A) | Protein folding | Hs99999904_m1 | (70) |
| RPLP0 | Large ribosomal protein P0 | Translation | Hs00420895_gH | (142) (143) (144) (145) |
| TBP | TATA box binding protein | Transcription initiation | Hs00427620_m1 | (141) (146) |
| TFRC | Transferrin receptor | Endocytosis of iron | Hs00951083_m1 | (147) (124) (148) |
| UBC | Ubiquitin C | Protein degradation | Hs00824723_m1 | (126) |

| | | | | |
|--------------|---|---------------------|---------------|-------|
| YWHAZ | Tyrosine 3-Monooxygenase/Tryptophan 5-Monooxygenase | Signal transduction | Hs01122445_g1 | (126) |
|--------------|---|---------------------|---------------|-------|

5.4 Results

5.4.1 Algorithm based reference genes stability

To determine the three most stable RGs when measuring ISGs in WB SLE and healthy donors, the Ct values of 15 RGs were inputted into RefFinder, and stability values were determined from each algorithm (Figure 9).

RefFinder comprehensive ranking determined *YWHAZ*, *PGK1* and *GUSB* to be the top three most stable RGs with previously used *PPIA* to be 10th stable and *ACTB* to be the least stable. Consensus of the top ranked genes, *YWHAZ* and *PGK1* were found to be 1st or 2nd in all algorithms, with the exception of BestKeeper, where *YWHAZ* is 4th ranked and *PGK1* is 7th ranked. *GUSB* is 3rd ranked in the RefFinder, BestKeeper and delta Ct algorithm however is ranked 5th in NormFinder and 7th in geNorm. Across all the different algorithms, there was agreement that *ACTB* is the least stable RG in the experimental conditions evaluated. *PPIA* displayed mid-stability ranking in all algorithms, identifying that the gene is neither the most nor least stable to use in the evaluation of ISG expression across healthy subjects and SLE patients in WB.

For the NormFinder and geNorm algorithms, a threshold of 0.15 or below identifies RGs to be stable. In analysis of SLE and healthy control WB ISG expression *YWHAZ*, *PGK1* and *GAPDH* achieved stability values below the cut off in the NormFinder algorithm but none of the genes are below the cut off in the geNorm algorithm. As no current single algorithm-based approach is considered definitive for rating stability among candidate RGs the consensus output across algorithms generated indicated the top three most stable RGs as *YWHAZ*, *PGK1* and *GUSB* (identified as 'SLE Ref' in Figure 11).

5.4.2 Comparing expression across candidate reference genes

To visualise algorithm ranking strategy, reflected Ct values were mapped for the three most stable genes: *YWHAZ*, *PGK1*, *GUSB*, previously used RG: *PPIA* and the least stable RG: *ACTB*, comparing healthy to SLE (Figure 10). A uniform mean expression Ct level is demonstrated between healthy and SLE patient groups for *YWHAZ* (Healthy mean=-8.058, (0.435), SLE mean=-8.864, (0.503)), *PGK1* (Healthy mean=-9.907, (0.313), SLE mean=-10.721, (0.666)), *GUSB* (Healthy mean=-11.208, (0.320) SLE mean=-11.638, (0.570)) and *PPIA* (Healthy mean=-8.191, (0.497), SLE mean=-8.694, (0.756)). Consistent with its lower stability ranking, *ACTB* displayed highly variable expression levels both within and between the two groups with Ct values ranging from an average of -9.823 Ct, (3.67) in healthy and an average of -11.946 Ct, (3.548) in SLE (Figure 10). The range of Ct values among candidate RGs demonstrates how bias could be readily introduced by insufficient RG stability.

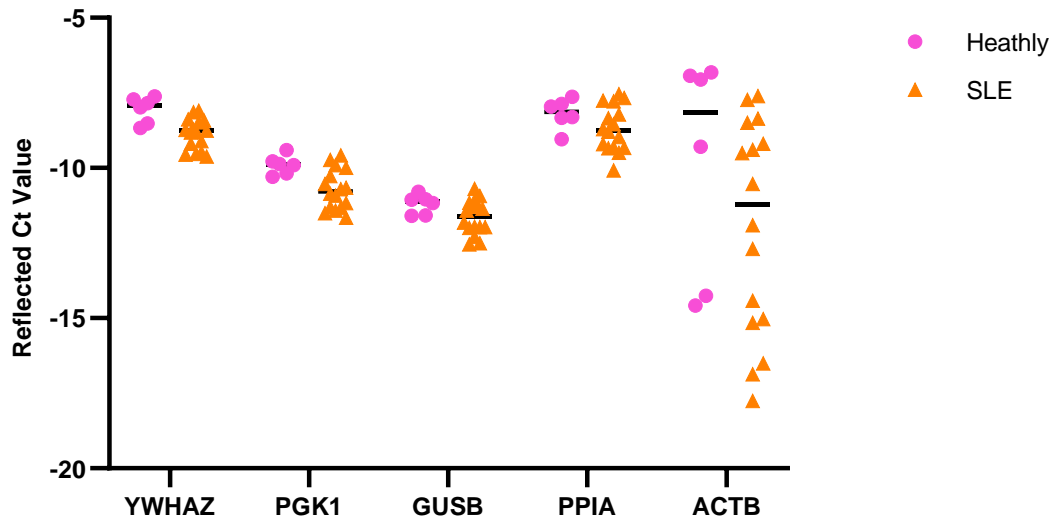


Figure 10. Reflected Ct values of top three most stable genes, previously used PPIA and least stable reference gene ACTB. Reflected Ct Values are demonstrated in order to represent larger values as higher expression and smaller values as lower expression. Healthy controls (circles) and SLE (triangle) are represented separately to indicate differences of reference gene between diagnostic groups and the mean value within groups is represented by a solid line.

5.4.3 Evaluating the effects of normalising IFN-Scores to different reference genes

To evaluate the impact of RG selection on calculated IFN-Scores, the ISGs were normalised first to the geometric mean of top three most stable RGs: *YWHAZ*, *PGK1* and *GAPDH*; next to the previously used *PPIA*; and lastly to the least stable RG, *ACTB*. SLE samples were compared to healthy donors to validate how different normalisation approaches can impact clinical IFN-Score results.

At first glance, a wider spread of data is seen in the delta Ct values *ACTB* compared to normalisation to the top three most stable RG, suggesting that normalisation to difference RGs introduces errors in the estimation of expression (Figure 11). Highly significant differences ($p < 0.0001$) were found when comparing healthy to SLE in all four IFN-Scores when normalising to the three most stable RGs and *PPIA*. However, lower significance (IFN-Score-A: $p = 0.02$, IFN-Score-C $p = 0.016$) or no significance (IFN-Score-B: $p = 0.167$, IFN-Score-D: $p = 0.157$) was found between the two groups when normalising to *ACTB*, which demonstrates the importance of selecting a validated RG (Figure 11).

To further assess how RGs can affect clinical comparisons, IFN-Scores of SLE samples and healthy donors were combined, and the different normalisation methods were compared (Figure 11). In all four IFN-Scores, strong positive correlations are observed between normalisation to the top three most stable RGs and *PPIA* (IFN-Score-A: $r = 0.98$; IFN-Score-B: $r = 0.85$; IFN-Score-C: $r = -0.97$; IFN-Score-D: $r = 0.86$, all with $p < 0.0001$). As expected, weaker correlations are found between normalisation to the top three most stable RGs and *ACTB* (IFN-Score-A: $r = 0.70$, $p = 0.0003$; IFN-Score-B: $r = 0.41$, $p = 0.585$; IFN-Score-C: $r = 0.57$, $p = 0.0054$; IFN-Score-D: $r = 0.42$, $p = 0.484$). These results suggest that normalisation of IFN-Scores to the top three most stable RG and *PPIA* produce similar results when calculated IFN-Scores.

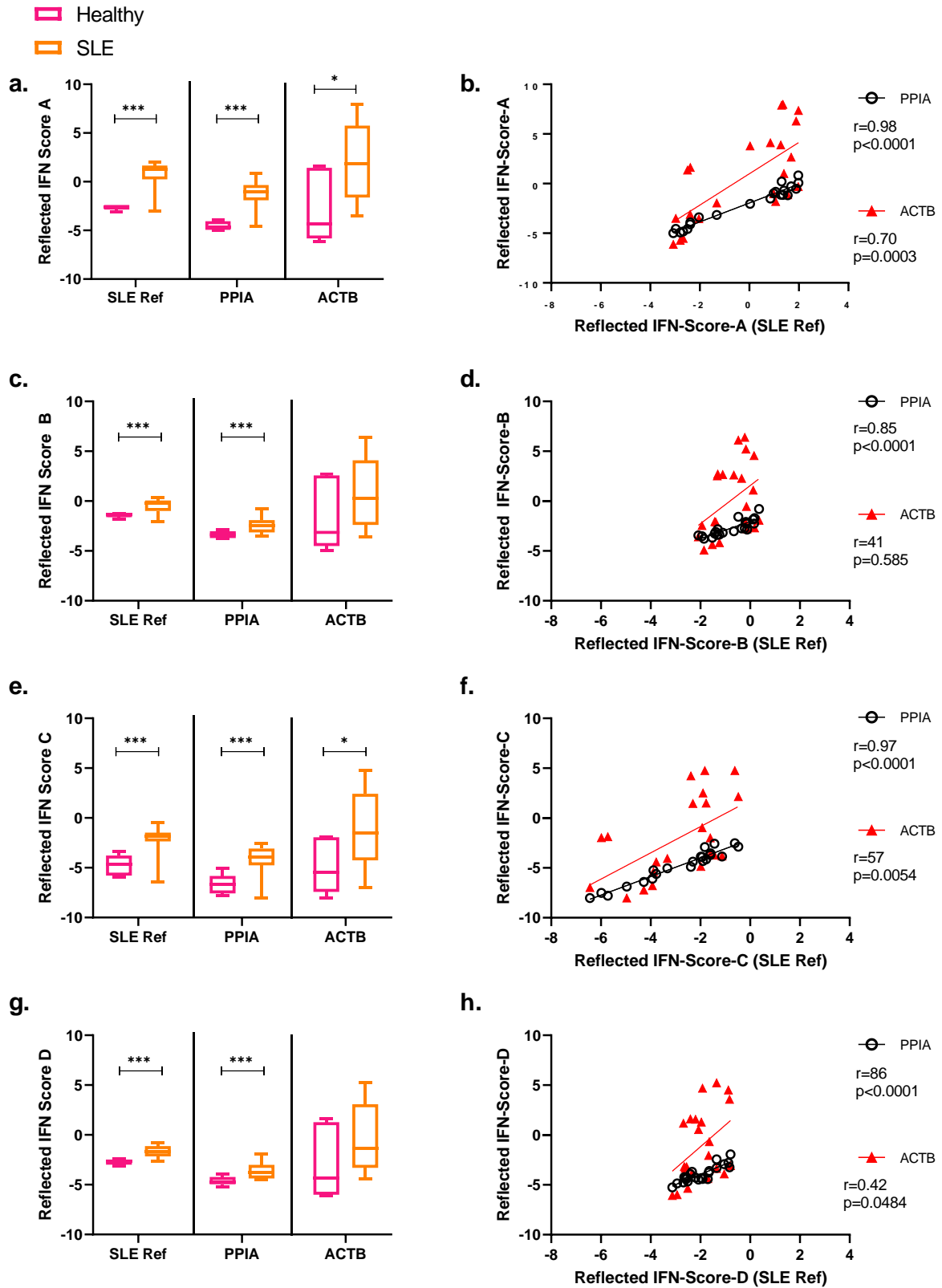


Figure 11. Evaluating the effects of normalising IFN-Scores to different reference genes in whole blood patient SLE samples and healthy controls. Box and whisker plots display median, interquartile ranges and minimum and maximum values of whole blood SLE ($n=16$) and controls

(n=6). Correlation of IFN-Scores when normalised to the IFN reference genes (*YWHAZ*, *PGK1*, *GUSB*) Vs *PPIA* and *ACTB*, includes all SLE and control samples. Independent t-test with unequal variance assumed and Pearson's correlation coefficient *r* and *p* values are represented for each plot. *** $p \leq 0.0001$, ** $p \leq 0.01$, * $p \leq 0.05$,

5.5 Discussion

Key results:

RT-PCR is a common way to measure and quantify gene expression and RGs are used to normalise quantification and compensates for technical differences such as, RNA quantity, enzyme efficiency, sample collection and preparation (107, 108). RGs should be uniformly expressed throughout all experimental conditions and MIQE guidelines suggest the use of three RGs, validated by an available software to minimise introduction of bias (109, 110).

Previously, publications of IFN-Score-A and -B in PBMCs have adopted *PPIA* as a RG, and prior work in the Leeds Lupus group has also made use of this in WB sample type (70, 74-77). While literature suggests that *PPIA* it is not regulated by IFN, it has so far lacked comprehensive validation as a RG for ISG expression studies, especially in blood. Although previous publications of IFN-Score-A and -B showed encouraging clinically significant results, a focused evaluation of performance of *PPIA* among other candidate RGs is essential to progress the development of IFN-Scores in with MIQE guidelines and the standard set out by the EULAR task force. Specifically, these state that RGs must be validated for each experimental conditions and at least three RGs should be used (110).

This work presents a detailed evaluation of 16 candidate RGs in a healthy and SLE WB cohort. *18s* was excluded from further analysis due to the abundance of the gene in each sample and was therefore not a suitable RG. The RefFinder algorithm ranked *YWHAZ*, *PGK1* and *GUSB* as the top three most stable RGs with *PPIA* 10th stable and *ACTB* least stable. Quantification of the leading candidate RGs in healthy and SLE WB show:

1. Normalisation of ISGs to *ACTB*, showed lower significance or even no significant difference between SLE and healthy and IFN-Scores were more dispersed within patient groups compared to normalising to *PPIA* and *YWHAZ*, *PGK1* and *GUSB*.
2. Normalisation of ISGs to *PPIA* showed similar significance between SLE and healthy compared to using the geometric mean of *YWHAZ*, *PGK1* and *GUSB*.

3. Correlation between normalising ISGs to *ACTB* and *YWHAZ*, *PGK1* and *GUSB* showed weak positive relationship and a strong positive correlation between *PPIA* and *YWHAZ*, *PGK1* and *GUSB*.

This work indicates promise that previous calculations of IFN-Scores with *PPIA* as a RG are still valid, however, RGs must be validated and the geometric mean of *YWHAZ*, *PGK1* and *GUSB* should be used to measure the ISGs reported here in a healthy and SLE WB cohort.

Context with other data:

It is interesting to note how some of the most commonly reported RGs *18s*, *GAPDH* and *ACTB* did not prove to be stable in this cohort (120-131). However, *GAPDH* could still be counted as stable in this cohort considering ranked 4th in the RefFinder algorithm. *18s* is an rRNA, and the quantity of *18s* was far too abundant to produce reliable results in this cohort (118). This is probably due to RNA extraction as the kit extracts total RNA from large mRNA, rRNA down to microRNA (miRNA) and small interfering RNA (siRNA).

Although the literature states the use of *18s* as a RG when measuring ISGs in autoimmunity, none of these publications justify the use of this gene (120-125). Furthermore, *ACTB* was the most stand out instable RG, which is an alarming finding since it is a commonly used RG in many RT-PCR studies (105). *ACTB* encodes for the β - actin protein, it is ubiquitously expressed and highly conserved. It's function is involved in cellular structure and motility, making it a common RG to use in experiments (119). However, *ACTB* expression has been reported to be dysregulated in several cancers and not suitable to use as a RG (149), therefore it could be assumed cell motility and cytoskeletal alterations are part of SLE, again reiterating the importance of validating a RG in every experiment.

Limitations:

A limitation of this study is that IFN-Scores -A, -B, -C and -D were designed with *PPIA* as the RG. As *YWHAZ*, *PGK1* and *GUSB* have now been established as the most stable RGs to normalise WB SLE and healthy samples, it would be unfeasible to re-run gene expression analysis with these RGs. Almost 400 WB samples have quantified ISG expression using *PPIA* as a RG and if this analysis was to be repeated with the 3 most stable RGs, it would come at a high cost and replicate unnecessary data. As a RG, *PPIA* has demonstrated the same clinical outcomes as newly verified top three most stable RGs *YWHAZ*, *PGK1* and *GUSB* and therefore reliability of previously calculated IFN-Scores normalised to *PPIA* can be assumed. However, to ensure guidelines are followed, it is now recommended to calculate WB IFN-Scores with the geometric mean of the newly validated RGs *YWHAZ*, *PGK1* and *GUSB*.

Future work:

Future work will utilise *YWHAZ*, *PGK1* and *GUSB* as RGs in WB SLE cohorts.

Conclusion:

RGs should be validated for every experimental design and MIQE guidelines should be followed. This does not appear to have been done for the majority of published papers reporting ISG assays. Choices of RGs may significantly influence the results of ISG-based assays. Notably, two of the RGs that showed to be unsuitable, are common in the published literature. *18s* is too abundant, and *ACTB* is too inconsistent in expression. Research into ISG expression in WB would benefit from using the most suitable genes identified here: *YWHAZ*, *PGK1*, *GUSB*.

Chapter 6: Discussion

Outline of report

The clinical heterogeneity of SLE and the increasing therapeutic options available for the disease place effective biomarker development as an increasing research priority. IFN pathway activation is recognised as a key component to SLE pathology, but ensuring validation and consistency, the most appropriate IFN assay should be justified (78). A two-score system for measuring ISG expression in PBMCs has been shown by the Leeds Lupus group to stratify autoimmune disease and predict disease progression in ANA positive individuals which builds upon identifying and addressing several key obstacles to applying these IFN-Scores in clinical settings (70, 74). The current work highlights the lack of transferability of ISG expression scores between PBMC and WB sample types, devises a novel two score system for WB samples drawn from the same elements of the SLE modular transcriptome and systematically identifies optimal RGs for use in this analysis.

Key results

Development of IFN-Scores: the impact of sample type

Previously, the Leeds Lupus group developed IFN-Score-A and -B based on a PBMC cohort (70, 74). However, processing PBMCs is time consuming, therefore a quicker, replicable method such as WB processing needs to be developed in order to reach clinical practice. This report begins to justify measurement of IFN-I pathway in WB sample type based on theoretical, experimental feasibility and clinical evidence.

Firstly, this work reveals the marked lack of transferability of IFN-Score across PBMC and WB sample types which vary in cellular composition. There was a striking lack of correlation between IFN-Scores measured in WB and PBMC samples. Notably this was not uniform for all ISGs. A better correlation was retained between sample types for IFN-Score-A than for IFN-Score-B, highlighting the complexity of ISG regulation. It is not clear what accounts for this difference; however, it could indicate that IFN-Score-A ISGs may be more stably expressed across cell populations than others.

A possible explanation of this could be IFN related epigenomic remodelling that sustains ISG regulation (150). The epigenome is regulated in response to external stimuli and alters gene expression without changes to DNA sequences. Instead, modifications including DNA methylation, histone modifications, chromatin accessibility can influence the regulation of effected genes. It has been reported that IFN induction of IFN regulatory factors mediate chromatin remodelling at ISG loci, making transcription of genes more accessible which can last for days or weeks even after IFN production has subsided therefore creating a sustained

expression of ISGs (39). Moreover, IFN-Score-A / M1.2 genes *IFIT1*, *IFI44L* and *RSAD2* were found to be hypomethylated in ANA positive compared to ANA negative in a study of 326 European females (150, 151), supporting the standardised regulation of M1.2 ISGs over time (73).

Existing data indicated IFN-Score-B in PBMCs to show more clinically useful associations and this may be underpinned by its more dynamic regulation across sample types and disease groups.

Development of WB IFN-Scores

To address the evident difficulties in applying existing IFN-Scores to WB. This work undertook factor analysis of 31 genes in a cohort of 395 WB samples including SLE, RA and healthy patients, to reveal a new two scoring system: IFN-Score-C (*CCL8*, *CXCL10*, *IFI27*, *ISG15* and *LAMP3*) and IFN-Score-D (*CASP1*, *CEACAM1*, *SOCS1* and *TRIM38*).

Interestingly IFN-Score-C shared similarities with IFN-Score-A and IFN-Score-D shared somewhat similar properties with IFN-Score-B. However, similarities did not appear to simply reflect the degree of overlap in gene composition or modular origin. Both new IFN-Scores showed weak relationships between PBMC and WB match samples, showing that IFN-Score-C and -D are more specific for WB and factor analysis was possibly driven by the specific cellular ISG expression such as neutrophils, basophils and eosinophils. However, when testing the clinical utility of all IFN-Scores in WB, IFN-Score-A was strongest when stratifying patient diagnosis, showing large effect sizes between sample groups, SLE, RA and healthy. Unfortunately, the performance of each IFN-Score by effect size across disease groups could not be directly compared in this work between PBMC and WB sample types as disease group samples for PBMCs were not sufficiently available for all groups. Thus, it remains to be determined whether IFN-Score-A measured in WB could outperform the originator score in PBMCs in terms of differentiating disease.

Addressing technical considerations for IFN assays

To comply with MIQE guidelines, RGs for qPCR based studies should be validated in all experimental conditions, and it is recommended to use at least three to normalise qPCR data (109, 110). The work presented here provides key evidence to support RG selection in the ongoing development and validation of IFN-Scores in WB for studies in SLE.

16 RGs were selected from the existing literature that are used to normalise ISG transcripts. Publicly available software, RefFinder ranked all RGs in order of stability, calculating *YWHAZ*, *PGK1* and *GUSB* to be the most stable in WB SLE. Interestingly, a commonly used RG, β -actin (*ACTB*) was ranked least stable which could suggest it is partially responsive to

IFN or other disease mediators in SLE. Comparison of normalising IFN-Scores to the three most stable RGs and *PPIA* reassuringly generated similar results. The current work however supports the use of *YWHAZ*, *PGK1* and *GUSB* when quantifying IFN-Scores in WB and future work should apply this style of analysis to WB based assays, particularly if considered for clinical development (152).

Context with other data and future work

The current work progresses the development of measuring IFN-signature in WB in line with current EULAR task force recommendations, and the findings could inform the wider field of IFN pathway quantification. Beyond practical relevance, the current work also further reveals the complexity of IFN dysregulation.

Understanding ISG regulation

IFNs signal and activate many different ISGs and cytokines and different types of IFN pathways overlap (34-39) and one of the EULAR points to consider is to determine stimuli of reported IFN-signatures (78).

3/5 genes in the newly described WB IFN-Score-C originate from IFN module M1.2, with 1/5 from M3.4 and 1/5 unclassified ISG. It may therefore be speculated that this IFN-Score is mainly upregulated by the presence of IFN-I and remains stable over time which was previously identified for M1.2 as a whole (73). IFN-Score-D includes 2 ISGs from M3.4 and 2 ISGs from 5.12 which were both described to be regulated by both IFN-I and IFN-II and so this score is likely to show distinct regulation to SLE disease activity (73).

Principal component analysis in a longitudinal study of SLE patient and healthy controls clustered 46 ISGs into 3 subsets. The first subset consisted of 36 commonly reported ISGs predominantly responding to IFN-I, the second subset consisted of underrepresented ISGs: *SERPING1*, *PARP9*, *CXCL10*, *SOCS1*, *C1QB* and *PDCD1CGS* and responding to IFN-I and IFN-II and the third subset, also underrepresented ISG: *S100A9*, *FCGRIA*, *S100A8*, *SOCS3* predominately respond to IFN-II. Only the third subset correlated with disease activity, suggesting that IFN-II response genes are a better solution for SLE biomarkers (153).

Alternatively, another approach to understanding the regulation of these genes is to test their expression under conditions of IFNAR blockade. In clinical trials of IFN-receptor blocking agent, Anifrolumab, an IFN-signature including ISGs, *IFI44*, *IFI44L*, *RSAD2* and *IFI27* was shown to be effectively suppressed by treatment. These ISGs fall predominately within M1.2 genes known to respond specifically to IFN-I and individuals with an elevated IFN-signature at baseline showed improved responses to Anifrolumab (57, 154).

The responsiveness of these transcripts has not yet been interrogated in experimental conditions, and it is possible that different regulatory mechanisms apply in different cell types, adding to the effects seen in comparison between WB and PBMCs. Work is already underway to test expression in response to IFN and other inflammatory cytokines both alone and in combination, in an effort to unpick the precise regulation of these ISGs in the context of SLE.

Understanding the expression of IFN-Score-A, -B, -C and -D following Anifrolumab treatment could help to reveal further similarities and differences between scores and assist their evaluation as clinical biomarkers.

Clinical associations

In this report, clinical relevance was determined by comparing diagnostic groups SLE, healthy and RA. It showed IFN-Score-A to have large effect sizes between all patient groups, IFN-Score-B and -D had large effect sizes between SLE Vs. healthy and RA Vs healthy but failed to distinguish between SLE and RA and IFN-Score-C had small effects sizes between all groups. However, sample sizes of the healthy and RA were a limitation in this study.

Previously in a PBMC sample type, IFN-Score-B was able to differentiate SLE and RA, and more importantly, predict progression to SLE in ANA-positive At-Risk individuals (70, 74). Therefore, IFN-Score-B was deemed the most promising for progression toward clinical practice. The ability of the novel IFN-Scores-C and -D, as well as existing IFN-Score-A and -B, now measured in WB, to distinguish clinically important end points remains to be determined and could ultimately guide the selection of candidate biomarkers for ongoing work.

It is possible that different selection of IFN-Scores will be required in different clinical contexts, for example in diagnostic use healthy Vs autoimmune disease compared with stratification between diseases. Moreover, interpretation of IFN-Scores alongside other cell type signatures in combination potentially will yield more useful clinical information (102, 152) (Carter et al. *In press*).

Experimental feasibility:

The IFN-pathway is large and complex therefore, there is no single assay that can evaluate the entirety of IFN pathway activation. The most commonly reported IFN biomarkers are measuring i) circulating IFN proteins by single-molecule array (SiMoA), ii) measuring IFN-stimulated proteins encoded by ISGs, iii) measuring cell-specific IFN response via cell specific markers and iv) ISG quantification. The latter method, ISG quantification via qPCR,

has been the most successful to quantify the IFN-pathway which this report aims to develop further (68).

Experimental feasibility begins with sample collection. Blood is easily accessible and can be drawn by any trained professional with minimal side effects for the patient. Working with WB is much more flexible compared to PBMCs and allows for options such as, room temperature sample transportation to other sites, whereas PBMCs would have to be transported on dry ice, incurring extra costs and risks.

It is important to note that many of the literature report IFN-signature utilise PAXgene™ blood collection methods (122, 123, 126-128, 130, 131, 142). PAXgene™ Blood RNA System (PreAnalytiX QIAGEN) is a similar system to Tempus™ in that they are both blood tubes that contain RNA stabilisers and ensure gene expression profiles that reflect the blood's state at the moment of sampling (155). Studies have shown that blood from the same individuals sampled into the PAXgene™ and Tempus™ tubes resulted in significantly different expression profiles of more than 2000 genes, hence the two WB collection methods cannot be used interchangeably (156). Therefore, it should be emphasised that the evaluation of the novel IFN-Scores-C and -D here is specific to using Tempus™ tube collection methods.

IFN-Scores have been quantified using the Fluidigm, Biomark platform: a high-throughput qPCR System that uses integrated fluidic circuits known as dynamic arrays and digital arrays. This innovative method allows researchers to gain data on 96 transcripts whilst only using minimal RNA material (1-3ul), consumables such as pipette tips and reagents like RT and Pre-amplification master mix. If the same amount of data were produced on 384 single gene-based qPCR plates, researchers would have to run 24 plates, therefore the Fluidigm system not only saves on materials but also save time in that whole experiment can be completed in 1 day. Moreover, TaqMan assay combinations are customisable allowing for development and variation of transcript quantification.

Development of IFN-Scores-C and -D has led to a smaller number of target genes compared to IFN-Score-A and -B and thus should incur cost saving on laboratory consumables. Specifically, previously, IFN-Score-A and -B would need to quantify 27 transcripts (12 for IFN-Score-A, 14 for IFN-Score-B and 1 RG (*PPIA*)) whereas newly developed IFN-Scores requires quantification of 12 transcripts (5 for IFN-Score-C, 4 for IFN-Score-D and 3 RGs (*YWHAZ*, *PGK1* and *GAPDH*)), less than half of the previous IFN-Scores. Therefore, running high through-put Fluidigm might be unnecessary. Other methods such as customised TaqMan® Array Custom Micro Fluidic Cards might ultimately be more suitable and can range from 12 to 384 assays and can run 1 to 8 samples. Pending further validation and

clinical evaluation, the development of the WB two-score system based on this limited gene set could be a substantial advance towards use as biomarkers in routine clinical practice.

Limitations

A limitation of this study is that accountability for age, sex, ethnicity, disease activity could indicate specific influences of ISG activity and inclusion of a more diverse range of diseases such as Sjogren's syndrome in factor analysis, could indicate other underlying causes of ISG expression which will be analysed in future projects.

The influence of WB IFN-Score-C and -D has not been determined. IFN-Score-D had the most promise in terms of becoming the new gold standard in SLE biomarker, considering its strong relationship with clinically relevant IFN-Score-B. However, the use of this is yet to be determined and needs to be studied closer. For example, using new RGs, comparison to drug response, new cohort with different clinical end point. Furthermore, variability of ISG expression in certain cell types has not been determined which could explained differences between sample types and help understanding of IFN-Score clustering.

Validation of RG stability was not determined before data analysis in chapters 3 and 4 were analysed using RG *PPIA*. However, as a RG, *PPIA* demonstrated the same clinical outcomes as newly verified most stable RGs *YWHAZ*, *PGK1* and *GUSB* and therefore reliability of previously calculated IFN-Scores normalised to *PPIA* can be assume.

Conclusion

This work adds to a body of emerging data indicating the complexity of ISG regulation in SLE and contributes to development of ISG expression scores towards clinical application. Previously denoted IFN-Score-A could potentially be interchanged between sample types and has proven to be useful to distinguish between SLE, RA and healthy in WB. However, for clinical utility, biomarkers must be able to answer significant clinical questions in real-world populations. For example, predicting response to treatment or disease progression in populations before these outcomes are known. So far, IFN-Score-B in PBMCs has shown greater promise for such questions. However, to move towards more feasible sample processing, establishment of a WB biomarker is desirable. The lack of agreement between IFN-Score-B between sample types called for reinterrogation of ISGs which creates two new WB scores IFN-Score-C and -D. IFN-Score-A shared overlapping genes with IFN-Score-C but WB calculations in the scores showed weak correlations. WB calculated IFN-Score-D showed strong correlation with IFN-Score-B suggesting it could act as a surrogate WB measurement of IFN-Score-B, and further work is justified to confirm whether it has the same, clinical utility. Furthermore, some of the RGs in the literature are shown to be

unsuitable for IFN assays. This work showed that WB IFN-Scores should utilise *YWHAZ*, *PGK1* and *GUSB* as RGs when quantifying ISGs.

References

1. Kaul, A., Gordon, C., Crow, M.K., Touma, Z., Urowitz, M.B., van Vollenhoven, R., Ruiz-Irastorza, G. and Hughes, G. Systemic lupus erythematosus. *Nature Reviews Disease Primers*. 2016, **2**(1), p.16039.
2. Wandstrat, A.E., Carr-Johnson, F., Branch, V., Gray, H., Fairhurst, A.-M., Reimold, A., Karp, D., Wakeland, E.K. and Olsen, N.J. Autoantibody profiling to identify individuals at risk for systemic lupus erythematosus. *Journal of Autoimmunity*. 2006, **27**(3), pp.153-160.
3. Barber, M.R.W., Drenkard, C., Falasinnu, T., Hoi, A., Mak, A., Kow, N.Y., Svenungsson, E., Peterson, J., Clarke, A.E. and Ramsey-Goldman, R. Global epidemiology of systemic lupus erythematosus. *Nat Rev Rheumatol*. 2021, **17**(9), pp.515-532.
4. Fanouriakis, A., Kostopoulou, M., Alunno, A., Aringer, M., Bajema, I., Boletis, J.N., Cervera, R., Doria, A., Gordon, C., Govoni, M., Houssiau, F., Jayne, D., Kouloumas, M., Kuhn, A., Larsen, J.L., Lerstrøm, K., Moroni, G., Mosca, M., Schneider, M., Smolen, J.S., Svenungsson, E., Tesar, V., Tincani, A., Troldborg, A., van Vollenhoven, R., Wenzel, J., Bertsias, G. and Boumpas, D.T. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. *Ann Rheum Dis*. 2019, **78**(6), pp.736-745.
5. Schmajuk, G., Hoyer, B.F., Aringer, M., Johnson, S.R., Daikh, D.I. and Dörner, T. Multicenter Delphi Exercise to Identify Important Key Items for Classifying Systemic Lupus Erythematosus. *Arthritis Care Res (Hoboken)*. 2018, **70**(10), pp.1488-1494.
6. Aringer, M., Costenbader, K., Daikh, D., Brinks, R., Mosca, M., Ramsey-Goldman, R., Smolen, J.S., Wofsy, D., Boumpas, D.T., Kamen, D.L., Jayne, D., Cervera, R., Costedoat-Chalumeau, N., Diamond, B., Gladman, D.D., Hahn, B., Hiepe, F., Jacobsen, S., Khanna, D., Lerstrøm, K., Massarotti, E., McCune, J., Ruiz-Irastorza, G., Sanchez-Guerrero, J., Schneider, M., Urowitz, M., Bertsias, G., Hoyer, B.F., Leuchten, N., Tani, C., Tedeschi, S.K., Touma, Z., Schmajuk, G., Anic, B., Assan, F., Chan, T.M., Clarke, A.E., Crow, M.K., Cziráj, L., Doria, A., Graninger, W., Halda-Kiss, B., Hasni, S., Izmirly, P.M., Jung, M., Kumánovics, G., Mariette, X., Padjen, I., Pego-Reigosa, J.M., Romero-Diaz, J., Rúa-Figueroa Fernández, Í., Seror, R., Stummvoll, G.H., Tanaka, Y., Tektonidou, M.G., Vasconcelos, C., Vital, E.M., Wallace, D.J., Yavuz, S., Meroni, P.L., Fritzler, M.J., Naden, R., Dörner, T. and Johnson, S.R. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. *Ann Rheum Dis*. 2019, **78**(9), p.1151.
7. Navarra, S.V., Guzmán, R.M., Gallacher, A.E., Hall, S., Levy, R.A., Jimenez, R.E., Li, E.K., Thomas, M., Kim, H.Y., León, M.G., Tanasescu, C., Nasonov, E., Lan, J.L., Pineda, L., Zhong, Z.J., Freimuth, W. and Petri, M.A. Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. *Lancet*. 2011, **377**(9767), pp.721-731.
8. Pearce, F.A., Rutter, M., Sandhu, R., Batten, R.L., Garner, R., Little, J., Narayan, N., Sharp, C.A., Bruce, I.N., Erb, N., Griffiths, B., Guest, H., Macphie, E., Packham, J., Hiley, C., Obrenovic, K., Rivett, A., Gordon, C. and Lanyon, P.C. BSR guideline on the management of adults with systemic lupus erythematosus (SLE) 2018: baseline multi-centre audit in the UK. *Rheumatology*. 2020, **60**(3), pp.1480-1490.

9. Smith, M.R. Rituximab (monoclonal anti-CD20 antibody): mechanisms of action and resistance. *Oncogene*. 2003, **22**(47), pp.7359-7368.
10. Ulf-Møller, C.J., Simonsen, J., Kyvik, K.O., Jacobsen, S. and Frisch, M. Family history of systemic lupus erythematosus and risk of autoimmune disease: Nationwide Cohort Study in Denmark 1977–2013. *Rheumatology*. 2017, **56**(6), pp.957-964.
11. Grumet, F.C., Coukell, A., Bodmer, J.G., Bodmer, W.F. and McDevitt, H.O. Histocompatibility (HL-A) antigens associated with systemic lupus erythematosus. A possible genetic predisposition to disease. *N Engl J Med*. 1971, **285**(4), pp.193-196.
12. Waters, H., Konrad, P. and Walford, R.L. The distribution of HL-A histocompatibility factors and genes in patients with systemic lupus erythematosus. *Tissue Antigens*. 1971, **1**(2), pp.68-73.
13. Ghodke-Puranik, Y. and Niewold, T.B. Genetics of the type I interferon pathway in systemic lupus erythematosus. *Int J Clin Rheumatol*. 2013, **8**(6).
14. Barbhaiya, M. and Costenbader, K.H. Environmental exposures and the development of systemic lupus erythematosus. *Curr Opin Rheumatol*. 2016, **28**(5), pp.497-505.
15. Yaniv, G., Twig, G., Shor, D.B.-A., Furer, A., Sherer, Y., Mozes, O., Komisar, O., Slonimsky, E., Klang, E., Lotan, E., Welt, M., Marai, I., Shina, A., Amital, H. and Shoenfeld, Y. A volcanic explosion of autoantibodies in systemic lupus erythematosus: A diversity of 180 different antibodies found in SLE patients. *Autoimmunity Reviews*. 2015, **14**(1), pp.75-79.
16. Orteu, C.H., Sontheimer, R.D. and Dutz, J.P. The pathophysiology of photosensitivity in lupus erythematosus. *Photodermatol Photoimmunol Photomed*. 2001, **17**(3), pp.95-113.
17. Malkiel, S., Barlev, A.N., Atisha-Fregoso, Y., Suurmond, J. and Diamond, B. Plasma Cell Differentiation Pathways in Systemic Lupus Erythematosus. *Front Immunol*. 2018, **9**, p.427.
18. Yap, D.Y.H. and Chan, T.M. B Cell Abnormalities in Systemic Lupus Erythematosus and Lupus Nephritis-Role in Pathogenesis and Effect of Immunosuppressive Treatments. *Int J Mol Sci*. 2019, **20**(24).
19. Deng, G.-M. and Tsokos, G.C. Pathogenesis and targeted treatment of skin injury in SLE. *Nature Reviews Rheumatology*. 2015, **11**(11), pp.663-669.
20. Isaacs, A. and Lindenmann, J. Virus interference. I. The interferon. *Proc R Soc Lond B Biol Sci*. 1957, **147**(927), pp.258-267.
21. Hooks, J.J., Moutsopoulos, H.M., Geis, S.A., Stahl, N.I., Decker, J.L. and Notkins, A.L. Immune Interferon in the Circulation of Patients with Autoimmune Disease. *New England Journal of Medicine*. 1979, **301**(1), pp.5-8.
22. Pascual, V., Farkas, L. and Banchereau, J. Systemic lupus erythematosus: all roads lead to type I interferons. *Curr Opin Immunol*. 2006, **18**(6), pp.676-682.
23. Banchereau, J. and Pascual, V. Type I interferon in systemic lupus erythematosus and other autoimmune diseases. *Immunity*. 2006, **25**(3), pp.383-392.

24. Bronson, P.G., Chaivorapol, C., Ortmann, W., Behrens, T.W. and Graham, R.R. The genetics of type I interferon in systemic lupus erythematosus. *Curr Opin Immunol.* 2012, **24**(5), pp.530-537.
25. Niewold, T.B., Hua, J., Lehman, T.J., Harley, J.B. and Crow, M.K. High serum IFN-alpha activity is a heritable risk factor for systemic lupus erythematosus. *Genes Immun.* 2007, **8**(6), pp.492-502.
26. Panousis, N.I., Bertias, G.K., Ongen, H., Gergianaki, I., Tektonidou, M.G., Trachana, M., Romano-Palumbo, L., Bielser, D., Howald, C., Pamfil, C., Fanouriakis, A., Kosmara, D., Repa, A., Sidiropoulos, P., Dermitzakis, E.T. and Boumpas, D.T. Combined genetic and transcriptome analysis of patients with SLE: distinct, targetable signatures for susceptibility and severity. *Ann Rheum Dis.* 2019, **78**(8), pp.1079-1089.
27. Crow, Y.J., Black, D.N., Ali, M., Bond, J., Jackson, A.P., Lefson, M., Michaud, J., Roberts, E., Stephenson, J.B., Woods, C.G. and Lebon, P. Cree encephalitis is allelic with Aicardi-Goutières syndrome: implications for the pathogenesis of disorders of interferon alpha metabolism. *J Med Genet.* 2003, **40**(3), pp.183-187.
28. d'Angelo, D.M., Di Filippo, P., Breda, L. and Chiarelli, F. Type I Interferonopathies in Children: An Overview. *Frontiers in Pediatrics.* 2021, **9**.
29. Alperin, J.M., Ortiz-Fernández, L. and Sawalha, A.H. Monogenic Lupus: A Developing Paradigm of Disease. *Front Immunol.* 2018, **9**, p.2496.
30. Rönnblom, L. and Eloranta, M.-L. The interferon signature in autoimmune diseases. *Current Opinion in Rheumatology.* 2013, **25**(2), pp.248-253.
31. Morand, E.F., Furie, R.A., Bruce, I.N., Vital, E.M., Dall'Era, M., Maho, E., Pineda, L. and Tummala, R. Efficacy of anifrolumab across organ domains in patients with moderate-to-severe systemic lupus erythematosus: a post-hoc analysis of pooled data from the TULIP-1 and TULIP-2 trials. *The Lancet Rheumatology.* 2022, **4**(4), pp.e282-e292.
32. Greenberg, S.A., Pinkus, J.L., Pinkus, G.S., Burtleson, T., Sanoudou, D., Tawil, R., Barohn, R.J., Saperstein, D.S., Briemberg, H.R., Ericsson, M., Park, P. and Amato, A.A. Interferon-alpha/beta-mediated innate immune mechanisms in dermatomyositis. *Ann Neurol.* 2005, **57**(5), pp.664-678.
33. Gottenberg, J.-E., Cagnard, N., Lucchesi, C., Letourneur, F., Mistou, S., Lazure, T., Jacques, S., Ba, N., Ittah, M., Lepajolec, C., Labetoulle, M., Ardizzone, M., Sibilia, J., Fournier, C., Chiochia, G. and Mariette, X. Activation of IFN pathways and plasmacytoid dendritic cell recruitment in target organs of primary Sjögren's syndrome. *Proceedings of the National Academy of Sciences of the United States of America.* 2006, **103**(8), pp.2770-2775.
34. Pestka, S., Krause, C.D. and Walter, M.R. Interferons, interferon-like cytokines, and their receptors. *Immunological Reviews.* 2004, **202**(1), pp.8-32.
35. Farrar, M.A. and Schreiber, R.D. The molecular cell biology of interferon-gamma and its receptor. *Annu Rev Immunol.* 1993, **11**, pp.571-611.
36. Lee, A.J. and Ashkar, A.A. The Dual Nature of Type I and Type II Interferons. *Frontiers in Immunology.* 2018, **9**(2061).

37. Andreakos, E., Zanoni, I. and Galani, I.E. Lambda interferons come to light: dual function cytokines mediating antiviral immunity and damage control. *Current Opinion in Immunology*. 2019, **56**, pp.67-75.
38. Stanifer, M.L., Pervolaraki, K. and Boulant, S. Differential Regulation of Type I and Type III Interferon Signaling. *International journal of molecular sciences*. 2019, **20**(6), p.1445.
39. Barrat, F.J., Crow, M.K. and Ivashkiv, L.B. Interferon target-gene expression and epigenomic signatures in health and disease. *Nature Immunology*. 2019, **20**(12), pp.1574-1583.
40. Rönnblom, L. and Alm, G.V. A Pivotal Role for the Natural Interferon α -producing Cells (Plasmacytoid Dendritic Cells) in the Pathogenesis of Lupus. *Journal of Experimental Medicine*. 2001, **194**(12), pp.F59-F64.
41. Ji, G., Beavis, R.C. and Novick, R.P. Cell density control of staphylococcal virulence mediated by an octapeptide pheromone. *Proceedings of the National Academy of Sciences*. 1995, **92**(26), pp.12055-12059.
42. Plataniias, L.C. Mechanisms of type-I- and type-II-interferon-mediated signalling. *Nature Reviews Immunology*. 2005, **5**(5), pp.375-386.
43. Cho, H.Y., Lee, S.W., Seo, S.K., Choi, I.W., Choi, I. and Lee, S.W. Interferon-sensitive response element (ISRE) is mainly responsible for IFN- α -induced upregulation of programmed death-1 (PD-1) in macrophages. *Biochim Biophys Acta*. 2008, **1779**(12), pp.811-819.
44. Plataniias, L.C. and Fish, E.N. Signaling pathways activated by interferons. *Exp Hematol*. 1999, **27**(11), pp.1583-1592.
45. Ivashkiv, L.B. and Donlin, L.T. Regulation of type I interferon responses. *Nature Reviews Immunology*. 2014, **14**(1), pp.36-49.
46. Kiefer, K., Oropallo, M.A., Cancro, M.P. and Marshak-Rothstein, A. Role of type I interferons in the activation of autoreactive B cells. *Immunology & Cell Biology*. 2012, **90**(5), pp.498-504.
47. Blanco, P., Palucka, A.K., Gill, M., Pascual, V. and Banchereau, J. Induction of Dendritic Cell Differentiation by IFN- γ ; in Systemic Lupus Erythematosus. *Science*. 2001, **294**(5546), pp.1540-1543.
48. Cucak, H., Yrlid, U., Reizis, B., Kalinke, U. and Johansson-Lindbom, B. Type I Interferon Signaling in Dendritic Cells Stimulates the Development of Lymph-Node-Resident T Follicular Helper Cells. *Immunity*. 2009, **31**(3), pp.491-501.
49. Reizis, B. Plasmacytoid Dendritic Cells: Development, Regulation, and Function. *Immunity*. 2019, **50**(1), pp.37-50.
50. Gilliet, M., Cao, W. and Liu, Y.-J. Plasmacytoid dendritic cells: sensing nucleic acids in viral infection and autoimmune diseases. *Nature Reviews Immunology*. 2008, **8**(8), pp.594-606.

51. Zhou, Z., Ma, J., Xiao, C., Han, X., Qiu, R., Wang, Y., Zhou, Y., Wu, L., Huang, X. and Shen, N. Phenotypic and functional alterations of pDCs in lupus-prone mice. *Scientific Reports*. 2016, **6**(1), p.20373.
52. Rodero, M.P., Decalf, J., Bondet, V., Hunt, D., Rice, G.I., Werneke, S., McGlasson, S.L., Alyanakian, M.A., Bader-Meunier, B., Barnerias, C., Bellon, N., Belot, A., Bodemer, C., Briggs, T.A., Desguerre, I., Frémond, M.L., Hully, M., van den Maagdenberg, A., Melki, I., Meyts, I., Musset, L., Pelzer, N., Quartier, P., Terwindt, G.M., Wardlaw, J., Wiseman, S., Rieux-Laucat, F., Rose, Y., Neven, B., Hertel, C., Hayday, A., Albert, M.L., Rozenberg, F., Crow, Y.J. and Duffy, D. Detection of interferon alpha protein reveals differential levels and cellular sources in disease. *J Exp Med*. 2017, **214**(5), pp.1547-1555.
53. Psarras, A., Alase, A., Antanaviciute, A., Carr, I.M., Md Yusof, M.Y., Wittmann, M., Emery, P., Tsokos, G.C. and Vital, E.M. Functionally impaired plasmacytoid dendritic cells and non-haematopoietic sources of type I interferon characterize human autoimmunity. *Nature Communications*. 2020, **11**(1), p.6149.
54. LaFleur, D.W., Nardelli, B., Tsareva, T., Mather, D., Feng, P., Semenuk, M., Taylor, K., Buergin, M., Chinchilla, D., Roshke, V., Chen, G., Ruben, S.M., Pitha, P.M., Coleman, T.A. and Moore, P.A. Interferon- κ , a Novel Type I Interferon Expressed in Human Keratinocytes*. *Journal of Biological Chemistry*. 2001, **276**(43), pp.39765-39771.
55. Sarkar, M.K., Hile, G.A., Tsoi, L.C., Xing, X., Liu, J., Liang, Y., Berthier, C.C., Swindell, W.R., Patrick, M.T., Shao, S., Tsou, P.S., Uppala, R., Beamer, M.A., Srivastava, A., Bielas, S.L., Harms, P.W., Getsios, S., Elder, J.T., Voorhees, J.J., Gudjonsson, J.E. and Kahlenberg, J.M. Photosensitivity and type I IFN responses in cutaneous lupus are driven by epidermal-derived interferon kappa. *Ann Rheum Dis*. 2018, **77**(11), pp.1653-1664.
56. Stannard, J.N., Reed, T.J., Myers, E., Lowe, L., Sarkar, M.K., Xing, X., Gudjonsson, J.E. and Kahlenberg, J.M. Lupus Skin Is Primed for IL-6 Inflammatory Responses through a Keratinocyte-Mediated Autocrine Type I Interferon Loop. *J Invest Dermatol*. 2017, **137**(1), pp.115-122.
57. Furie, R., Khamashta, M., Merrill, J.T., Werth, V.P., Kalunian, K., Brohawn, P., Illei, G.G., Drappa, J., Wang, L., Yoo, S. and Investigators, f.t.C.S. Anifrolumab, an Anti-Interferon- α Receptor Monoclonal Antibody, in Moderate-to-Severe Systemic Lupus Erythematosus. *Arthritis & Rheumatology*. 2017, **69**(2), pp.376-386.
58. Morand, E.F., Furie, R., Tanaka, Y., Bruce, I.N., Askanase, A.D., Richez, C., Bae, S.-C., Brohawn, P.Z., Pineda, L., Berglind, A. and Tummala, R. Trial of Anifrolumab in Active Systemic Lupus Erythematosus. *New England Journal of Medicine*. 2019, **382**(3), pp.211-221.
59. Vital, E.M., Merrill, J.T., Morand, E.F., Furie, R.A., Bruce, I.N., Tanaka, Y., Manzi, S., Kalunian, K.C., Kalyani, R.N., Streicher, K., Abreu, G. and Tummala, R. Anifrolumab efficacy and safety by type I interferon gene signature and clinical subgroups in patients with SLE: post hoc analysis of pooled data from two phase III trials. *Ann Rheum Dis*. 2022, **81**(7), pp.951-961.
60. Henry, N.L. and Hayes, D.F. Cancer biomarkers. *Molecular Oncology*. 2012, **6**(2), pp.140-146.

61. Robb, M.A., McInnes, P.M. and Califf, R.M. Biomarkers and Surrogate Endpoints: Developing Common Terminology and Definitions. *JAMA*. 2016, **315**(11), pp.1107-1108.
62. Hsu, W.H., Yang, J.C.H., Mok, T.S. and Loong, H.H. Overview of current systemic management of EGFR-mutant NSCLC. *Annals of Oncology*. 2018, **29**, pp.i3-i9.
63. Lu, H., Zhang, J., Jiang, Z., Zhang, M., Wang, T., Zhao, H. and Zeng, P. Detection of Genetic Overlap Between Rheumatoid Arthritis and Systemic Lupus Erythematosus Using GWAS Summary Statistics. *Front Genet*. 2021, **12**, p.656545.
64. Fernandez, S.A., Lobo, A.Z., Oliveira, Z.N., Fukumori, L.M., AM, P.r. and Rivitti, E.A. Prevalence of antinuclear autoantibodies in the serum of normal blood donors. *Rev Hosp Clin Fac Med Sao Paulo*. 2003, **58**(6), pp.315-319.
65. Arriens, C., Wren, J.D., Munroe, M.E. and Mohan, C. Systemic lupus erythematosus biomarkers: the challenging quest. *Rheumatology*. 2016, **56**(suppl_1), pp.i32-i45.
66. Carter, L.M., Isenberg, D.A. and Ehrenstein, M.R. Elevated serum BAFF levels are associated with rising anti-double-stranded DNA antibody levels and disease flare following B cell depletion therapy in systemic lupus erythematosus. *Arthritis Rheum*. 2013, **65**(10), pp.2672-2679.
67. Fujio, K., Takeshima, Y., Nakano, M. and Iwasaki, Y. Review: transcriptome and trans-omics analysis of systemic lupus erythematosus. *Inflamm Regen*. 2020, **40**, p.11.
68. Burska, A., Rodriguez Carrio, J., Conaghan, P.G., Dik, W.A., Biesen, R., Eloranta, M.L., Cavalli, G., Visser, M., Boumpas, D., Bertsias, G., Wahren-Herlenius, M., Rehwinkel, J., Frémond, M.L., Crow, M.K., Ronnblom, L., Vital, E. and Versnel, M. POS0370 TYPE I INTERFERON PATHWAY ASSAYS IN PATIENTS WITH RHEUMATIC AND MUSCULOSKELETAL DISEASES - SYSTEMATIC LITERATURE REVIEW (SLR) AND DEVELOPMENT OF CONSENSUS TERMINOLOGY FROM A EULAR TASKFORCE. *Ann Rheum Dis*. 2021, **80**(Suppl 1), pp.415-415.
69. Baechler, E.C., Batliwalla, F.M., Karypis, G., Gaffney, P.M., Ortmann, W.A., Espe, K.J., Shark, K.B., Grande, W.J., Hughes, K.M., Kapur, V., Gregersen, P.K. and Behrens, T.W. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc Natl Acad Sci U S A*. 2003, **100**(5), pp.2610-2615.
70. El-Sherbiny, Y.M., Psarras, A., Md Yusof, M.Y., Hensor, E.M.A., Tooze, R., Doody, G., Mohamed, A.A.A., McGonagle, D., Wittmann, M., Emery, P. and Vital, E.M. A novel two-score system for interferon status segregates autoimmune diseases and correlates with clinical features. *Scientific Reports*. 2018, **8**(1), p.5793.
71. Kirou, K.A., Lee, C., George, S., Louca, K., Peterson, M.G. and Crow, M.K. Activation of the interferon-alpha pathway identifies a subgroup of systemic lupus erythematosus patients with distinct serologic features and active disease. *Arthritis Rheum*. 2005, **52**(5), pp.1491-1503.
72. Chaussabel, D., Quinn, C., Shen, J., Patel, P., Glaser, C., Baldwin, N., Stichweh, D., Blankenship, D., Li, L., Munagala, I., Bennett, L., Allantaz, F., Mejias, A., Ardura, M., Kaizer, E., Monnet, L., Allman, W., Randall, H., Johnson, D., Lanier, A., Punaro, M., Wittkowski, K.M., White, P., Fay, J., Klintmalm, G., Ramilo, O., Palucka, A.K.,

- Banchereau, J. and Pascual, V. A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. *Immunity*. 2008, **29**(1), pp.150-164.
73. Chiche, L., Jourde-Chiche, N., Whalen, E., Presnell, S., Gersuk, V., Dang, K., Anguiano, E., Quinn, C., Burtey, S., Berland, Y., Kaplanski, G., Harle, J.R., Pascual, V. and Chaussabel, D. Modular transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal distinct type I and type II interferon signatures. *Arthritis Rheumatol*. 2014, **66**(6), pp.1583-1595.
 74. Md Yusof, M.Y., Psarras, A., El-Sherbiny, Y.M., Hensor, E.M.A., Dutton, K., Ul-Hassan, S., Zayat, A.S., Shalhaf, M., Alase, A., Wittmann, M., Emery, P. and Vital, E.M. Prediction of autoimmune connective tissue disease in an at-risk cohort: prognostic value of a novel two-score system for interferon status. *Ann Rheum Dis*. 2018, **77**(10), p.1432.
 75. Garcia-Montoya, L., Wigston, Z., Burska, A., Mankia, K., Vital, E. and Emery, P. THU0015 TYPE I INTERFERON SIGNATURE PREDICTS PROGRESSION TO INFLAMMATORY ARTHRITIS IN ACPA+ AT-RISK INDIVIDUALS WITHOUT CLINICAL SYNOVITIS. *Ann Rheum Dis*. 2020, **79**(Suppl 1), pp.220-221.
 76. Alase, A., Wigston, Z., Burska, A., Yusof, Y.M., Reynolds, J., Consortium, T.M., Wittmann, M., Bruce, I. and Vital, E.M. P175 MASTERPLANS: prediction of response to rituximab in SLE using a validated two-score system for interferon. *Rheumatology*. 2020, **59**(Supplement_2).
 77. Wigston, Z., Burska, A., Alase, A., Mahmoud, K. and Vital, E. OP0091 A TWO-SCORE INTERFERON SIGNATURE AND MUSCULOSKELETAL IMAGING EXPLAIN THE ASSOCIATION BETWEEN INTERFERON AND ARTHRITIS IN SLE. *Ann Rheum Dis*. 2020, **79**(Suppl 1), pp.60-60.
 78. Rodríguez-Carrio, J., Burska, A.N., Conaghan, P.G., Dik, W.A., Biesen, R., Eloranta, M.-L., Cavalli, G., Visser, M., Boumpas, D., Bertsias, G., Wahren-Herlenius, M., Rehwinkel, J., Frémond, M.-L., Crow, M.K., Ronnblom, L., Versnel, M.A. and Vital, E.M. 2022 EULAR Points to Consider for the measurement and reporting of IFN-I assays in clinical research and practice. [Press release]. 2022. Available
 79. Duale, N., Lipkin, W.I., Briese, T., Aarem, J., Rønningen, K.S., Aas, K.K., Magnus, P., Harbak, K., Susser, E. and Brunborg, G. Long-term storage of blood RNA collected in RNA stabilizing Tempus tubes in a large biobank – evaluation of RNA quality and stability. *BMC Research Notes*. 2014, **7**(1), p.633.
 80. Yang, J., Diaz, N., Adelsberger, J., Zhou, X., Stevens, R., Rupert, A., Metcalf, J.A., Baseler, M., Barbon, C., Imamichi, T., Lempicki, R. and Cosentino, L.M. The effects of storage temperature on PBMC gene expression. *BMC Immunol*. 2016, **17**, p.6.
 81. Altman, D.G. and Bland, J.M. Measurement in medicine: the analysis of method comparison studies. *Journal of the Royal Statistical Society: Series D (The Statistician)*. 1983, **32**(3), pp.307-317.
 82. Yu, Y. and Su, K. Neutrophil Extracellular Traps and Systemic Lupus Erythematosus. *J Clin Cell Immunol*. 2013, **4**.
 83. Whitney, A.R., Diehn, M., Popper, S.J., Alizadeh, A.A., Boldrick, J.C., Relman, D.A. and Brown, P.O. Individuality and variation in gene expression patterns in human

- blood. *Proceedings of the National Academy of Sciences of the United States of America*. 2003, **100**(4), pp.1896-1901.
84. Ellis, T.N. and Beaman, B.L. Interferon-gamma activation of polymorphonuclear neutrophil function. *Immunology*. 2004, **112**(1), pp.2-12.
 85. Coit, P., Yalavarthi, S., Ognenovski, M., Zhao, W., Hasni, S., Wren, J.D., Kaplan, M.J. and Sawalha, A.H. Epigenome profiling reveals significant DNA demethylation of interferon signature genes in lupus neutrophils. *J Autoimmun*. 2015, **58**, pp.59-66.
 86. Bodewes, I.L.A., Al-Ali, S., van Helden-Meeuwsen, C.G., Maria, N.I., Tarn, J., Lendrem, D.W., Schreurs, M.W.J., Steenwijk, E.C., van Daele, P.L.A., Both, T., Bowman, S.J., Griffiths, B., Ng, W.-F., registry, U.P.S.s.S. and Versnel, M.A. Systemic interferon type I and type II signatures in primary Sjögren's syndrome reveal differences in biological disease activity. *Rheumatology*. 2018, **57**(5), pp.921-930.
 87. Berry, M.P., Graham, C.M., McNab, F.W., Xu, Z., Bloch, S.A., Oni, T., Wilkinson, K.A., Banchereau, R., Skinner, J., Wilkinson, R.J., Quinn, C., Blankenship, D., Dhawan, R., Cush, J.J., Mejias, A., Ramilo, O., Kon, O.M., Pascual, V., Banchereau, J., Chaussabel, D. and O'Garra, A. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature*. 2010, **466**(7309), pp.973-977.
 88. Martinelli, S., Urosevic, M., Daryadel, A., Oberholzer, P.A., Baumann, C., Fey, M.F., Dummer, R., Simon, H.-U. and Yousefi, S. Induction of Genes Mediating Interferon-dependent Extracellular Trap Formation during Neutrophil Differentiation *. *Journal of Biological Chemistry*. 2004, **279**(42), pp.44123-44132.
 89. Cooles, F.A.H., Anderson, A.E., Lendrem, D.W., Norris, J., Pratt, A.G., Hilkens, C.M.U. and Isaacs, J.D. The interferon gene signature is increased in patients with early treatment-naïve rheumatoid arthritis and predicts a poorer response to initial therapy. *J Allergy Clin Immunol*. 2018, **141**(1), pp.445-448.e444.
 90. Icen, M., Nicola, P.J., Maradit-Kremers, H., Crowson, C.S., Thorneau, T.M., Matteson, E.L. and Gabriel, S.E. Systemic lupus erythematosus features in rheumatoid arthritis and their effect on overall mortality. *J Rheumatol*. 2009, **36**(1), pp.50-57.
 91. Team, R.C. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2021.
 92. McCall, M.N., McMurray, H.R., Land, H. and Almudevar, A. On non-detects in qPCR data. *Bioinformatics*. 2014, **30**(16), pp.2310-2316.
 93. Kaiser, H.F. An index of factorial simplicity. *psychometrika*. 1974, **39**(1), pp.31-36.
 94. A, K. ggcorrplot: Visualization of a Correlation Matrix using 'ggplot2'_. R package version 0.1.3. 2019.
 95. Revelle, W. psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA. 2022.
 96. A, D. paran: Horn's Test of Principal Components/Factors_. R package version 1.5.2. 2018.

97. Delacre., M., Lakens., D., Ley., C., Liu., L. and Leys., C. Why Hedges' g^* s based on the non-pooled standard deviation should be reported with Welch's t-test. 2021.
98. Monaco, G., Lee, B., Xu, W., Mustafah, S., Hwang, Y.Y., Carré, C., Burdin, N., Visan, L., Ceccarelli, M., Poidinger, M., Zippelius, A., Pedro de Magalhães, J. and Larbi, A. RNA-Seq Signatures Normalized by mRNA Abundance Allow Absolute Deconvolution of Human Immune Cell Types. *Cell Rep.* 2019, **26**(6), pp.1627-1640.e1627.
99. Palanichamy, A., Bauer, J.W., Yalavarthi, S., Meednu, N., Barnard, J., Owen, T., Cistrone, C., Bird, A., Rabinovich, A., Nevarez, S., Knight, J.S., Dedrick, R., Rosenberg, A., Wei, C., Rangel-Moreno, J., Liesveld, J., Sanz, I., Baechler, E., Kaplan, M.J. and Anolik, J.H. Neutrophil-Mediated IFN Activation in the Bone Marrow Alters B Cell Development in Human and Murine Systemic Lupus Erythematosus. *The Journal of Immunology.* 2014, **192**(3), pp.906-918.
100. Garcia-Romo, G.S., Caielli, S., Vega, B., Connolly, J., Allantaz, F., Xu, Z., Punaro, M., Baisch, J., Guiducci, C., Coffman, R.L., Barrat, F.J., Banchereau, J. and Pascual, V. Netting neutrophils are major inducers of type I IFN production in pediatric systemic lupus erythematosus. *Sci Transl Med.* 2011, **3**(73), p.73ra20.
101. Kaplan, M.J. Neutrophils in the pathogenesis and manifestations of SLE. *Nat Rev Rheumatol.* 2011, **7**(12), pp.691-699.
102. Wahadat, M.J., Schonenberg-Meinema, D., van Helden-Meeuwsen, C.G., van Tilburg, S.J., Groot, N., Schatorjé, E.J.H., Hoppenreijs, E.P.A.H., Hissink Muller, P.C.E., Brinkman, D.M.C., Dvorak, D., Verkaaik, M., van den Berg, J.M., Bouchalova, K., Kamphuis, S. and Versnel, M.A. Gene signature fingerprints stratify SLE patients in groups with similar biological disease profiles: a multicentre longitudinal study. *Rheumatology.* 2022.
103. Green, M.R. and Sambrook, J. Quantification of RNA by Real-Time Reverse Transcription-Polymerase Chain Reaction (RT-PCR). *Cold Spring Harbor protocols.* 2018, **2018**(10), p.pdb.prot095042.
104. Schmittgen, T.D. and Livak, K.J. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc.* 2008, **3**(6), pp.1101-1108.
105. Suzuki, T., Higgins, P.J. and Crawford, D.R. Control selection for RNA quantitation. *Biotechniques.* 2000, **29**(2), pp.332-337.
106. Chapman, J.R. and Waldenström, J. With Reference to Reference Genes: A Systematic Review of Endogenous Controls in Gene Expression Studies. *PLoS One.* 2015, **10**(11), p.e0141853.
107. Stamova, B.S., Apperson, M., Walker, W.L., Tian, Y., Xu, H., Adamczyk, P., Zhan, X., Liu, D.Z., Ander, B.P., Liao, I.H., Gregg, J.P., Turner, R.J., Jickling, G., Lit, L. and Sharp, F.R. Identification and validation of suitable endogenous reference genes for gene expression studies in human peripheral blood. *BMC Med Genomics.* 2009, **2**, p.49.
108. Guénin, S., Mauriat, M., Pelloux, J., Van Wuytswinkel, O., Bellini, C. and Gutierrez, L. Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. *J Exp Bot.* 2009, **60**(2), pp.487-493.

109. Jacob, F., Guertler, R., Naim, S., Nixdorf, S., Fedier, A., Hacker, N.F. and Heinzelmann-Schwarz, V. Careful Selection of Reference Genes Is Required for Reliable Performance of RT-qPCR in Human Normal and Cancer Cell Lines. *PLoS One*. 2013, **8**(3), p.e59180.
110. Bustin, S.A., Beaulieu, J.F., Huggett, J., Jaggi, R., Kibenge, F.S., Olsvik, P.A., Penning, L.C. and Toegel, S. MIQE précis: Practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments. *BMC Mol Biol*. 2010, **11**, p.74.
111. Andersen, C.L., Jensen, J.L. and Ørntoft, T.F. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer research*. 2004, **64**(15), pp.5245-5250.
112. Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. and Speleman, F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*. 2002, **3**(7), pp.1-12.
113. Pfaffl, M.W., Tichopad, A., Prgomet, C. and Neuvians, T.P. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper—Excel-based tool using pair-wise correlations. *Biotechnology Letters*. 2004, **26**(6), pp.509-515.
114. Silver, N., Best, S., Jiang, J. and Thein, S.L. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Molecular Biology*. 2006, **7**(1), pp.1-9.
115. Xie, F., Xiao, P., Chen, D., Xu, L. and Zhang, B. miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. *Plant Mol Biol*. 2012.
116. Riemer, A.B., Keskin, D.B. and Reinherz, E.L. Identification and validation of reference genes for expression studies in human keratinocyte cell lines treated with and without interferon- γ - a method for qRT-PCR reference gene determination. *Experimental dermatology*. 2012, **21**(8), pp.625-629.
117. Bednarz-Misa, I., Neubauer, K., Zacharska, E., Kapturkiewicz, B. and Krzystek-Korpacka, M. Whole blood ACTB, B2M and GAPDH expression reflects activity of inflammatory bowel disease, advancement of colorectal cancer, and correlates with circulating inflammatory and angiogenic factors: Relevance for real-time quantitative PCR. *Adv Clin Exp Med*. 2020, **29**(5), pp.547-556.
118. Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. and Speleman, F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*. 2002, **3**(7), p.research0034.0031.
119. *Gene Cards*. [Online]. [Accessed]. Available from: <https://www.genecards.org/>
120. Thurlings, R.M., Boumans, M., Tekstra, J., van Roon, J.A., Vos, K., van Westing, D.M., van Baarsen, L.G., Bos, C., Kirou, K.A., Gerlag, D.M., Crow, M.K., Bijlsma, J.W., Verweij, C.L. and Tak, P.P. Relationship between the type I interferon signature and the response to rituximab in rheumatoid arthritis patients. 2010, **62**(12), pp.3607-3614.

121. van Baarsen, L.G.M., Wijbrandts, C.A., Rustenburg, F., Cantaert, T., van der Pouw Kraan, T.C.T.M., Baeten, D.L., Dijkmans, B.A.C., Tak, P.P. and Verweij, C.L. Regulation of IFN response gene activity during infliximab treatment in rheumatoid arthritis is associated with clinical response to treatment. *Arthritis Research & Therapy*. 2010, **12**(1), p.R11.
122. van Baarsen, L.G.M., Bos, W.H., Rustenburg, F., van der Pouw Kraan, T.C.T.M., Wolbink, G.J.J., Dijkmans, B.A.C., van Schaardenburg, D. and Verweij, C.L. Gene expression profiling in autoantibody-positive patients with arthralgia predicts development of arthritis. 2010, **62**(3), pp.694-704.
123. Li, Q.-Z., Zhou, J., Lian, Y., Zhang, B., Branch, V.K., Carr-Johnson, F., Karp, D.R., Mohan, C., Wakeland, E.K. and Olsen, N.J. Interferon signature gene expression is correlated with autoantibody profiles in patients with incomplete lupus syndromes. 2010, **159**(3), pp.281-291.
124. Yao, Y., Higgs, B.W., Morehouse, C., De los Reyes, M., Trigona, W., Brohawn, P., White, W., Zhang, J., White, B., Coyle, A.J.J.H.g. and HGP, p. Development of potential pharmacodynamic and diagnostic markers for anti-IFN- α monoclonal antibody trials in systemic lupus erythematosus. 2009, **2009**.
125. vitalLamot, L., Niemietz, I. and Brown, K.L. Comparable type I interferon score determination from PAXgene and Tempus whole blood RNA collection and isolation systems. *BMC Res Notes*. 2019, **12**(1), p.511.
126. Furie, R., Werth, V.P., Merola, J.F., Stevenson, L., Reynolds, T.L., Naik, H., Wang, W., Christmann, R., Gardet, A., Pellerin, A., Hamann, S., Auluck, P., Barbey, C., Gulati, P., Rabah, D. and Franchimont, N. Monoclonal antibody targeting BDCA2 ameliorates skin lesions in systemic lupus erythematosus. *J Clin Invest*. 2019, **129**(3), pp.1359-1371.
127. Raterman, H.G., Vosslamber, S., de Ridder, S., Nurmohamed, M.T., Lems, W.F., Boers, M., van de Wiel, M., Dijkmans, B.A.C., Verweij, C.L. and Voskuyl, A.E. The interferon type I signature towards prediction of non-response to rituximab in rheumatoid arthritis patients. *Arthritis Research & Therapy*. 2012, **14**(2), p.R95.
128. de Jong, T.D., Vosslamber, S., Blits, M., Wolbink, G., Nurmohamed, M.T., van der Laken, C.J., Jansen, G., Voskuyl, A.E., Verweij, C.L.J.A.r. and therapy. Effect of prednisone on type I interferon signature in rheumatoid arthritis: consequences for response prediction to rituximab. 2015, **17**(1), p.78.
129. Lübbers, J., Brink, M., van de Stadt, L.A., Vosslamber, S., Wesseling, J.G., van Schaardenburg, D., Rantapää-Dahlqvist, S. and Verweij, C.L. The type I IFN signature as a biomarker of preclinical rheumatoid arthritis. 2013, **72**(5), pp.776-780.
130. Bilgic, H., Ytterberg, S.R., Amin, S., McNallan, K.T., Wilson, J.C., Koeuth, T., Ellingson, S., Newman, B., Bauer, J.W., Peterson, E.J., Baechler, E.C. and Reed, A.M. Interleukin-6 and type I interferon-regulated genes and chemokines mark disease activity in dermatomyositis. 2009, **60**(11), pp.3436-3446.
131. Landolt-Marticorena, C., Bonventi, G., Lubovich, A., Ferguson, C., Unnithan, T., Su, J., Gladman, D.D., Urowitz, M., Fortin, P.R. and Wither, J. Lack of association between the interferon- α signature and longitudinal changes in disease activity in systemic lupus erythematosus. 2009, **68**(9), pp.1440-1446.

132. van den Hoogen, L.L., Fritsch-Stork, R.D., Versnel, M.A., Derksen, R.H., van Roon, J.A. and Radstake, T.R. Monocyte type I interferon signature in antiphospholipid syndrome is related to proinflammatory monocyte subsets, hydroxychloroquine and statin use. *Ann Rheum Dis.* 2016, **75**(12), p.e81.
133. van den Hoogen, L.L., van Roon, J.A.G., Mertens, J.S., Wienke, J., Lopes, A.P., de Jager, W., Rossato, M., Pandit, A., Wichers, C.G.K., van Wijk, F., Fritsch-Stork, R.D.E. and Radstake, T. Galectin-9 is an easy to measure biomarker for the interferon signature in systemic lupus erythematosus and antiphospholipid syndrome. *Ann Rheum Dis.* 2018, **77**(12), pp.1810-1814.
134. Blokland, S.L.M., van den Hoogen, L.L., Leijten, E.F.A., Hartgring, S.A.Y., Fritsch, R., Kruize, A.A., van Roon, J.A.G. and Radstake, T. Increased expression of Fas on group 2 and 3 innate lymphoid cells is associated with an interferon signature in systemic lupus erythematosus and Sjögren's syndrome. *Rheumatology (Oxford)*. 2019, **58**(10), pp.1740-1745.
135. Ahn, H.R., Baek, G.O., Yoon, M.G., Son, J.A., You, D., Yoon, J.H., Cho, H.J., Kim, S.S., Cheong, J.Y. and Eun, J.W. HMBS is the most suitable reference gene for RT-qPCR in human HCC tissues and blood samples. *Oncol Lett.* 2021, **22**(5), p.791.
136. Hua, J., Kirou, K., Lee, C. and Crow, M.K. Functional assay of type I interferon in systemic lupus erythematosus plasma and association with anti-RNA binding protein autoantibodies. *Arthritis Rheum.* 2006, **54**(6), pp.1906-1916.
137. Mathian, A., Mouries-Martin, S., Dorgham, K., Devilliers, H., Barnabei, L., Ben Salah, E., Cohen-Aubart, F., Garrido Castillo, L., Haroche, J., Hie, M., Pineton de Chambrun, M., Miyara, M., Sterlin, D., Pha, M., D, L.T.H., Rieux-Laucat, F., Rozenberg, F., Gorochov, G. and Amoura, Z. Monitoring Disease Activity in Systemic Lupus Erythematosus With Single-Molecule Array Digital Enzyme-Linked Immunosorbent Assay Quantification of Serum Interferon- α . *Arthritis Rheumatol.* 2019, **71**(5), pp.756-765.
138. Reynolds, J.A., Briggs, T.A., Rice, G.I., Darmalinggam, S., Bondet, V., Bruce, E., Khan, M., Haque, S., Chinoy, H., Herrick, A.L., McCarthy, E.M., Zeef, L., Hayes, A., Duffy, D., Parker, B. and Bruce, I.N. Type I interferon in patients with systemic autoimmune rheumatic disease is associated with haematological abnormalities and specific autoantibody profiles. *Arthritis Res Ther.* 2019, **21**(1), p.147.
139. Ledderose, C., Heyn, J., Limbeck, E. and Kreth, S. Selection of reliable reference genes for quantitative real-time PCR in human T cells and neutrophils. *BMC Res Notes.* 2011, **4**, p.427.
140. Allenbach, Y., Leroux, G., Suárez-Calvet, X., Preusse, C., Gallardo, E., Hervier, B., Rigolet, A., Hie, M., Pehl, D., Limal, N., Hufnagl, P., Zerbe, N., Meyer, A., Aouizerate, J., Uzunhan, Y., Maisonobe, T., Goebel, H.H., Benveniste, O. and Stenzel, W. Dermatomyositis With or Without Anti-Melanoma Differentiation-Associated Gene 5 Antibodies: Common Interferon Signature but Distinct NOS2 Expression. *Am J Pathol.* 2016, **186**(3), pp.691-700.
141. Oliveira, J.J., Karrar, S., Rainbow, D.B., Pinder, C.L., Clarke, P., Rubio García, A., Al-Assar, O., Burling, K., Morris, S., Stratton, R., Vyse, T.J., Wicker, L.S., Todd, J.A. and Ferreira, R.C. The plasma biomarker soluble SIGLEC-1 is associated with the type I interferon transcriptional signature, ethnic background and renal disease in systemic lupus erythematosus. *Arthritis Res Ther.* 2018, **20**(1), p.152.

142. Feng, X., Wu, H., Grossman, J.M., Hanvivadhanakul, P., FitzGerald, J.D., Park, G.S., Dong, X., Chen, W., Kim, M.H., Weng, H.H., Furst, D.E., Gorn, A., McMahon, M., Taylor, M., Brahn, E., Hahn, B.H. and Tsao, B.P. Association of increased interferon-inducible gene expression with disease activity and lupus nephritis in patients with systemic lupus erythematosus. 2006, **54**(9), pp.2951-2962.
143. Feng, X., Chen, W., Xiao, L., Gu, F., Huang, J., Tsao, B.P. and Sun, L. Artesunate inhibits type I interferon-induced production of macrophage migration inhibitory factor in patients with systemic lupus erythematosus. *Lupus*. 2017, **26**(1), pp.62-72.
144. Feng, X., Huang, J., Liu, Y., Xiao, L., Wang, D., Hua, B., Tsao, B.P. and Sun, L. Identification of interferon-inducible genes as diagnostic biomarker for systemic lupus erythematosus. *Clin Rheumatol*. 2015, **34**(1), pp.71-79.
145. Feng, X., Wu, H., Grossman, J.M., Hanvivadhanakul, P., FitzGerald, J.D., Park, G.S., Dong, X., Chen, W., Kim, M.H., Weng, H.H., Furst, D.E., Gorn, A., McMahon, M., Taylor, M., Brahn, E., Hahn, B.H. and Tsao, B.P. Association of increased interferon-inducible gene expression with disease activity and lupus nephritis in patients with systemic lupus erythematosus. *Arthritis Rheum*. 2006, **54**(9), pp.2951-2962.
146. Wither, J., Johnson, S.R., Liu, T., Noamani, B., Bonilla, D., Lisnevskaja, L., Silverman, E., Bookman, A. and Landolt-Marticorena, C. Presence of an interferon signature in individuals who are anti-nuclear antibody positive lacking a systemic autoimmune rheumatic disease diagnosis. *Arthritis Res Ther*. 2017, **19**(1), p.41.
147. Tan, F.K., Zhou, X., Mayes, M.D., Gourh, P., Guo, X., Marcum, C., Jin, L. and Arnett, F.C., Jr. Signatures of differentially regulated interferon gene expression and vasculotrophism in the peripheral blood cells of systemic sclerosis patients. *Rheumatology*. 2006, **45**(6), pp.694-702.
148. Kennedy, W.P., Maciuga, R., Wolslegel, K., Tew, W., Abbas, A.R., Chaivorapol, C., Morimoto, A., McBride, J.M., Brunetta, P., Richardson, B.C., Davis, J.C., Jr., Behrens, T.W. and Townsend, M.J. Association of the interferon signature metric with serological disease manifestations but not global activity scores in multiple cohorts of patients with SLE. *Lupus Sci Med*. 2015, **2**(1), p.e000080.
149. Sikand, K., Singh, J., Ebron, J.S. and Shukla, G.C. Housekeeping Gene Selection Advisory: Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) and β -Actin Are Targets of miR-644a. *PLoS One*. 2012, **7**(10), p.e47510.
150. Lanata, C.M., Chung, S.A. and Criswell, L.A. DNA methylation 101: what is important to know about DNA methylation and its role in SLE risk and disease heterogeneity. *Lupus Science & Medicine*. 2018, **5**(1), p.e000285.
151. Chung, S.A., Nititham, J., Elboudwarej, E., Quach, H.L., Taylor, K.E., Barcellos, L.F. and Criswell, L.A. Genome-Wide Assessment of Differential DNA Methylation Associated with Autoantibody Production in Systemic Lupus Erythematosus. *PLoS One*. 2015, **10**(7), p.e0129813.
152. Carter, L.M., Alase, A., Wigston, Z., Psarras, A., Burska, A., Yusof, M.Y.M., Reynolds, J.A., Consortium, M., Wittmann, M., Bruce, I. and Vital, E.M. OA09 Gene expression analysis reveals distinct ancestry-specific profiles associated with response to rituximab in SLE. *Rheumatology*. 2022, **61**(Supplement_1).

153. Siddiqi, K.Z., Zinglensen, A.H., Iversen, K.K., Rasmussen, N.S., Nielsen, C.T. and Jacobsen, S. A cluster of type II interferon-regulated genes associates with disease activity in patients with systemic lupus erythematosus. *Journal of Autoimmunity*. 2022, **132**, p.102869.
154. Yao, Y., Higgs, B.W., Richman, L., White, B. and Jallal, B. Use of type I interferon-inducible mRNAs as pharmacodynamic markers and potential diagnostic markers in trials with sifalimumab, an anti-IFN α antibody, in systemic lupus erythematosus. *Arthritis Research & Therapy*. 2010, **12**(1), p.S6.
155. Rainen, L., Oelmueller, U., Jurgensen, S., Wyrich, R., Ballas, C., Schram, J., Herdman, C., Bankaitis-Davis, D., Nicholls, N., Trollinger, D. and Tryon, V. Stabilization of mRNA Expression in Whole Blood Samples. *Clinical Chemistry*. 2002, **48**(11), pp.1883-1890.
156. Skogholt, A.H., Ryeng, E., Erlandsen, S.E., Skorpen, F., Schønberg, S.A. and Sætrum, P. Gene expression differences between PAXgene and Tempus blood RNA tubes are highly reproducible between independent samples and biobanks. *BMC Research Notes*. 2017, **10**(1), p.136.