Article:
Carter, LM orcid.org/0000-0002-8724-5937, Alase, A, Wigston, Z et al. (10 more authors) (2023) Gene expression and autoantibody analysis reveals distinct ancestry-specific profiles associated with response to rituximab in refractory systemic lupus erythematosus. Arthritis & Rheumatology, 75 (5). pp. 697-710. ISSN 2326-5191

https://doi.org/10.1002/art.42404

Article Type Full length

Title

Gene expression and autoantibody analysis reveals distinct ancestry-specific profiles associated with response to rituximab in refractory systemic lupus erythematosus

Authors

Lucy M Carter PhD MRCP*[1,2], Adewonuola Alase PhD*[1], Zoe Wigston BSc[1], Antony Psarras MBChB PhD[1], Agata Burska PhD[1], Emily Sutton PhD[3], Md Yuzaiful Md Yusof PhD MRCP[1,2], John A Reynolds PhD MRCP[4,5], The Masterplans Consortium, Neil McHugh MD FRACP FRCPath[6], Paul Emery MD FRCP[1,2], Miriam Wittmann MD[7], Ian N. Bruce MD FRCP[3], Edward M Vital PhD MRCP[1,2].
*Equal contribution

Affiliations

1. Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, UK
2. NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds, UK
3. Centre for Epidemiology Versus Arthritis, Division of Musculoskeletal and Dermatological Sciences, University of Manchester, Manchester, UK
4. Institute of Inflammation and Ageing, University of Birmingham, UK
5. Sandwell and West Birmingham NHS Trust, Birmingham, UK
6. Department of Pharmacy and Pharmacology, University of Bath, Claverton, Bath, UK
7. Department of Dermatology, University Medical Centre, Johannes Gutenberg-University, Mainz, Germany

Correspondence

Dr Edward Vital
Chapel Allerton Hospital, Chapeltown Road, Leeds, LS7 4SA, United Kingdom
Email: e.m.vital@leeds.ac.uk

Key Words Biologicals, B-lymphocytes, Biomarkers, Systemic Lupus Erythematosus

Word Count 4295

Funding

This research was funded/supported by the Medical Research Council of MASTERPLANS (grant code MR/M01665X/1). The BILAG-BR has received funding support from Lupus UK, and unrestricted grants from Roche and GSK paid to University of Manchester. This article/paper/report also presents independent research funded/supported by the NIHR Leeds Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Financial Disclosures

YY has received consultancy fees from Aurinia Pharmaceuticals and UCB. PE has received consultancy fees from BMS, Abbvie, Abbott, Pfizer, MSD, Novartis, Roche, Janssen and UCB. He has received research grants paid to his employer from Abbvie, BMS, Pfizer, MSD and Roche. MW has received honoraria for advisory board contributions and educational lectures and research grants paid to her employer from AstraZeneca, Pfizer, Sanofi, Novartis, Leo, UCB and Abbvie. EMV has received consultancy fees from Roche, GSK, AstraZeneca, Aurinia Pharmaceuticals, Lilly and Novartis. He has also received research grants paid to his employer from Roche, AstraZeneca and Sandoz. INB has received grant/research support from Genzyme/Sanofi, GlaxoSmithKline, Roche, and UCB; received consulting fees from Eli Lilly, GlaxoSmithKline, ILTOO, Merck Serono, and UCB; and was a speaker for AstraZeneca, GlaxoSmithKline, and UCB. All other authors have declared no competing interests.

Abstract (246 words)

Objective

Gene expression profiles are associated with the clinical heterogeneity of SLE but are not well studied as biomarkers for therapy. Many clinical and demographic features influence treatment responses. We studied gene expression and response to rituximab in a multi-ethnic UK cohort refractory to standard therapy.

Methods

Baseline expression of transcripts known to associate with clinical features of SLE was evaluated in whole blood by 96-probe Taqman® array in patients (n=213) with active SLE, prospectively enrolled in British Isles Lupus Assessment Group (BILAG) Biologics Registry. Autoantibodies were measured using immunoprecipitation and ELISA. Response to first cycle rituximab (n=110) was determined by BILAG-2004 criteria at 6 months.

Results

Interferon scores were lower in European ancestry patients than all other groups. The relationship between blood interferon scores and plasmablast, neutrophil, myeloid, inflammation and erythropoiesis-annotated scores differed between patients of European and non-European ancestries. Hierarchical clustering revealed 3 distinct non-European ancestry patient subsets with stratified response to rituximab which was not explained by sociodemographic and clinical variables. Response was lowest in an interferon-low, neutrophil-high cluster and highest in a cluster with high expression across all signatures ($p<0.001$). Clusters within European ancestry patients did not predict response to rituximab but segregated patients by global disease activity and renal involvement. In both ancestral groups, interferon-high clusters associated with U1RNP-Sm antibodies.

Conclusion

Ancestry appears central to the immunological and clinical heterogeneity in SLE. These results suggest that ancestry, disease activity and transcriptional signatures could each assist predict the effectiveness of B-cell depletion.

Introduction

Systemic lupus erythematosus (SLE) is a complex multisystem disease in which immune dysregulation culminates in autoantibodies to nuclear antigens, immune complex deposition, complement activation and tissue injury (1). Underlying immunopathological diversity contributes to variability in disease severity, response to therapy and clinical outcomes which is incompletely understood. Heterogeneity between ancestral groups appears particularly important, though minority ancestral groups remain under-represented in most clinical studies (2). Non-European ancestry populations frequently show higher prevalence and younger onset than populations of European Ancestry (3) with greater renal involvement and damage accrual (4, 5). Improved stratification between and within ancestral groups could therefore be a crucial strategy to improve treatment selection and achieve greater parity in clinical outcomes.

B-cell depletion using the anti-CD20 monoclonal antibody rituximab is an important therapeutic strategy in refractory SLE (6). Despite its widespread use, initial clinical trials failed to meet primary endpoints and response can vary markedly (7, 8). Patients of African ancestry showed greater response to rituximab in one major trial (9) but appeared less responsive to B-cell directed therapy with BAFF neutralizing monoclonal antibody belimumab (10). Differential efficacy across other minority ancestral groups such as subcontinental Asian patients is less well characterized, and it is not always clear how much differences in outcomes are influenced by geographic and social factors influencing access to healthcare.

Gene expression profiles show potential to assist SLE stratification. The blood transcriptome of SLE has been comprehensively evaluated by microarray (11, 12) permitting the assembly of co-expressed transcripts into functionally annotated modules which distinguish by disease activity, auto-antibody status (11), renal involvement (12, 13) and cutaneous manifestations (14). Well characterized autoantibody clusters to RNA-binding proteins have been associated with certain clinical phenotypes, interferon signatures and less favourable response to B-cell depletion (15-17). We have previously validated two continuous interferon (IFN) gene expression scores, IFN-Score-A and IFN-Score-B, derived from factor analysis of IFN annotated modules, which yielded stronger clinical associations than a more global IFN signature (18, 19). Ancestral background significantly influences IFN signatures (20) and other transcriptional profiles in SLE (21). Although B-cell dynamics after rituximab therapy can predict subsequent outcomes, pre-treatment biomarkers that predict response are lacking (22-24). Gene expression profiles associated with response to rituximab in SLE have not been evaluated and ancestry-specific effects have not been explored.

The present study examines the relationship between ancestry, whole blood gene expression signatures, auto-antibody status and response to first cycle rituximab therapy in a multi-ethnic UK SLE cohort with disease refractory to standard therapy. MASTERPLANS is a consortium aiming to stratify therapy in SLE.

Methods

Patients

The British Isles Lupus Assessment Group Biologics Register (BILAG-BR) is a prospective UK-wide registry evaluating the safety and efficacy of biologics in SLE. Research ethics approval was obtained from North West–Greater Manchester West NRES Committee (REC: 09/H1014/64) and UK Health Research Authority (IRAS ref. 24407). Eligibility for rituximab in England requires cyclophosphamide and/or mycophenolate mofetil treatment failure, active SLE (at least 1xBILAG A and/or 2xBILAG B, or SLEDAI-2K≥6), or unacceptably high dose glucocorticoids to control disease (25, 26). Comprehensive clinical and demographic data, including clinical haematology and immunology obtained through local diagnostic laboratories, were captured prospectively. Patient self-identified ancestry was recorded according to UK 2011 Census categories. Socioeconomic deprivation was measured by 2019 Index of Multiple Deprivation rank of the statistical geography of postal address on enrolment (See Supplementary Appendix 1). The primary rituximab response criterion was evaluated in patients with either BILAG-2004 ≥1xA or ≥2xB at baseline (n=110; Supplementary Fig. S1).

Clinical outcomes

Disease activity was assessed using the BILAG-2004 Index (27, 28). Response was defined as improvement in all BILAG A scores and no more than one persisting BILAG B score at six months post treatment, with no new BILAG grade A/B flares (22).

Whole blood gene expression analysis

Gene expression was evaluated in pre-rituximab whole blood TEMPUS^{TM} samples (n=213), blind to participant's clinical status, using a customised 96.96 Taqman® array as previously described (18). Ct values were normalised to reference gene peptidylprolyl isomerase A (PPIA) and ΔCt was reflected such that higher values indicate greater gene expression.

Gene selection and gene expression scores

IFN- annotated transcripts comprised two validated continuous ISG expression scores (18) and 7 additional well characterized interferon stimulated genes (ISGs). IFN-Score-A includes transcripts most frequently reported in global IFN-I signatures. IFN-Score-B includes additional ISGs that may be dynamically responsive to multiple IFN subtypes. Genes annotated to plasmablast (n=4, M4.11, M7.7), neutrophil (n=15, M5.15), myeloid lineage (n=17, M3.2, M5.7), inflammation (n=13, M4.2) and erythropoesis (n=11, M2.3, M3.1) were selected from previously described modules based on known molecular function and attributes (11). Supplementary Table S1 shows a complete listing of transcripts and corresponding Taqman® ID. Gene expression scores for each annotation were represented by median reflected ΔCt of the relevant transcripts.

Immunoprecipitation and ELISA

Autoantibody analysis was performed at the specialist autoimmune serology laboratory at the University of Bath for a subset of patients using serum contemporaneous to gene expression. Serotyping for Ro60,

La and Sm-RNP was performed by radio-labelled protein immunoprecipitation as previously described (29, 30). Anti-SS-A 52 IgG (Abnova, Taiwan), anti-Cardiolipin IgG III and anti-dsDNA IgG (both Quanta Lite®, Inova Diagnostics, USA) were evaluated by ELISA (Supplementary Appendix 1).

Statistical analysis

Statistical analyses were performed in R version 4.1.1 and R Studio v1.3.1093. PCR Ct values falling below the prespecified minimum signal intensity were imputed using the *nondetects* package (31). Heirarchical clustering was undertaken using the complete linkage method in package *hclust*. Data visualisation was performed using *ggplot2*, *heatmap*, *Corrplot* and *ComplexUpset* (32). Correlation was assessed by Spearman's correlation co-efficient. Normally distributed continuous variables were compared by t-test or ANOVA and Tukey HSD post-hoc test. Kruskal-Wallis and post-Hoc Dunn's test was used on non-parametric variables. Categorical variables were compared by Chi-Square test. Statistical significance was defined as $p \leq 0.05$ throughout.

Results

Study population

213 patients enrolled in BILAG-BR had pre-treatment whole blood available for gene expression analysis. 162/213 (76%) patients were enrolled on commencing biologic therapy with rituximab (Supplementary Fig. S1). 128/213 (60%) patients were of White European ancestry, specifically white British (55%) or Irish (5%). Minority ancestral groups were of African ancestry (n=27; 13%), Subcontinental Asian (n=27; 13%), Chinese and Other Asian heritage (n=11; 5%). Other ancestral backgrounds including mixed ethnicity accounted for the remaining 9%. Compared with patients of European ancestry (n=128), patients of non-European ancestry collectively (n=85) were significantly younger (*37 vs 43 years*; t= -3.4, p=0.001), with lower prevalence of cigarette smoking (*19% vs 44%*; X^2= 10.4, p=0.001) and resided in areas of significantly higher overall relative deprivation (Index of Multiple Deprivation (IMD) *rank 17526 vs 11311*, t= 3.1, p=0.002) and higher relative deprivation in 6/7 composite IMD domains (Supplementary Fig.S2). They demonstrated higher rates of hypocomplementaemia (*57% vs 43.0%*; X^2= 4.0, p=0.045), higher total IgG (*16.2g/L vs 10.9g/L*; t= 4.8, p<0.000) and higher seropositivity for U1RNP-Sm (*50% vs 12%*; X^2= 28.6, p <0.000), Ro-60 (*45% vs 29%; X^2*= 4.2, p=0.040) and anti-dsDNA antibodies (*68% vs 46%; X^2*= 7.3, p=0.006). No substantive differences in disease activity (SLEDAI-2K and numerical BILAG), registration therapy or concomitant SLE therapies were apparent. Full clinical and demographic characteristics of are summarised in Supplementary Table S2.

Relationships between annotated gene expression scores varies by patient ancestry

IFN-Score-A, IFN-Score-B and gene expression scores annotated to plasmablast, neutrophil, myeloid lineage, inflammation and erythropoiesis showed distinct profiles associated with patient ancestry. Consistent with previous literature, the IFN signature, measured by IFN-Score-A showed marked separation between European and non-European UK ancestries (Supplementary Fig.S3). Patients of European ancestry showed lower median expression of IFN-Score-A (*-1.72 vs -0.77*; W= 3827, p=0.0002) with a bimodal distribution which was not apparent among patients of non-European ancestry (Fig. 1A; Fig.S3). European ancestry patients also displayed lower IFN-Score-B (*-2.62 vs -2.19*; t = 2.13, p=0.034) and plasmablast Score (*-6.15 vs -5.15*; t= 3.73, p=0.0002; Fig.1A).

Among patients of European ancestry, gene expression scores across all annotations were closely aligned with IFN pathway activation. IFN-Score-B, which comprises ISGs sensitive to multiple IFN subtypes, showed significant positive correlation with Plasmablast (n=128, R^2= 0.265, p=0.002), neutrophil (R^2= 0.530, p<0.000), myeloid (R^2= 0.714, p<0.000), inflammation (R^2= 0.598, p<0.000) and erythropoiesis (R^2= 0.437, p<0.000) annotated scores (Fig. 1B, D).

In contrast, among patients of non-European ancestry, plasmablast and neutrophil gene expression scores were completely dissociated from IFN status. There was no significant correlation between IFN-Score-B and plasmablast (n=85, R^2= 0.001, p=0.990) or neutrophil Score (R^2= 0.109, p=0.318; Fig.1C,E) but, as observed in patients of European ancestry, a strong positive correlation was retained between IFN-Score-B and myeloid (R^2= 0.716, p<0.000), inflammation (R^2= 0.445, p<0.000) and

erythropoesis Scores (R^2= 0.296, p=0.006; Fig.1C,E). Similar relationships were observed with IFN-Score-A though the strength of correlation, where present, was weaker than for IFN-Score-B (Fig.1B,C and Supplementary Fig.S4). The same pattern was consistent across African ancestry and Subcontinental Asian patients when evaluated discretely (Supplementary Fig. S5).

Among patients of European ancestry both IFN Score-A and -B were both positively correlated with overall disease activity, with a stronger relationship for SLEDAI-2K (IFN-Score-A R^2= 0.366, p<0.000; IFN-Score-B R^2= 0.333, p<0.000) than numerical BILAG (IFN-Score-A R^2= 0.282, p=0.002; IFN-Score-B R^2= 0.224, p=0.013). In contrast, among patients of non-European ancestry, IFN status was not related to overall disease activity with no significant correlation between either IFN-Score and SLEDAI-2K (IFN-Score-A R^2= 0.159, p=0.156; IFN-Score-B R^2= 0.194, p=0.083) or numerical BILAG (IFN-Score-A R^2= 0.174, p=0.128; IFN-Score-B R^2= 0.133, p=0.247).

Several transcriptomic features were common to both ancestral groups. Specifically there was significant positive correlation between neutrophil, myeloid and inflammation annotated scores (Fig.1B,C) in both European ancestry and non-European ancestry subjects. Similarly in both ancestral groups IFN-score A showed significant positive correlation with myeloid, inflammation scores and erythropoesis scores (Supplementary Fig.S4) No score showed significant relationship with disease duration or current glucocorticoid doses (Fig.1B,C).

Gene expression scores show ancestry restricted associations with organ domain involvement

The relationship between gene expression scores and active BILAG-2004 grade A/B disease (compared with grade C or lower) varied between European and non-European ancestral groups. Among patients of European ancestry, active mucocutaneous and renal disease was associated with significantly higher mean IFN-Score-A (mucocutaneous domain *1.793 vs -2.930,* t= -2.65, p=0.008; renal domain *-1.375 vs -2.794*, t= -3.45, p=0.001) and IFN-Score-B (mucocutaneous domain *-2.308 vs -2.708*, t= -2.08, p= 0.040; renal domain *-0.196 vs -0.938,* t= -3.0, p=0.003). Active renal disease was strongly associated with higher neutrophil score *(-4.899 vs -5.994*, t= -3.5, p=0.000) in European ancestry patients. Active musculoskeletal disease was not distinguished by any score among European ancestry patients (Fig.2B).

In contrast, among patients of non-European ancestry IFN scores were not associated with active disease in any of these organ domains (Fig.2). However, active mucocutaneous disease in non-European ancestry patients was associated with higher mean plasmablast score (*-4.940 vs -5.581*, t= -2.2, p=0.033; Fig.2A). Active renal disease among non-European ancestry patients was conversely associated with lower plasmablast score (*-5.668 vs -5.031*, t= 2.5, p=0.028; Fig.2C). Unlike among European ancestry patients neutrophil score did not distinguish active and inactive renal involvement (Fig.2C) but active musculoskeletal disease was associated with lower neutrophil score (*-5.892 vs -4.958*, t= 2.1, p=0.018; Fig.2B).

The co-occurrence of BILAG-2004 grade A/B involvement across organ systems is shown in Fig.2.

Among patients of European ancestry, mucocutaneous disease was most prevalent overall and the most frequent patterns of organ involvement were; co-occurring mucocutaneous and musculoskeletal disease followed by single organ renal disease and single domain mucocutaneous disease (Fig.2D). Among non-European ancestry patients renal disease was most prevalent overall and single organ renal disease was most the frequent pattern of involvement, followed by single domain musculoskeletal disease and concurrently active renal and mucocutaneous activity (Fig.2E).

Transcript level clustering define distinct disease profiles in European and non-European Ancestry Patients

Unsupervised hierarchical clustering of expression across the 94 individual genes was performed for European ancestry and non-European ancestry patient subsets undergoing first cycle rituximab. Three patient clusters were each apparent in European and non-European ancestries, but the disease characteristics associated with transcriptional clusters varied by ancestry.

Non-European Ancestry Clusters

Among non-European subjects three clusters were observed (Fig.3). These were NEA-1: *interferon low, neutrophil-myeloid-inflammation high*; NEA-2: *interferon high, neutrophil-myeloid -inflammation low* and, NEA-3: *all signatures high*. Plasmablast and erythropoiesis scores were similar in all clusters. Age, disease duration, and anti-malarial use did not significantly differ between clusters. Moreover, there were no significant differences in baseline disease activity (SLEDAI-2K and numerical BILAG) or BILAG-2004 organ domain involvement between the three clusters (Table 1). The NEA-1 *IFN-low neutrophil-myeloid-inflammation high* cluster was most clinically and serologically distinct while NEA-2 and -3 were clinically and serologically similar despite markedly different transcriptional profile. Ancestral subgroups did not fully explain these clusters. Subcontinental Asian patients were equally represented across all three clusters. Patients of Chinese / Other Asian ancestry, though fewest in number, were exclusively located in the NEA-3 *all signatures high* cluster. African ancestry patients were found in all clusters but concentrated in NEA-2 *interferon high, neutrophil-myeloid -inflammation low* cluster (X^2= 13.9, p=0.029). Use of concurrent conventional immunosuppressants was lowest in NEA-1, highest among NEA-3 and intermediate in cluster NEA-2 (X^2= 6.08, p=0.047). There was a trend toward higher glucocorticoid requirement among patients in NEA-1 which was not statistically significant. There were differences in total peripheral white count (F= 6.6, p=0.007) and neutrophil count (F= 5.0, p=0.011) between clusters with significantly higher counts observed in NEA-1 than NEA-2, and NEA-3 lying between. There was a trend towards anaemia in NEA-1 (F=3.1, p=0.055) and a trend toward lymphopaenia in cluster NEA-3 (F= 3.2, p=0.051). The NEA-1 *IFN-low, neutrophil-myeloid-inflammation high* cluster was characterized by lower total IgG and lower seropositivity for U1RNP-Sm (X^2= 9.1, p=0.010) compared with the other two clusters. There were no significant differences in IgG level or autoantibody status between NEA-2 and NEA-3.

European Ancestry Clusters

Among patients of European ancestry three clusters were also evident. These were EA-1: *all signatures high*, EA-2: *IFN high, neutrophil-myeloid-inflammation-erythropoesis low*, and, EA-3 *all signatures low*

(Fig.4). There were no significant differences in plasmablast score between European ancestry gene expression clusters (F= 1.4, p=0.238) but unlike among non-European ancestry clusters, there were significant differences in erythropoesis score which paralleled those observed in the neutrophil, myeloid, inflammation annotated transcripts (Fig.4). There were no significant differences in age, disease duration or concurrent use of conventional immunosuppressants and antimalarials between European ancestry clusters (Table 2). However, in contrast to non-European Ancestry patients, clusters derived from European ancestry patients were significantly separated by disease activity as measured by SLEDAI (F= 4.2, p=0.018) and numerical BILAG (F= 4.4, p=0.014) and by BILAG-2004 organ domain involvement (Table 2). While mucocutaneous and musculoskeletal disease were similarly distributed across all clusters, BILAG-2004 grade A/B renal disease was highly concentrated in the EA-1 *all annotations high* cluster (X^2= 15.5, p< 0.000). Global disease activity was lowest in the EA-3 *all annotations low* cluster but did not significantly differ between clusters EA-1 and EA-2 despite their differential renal involvement. EA-3 who displayed lower disease activity, also showed significantly lower frequency of U1RNP-Sm seropositivity and a trend toward lower rates of hypocomplementaemia and dsDNA antibody positivity which was short of statistical significance. Clusters differed in mean neutrophil count (F= 4.9, p=0.010) and lymphocyte count (F= 5.5, p=0.006). The EA-1 *all signatures high cluster* characterized by high disease activity and renal involvement was the only cluster demonstrating lymphopenia <1.0 x10^9/L and also demonstrated higher neutrophil count than EA-2 (p=0.050) and EA-3 (p=0.013).

Of the transcriptional profiles identified, only the *all signatures high* profile was the only common to both ancestral groups (NEA-3 and EA-1) but its clinical associations were distinct between the two. *IFN high, neutrophil-myeloid-inflammation* low clusters could be distinguished between ancestries by Erythropoiesis-annotated transcripts while the *all signatures low* profile was unique to European ancestry patients. Supplementary Table S3 summarises key clusters characteristics.

Transcriptional profile differentially associates with response to rituximab in European and non-European ancestry patients

110 patients had evaluable follow-up data at six months following cycle 1 of rituximab (Supplementary Fig.S1). 70/110 (63%) achieved an overall treatment response by BILAG-2004 criteria. Response rate did not significantly differ between European (45/65, 69%) and non-European patients (25/45, 56%; X^2= 2.1, p=0.142) and response was not associated with socioeconomic deprivation (t= -0.1, p=0.936; Supplementary Fig.S6). Response was associated with reduction in median (IQR) oral glucocorticoid dose from 10 (5, 14) mg to 5 (0, 10) mg daily and no additional conventional immunosuppressant therapy was registered for any patient between baseline and 6 months. Supplementary Table 4 summarises the characteristics of responders and non-responders in both European ancestry and non-European ancestry subsets. Response by UK Census ancestral category is detailed in Supplementary Table 5.

Among non-European ancestry patients, transcriptional cluster membership significantly stratified response to rituximab (X^2= 14.5, p<0.001; Table 1). Non-European ancestry patients within the NEA-1

IFN-low, neutrophil-myeloid-inflammation high cluster, though fewest in number, showed distinctly poorer BILAG response to rituximab with only 12.5% achieving overall response. Both of the IFN-high non-European ancestry clusters (NEA-2 and NEA-3) achieved more favourable therapeutic response. The *all signatures high* NEA-3 cluster had highest response rate at 17/20 (85%). Cluster NEA-2 characterised by *IFN high, neutrophil-myeloid-inflammation low* signatures, though clinically and serologically similar to NEA-3 achieved a significantly lower rate of overall response at 41.2%. This distinctive rituximab response profile between NEA-1 and NEA-3 was maintained for each composite ancestral group. Heterogeneity in rituximab response between ancestral groups was most pronounced in NEA-2 (Supplementary Table S6).

In contrast, no significant difference in overall BILAG-2004 response was observed between European ancestry clusters (X^2= 2.1, p=0.353; Table 2). While the IFN-low cluster was adversely associated with treatment response among non-European ancestry cluster NEA-1, among European ancestry patients the IFN-low cluster EA-3 identified patients with lowest serological and clinical disease activity who in fact showed a trend towards more favourable response.

Discussion

Stratification incorporating interacting demographic, clinical and immunophenotypic features has the potential to assist individualized selection of therapies and improve overall outcomes for SLE patients. Using prospective registry evaluation of a multiethnic UK SLE cohort, we demonstrate that transcriptomic signatures differ between ancestral groups and differentially associate with response to rituximab. These results have implications for understanding the pathogenesis of SLE and improving stratification approaches for evaluating therapeutic interventions.

Epidemiological studies consistently demonstrate ethnic and geographic differences in the incidence and prevalence of SLE with disproportionate rates among Black and African American, Hispanic and Asian populations compared with White European ancestral groups (3). Patients from non-European ancestries demonstrate younger onset, greater renal involvement and, among African ancestry patients in particular, higher rates of secondary damage including atherosclerotic cardio- and cerebrovascular disease. (4, 33). Furthermore, racial and ethnic disparities in mortality appear only partially attenuated by socioeconomic and geographic factors (34).

Genetic and immunological studies suggest potential explanations for ancestral differences. More than 100 SLE susceptibility loci have been identified, with varied roles ranging from nucleic acid processing, IFN pathway and adaptive immune responses (35). Several genetic risk variants for SLE are not shared between ancestral groups, pointing to diverging heritable immunopathological mechanisms in different ancestral groups. For example, polymorphism in PTPN22, a negative T-cell regulator, associates with heightened risk of SLE in Hispanic and European populations but not among African ancestry groups (35, 36). Distinct genes and single nucleotide polymorphisms also associate with lupus nephritis risk among SLE patients of different ancestries (37). Notably, genetic variants in interferon regulatory factor (IRF) transcription factors IRF5 and IRF7 are all associated with SLE and risk haplotypes appear to exert ancestry-specific effects which are closely linked to serum IFN activity and auto-antibody profile (35, 38). Ancestral differences in DNA methylation associated with several ISGs has also been observed. (39). This heterogeneity may help explain why the relationship between IFN pathway activation and other transcriptomic annotations differed between clusters of SLE patients in this study. This observation supports previous analyses. Using a machine learning approach, Catalina et al., (21) found ancestry was the dominant influence on whole blood gene expression profiles in SLE, above sex, disease characteristics and therapeutics. Importantly, many modular signatures consistently differed between healthy individuals of different ancestries, with enrichment of granulocyte, inflammasome and monocyte scores among European ancestry subjects and activated T-cell and B-cell dominant signatures among African ancestry subjects.

The relationship between gene expression profiles and response to SLE therapies has not previously been investigated. Here we show that a selected transcriptomic profile associates with organ domain activity and predicts response to rituximab in an ancestry-specific fashion. While IFN signatures have been described as predictors of outcomes in SLE (40), our data indicate these are more informative when evaluated in combination with gene expression scores representing other key areas of the SLE

transcriptome, as has also recently been explored in juvenile SLE (41). Moreover, apparently similar transcriptional profiles yield distinct disease and prognostic associations for rituximab treatment dependent upon ancestral group. An *all signatures high* profile associated with highly rituximab responsive disease among non-European ancestry patients, while among European ancestry patients it was associated with greater renal involvement. In contrast, an *all signatures low* profile observed among European ancestry subjects lacked an equivalent among non-European ancestry patients. Meanwhile, other transcriptomic features were shared between ancestral groups, such as the correlations between IFN-Score-A and myeloid / inflammation signatures. These profiles could ultimately guide more optimised use of rituximab and may indicate a greater or lesser role for B-cells in these immunological subtypes but interpretation in an ancestry specific context appears critical.

In stratification studies it is often unclear whether biomarkers predict response to specific therapies or overall favorable disease natural history. While we do not have outcome data on other therapies or placebo, eligibility for rituximab in this study did require prior failure of either mycophenolate or cyclophosphamide. Another challenge in stratification studies is understanding the relationship between multiple interacting factors which influence response. Ancestry, autoantibody status, social deprivation and gene expression all have plausible impact on therapeutic response to rituximab. Indeed biobehavioural factors linked with sociodemographic conditions may also influence inflammation-related gene expression (42). Here, we show that stratification of response by gene expression profile was not influenced by major domains of social deprivation and could distinguish clusters not wholly explained by autoantibody status.

Among non-European ancestry patients we have identified a small but very distinctive cluster, NEA-1: *interferon low*, *neutrophil-myeloid-inflammation high* which demonstrated poorest response to rituximab. These patients showed high disease activity, including significant rates of active renal involvement but high rates of rituximab failure. Elevated B-cell and plasmablast activity, associated with RNP and dsDNA seropositivity appear more characteristic of non-European ancestry SLE patients, particularly those of African heritage (21). Indeed, in vitro evidence indicates IFN-I promotes differentiation of B-cells towards plasmablasts and plasma cells (43) and their polarization towards pro-inflammatory phenotypes (44). Expression of BAFF, a key mediator of B-cell dynamics can also be predicted by serum IFN activity and shows higher expression among African American SLE patients (45). Thus low-IFN, low antibody burden disease in patients of non-European ancestry as in cluster NEA-1 isolates a rituximab resistant patient subset, potentially with least B-cell dominant disease. This small cluster comprised substantial number of Subcontinental Asian patients, who have been sparsely evaluated in existing literature (46).

The relationship between autoantibodies, ancestry and interferon status is complex. Consistent with existing literature we observed RNP-Sm positivity, was enriched within both EA and NEA IFN-high clusters, though with higher prevalence and stronger associations in those of non-European ancestry (47). Previous studies reveal the IFN signature among European ancestry patients also associates with dsDNA seropositivity, and may be apparent independent of auto-antibodies. (48). Our data extends

understanding of this further by showing that gene expression scores outside of the IFN signature refine the clinical associations of the RNP -IFN interaction, particularly on rituximab responsive.

The current work has some limitations. Importantly, replication in a validation cohort is still required to verify the transcriptional clusters identified. Additionally, due to relatively lower numbers of minority ethnic groups our analysis focused on the non-European ancestries collectively. This work was therefore not able to fully explore heterogeneity within the non-European ancestry population and may be underpowered to detect specific features within our less represented groups. Similarly, as distribution of ancestral groups across clusters was not uniform, the influence of individual ancestries to cluster characteristics could not be fully delineated. Further efforts to evaluate ancestral groups discretely are needed. One further consideration is that this work made use of whole blood transcriptomic profile which has the advantage of relative simplicity for development as a clinically applicable platform, but does not permit interrogation of effects driven by differing immune cell population size which vary between ancestries. Similarly, this work used a specifically selected subset of transcripts pre-defined from microarray studies and thus may not capture the effect of other important transcripts which could influence response to rituximab. As this work did not include a placebo arm, it cannot account for differences in treatment response which are attributable to differences in the natural history of disease. We were also unable to account for differential depth of B-cell depletion between groups.

In conclusion, in a UK multiethnic refractory SLE cohort we observe distinct transcriptomic signatures in SLE which are differentiated by ancestral background and the relationship between IFN pathway activation and other annotated components of the SLE transcriptome. These profiles stratified response to rituximab in an ancestry-specific fashion and this relationship was not attributable to social deprivation or auto-antibody status. Finally, we observed a small subset of patients with active SLE, poor response rituximab who may have significant unmet needs not addressed by existing SLE therapies. The gene expression panel employed in this study should be further validated for prediction of response to rituximab. Other studies which aim to stratify lupus trials and develop biomarkers should consider ancestry, other demographic variables and patterns of organ involvement alongside overall response. This study adds to a body of work suggesting that there may be subtypes of SLE with less critical roles for B-cell as a therapeutic target.

Acknowledgements

We thank all the patients who have contributed to this research, clinical staff who supported patient recruitment and laboratory staff who undertook sample processing. MYMY's work was supported by NIHR Doctoral Research Fellowship (DRF-2014-07-155) and Wellcome Trust Institutional Strategic Fund Fellowship (204825/Z/16/Z). PE is Versus Arthritis Professor of Rheumatology. INB is funded by the NIHR Manchester BRC. EV's work was supported by NIHR Clinician Scientist award (CS-2013-13-032).

Data availability statement

The data underlying this article will be shared on reasonable request to the corresponding author.

Author Contributions

Lucy M Carter:	1a, 1b, 1c, 2, 3
Adewonuola Alase:	1a, 1b, 1c, 2, 3
Zoe Wigston:	1b, 1c, 2, 3
Antony Psarras:	1b, 1c, 2, 3
Agata Burska:	1b, 1c, 2, 3
Emily Sutton:	1b, 1c, 2, 3
Md Yuzaiful Md Yusof:	1b, 1c, 2, 3
John A Reynolds:	1a, 1b, 1c, 2, 3
Neil McHugh:	1a, 1b, 1c, 2, 3
Paul Emery:	1a, 1b, 1c, 2, 3
Miriam Wittmann:	1a, 1b, 1c, 2, 3
Ian N. Bruce:	1a, 1b, 1c, 2, 3
Edward M Vital:	1a, 1b, 1c, 2, 3

All authors have agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Tsokos GC, Lo MS, Costa Reis P, Sullivan KE. New insights into the immunopathogenesis of systemic lupus erythematosus. Nat Rev Rheumatol. 2016;12(12):716-30.

2. Williams JN, Dall'Era M, Lim SS, Feldman CH, Arntsen KA, Blazer AD, et al. Increasing Ancestral Diversity in Systemic Lupus Erythematosus Clinical Studies. Arthritis Care Res (Hoboken). 2020.

3. Barber MRW, Drenkard C, Falasinnu T, Hoi A, Mak A, Kow NY, et al. Global epidemiology of systemic lupus erythematosus. Nat Rev Rheumatol. 2021;17(9):515-32.

4. Bruce IN, O'Keeffe AG, Farewell V, Hanly JG, Manzi S, Su L, et al. Factors associated with damage accrual in patients with systemic lupus erythematosus: results from the Systemic Lupus International Collaborating Clinics (SLICC) Inception Cohort. Ann Rheum Dis. 2015;74(9):1706-13.

5. Plantinga L, Lim SS, Patzer R, McClellan W, Kramer M, Klein M, et al. Incidence of End-Stage Renal Disease Among Newly Diagnosed Systemic Lupus Erythematosus Patients: The Georgia Lupus Registry. Arthritis Care Res (Hoboken). 2016;68(3):357-65.

6. Fanouriakis A, Kostopoulou M, Alunno A, Aringer M, Bajema I, Boletis JN, et al. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. Annals of the Rheumatic Diseases. 2019;78(6):736-45.

7. Freitas S, Mozo Ruiz M, Costa Carneiro A, Isenberg DA. Why do some patients with systemic lupus erythematosus fail to respond to B-cell depletion using rituximab? Clin Exp Rheumatol. 2020;38(2):262-6.

8. McCarthy EM, Sutton E, Nesbit S, White J, Parker B, Jayne D, et al. Short-term efficacy and safety of rituximab therapy in refractory systemic lupus erythematosus: results from the British Isles Lupus Assessment Group Biologics Register. Rheumatology (Oxford). 2018;57(3):470-9.

8. McCarthy EM, Sutton E, Nesbit S, White J, Parker B, Jayne D, et al. Short-term efficacy and safety of rituximab therapy in refractory systemic lupus erythematosus: results from the British Isles Lupus Assessment Group Biologics Register. Rheumatology (Oxford). 2018;57(3):470-9.

9. Merrill JT, Neuwelt CM, Wallace DJ, Shanahan JC, Latinis KM, Oates JC, et al. Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: the randomized, double-blind, phase II/III systemic lupus erythematosus evaluation of rituximab trial. Arthritis Rheum. 2010;62(1):222-33.

10. D'Cruz D, Maksimowicz-McKinnon K, Oates J, Santiago MB, Bass D, Burriss S, et al. 200 Efficacy and safety of belimumab in patients of black race with systemic lupus erythematosus: results from the EMBRACE study. Lupus Science & Medicine. 2019;6(Suppl 1):A149-A50.

11. Chiche L, Jourde-Chiche N, Whalen E, Presnell S, Gersuk V, Dang K, et al. Modular transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal distinct type I and type II interferon signatures. Arthritis Rheumatol. 2014;66(6):1583-95.

12. Banchereau R, Hong S, Cantarel B, Baldwin N, Baisch J, Edens M, et al. Personalized Immunomonitoring Uncovers Molecular Networks that Stratify Lupus Patients. Cell. 2016;165(3):551-65.

13. Jourde-Chiche N, Whalen E, Gondouin B, Speake C, Gersuk V, Dussol B, et al. Modular transcriptional repertoire analyses identify a blood neutrophil signature as a candidate biomarker for lupus nephritis. Rheumatology (Oxford). 2017;56(3):477-87.

14. Zhu JL, Tran LT, Smith M, Zheng F, Cai L, James JA, et al. Modular gene analysis reveals distinct molecular signatures for subsets of patients with cutaneous lupus erythematosus. Br J Dermatol. 2021;185(3):563-72.

15. Cambridge G, Isenberg DA, Edwards JC, Leandro MJ, Migone TS, Teodorescu M, et al. B cell depletion therapy in systemic lupus erythematosus: relationships among serum B lymphocyte stimulator levels, autoantibody profile and clinical response. Ann Rheum Dis. 2008;67(7):1011-6.

16. Weckerle CE, Franek BS, Kelly JA, Kumabe M, Mikolaitis RA, Green SL, et al. Network analysis of associations between serum interferon-α activity, autoantibodies, and clinical features in systemic lupus erythematosus. Arthritis Rheum. 2011;63(4):1044-53.

17. Rai R, Chauhan SK, Singh VV, Rai M, Rai G. RNA-seq Analysis Reveals Unique Transcriptome Signatures in Systemic Lupus Erythematosus Patients with Distinct Autoantibody Specificities. PLoS One. 2016;11(11):e0166312.

18. El-Sherbiny YM, Psarras A, Md Yusof MY, Hensor EMA, Tooze R, Doody G, et al. A novel two-score system for interferon status segregates autoimmune diseases and correlates with clinical features. Sci Rep. 2018;8(1):5793-.

19. Md Yusof MY, Psarras A, El-Sherbiny YM, Hensor EMA, Dutton K, Ul-Hassan S, et al. 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):1432-9.

20. Oliveira JJ, Karrar S, Rainbow DB, Pinder CL, Clarke P, Rubio García A, et al. 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):152.

21. Catalina MD, Bachali P, Yeo AE, Geraci NS, Petri MA, Grammer AC, et al. Patient ancestry significantly contributes to molecular heterogeneity of systemic lupus erythematosus. JCI Insight. 2020;5(15).

22. Md Yusof MY, Shaw D, El-Sherbiny YM, Dunn E, Rawstron AC, Emery P, et al. Predicting and managing primary and secondary non-response to rituximab using B-cell biomarkers in systemic lupus erythematosus. Ann Rheum Dis. 2017;76(11):1829-36.

23. Vital EM, Dass S, Buch MH, Henshaw K, Pease CT, Martin MF, et al. B cell biomarkers of rituximab responses in systemic lupus erythematosus. Arthritis Rheum. 2011;63(10):3038-47.

24. Carter LM, Isenberg DA, Ehrenstein MR. 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):2672-9.

25. NHSEngland. Clinical Commissioning Policy Rituximab for refractory Systemic Lupus Erythematosus (SLE) in adults and post-pubescent children [200402P]. 2019.

26. Gordon C, Amissah-Arthur M-B, Gayed M, Brown S, Bruce IN, D'Cruz D, et al. The British Society for Rheumatology guideline for the management of systemic lupus erythematosus in adults: Executive Summary. Rheumatology. 2017;57(1):14-8.

27. Isenberg DA, Rahman A, Allen E, Farewell V, Akil M, Bruce IN, et al. BILAG 2004. Development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. Rheumatology (Oxford). 2005;44(7):902-6.

28. Yee CS, Cresswell L, Farewell V, Rahman A, Teh LS, Griffiths B, et al. Numerical scoring for the BILAG-2004 index. Rheumatology (Oxford). 2010;49(9):1665-9.

29. Betteridge Z, Gunawardena H, North J, Slinn J, McHugh N. Anti-synthetase syndrome: a new autoantibody to phenylalanyl transfer RNA synthetase (anti-Zo) associated with polymyositis and interstitial pneumonia. Rheumatology (Oxford). 2007;46(6):1005-8.

30. Tansley SL, Simou S, Shaddick G, Betteridge ZE, Almeida B, Gunawardena H, et al. Autoantibodies in juvenile-onset myositis: Their diagnostic value and associated clinical phenotype in a

large UK cohort. J Autoimmun. 2017;84:55-64.

31. McCall MN, McMurray HR, Land H, Almudevar A. On non-detects in qPCR data. Bioinformatics (Oxford, England). 2014;30(16):2310-6.

32. Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. UpSet: Visualization of Intersecting Sets. IEEE Trans Vis Comput Graph. 2014;20(12):1983-92.

33. Barbhaiya M, Feldman CH, Guan H, Gómez-Puerta JA, Fischer MA, Solomon DH, et al. Race/Ethnicity and Cardiovascular Events Among Patients With Systemic Lupus Erythematosus. Arthritis & Rheumatology. 2017;69(9):1823-31.

34. Yen EY, Shaheen M, Woo JMP, Mercer N, Li N, McCurdy DK, et al. 46-Year Trends in Systemic Lupus Erythematosus Mortality in the United States, 1968 to 2013: A Nationwide Population-Based Study. Ann Intern Med. 2017;167(11):777-85.

35. Goulielmos GN, Zervou MI, Vazgiourakis VM, Ghodke-Puranik Y, Garyfallos A, Niewold TB. The genetics and molecular pathogenesis of systemic lupus erythematosus (SLE) in populations of different ancestry. Gene. 2018;668:59-72.

36. Namjou B, Kim-Howard X, Sun C, Adler A, Chung SA, Kaufman KM, et al. PTPN22 association in systemic lupus erythematosus (SLE) with respect to individual ancestry and clinical sub-phenotypes. PLoS One. 2013;8(8):e69404.

37. Lanata CM, Nititham J, Taylor KE, Chung SA, Torgerson DG, Seldin MF, et al. Genetic contributions to lupus nephritis in a multi-ethnic cohort of systemic lupus erythematous patients. PLOS ONE. 2018;13(6):e0199003.

38. Niewold TB, Kelly JA, Kariuki SN, Franek BS, Kumar AA, Kaufman KM, et al. IRF5 haplotypes demonstrate diverse serological associations which predict serum interferon alpha activity and explain the majority of the genetic association with systemic lupus erythematosus. Ann Rheum Dis. 2012;71(3):463-8.

39. Joseph S, George NI, Green-Knox B, Treadwell EL, Word B, Yim S, et al. Epigenome-wide association study of peripheral blood mononuclear cells in systemic lupus erythematosus: Identifying DNA methylation signatures associated with interferon-related genes based on ethnicity and SLEDAI. J Autoimmun. 2019;96:147-57.

40. Psarras A, Emery P, Vital EM. Type I interferon-mediated autoimmune diseases: pathogenesis, diagnosis and targeted therapy. Rheumatology (Oxford). 2017;56(10):1662-75.

41. Wahadat MJ, Schonenberg-Meinema D, van Helden-Meeuwsen CG, van Tilburg SJ, Groot N, Schatorjé EJH, et al. Gene signature fingerprints stratify SLE patients in groups with similar biological disease profiles: a multicentre longitudinal study. Rheumatology. 2022.

42. Cole SW, Shanahan MJ, Gaydosh L, Harris KM. Population-based RNA profiling in Add Health finds social disparities in inflammatory and antiviral gene regulation to emerge by young adulthood. Proc Natl Acad Sci U S A. 2020;117(9):4601-8.

43. Jego G, Palucka AK, Blanck JP, Chalouni C, Pascual V, Banchereau J. Plasmacytoid dendritic cells induce plasma cell differentiation through type I interferon and interleukin 6. Immunity. 2003;19(2):225-34.

44. Care MA, Stephenson SJ, Barnes NA, Fan I, Zougman A, El-Sherbiny YM, et al. Network Analysis Identifies Proinflammatory Plasma Cell Polarization for Secretion of ISG15 in Human Autoimmunity. J Immunol. 2016;197(4):1447-59.

45. Ritterhouse LL, Crowe SR, Niewold TB, Merrill JT, Roberts VC, Dedeke AB, et al. B lymphocyte stimulator levels in systemic lupus erythematosus: higher circulating levels in African American patients and increased production after influenza vaccination in patients with low baseline levels. Arthritis

Rheum. 2011;63(12):3931-41.

46. Yee CS, Su L, Toescu V, Hickman R, Situnayake D, Bowman S, et al. Birmingham SLE cohort: outcomes of a large inception cohort followed for up to 21 years. Rheumatology (Oxford). 2015;54(5):836-43.

47. Hubbard EL, Pisetsky DS, Lipsky PE. Anti-RNP antibodies are associated with the interferon gene signature but not decreased complement levels in SLE. Ann Rheum Dis. 2022;81(5):632-43.

48. Ko K, Koldobskaya Y, Rosenzweig E, Niewold T. Activation of the Interferon Pathway is Dependent Upon Autoantibodies in African-American SLE Patients, but Not in European-American SLE Patients. Frontiers in Immunology. 2013;4(309).

Figure Legends

Figure 1. Expression and inter-relationship of annotated gene expression scores varies by patient ancestry

Violin plot series (A) with overlay jitter points show distribution of IFN-Score-A, IFN-Score-B, plasmablast, neutrophil, myeloid lineage, inflammation and erythropoiesis -annotated expression scores in patients of European ancestry (EA) and non-European ancestry (NEA). Bar indicates median expression. Matrix correlograms show by size and colour intensity, the strength of positive (blue) or negative (red) Spearman's correlation coefficient between annotated gene expression scores and selected clinical variables for patients of European ancestry (B) and non-European ancestry (C). Scatterplot series show relationship between plasmablast, neutrophil, myeloid lineage, inflammation and erythropoiesis annotated expression scores and interferon pathway activation as measured by IFN-Score-B in patients of European (D) and non-European ancestry (E). Regression line shown in blue and standard error in grey. All gene expression scores are shown as ΔCt from reference gene PPIA reflected across zero such that higher values indicate higher expression.

Figure 2. Gene expression scores differentially associated with BILAG-2004 domain activity in European and Non-European ancestry patients

Boxplot series (A-C) show IFN-Score-A, IFN-Score-B, plasmablast and neutrophil scores in patients of European ancestry (EA) and non-European ancestry (NEA) with BILAG-2004 grade A/B disease activity in the mucocutaneous (A, red), musculoskeletal (B, teal) and renal domain (C, purple) as compared with patients with lower (grade C or below) domain activity (grey). Upset plots for European ancestry (D) and on-European ancestry (E) patients show frequency of BILAG-2004 grade A/B activity (bar chart, intersection size) according to BILAG-2004 domain co-involvement (dot-connectivity plot, group). Upset plot horizontal bar chart (set size) shows the frequency of grade A/B activity by each individual BILAG-2004 domain.

Figure 3. Transcript level clustering in non-European ancestry patients identifies differential response to rituximab

Heatmap (A) shows expression (reflected ∆Ct) of 94 transcripts organized by module annotation, for 55 patients of non-European ancestry commencing cycle 1 of rituximab for active SLE within BILAG-BR. Side column identifies BILAG-2004 response to rituximab as responder (rose), non-responder (grey) or response undetermined (white). Dendrogram indicates by color, 3 clusters identified by unsupervised hierarchical clustering at transcript level. Boxplots (B-H) show significant differences in gene expression scores annotated to IFN-Score-A (B), IFN-Score-B (C), plasmablast (D), neutrophil (E), myeloid lineage (F), inflammation (G) and erythropoiesis (H) transcripts according to patient clusters derived from dendrogram. Cluster 1 (NEA-1; pale violet red) are IFN-low, neutrophil-myeloid-inflammation high; Cluster 2 (NEA-2; sea green) are IFN-high, neutrophil-myeloid-inflammation low, Cluster 3 (NEA-3; royal blue) show high expression across all annotations. Erythropoiesis-annotated expression did not differ between clusters. Response to rituximab (rose; heatmap side column) is highest in Cluster 3, lowest in Cluster 1 and intermediate in Cluster 2. Heatmaps are displayed centred and scaled by column

(transcript).

Figure 4. Transcript level clustering in European ancestry segregates patients by disease activity and renal involvement

Heatmap (A) shows expression (reflected ∆Ct) of 94 transcripts organized by module annotation, for 82 patients of European ancestry commencing cycle 1 of rituximab for active SLE within BILAG-BR. Side column identifies BILAG-2004 response to rituximab as responder (rose), non-responder (grey) or response undetermined (white). Dendrogram indicates by color, 3 clusters identified by transcript level unsupervised hierarchical clustering of gene expression. Boxplots (B-H) show significant differences in gene expression scores annotated to IFN-Score-A (B), IFN-Score-B (C), plasmablasts (D), neutrophil (E), myeloid lineage (F), inflammation (G) and erythropoiesis (H) transcripts according to patient clusters indicated on dendrogram. Cluster 1 (EA-1; purple) show high expression across all annotations, Cluster 2 (EA-2; grey) are IFN-high, neutrophil-myeloid-inflammation-erythropoesis low and Cluster 3 (EA-3; violet red) show low expression across all annotations. Heatmaps are displayed centred and scaled by column (transcript).

Tables

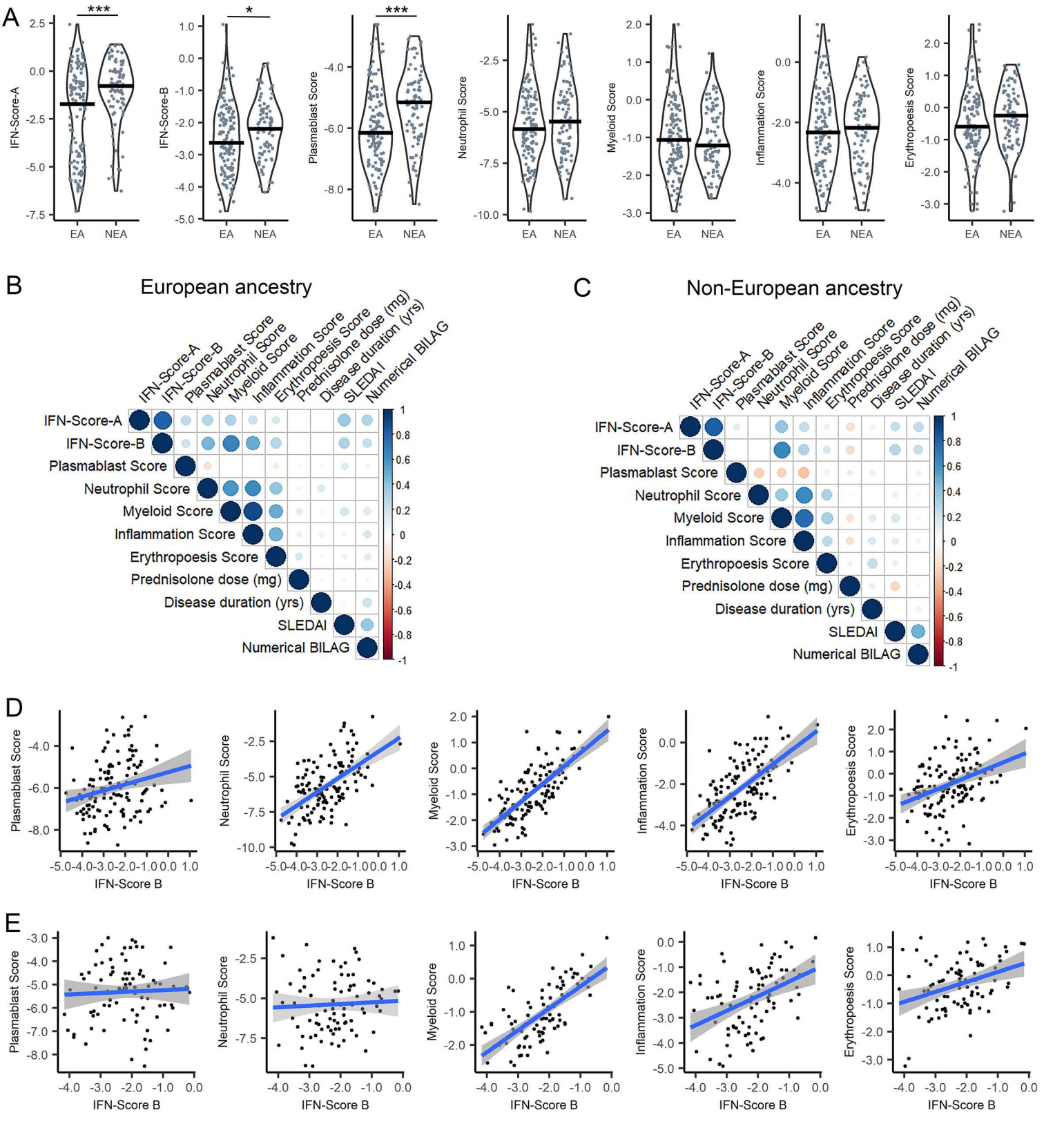
Table 1: Clusters derived from gene expression profile of non European ancestry patients commencing rituximab				
Clinical characteristics	**NEA Cluster 1 n = 9 IFN low, neutrophil-myeloid-inflammation high**	**NEA Cluster 2 n = 21 IFN high, neutrophil-myeloid-inflammation low**	**NEA Cluster 3 n = 25 All signatures high**	**p value**
Ancestry, n/N (% of ancestral group)				
African	1/18 (6)	11/18 (61)	6/18 (33)	
South Asian	5 (31)	5 (31)	6 (38)	**0.029**
Other Asian (incl. Chinese)	0 (0)	0 (0)	8 (100)	
Other (inc. Mixed)	3 (24)	5 (38)	5 (38)	
Female patient, (%)	8 (89)	17 (81)	22 (88)	0.756
Age (years), median (IQR)	45 (32, 50)	38 (27, 47)	32 (22, 39)	0.345
Disease duration (years), mean (95% CI)	9 (8, 21)	12 (7, 16)	13 (8, 20)	0.498
Current smoker (n/N, %)	4 (44)	2/11 (18)	1/19	0.104
Index of multiple deprivation (Rank), median (IQR)	7964 (4972, 23459)	7126 (2148, 13019)	16112 (7624, 23027)	0.079
BILAG A or B score, n (%)				
Constitutional	0 (0)	2	1 (4)	-
Mucocutaneous	2 (22)	9 (43)	9 (36)	0.559
Neuropsychiatric	0 (0)	4 (19)	2 (8)	-
Musculoskeletal	2 (22)	9 (43)	8 (32)	0.517
Cardiorespiratory	2 (22)	3 (14)	4 (16)	-
Gastroenterology	0 (0)	1 (4)	3 (12)	-
Ophthalmic	0 (0)	1 (4)	1 (4)	-
Renal	6 (67)	10 (48)	11 (44)	0.499
Haematology	1 (11)	2 (2)	1 (4)	-
BILAG numerical score, median (IQR)	15 (13, 20)	21 (13, 29)	14 (13, 21)	0.381
SLEDAI Score, median (IQR)	8 (4, 12)	8 (5, 14)	8 (4, 12)	0.674
SLICC damage index, median (IQR)	1 (0, 2)	1 (0, 2)	0 (0, 1)	0.281
Full blood count, median (IQR)				
Hb (g/L)	107.5 (100.2, 119.2)	124.5 (103.8, 138.2)	114.0 (102.0, 116.0)	0.055
WCC ($x10^9$ / L)	9.6 (9.0, 11.5)#	6.1 (3.7, 7.3)*	6.8 (4.1, 10.5)	**0.007**
Neutrophils ($x10^9$ / L)	7.6 (6.3, 9.2)#	4.5 (2.6, 5.6)*	6.1 (3.1, 9.0)	**0.011**
Lymphocytes ($x10^9$ / L)	1.9 (1.2, 2.3)	1.0 (0.8, 1.7)	0.8 (0.5, 1.0)	0.051
Platelets ($x10^9$ / L)	256 (131, 299)	230 (205, 318)	233 (203, 286)	0.707
Total IgG (g/L), median (IQR)	8.0 (6.9, 10.6)#	15.1 (12.3, 16.7)*	16.8 (12.5, 20.5)*	**0.033**
Low C3 or C4, n (%)	5 (56)	8 (38)	15 (50)	0.255
Concurrent Immunosuppressant, n (%)				

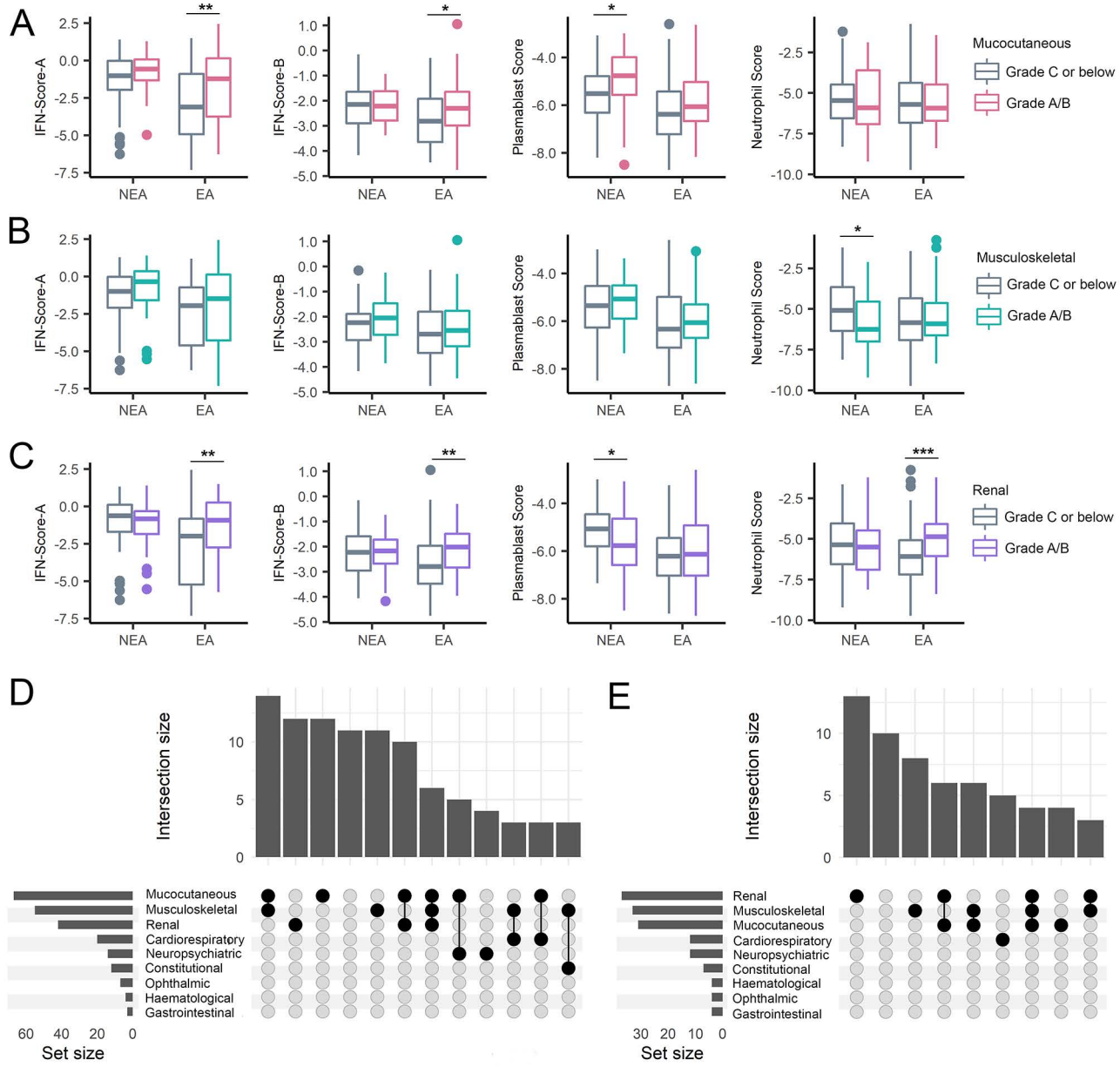
Table 1: Clusters derived from gene expression profile of non European ancestry patients commencing rituximab				
Clinical characteristics	**NEA Cluster 1** **n = 9** **IFN low, neutrophil-myeloid-inflammation high**	**NEA Cluster 2** **n = 21** **IFN high, neutrophil-myeloid-inflammation low**	**NEA Cluster 3** **n = 25** **All signatures high**	**p value**
Any agent (MMF, MTX, CNI, AZA)	0 (0)	6 (28)	11 (44)	**0.047**
Mycophenolate mofetil	0 (0)	4 (19)	9 (36)	0.076
Anti-malarial, n (%)	5 (56)	11 (52)	13 (52)	0.982
Oral glucocorticoid dose (mg), mean (95% CI)	20 (5, 20)	10 (9.25, 15)	10, (6, 10)	0.414
Immunoprecipitation and ELISA	**n = 8**	**n = 16**	**n = 19**	
U1RNP-Sm positive n (%)	0 (0)	8 (50)	12 (63)	**0.010**
Ro-60 n (%)	4 (50)	4 (25)	9 (47)	0.321
La	2 (25)	0 (0)	1 (5)	0.071
Ro-52 ELISA n (%)	2 (25)	1 (6)	5 (26)	0.276
dsDNA ELISA n (%)	4 (50)	10 (63)	15 (79)	0.296
Cardiolipin ELISA n (%)	2 (25)	1 (6)	2 (11)	0.393
Response to rituximab 6 months	**n = 8**	**n = 17**	**n = 20**	
BILAG responder (complete or partial), n (%)	1 (12.5)	7 (41.2)	17 (85.0)	**<0.001**
Post hoc analyses: * denotes significant difference from NEA-1, # denotes significant difference from NEA-2, p < 0.05				

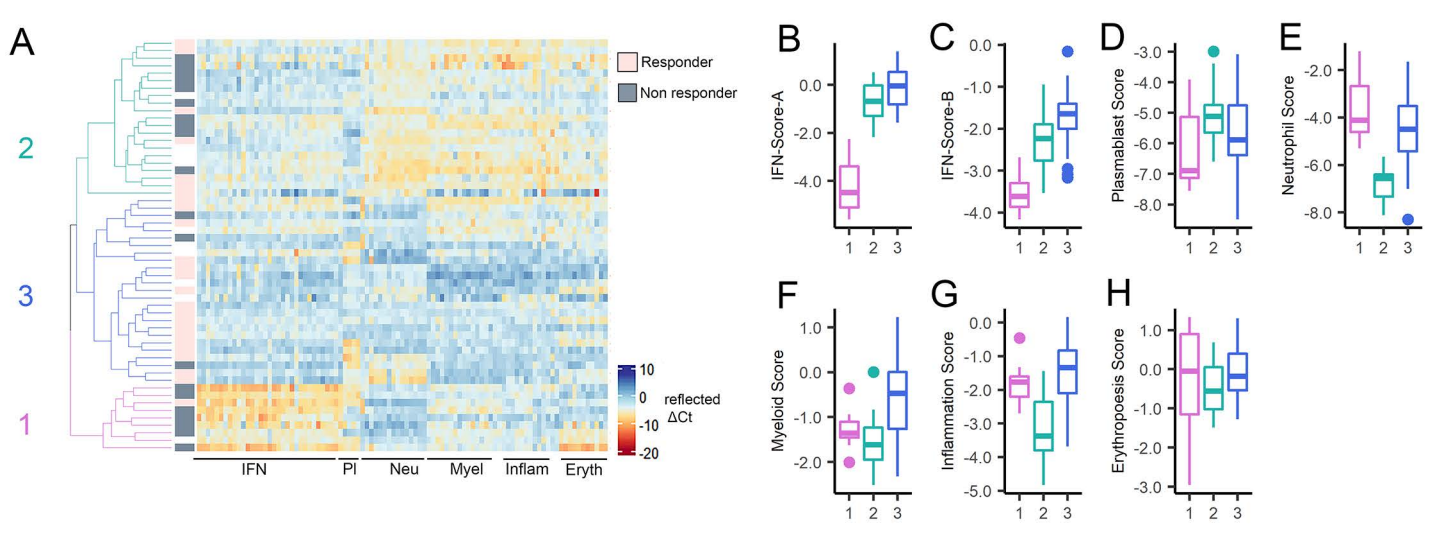
Table 2: Clusters derived from gene expression profile of European ancestry patients commencing rituximab

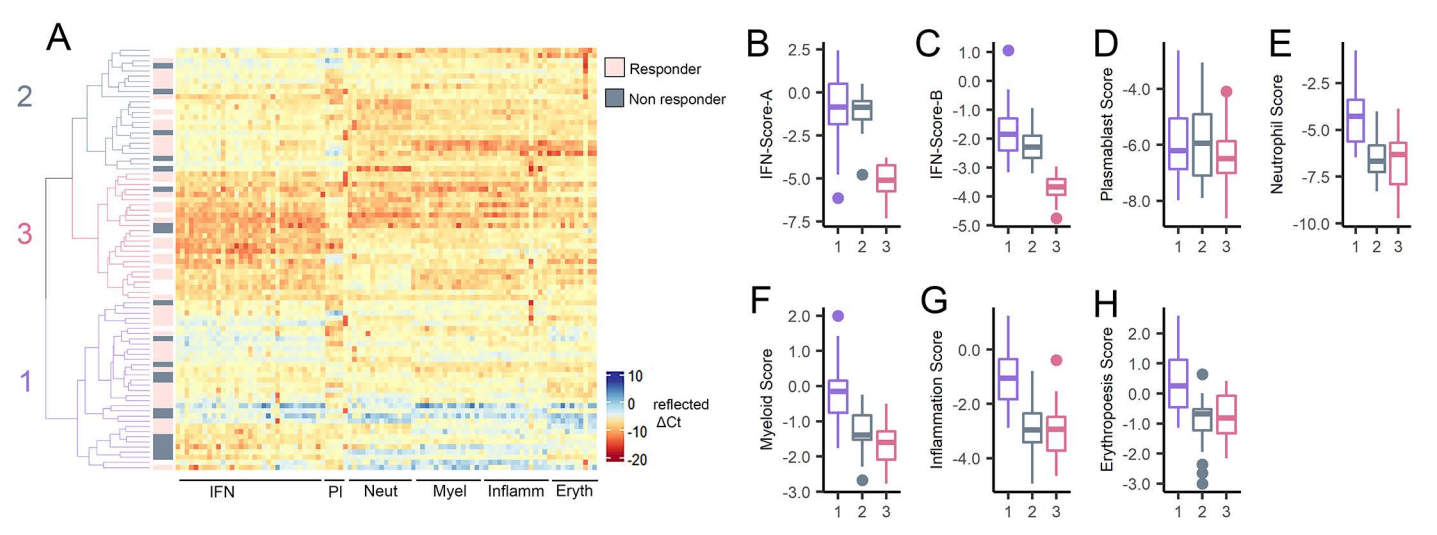
Clinical characteristics	European Cluster 1 n = 33 All signatures high	European Cluster 2 n = 24 IFN high, eeutrophil-myeloid-inflammation-erythropoiesis low	European Cluster 3 n = 25 All signatures low	p value
Ancestry, n/N (% of ancestral group)				
British	32 (97)	23 (96)	23 (92)	
Irish	0 (0)	1 (4)	2 (8)	0.402
Other	1 (3)	0 (0)	0 (0)	
Female patient, (%)	29 (88)	24 (100)	23 (92)	0.291
Age (years), median (IQR)	41 (33, 52)	43 (37, 50)	40 (32, 46)	0.482
Disease duration (years), mean (95% CI)	9 (6, 21)	14 (10, 17)	9 (7, 16)	0.618
Current smoker, n/N (%)	10/21 (48)	7/ 20 (35)	10/ 17 (58.5)	0.166
Index of multiple deprivation (Rank), median (IQR)	13051 (6083, 20015)	14353 (9926, 22000)	19709 (15340, 24186)	0.090
BILAG A or B score, n (%)				
Constitutional	5 (15)	2 (8)	3 (12)	-
Mucocutaneous	19 (58)	16 (67)	11 (44)	0.272
Neuropsychiatric	5 (15)	4 (17)	5 (20)	-
Musculoskeletal	17 (52)	12 (50)	8 (32)	0.284
Cardiorespiratory	6 (18)	3 (13)	3 (12)	-
Gastroenterology	0 (0)	2 (8)	1 (4)	-
Ophthalmic	0 (0)	1 (4)	2 (8)	-
Renal	20 (61)	5 (21)	4 (16)	**0.000**
Haematology	1 (3)	1 (4)	1 (4)	-
BILAG numerical score, mean (95% CI)	22 (16, 28)	21 (13, 24)	13 (9, 20)*	**0.014**
SLEDAI Score, median (IQR)	12 (8, 14)	8 (4, 11)	6 (2, 10)*	**0.018**
SLICC damage index, median (IQR)	1 (0, 2)	1 (0, 2)	0 (0, 1)	0.381
Full blood count, median (IQR)				
Hb (g/L)	121.0 (107.8, 131.8)	127.0 (116.0, 134.0)	126.0 (114.6, 139.0)	0.618
WCC ($x10^9$ / L)	7.2 (5.2, 11.7)	5.8 (5.4, 7.6)	6.3 (5.2, 7.2)	0.083
Neutrophils ($x10^9$ / L)	6.1 (3.6, 9.5)#	4.1 (3.3, 6.1)*	3.7 (2.9, 5.1)*	**0.010**
Lymphocytes ($x10^9$ / L)	0.9 (0.5, 1.3)	1.0 (0.8, 1.5)	1.5 (1.1, 2.1)*	**0.006**
Platelets ($x10^9$ / L)	279 (228, 397)#	250 (196, 278)*	276 (222, 355)	**0.044**
Total IgG (g/L), median (IQR)	10.8 (8.3, 12.8)	11.4 (9.5, 14.6)	10.15 (8.0, 12.0)	0.224
Low C3 or C4, n (%)	16 (48)	11 (46)	5 (20)	0.063
Concurrent Immunosuppressant, n (%)				
Any agent (MMF, MTX, CNI, AZA)	13 (39)	10 (42)	13 (52)	0.616

Table 2: Clusters derived from gene expression profile of European ancestry patients commencing rituximab				
Clinical characteristics	**European Cluster 1** **n = 33** **All signatures high**	**European Cluster 2** **n = 24** **IFN high, eeutrophil-myeloid-inflammation-erythropoiesis low**	**European Cluster 3** **n = 25** **All signatures low**	**p value**
Mycophenolate mofetil	11 (33)	6 (25)	7 (28)	0.611
Anti-malarial, n (%)	17 (74)	12 (50)	14 (56)	0.907
Oral glucocorticoid dose (mg), mean (95% CI)	11 (6, 20)	10 (7, 12)	11 (8, 25)	0.456
Immunoprecipitation and ELISA	**n = 30**	**n = 21**	**n = 22**	
U1RNP-Sm positive, n (%)	3 (10)	7 (33)	0 (0)	**0.004**
Ro-60 n (%)	8 (27)	7 (33)	5 (23)	0.733
La	1 (3)	1 (4)	1 (4)	0.932
Ro-52 ELISA n (%)	3 (10)	6 (25)	1 (4)	**0.038**
dsDNA ELISA n (%)	17 (56)	11 (52)	5 (23)	0.054
Cardiolipin ELISA n (%)	2 (7)	1 (4)	3 (14)	0.526
Response to rituximab 6 months	**n = 31**	**n = 18**	**n =16**	
BILAG responder (complete or partial), n (%)	19 (61)	13 (72)	13 (81)	0.353
Post hoc analyses: * denotes significant difference from EA-1, # denotes significant difference from EA-2, p < 0.05				


A
IFN-Score-A
IFN-Score-B
Plasmablast Score
Neutrophil Score
Myeloid Score
Inflammation Score
Erythropoesis Score
EA
NEA
***
*
B
European ancestry
C
Non-European ancestry
Prednisolone dose (mg)
Disease duration (yrs)
SLEDAI
Numerical BILAG
D
E
IFN-Score B


A
IFN-Score-A
IFN-Score-B
Plasmablast Score
Neutrophil Score
NEA
EA
Mucocutaneous
Grade C or below
Grade A/B
B
Musculoskeletal
Grade C or below
Grade A/B
C
Renal
Grade C or below
Grade A/B
D
Intersection size
Mucocutaneous
Musculoskeletal
Renal
Cardiorespiratory
Neuropsychiatric
Constitutional
Ophthalmic
Haematological
Gastrointestinal
Set size
E
Intersection size
Renal
Musculoskeletal
Mucocutaneous
Cardiorespiratory
Neuropsychiatric
Constitutional
Haematological
Ophthalmic
Gastrointestinal
Set size


A
Responder
Non responder
2
3
1
10
0
-10
-20
reflected
ΔCt
IFN
PI
Neu
Myel
Inflam
Eryth
B
IFN-Score-A
C
IFN-Score-B
D
Plasmablast Score
E
Neutrophil Score
F
Myeloid Score
G
Inflammation Score
H
Erythropoesis Score
1
2
3


A
2
3
1
Responder
Non responder
10
0
-10
-20
reflected
ΔCt
IFN
PI
Neut
Myel
Inflamm
Eryth
B
IFN-Score-A
C
IFN-Score-B
D
Plasmablast Score
E
Neutrophil Score
F
Myeloid Score
G
Inflammation Score
H
Erythropoesis Score
1
2
3

Gene expression and autoantibody analysis reveals distinct ancestry-specific profiles associated with response to rituximab in refractory SLE.

Supplementary Material

Contents:

Appendix 1: Supplementary Methods

Deprivation indices

Index of Multiple Deprivation is the official measure of relative deprivation in England assigned to a standard statistical geography each covering approximately 1,500 residents or 650 households. 2019 IMD combines seven deprivations domains (Income Deprivation, Employment Deprivation, Education, Skills and Training Deprivation, Health Deprivation and Disability, Crime, Barriers to Housing and Services, Living Environment Deprivation). Relative deprivation rank for 2019 IMD and it's composite domains from 1 (most deprived area) to 32,844 (least deprived area) were derived for patients based on postal address on enrollment to BILAG-Biologics Registry. The 2019 English indices of deprivation datasets, summaries and explanatory material are publicly available from UK Government Ministry of Housing, Communities & Local Government online at:

https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019 [accessed 01 February 2022]

ELISA

Analysis of all autoantibodies was performed by the specialist autoimmune serology laboratory at the University of Bath. Anti-dsDNA IgG, anti-Ro52 IgG and anti-cardiolipin IgG were measured using the following commercially available kits and according to the manufacturers instructions.

Target	Kit	Specificity	Catalogue no.
Anti-dsDNA IgG	Inova Diagnostics QUANTA Lite® dsDNA SC ELISA	IgG	708510
Anti-Ro52	Abnova SS-A 52 Ab ELISA Kit	IgG	KA1113
Anti-Cardiolipin IgG	Inova Diagnostics QUANTA Lite® ACA IgG III	IgG	708625

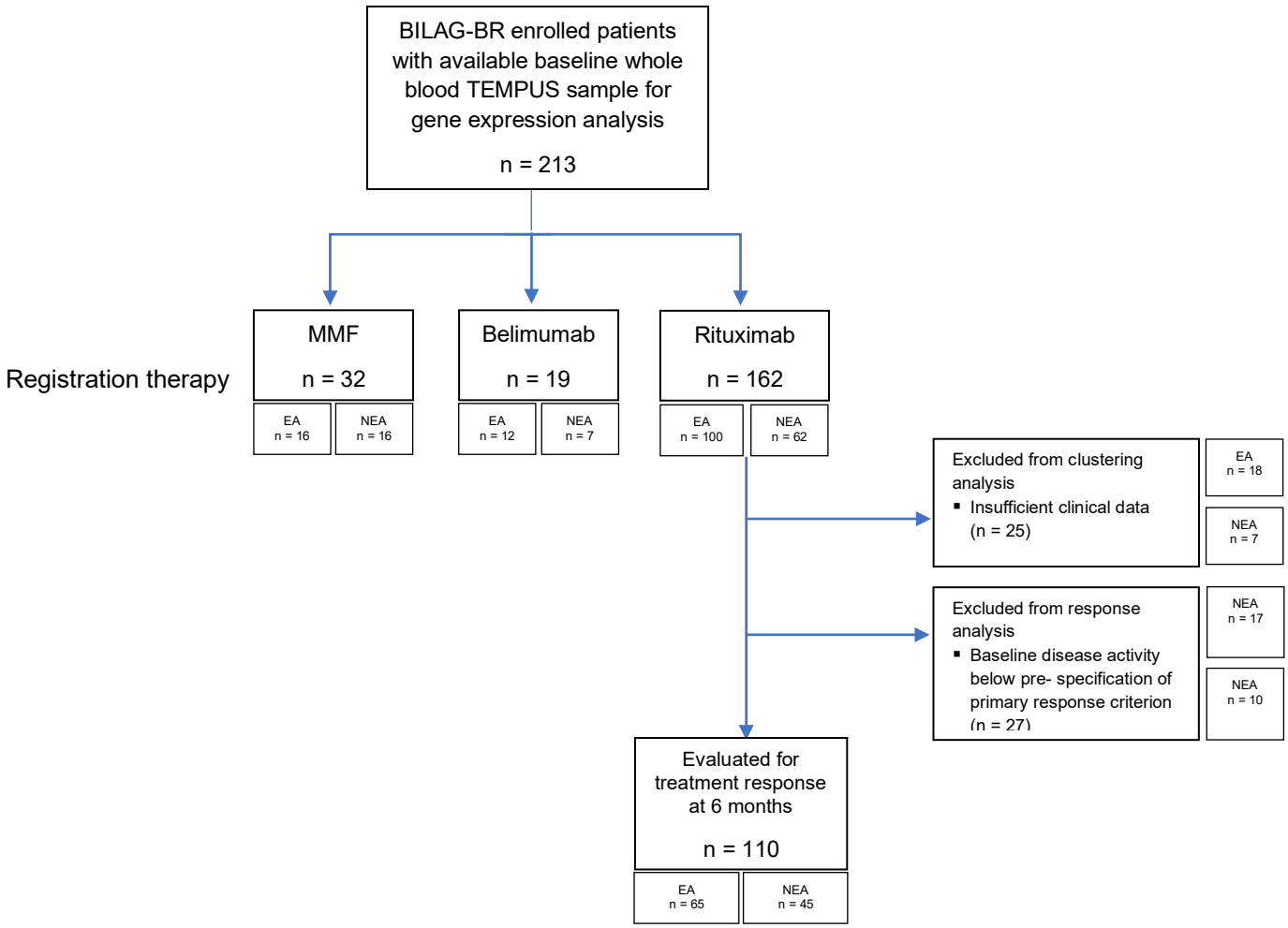


Figure S1. STROBE flow chart of BILAG-BR patients evaluated with gene expression and for response to cycle 1 of rituximab

MMF, mycophenolate mofetil; EA, European ancestry; NEA, Non-European ancestry

Table S1. Module assignment and Taqman Probe ID for the 94-gene expression panel

Gene Symbol	Name	Module	Annotation	Probe ID
IFI44L	Interferon induced protein 44 like	1.2	IFN Score A	hs00915292_m1
EIF2AK2	Eukaryotic Translation Initiation Factor 2 Alpha Kinase 2	3.4	ISG	hs00169345_m1
CCL8	C-C Motif Chemokine Ligand 8	3.4	IFN Score A	hs04187715_m1
ISG15	Interferon-stimulated gene 15	1.2	IFN Score A	hs00192713_m1
XAF1	X-linked inhibitor of apoptosis (XIAP)-associated factor 1	1.2	IFN Score A	hs01550142_m1
IFI44	Interferon Induced Protein 44	1.2	IFN Score A	hs00951349_m1
GBP1	Guanylate Binding Protein 1	3.4	IFN Score A	hs00977005_m1
IFI27	Interferon Alpha Inducible Protein 27	-	IFN Score A	hs01086373_g1
IRF7	Interferon Regulatory Factor 7	3.4	IFN Score A	hs01014809_g1
CXCL10	C-X-C motif chemokine ligand 10	1.2	IFN Score A	hs01124251_g1
CEACAM1	Carcinoembryonic antigen-related (CEA) cell adhesion molecule 1	3.4	IFN Score A	hs00989786_m1
RSAD2	Radical S-Adenosyl Methionine Domain Containing 2	1.2	IFN Score A	hs00369813_m1
IFIT1	Interferon Induced Protein With Tetratricopeptide Repeats 1	1.2	IFN Score A	hs01911452_s1
CXCL11	C-X-C motif chemokine ligand 11	-	ISG	hs00171138_m1
SIGLEC 1	Sialic Acid Binding Ig Like Lectin 1	-	ISG	hs00988063_m1
MX1	MX Dynamin Like GTPase 1)	1.2	ISG	hs00895608_m1
IFI6	IFI6 interferon alpha inducible protein 6	-	ISG	hs00242571_m1
HERC5	HECT And RLD Domain Containing E3 Ubiquitin Protein Ligase 5	1.2	ISG	hs00180943_m1
IFIH1	Interferon induced with helicase C domain 1	3.4	IFN Score B	hs01070332_m1
CASP1	Caspase 1	5.12	ISG	hs00354836_m1
SOCS1	Suppressor of cytokine signaling 1	3.4	IFN Score B	hs00705164_s1
SERPING1	Serpin Family G Member 1	1.2	IFN Score B	hs00163781_m1
NT5C3B	5'-Nucleotidase, Cytosolic IIIB	5.12	IFN Score B	hs00369454_m1
UNC93B1	Unc-93 Homolog B1	5.12	IFN Score B	hs00276771_m1
SP100	SP100 Nuclear Antigen	5.12	IFN Score B	hs00162109_m1
TRIM38	Tripartite Motif Containing 38	5.12	IFN Score B	hs00197164_m1
IFI16	Interferon Gamma Inducible Protein 16	5.12	IFN Score B	hs00194261_m1
BST2	Bone Marrow Stromal Cell Antigen 2	5.12	IFN Score B	hs01561315_m1
TAP1	Transporter 1, ATP Binding Cassette Subfamily B Member	5.12	IFN Score B	hs00388675_m1
STAT1	Signal transducer and activator of transcription 1	3.4	IFN Score B	hs01013996_m1
UBE2L6	Ubiquitin Conjugating Enzyme E2 L6	3.4	IFN Score B	hs01125548_m1
LAMP3	Lysosome-associated membrane glycoprotein 3	1.2	IFN Score B	hs00180880_m1
PHF11	PHD finger protein 11	5.12	IFN Score B	hs00211573_m1
DERL3	Derlin 3	7.7	Plasmablast	hs00405322_m1
TNFRSF17	TNF Receptor Superfamily Member 17	4.11	Plasmablast	hs00171292_m1
TXNDC5	Thioredoxin Domain Containing 5	4.11	Plasmablast	hs00229373_m1
IGJ	Immunoglobulin J Chain	4.11	Plasmablast	hs00376160_m1
HP	Haptoglobin	5.15	Neutrophil	hs00978377_m1
TCN1	Transcobalamin 1	5.15	Neutrophil	hs01055542_m1
ELA2	Neutrophil elastase	5.15	Neutrophil	hs00236952_m1
OLR1	Oxidized Low Density Lipoprotein Receptor 1	5.15	Neutrophil	hs00234028_m1

CEACAM6	CEA Cell Adhesion Molecule 6	5.15	Neutrophil	hs00366002_m1
AZU1	Azurocidin 1	5.15	Neutrophil	hs01106962_m1
ARG1	Arginase 1	5.15	Neutrophil	hs00968978_m1
CEACAM8	CEA Cell Adhesion Molecule 8	5.15	Neutrophil	hs00266198_m1
CAMP	Cathelicidin Antimicrobial Peptide	5.15	Neutrophil	hs00189038_m1
MMP8	Matrix metalloproteinase-8	5.15	Neutrophil	hs01029060_m1
MPO	Myeloperoxidase	5.15	Neutrophil	hs00924296_m1
CKAP4	Cytoskeleton Associated Protein 4	3.2	Myeloid lineage	hs00199135_m1
DEFA3	Defensin Alpha 3	5.15	Neutrophil	hs00414018_m1
DEFA1	Defensin Alpha 1	5.15	Neutrophil	hs00234383_m1
LTF	Lactotransferrin	5.15	Neutrophil	hs00914330_m1
DEFA4	Defensin Alpha 4	5.15	Neutrophil	hs010566650_m1
LY96	Lymphocyte Antigen 96	3.2	Myeloid lineage	hs01026734_m1
BST1	Bone Marrow Stromal Cell Antigen 1	3.2	Myeloid lineage	hs01070189_m1
BCL6	B-Cell Lymphoma 6	3.2	Myeloid lineage	hs01115889_m1
ARHGAP9	Rho GTPase Activating Protein 9	5.7	Myeloid Lineage	hs01037142_m1
MKNK1	MAPK Interacting Serine/Threonine Kinase 1	3.2	Myeloid lineage	hs00374375_m1
NCF4	Neutrophil Cytosolic Factor 4	3.2	Myeloid lineage	hs01055674_m1
HCK	HCK Proto-Oncogene, Src Family Tyrosine Kinase	3.2	Myeloid lineage	hs01067412_m1
ITGAM	Integrin Subunit Alpha M	3.2	Myeloid lineage	hs01064804_m1
CD63	CD63 antigen	3.2	Myeloid lineage	hs00156390_m1
IL17R(IL17RA)	Interleukin 17A receptor	3.2	Myeloid lineage	hs00234888_m1
GPR97	Adhesion G Protein-Coupled Receptor G3	3.2	Myeloid lineage	hs00416888_m1
CD55	CD55 molecule	3.2	Myeloid lineage	hs00892618_m1
ANXA3	Annexin A3	4.2	Inflammation	hs00971411_m1
CR1	Complement C3b/C4b Receptor 1	4.2	Inflammation	hs01079080_m1
PBEF1	Pre-B-Cell Colony-Enhancing Factor 1	3.2	Myeloid lineage	hs00237184_m1
H3F3B	H3.3 Histone B	5.7	Myeloid Lineage	hs00855159_g1
HMGB2	High Mobility Group Box 2	3.2	Myeloid lineage	hs01128615_m1
PIK3CD	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Delta	5.7	Myeloid Lineage	hs00192399_m1
SLPI	Secretory Leukocyte Peptidase Inhibitor	4.2	Inflammation	hs00268204_m1
FCAR	Fc Fragment Of IgA Receptor	4.2	Inflammation	hs00370197_m1
MAPK14	Mitogen-Activated Protein Kinase 14	4.2	Inflammation	hs01051152_m1
IL18RAP	Interleukin 18 Receptor Accessory Protein	4.2	Inflammation	hs00977702_m1
CLEC4D	C-Type Lectin Domain Family 4 Member D	4.2	Inflammation	hs01073581_m1
MCEMP1	Mast Cell Expressed Membrane Protein 1	4.2	Inflammation	hs00545332_m1
NFIX	Nuclear factor I/X	2.3	Erythropoiesis	hs00958850_m1
S100A12	S100 Calcium Binding Protein A12	4.2	Inflammation	hs00194525_m1
S100P	S100 Calcium Binding Protein P	4.2	Inflammation	hs00195584_m1
MMP9	Matrix Metalloproteinase 9	4.2	Inflammation	hs00957562_m1
SERPINA1	Serpin Family A Member 1	4.2	Inflammation	hs00165475_m1
PGLYRP1	Peptidoglycan Recognition Protein 1	4.2	Inflammation	hs00175475_m1
GUK1	Guanylate Kinase 1	3.1	Erythropoiesis	hs00176133_m1
MAP2K3	Mitogen-Activated Protein Kinase Kinase 3	3.1	Erythropoiesis	hs03005115_m1
ADIPOR1	Adiponectin Receptor 1	3.1	Erythropoiesis	hs00360422_m1

FLCN	Folliculin	3.1	Erythropoiesis	hs00376065_m1
IGF2BP2	Insulin Like Growth Factor 2 MRNA Binding Protein 2	3.1	Erythropoiesis	hs01118006_m1
PLEK2	Pleckstrin 2	2.3	Erythropoiesis	hs01026898_m1
PIP5K2A	Phosphatidylinositol-4-Phosphate 5-Kinase Type II Alpha	2.3	Erythropoiesis	hs01124167_m1
MXI1	MAX Interactor 1, Dimerization Protein	2.3	Erythropoiesis	hs00365651_m1
BCL2L1	BCL2 Like 1	2.3	Erythropoiesis	hs04408449_m1
GSPT1	G1 To S Phase Transition 1	3.1	Erythropoiesis	hs01093019_m1
PPIA	Peptidylprolyl Isomerase A	-	Reference	hs99999904_m1

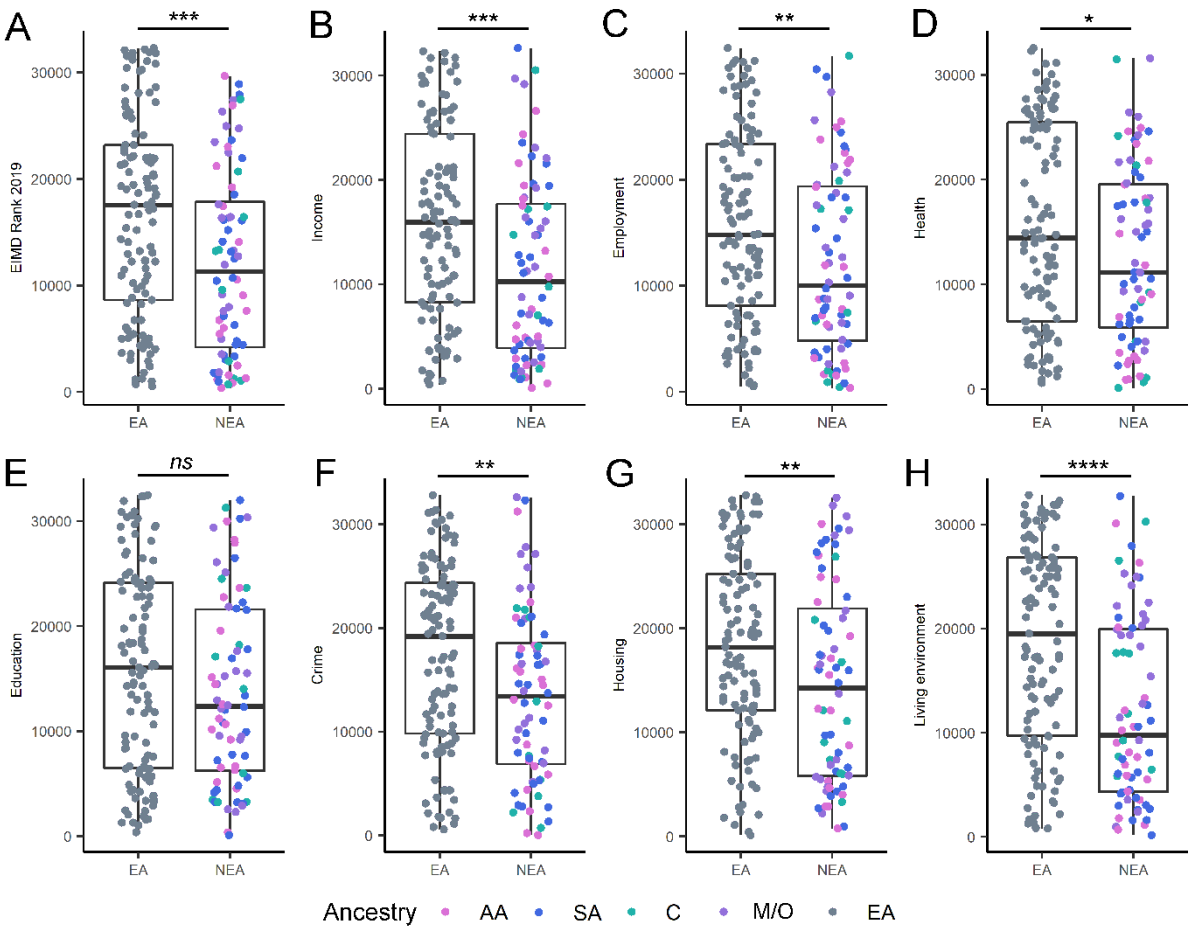


Figure S2. BILAG-BR patients of non-European ancestry reside in areas of greater relative deprivation.

Boxplot series with overlay jitter points color coded by ancestral group, show ranked deprivation indices from 1 (most deprived area) to 32,844 (least deprived area). Compared with patients of European ancestry (EA), patients of non-European ancestry (NEA) reside in areas of significantly higher overall deprivation as measured by the 2019 English Index of Multiple Deprivation (EIMD, **A**). Composite domains of deprivation Income (**B**), Employment (**C**), Health (**D**), Education (**E**), Crime (**F**), Barriers to Housing and Services (**G**) and Living Environment (**H**) show NEA patients reside in areas of significantly higher relative deprivation across all domains with the exception of Education (**E**). * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$, *ns* non significant. AA, African ancestry; SA, Subcontinental Asian; C, Chinese and other Asian; O/M Other or Mixed heritage; EA, European ancestry.

Table S2. Baseline characteristics of BILAG-BR patients of European and non-European ancestry.				
Clinical characteristics	**Overall** **n = 213**	**European Ancestry** **n = 128**	**Non-European Ancestry** **n = 85**	**p value**
Female patient, (%)	196 (92)	121 (94%)	75 (88.2)	0.09
Age (years), median (IQR)	39 (30, 50)	41 (34, 52)	36 (26, 47)	**0.001**
Disease duration (years), median (IQR)	11 (7, 18)	12 (7, 18)	11 (7, 16)	0.569
Current smoker (n/N, %)	47/142 (33)	37/85 (44)	10/52 (19)	**0.001**
Index of multiple deprivation (Rank), median (IQR)	14100 (5993, 21990)	17526 (8633, 23176)	11311 (4197, 17847)	**0.002**
BILAG-BR registration therapy				
Belimumab	19 (9)	12 (9)	7 (8)	
Mycophenolate mofetil	32 (15)	16 (13)	16 (19)	0.549
Rituximab	162 (76)	100 (78)	62 (73)	
BILAG A or B score, n (%)				
Constitutional	19 (9.0)	12 (9.4)	7 (8.2)	
Mucocutaneous	98 (46.0)	67 (52.3)	31 (24.2)	**0.029**
Neuropsychiatric	26 (12.2)	14 (10.9)	12 (14.1)	
Musculoskeletal	88 (41.3)	55 (42.9)	33 (38.8)	
Cardiorespiratory	32 (15.0)	20 (15.6)	12 (14.1)	0.644
Gastroenterology	7 (3.2)	3 (2.3)	4 (4.7)	
Ophthalmic	11 (5.1)	7 (5.4)	4 (4.7)	
Renal	79 (37.1)	42 (32.8)	37 (44)	0.078
Haematology	8 (3.7)	4 (3)	4 (4.7)	
BILAG numerical score, median (IQR)	18 (12, 24)	17 (13, 25)	19 (12, 24)	0.642
SLEDAI Score, median (IQR)	8 (4, 13)	8 (4, 12)	8 (4, 14)	0.274
SLICC damage index, median (IQR)	0 (0,1)	0 (0,1)	0 (0, 1)	0.695
Full blood count, median (IQR)				
Hb (g/L)	121.0 (108.0, 133.0)	124.0 (115.0, 134.5)	115.5 (102.2, 127.0)	**0.004**
WCC (x10^9 / L)	6.5 (4.5, 9.0)	6.5 (4.8, 8.9)	6.5 (4.0, 9.3)	0.976
Neutrophils (x10^9 / L)	4.6 (2.9, 7.0)	4.6 (2.9, 6.7)	4.7 (2.9,7.5)	0.866
Lymphocytes (x10^9 / L)	1.0 (0.7, 1.6)	1.1 (0.7, 1.6)	1.0 (0.7, 1.5)	0.311
Platelets (x10^9 / L)	261 (201, 331)	268 (205, 345)	253 (200, 307)	0.080
Total IgG (g/L), median (IQR)	12.9 (9.5, 16.7)	10.9 (8.5, 14.25)	16.2 (12.5, 20.9)	**<0.000**
Low C3 or C4, n (%)	103 (48)	55 (43.0)	48 (56.5)	**0.045**
Concurrent Immunosuppressant, n (%)				
Any agent (MMF, MTX, CNI, AZA)	104 (49)	65 (50.7)	39 (45.8)	0.483
Mycophenolate mofetil	77 (36)	45 (35.2)	32 (37.6%)	0.711
Anti-malarial, n (%)	109 (51)	65 (50.7)	44 (51.8)	0.888
Oral glucocorticoid dose (mg), median (IQR)	10 (7.5, 20)	10 (5, 20)	10 (7.5, 20)	0.180

Table S2. Baseline characteristics of BILAG-BR patients of European and non-European ancestry.				
Clinical characteristics	**Overall** **n = 213**	**European Ancestry** **n = 128**	**Non-European Ancestry** **n = 85**	**p value**
Immunoprecipitation and ELISA	**n = 158**	**n = 96**	**n = 62**	
U1RNP-Sm positive, n (%)	42 (27)	11 (11.5)	31 (50.0)	**<0.000**
Ro-60, n (%)	56 (35)	28 (29.2)	28 (45.1)	**0.040**
La,	11 (7)	7 (7.3)	4 (6.5)	0.839
Ro-52 ELISA, n (%)	30 (19)	16 (16.7)	14 (22.6)	0.354
dsDNA ELISA n (%)	86 (54)	44 (45.8)	42 (67.7)	**0.006**
Cardiolipin ELISA n (%)	20 (13)	12 (12.5)	8 (12.9)	0.960
Response to rituximab 6 months	**n = 110**	**n = 65**	**n = 45**	
BILAG responder (complete or partial), n (%)	70 (63)	45 (69)	25 (56)	0.142

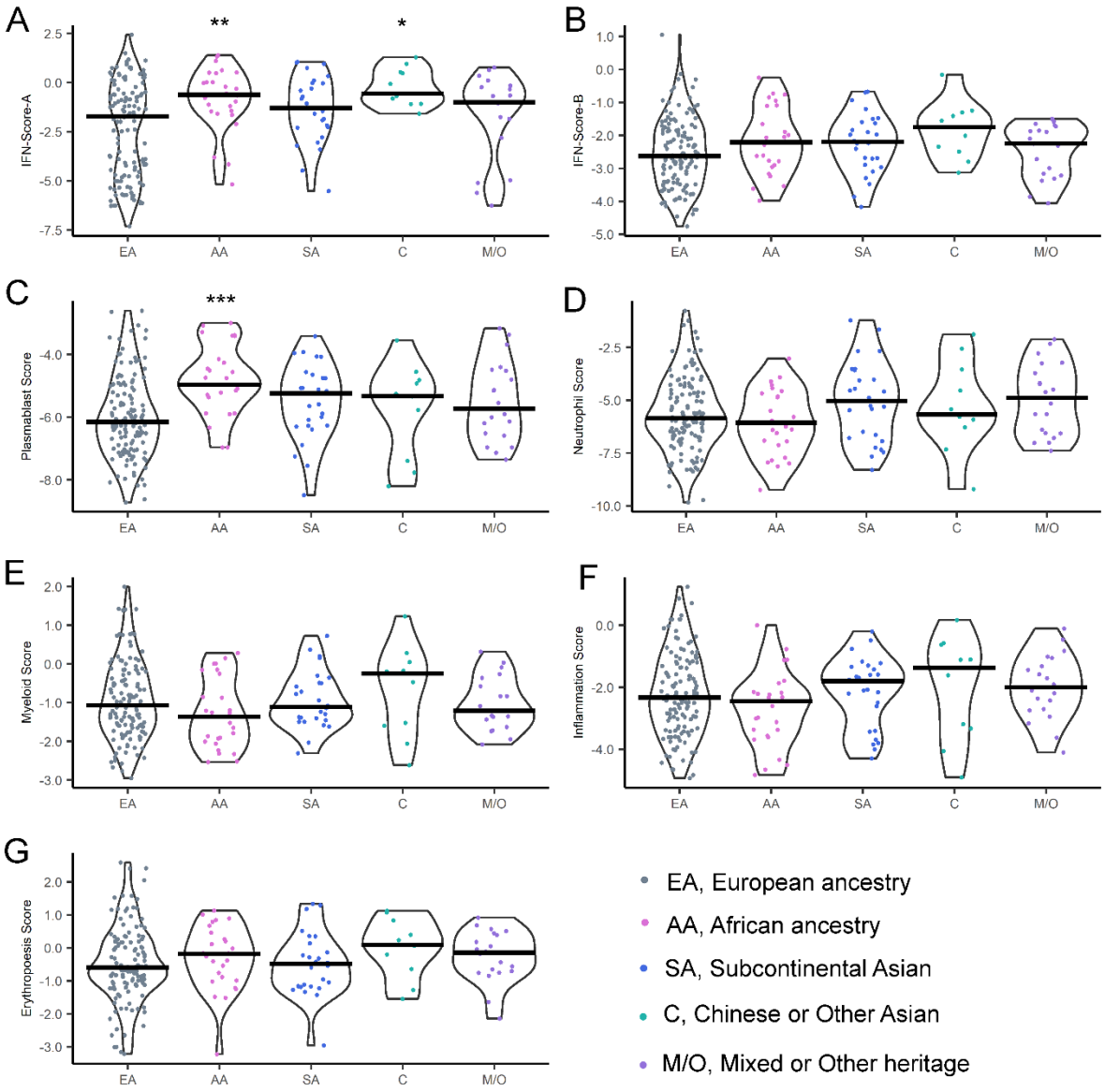


Figure S3. Interferon and Plasmablast annotated gene expression scores vary by ancestral background.

Violin plot series with overlay jitter points color coded by ancestral group, show distribution and median (horizontal bar) expression of IFN-Score-A (**A**) which comprises interferon stimulated genes (ISGs) typical of the global type-I IFN signature, IFN-Score-B (**B**) comprising more diversely regulated ISGs, Plasmablast (**C**), Neutrophil (**D**), Myeloid (**E**), Inflammation (**F**) and Eyrthropoesis (**G**) - annotated gene expression scores. Consistent with existing literature, patients of European ancestry show lower IFN-Score-A expression, but also display a bimodal distribution of IFN-Score-A which is not observed in patients of non-European ancestries. Significantly higher Plasmablast Score is evident among patients of African Ancestry. * $p \leq 0.05$, ** $p \leq 0.01$ for comparison against European ancestry (EA) group by Kruskal Wallace test and post-hoc pairwise Dunn's test.

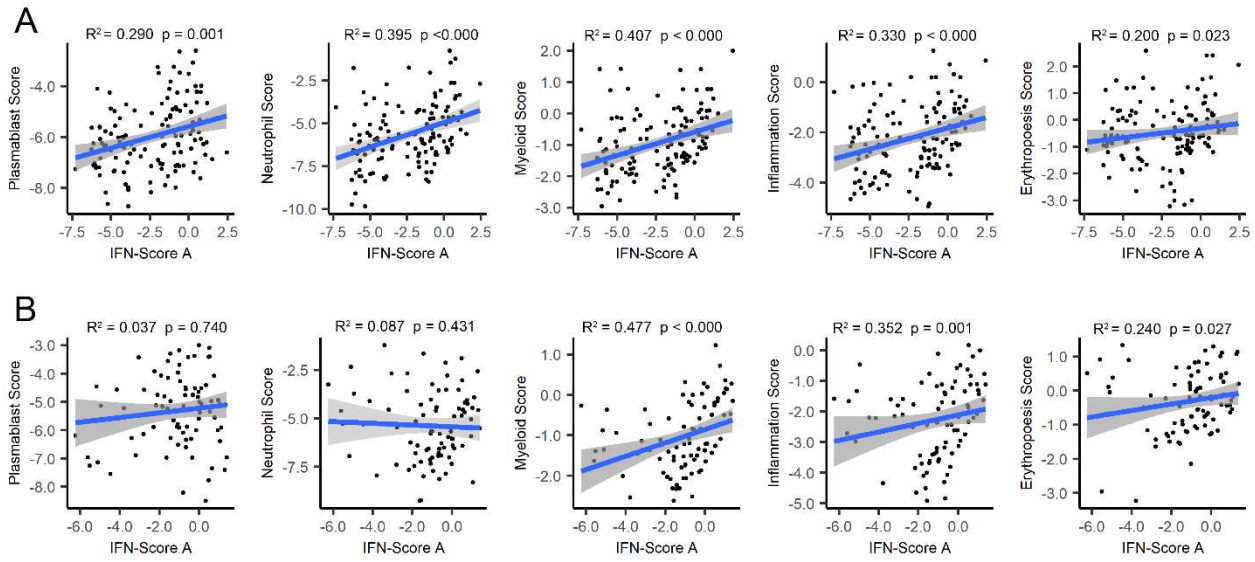


Figure S4. Relationship between IFN-Score-A and other functionally annotated gene expression scores varies by patient ancestry.

Scatterplot series show relationship between Plasmablast, Neutrophil, Myeloid lineage, Inflammation and Eyrthropoesis annotated expression scores and interferon pathway activation as measured by IFN-Score-A in patients of European (**A**) and non-European ancestry (**B**). Regression line shown in blue and standard error in grey. All gene expression scores are shown as ΔCt from reference gene PPIA reflected across zero such that higher values indicate higher expression.

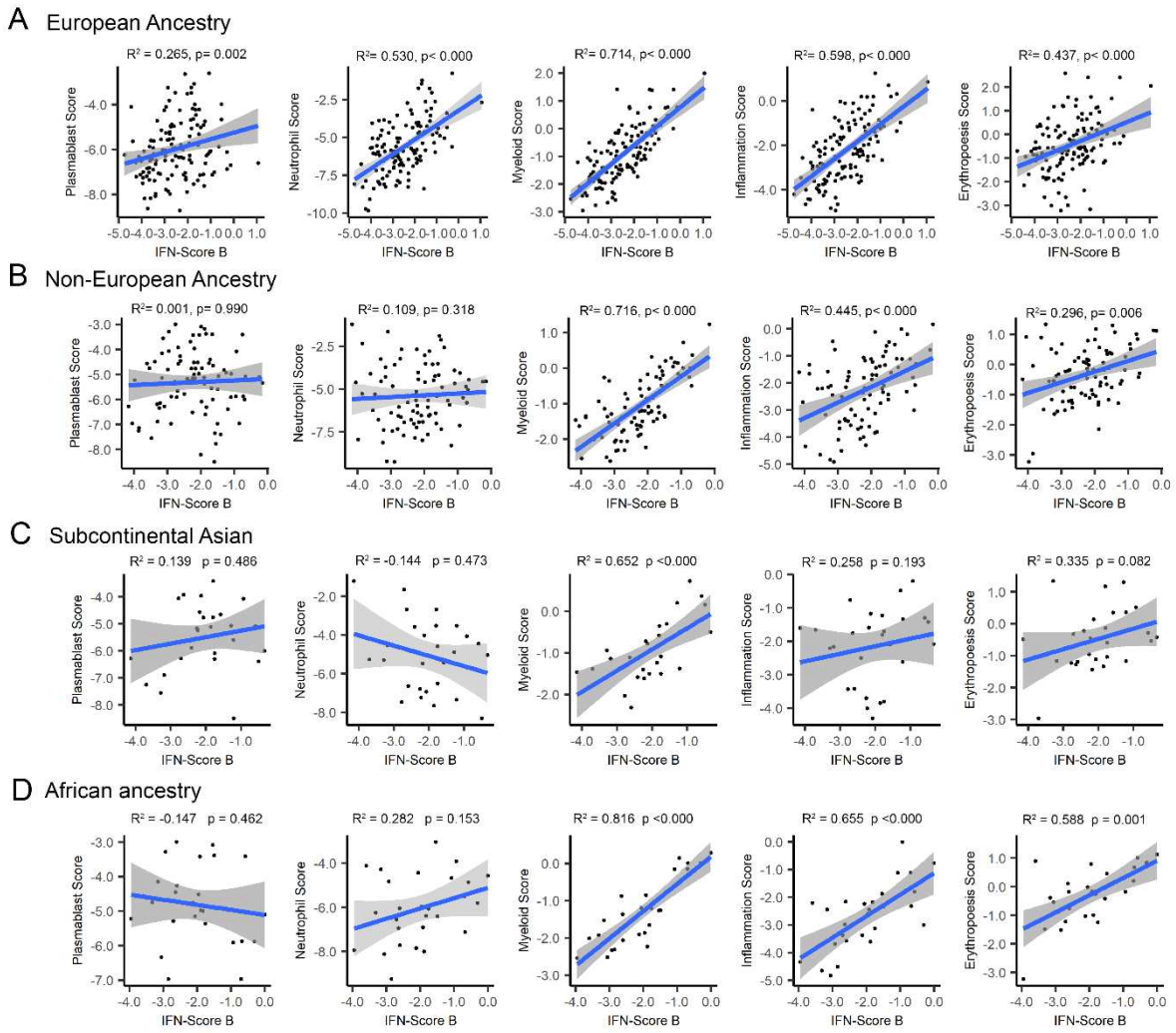


Figure S5. Commonalities in the relationship between IFN-Score-B and other functionally annotated gene expression scores varies displays are shared between Subcontinental Asian and African ancestry patients.

Scatterplot series show relationship between Plasmablast, Neutrophil, Myeloid lineage, Inflammation and Eyrthropoesis annotated expression scores and interferon pathway activation as measured by IFN-Score-B in patients of European (**A**) ancestry, non-European ancestry collectively (**B**) and separately the two larger composite non-European ancestry populations; Subcontinental Asian (**C**) and African ancestry (**D**) patients. The dissociation between IFN-Score-B and Plasmablast and Neutrophil- annotated scores observed collectively in non-European ancestry patients is apparent in both Subcontinental Asian and African ancestry patients independently. Regression line shown in blue and standard error in grey. All gene expression scores are shown as ΔCt from reference gene PPIA reflected across zero such that higher values indicate higher expression.

Table S3: Summary of similarities and distinguishing features between gene expression clusters							
	Clinical characteristics	**NEA-3** **All signatures high**	**EA-1** **All signatures high**	**NEA-2** **IFN-high, Neutrophil-myeloid-inflammation low**	**EA-2** **IFN-high, Neutrophil-myeloid-inflammation-erythropoesis low**	**NEA-1** **IFN low, Neutrophil-myeloid-inflammation high**	**EA-3** **All signatures low**
Ancestry	**European ancestry**	-	+	-	+	-	+
	Non-European ancestry	+	-	+	-	+	-
Gene expression signature	**IFN**	+	+	+	+	-	-
	Neutrophil-Myeloid-Inflammation	+	+	-	-	+	-
	Erythropoesis	+	+	+	-	+	-
Organ domain involvement	**Mucocutaneous (%)**	36	58	43	67	22	44
	Renal (%)	44	61	48	21	67	16
Seropositivity	**RNP**	63	10	50	33	0	0
	dsDNA	79	56	63	52	50	23
Response to RTX	**BILAG response 6 months**	85	61	41	72	13	81

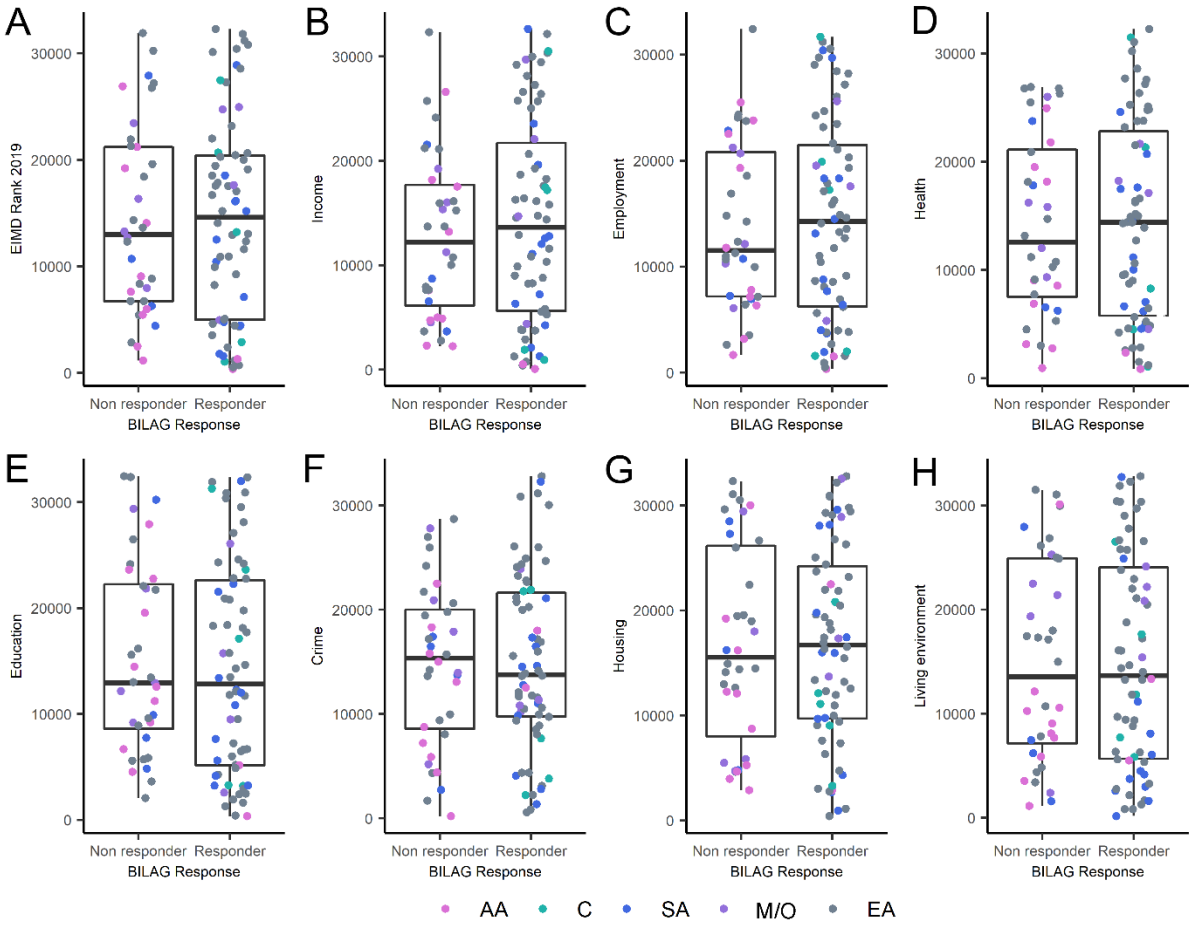


Figure S6. BILAG response following rituximab therapy is not associated with indices of socioeconomic deprivation.

Boxplot series with overlay jitter points color coded by ancestral group, show ranked deprivation indices from 1 (most deprived area) to 32,844 (least deprived area). Overall deprivation as measured by the 2019 English Index of Multiple Deprivation (EIMD, **A**) and composite domains of deprivation Income (**B**), Employment (**C**), Health (**D**), Education (**E**), Crime (**F**), Barriers to Housing and Services (**G**) and Living Environment (**H**) did not significantly differ between patients achieving a BILAG response to rituximab at 6 months (Responders) and patients not achieving response (Non Responders).

Table S4: Characteristics of BILAG responders and non-responders following cycle 1 rituximab				
	European Ancestry		**Non-European Ancestry**	
Clinical characteristics	**Responders n = 45**	**Non-Responders n = 20**	**Responders n = 25**	**Non-Responders n = 20**
Female patient, (%)	43 (95)	16 (80)	20 (80)	18 (90)
Age (years), median (IQR)	41 (33, 52)	42 (38, 58)	38 (26, 44)	38 (25, 50)
Disease duration (yrs), median (IQR)	13, (8,21)	10 (5, 18)	12 (7, 16)	12 (8, 20)
Current smoker, n/N (%)	14/30 (47)	8/15 (53)	1/16 (6)	4/15 (27)
Index of multiple deprivation (Rank), median (IQR)	16679 (8756, 20,535)	14353 (8746, 20535)	12518 (3650, 18075)	10699 (6126, 17769) 0.484
BILAG A or B score, n (%)				
Constitutional	5 (11)	5 (25)	9 (36)	2 (10)
Mucocutaneous	27 (60)	13 (65)	11 (44)	7 (35)
Neuropsychiatric	6 (13)	6 (30)	1 (4)	5 (25)
Musculoskeletal	24 (53)	11 (55)	9 (36)	9 (45)
Cardiorespiratory	7 (16)	4 (20)	4 (16)	4 (20)
Gastroenterology	3 (7)	0 (0)	3 (12)	1 (5)
Ophthalmic	1 (2)	2 (10)	1 (4)	1 (5)
Renal	19 (42)	8 (40)	13 (52)	11 (55)
Haematology	1 (2)	1 (6)	0 (0)	4 (20)
Numerical BILAG, median (IQR)	21 (16, 25)	22 (20, 28)	15 (13,22)	21 (20, 30)
SLEDAI Score, median (IQR)	10 (8, 14)	8 (7, 11)	8 (4, 14)	8 (6,13)
SLICC damage index, median (IQR)	0 (0, 2)	1 (0, 4)	0 (0,1)	1 (0, 2)
Total IgG (g/L), median (IQR)	10.6 (8.4, 13.9)	11.7 (9.5, 15.2)	12.7 (10.6, 18.0)	15.0 (9.5, 16.5)
Low C3 or C4, n (%)	23 (51)	7 (35)	14 (56)	12 (60)
Concurrent Immunosuppression n (%)				
Any agent (MMF, MTX, CNI, AZA)	22 (48)	8 (40)	10 (40)	5 (25)
Mycophenolate mofetil	15 (33)	6 (30)	8 (32)	4 (20)
Anti-malarial, n (%)	26 (58)	10 (50)	14 (56)	11 (55)
Oral glucocorticoid dose at baseline (mg), median (IQR)	10 (7, 18)	10 (5, 20)	10 (8, 10)	15 (10,20)
Oral glucocorticoid dose at 6 months (mg), median (IQR)	5 (0, 10)	5 (0, 10)	5 (3, 9)	10 (6, 13)
Change in oral glucocorticoid dose at 6 months (mg), median (IQR)	-1 (-13, 0)	0 (-5, 0)	-4 (-10, 0)	0 (-9, +2)
Immunoprecipitation and ELISA	**n = 37**	**n = 19**	**n = 20**	**n = 13**
U1RNP-Sm positive, n (%)	7 (19)	2 (11)	9 (45)	5 (38)
Ro-60 n (%)	12 (32)	4 (21)	8 (32)	5 (38)
La	2 (5)	1 (5)	0 (0)	1 (7)
Ro-52 ELISA n (%)	6 (16)	2 (10)	2 (10)	2 (15)
dsDNA ELISA n (%)	20 (54)	9 (47)	11 (55)	9 (69)

Table S4: Characteristics of BILAG responders and non-responders following cycle 1 rituximab				
	European Ancestry		Non-European Ancestry	
Clinical characteristics	**Responders n = 45**	**Non-Responders n = 20**	**Responders n = 25**	**Non-Responders n = 20**
Cardiolipin ELISA n (%)	3 (8)	3 (15)	1 (5)	3 (23)

Table S5. Response to rituximab by 2011 UK Census ancestral group	
Ancestral background **UK 2011 Census category**	**BILAG response** **6 months post RTX**
Black, African, Caribbean or Black British, n/N (%)	**3/14 (21)**
Black African	1/7
Black Caribbean	2/6
Other Black background	0/1
Asian or British Asian, n/N (%)	**18/22 (82)**
Indian	6/9
Pakistani	6/7
Chinese	3/3
Other Asian background	3/3
Mixed or multiple ethnic groups, n/N (%)	**3/5 (60)**
White and Black Caribbean	2/3
White and Black African	0/1
Any other Mixed or Multiple ethnic background	1/1
White, n/N (%)	**45/65 (69)**
English, Welsh, Scottish, Northern Irish or British	42/62
Irish	2/2
Any other White background	1/1
Other, n/N (%)	**1/4 (25)**

Table S6: Differential response to rituximab between non-European ancestry gene expression clusters itemized by patient ancestry				
Response to rituximab 6 months	**All patients**	**NEA Cluster 1 n = 8 IFN low, Neutrophil-myeloid-inflammation high**	**NEA Cluster 2 n = 17 IFN high, Neutrophil-myeloid-inflammation low**	**NEA Cluster 3 n = 20 All signatures high**
BILAG responder (complete or partial), n (%)	**25/45 (56)**	**1 (12.5)**	**7 (41.2)**	**17 (85.0)**
African Ancestry	3/14 (21)	0/1 (0)	2/10 (20)	1/3 (33)
South Asian	12/16 (75)	0/4 (0)	5/5 (100)	7/7 (100)
Other Asian (incl. Chinese)	6/6 (100)	-	-	6/6 (100)
Other (inc. Mixed)	4/9 (44)	1/3 (33)	0/2 (0)	3/4 (75)

Appendix 2: Contributors to the MASTERPLANS Consortium

The University of Manchester: Prof Katherine Payne; Dr Mark Lunt; Prof Niels Peek; Dr Nophar Geifman; Dr Sean Gavan; Dr Gillian Armitt; Dr Patrick Doherty; Dr Jennifer Prattley; Dr Narges Azadbakht; Angela Papazian; Dr Helen Le Sueur; Carmen Farrelly; Clare Richardson; Zunnaira Shabbir; Lauren Hewitt; Dr Emily Sutton; Alison Fountain, Patrick Doherty.

University of Bath: Prof Neil McHugh.

University of Birmingham: Prof Caroline Gordon; Prof Stephen Young.

University of Cambridge: Prof David Jayne; Prof Vern Farewell; Dr Li Su.

Imperial College London: Prof Matthew Pickering; Prof Elizabeth Lightstone; Dr Alyssa Gilmore; Prof Marina Botto.

King's College London: Prof Timothy Vyse; Dr David Lester Morris; Prof David D'Cruz.

University of Liverpool: Prof Michael Beresford; Prof Christian Hedrich; Dr Angela Midgley; Dr Jenna Gritzfeld.

University College London: Prof Michael Ehrenstein; Prof David Isenberg; Mariea Parvaz.

MASTERPLANS Patient and Public Involvement Group: Jane Dunnage; Jane Batchelor Elaine Holland; Pauline Upsal