RMD Open
Rheumatic & Musculoskeletal Diseases

ORIGINAL RESEARCH

Shared genetic susceptibility between idiopathic inflammatory myopathies and common B cell lymphoma subtypes found primarily in the human leucocyte antigen region

Weng Ian Che ,[1] Anton Öberg Sysojev ,[2] Catherine Zhu,[3] Karina Patasova,[2] International Lymphoma Epidemiology Consortium (InterLymph), IMACS Myositis Genetics Scientific Interest Group (MYOGEN), Karin E Smedby,[2,4] Ingrid E Lundberg ,[5,6] Helga Westerlind ,[2] Marie Holmqvist [2,6]

To cite: Che WI, Öberg Sysojev A, Zhu C, *et al*. Shared genetic susceptibility between idiopathic inflammatory myopathies and common B cell lymphoma subtypes found primarily in the human leucocyte antigen region. *RMD Open* 2025;**11**:e006035. doi:10.1136/rmdopen-2025-006035

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/rmdopen-2025-006035).

Received 17 June 2025
Accepted 22 July 2025

For numbered affiliations see end of article.

Correspondence to
Dr Marie Holmqvist;
marie.holmqvist@ki.se

ABSTRACT

Objectives To estimate shared genetic susceptibility between major subtypes of idiopathic inflammatory myopathies (IIM) and B cell lymphomas.

Methods We paired summary statistics from genome-wide association studies (GWASs) of diffuse large B cell lymphoma, follicular lymphoma (FL), chronic lymphocytic leukaemia (CLL) and marginal zone lymphoma with those of dermatomyositis (DM) and polymyositis (PM) from a GWAS and an ImmunoChip study. We estimated local genetic correlation (r_g) for each disease pair using local analysis of (co)variant association (Bonferroni-corrected p value<0.05) and identified genetic variants jointly associated with both diseases using pleiotropy-informed false discovery rate (conjunctional false discovery rate <0.05). Functional mapping and annotation analyses were also performed.

Results We identified significant r_g (ranging from −0.50 to 0.84) across 16 loci, with half located in the human leucocyte antigen (HLA) region, for the disease pairs of IIM and B cell lymphoma subtypes. Furthermore, jointly associated single-nucleotide polymorphisms were predominantly found in the HLA region. Specifically, all disease pairs showed shared genetic susceptibility in the HLA class I regions, while additional correlations in class III and class II regions were specific to DM and PM disease pairs, respectively. For some non-HLA loci with significant r_g, functional analyses revealed immune-related responses potentially overlapping between DM and FL, DM and CLL, and PM and CLL.

Conclusion We revealed that DM and PM share genetic susceptibility with common B cell lymphoma subtypes in both immune-related loci and loci with unclear biological functions. These novel findings improve our understanding of the pathological link between IIM and B cell lymphomas.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Although an overlap in biological mechanisms between idiopathic inflammatory myopathies (IIM) and B cell lymphomas has been suggested, no genetic-level studies have been conducted to date.

WHAT THIS STUDY ADDS

⇒ Dermatomyositis (DM) and polymyositis (PM) share genetic susceptibility with common B cell lymphoma subtypes at several loci across the whole genome, particularly within the human leucocyte antigen (HLA) region, which generally demonstrates stronger associations than non-HLA loci.

⇒ While both DM and PM are associated with non-Hodgkin's lymphoma, distinct genetic correlations at HLA and non-HLA loci suggest differing underlying genetic mechanisms.

⇒ Functional annotation of non-HLA loci with significant genetic correlations between IIM and B cell lymphomas suggests overlapping interferon-regulated inflammatory response, specifically between DM and follicular lymphoma, DM and chronic lymphocytic leukaemia (CLL), and PM and CLL.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The findings offer valuable insights into the genetic basis of IIM and B cell lymphoma co-occurrence, informing future aetiological research and facilitating risk communication with affected patients.

INTRODUCTION

Idiopathic inflammatory myopathies (IIM) are a group of rare autoimmune diseases characterised by chronic inflammation primarily affecting proximal muscles.[1] The diseases are often accompanied with extra-muscular involvement in skin, joints and lungs. IIM is highly heterogeneous and can be classified into different subtypes based on the clinical, serological, histological and immunogenic features: dermatomyositis (DM), polymyositis

(PM), immune-mediated necrotising myopathy (IMNM), antisynthetase syndrome, inclusion body myositis, overlap myositis and juvenile myositis.[1]

B cell lymphomas are malignancies that develop from different stages of B lymphocyte maturation.[2] Diffuse large B cell lymphoma (DLBCL), the most common and aggressive subtype, is thought to arise from centroblasts or plasmablasts within the germinal centre. Other common, mostly indolent subtypes include follicular lymphoma (FL), chronic lymphocytic leukaemia (CLL) and marginal zone lymphoma (MZL), which are believed to originate from centrocytes, mature or memory B cells and marginal zone B cells, respectively.[2] However, the precise cellular origins remain under debate, as accumulating evidence suggests that key genetic lesions may occur early in B cell development.[3] Furthermore, these subtypes can be further classified into more homogeneous subgroups based on immunophenotypic profiles and cytogenetic alterations.[4]

Patients with IIM are at risk of having B cell lymphomas, particularly around the time of IIM diagnosis.[5–9] Further investigations by B cell lymphoma subtypes suggest associations with DLBCL and FL in patients with IIM,[10] whereas evidence supporting elevated risks of CLL and MZL after IIM diagnosis is sparse.[11] When IIM coexists with B cell lymphomas, the prognosis tends to be poor, with a 5-year survival of 51%.[12]

The pathological link between IIM and B cell lymphomas is unclear and expected to be complex, given the bidirectional associations observed in previous work.[5] Although little is known about the biological mechanism behind the co-occurrence of IIM and B cell lymphomas, common features associated with disease development may suggest shared susceptibility. First, the development of IIM and B cell lymphomas has been shown to involve antigen stimulation and chronic B cell activation.[1 2] Second, previous studies found that first-degree relatives of patients with IIM were more likely to have B cell lymphomas overall, DLBCL and FL, although these associations were statistically non-significant, and the association was estimated with fewer than 10 cases for DLBCL and fewer than 5 cases for FL.[13 14] Several human leucocyte antigen (HLA) alleles have also been associated with both IIM and B cell lymphomas.[15–21] These findings together suggest that IIM may share genetic components with B cell lymphomas. This assumption has however not been investigated at a genetic level.

To understand the importance of genetics in the development of both diseases, we aimed to quantify the shared genetic susceptibility, in terms of locus-specific/local genetic correlation and genetic variants jointly associated with both diseases, between IIM and common B cell lymphoma subtypes using summary statistics from published genome-wide association studies (GWAS).

MATERIALS AND METHODS

Data sources

The Genetics Scientific Interest Group (MYOGEN) of the International Myositis Assessment & Clinical Studies Group is an international network aiming to investigate the genetics of IIM. Between 2013 and 2015, they performed GWASs in IIM (DM and PM) with genotyping data for the whole genome (online supplemental table 1),[20 21] followed by a study covering 186 loci related to 12 autoimmune diseases (the 'ImmunoChip') in a larger cohort of IIM conducted in 2016.[19] Each of the GWAS and the ImmunoChip study included data on genotyped single-nucleotide polymorphisms (SNP) for patients with adult or juvenile DM and PM and geographically matched controls, which were shared across the IIM subtypes. Patient inclusion and exclusion criteria of the MYOGEN and information on genotyping have been described in detail in previous studies.[19–21] Recently, data from 657 and 1793 patients with adult-onset IIM of European descent from the GWAS and the ImmunoChip studies, respectively, along with their controls, were imputed using the Trans-Omics for Precision Medicine (TOPMed) Imputation Server against the TOPMed reference panel version R2 on the human genome build GRCh38.[22 23] Detailed information on quality control (QC), imputation and association testing of SNPs is presented in the online supplemental methods and elsewhere.[24 25]

The International Lymphoma Epidemiology Consortium (InterLymph) is an international consortium of investigators working with epidemiological studies of lymphoid neoplasms. This consortium includes a wide range of genetic and environmental risk factor data on non-Hodgkin's lymphoma (NHL) subtypes and control subjects. In our study, we used summary statistics from previously performed meta-analyses of genotyped and imputed SNPs from the InterLymph for four common B cell lymphoma subtypes, namely DLBCL, FL, CLL and MZL, among individuals of European descent.[15–18]

The total number of SNPs imputed and numbers of patients and controls included for each disease subtype are presented in online supplemental table 2.

Data preparation

Summary statistics from MYOGEN and InterLymph were quality controlled and aligned before the statistical analyses (online supplemental methods and online supplemental table 2). We paired each B cell lymphoma subtype to the GWAS DM, the GWAS PM, the ImmunoChip DM and the ImmunoChip PM data and kept only the overlapping SNPs in each pairing, resulting in a total of eight GWAS and eight ImmunoChip disease pairs.

Statistical methods

Local genetic correlations

We estimated the local genetic correlation at non-overlapping loci for each disease pair, with local analysis of (co)variant association (LAVA).[26–28] Compared with other alternative methods, LAVA has better performance

regarding type 1 error, statistical power and estimation stability when target population matches to the one of a reference panel.[26 29] In the LAVA analyses, the whole genome was partitioned into 2495 genomic loci, predefined by the authors based on a specific algorithm, making these loci approximately equal size of about 1 Mb (2500 SNPs) and semi-independent from each other, as indicated by a metric value defined as a function of the linkage disequilibrium (LD) in a pair of SNPs, less than 0.25. Details of genome partitioning are described elsewhere (https://github.com/cadeleeuw/lava-partitioning).[26] The HLA region was partitioned into 21 distinct loci. There were two steps involved in the LAVA analyses. First, we computed the local heritability (the proportion of the phenotypic variance of a disease explained by all SNPs present in a locus) for the 2495 loci for each disease in a pair (univariate analysis). Only loci showing significant heritability in the univariate analyses of both diseases in a pair were taken forward for genetic correlation estimation (bivariate analysis). Multiple logistic regression models were fitted in both steps where heritability and genetic correlation were estimated by regressing the SNP test statistics as well as the products of paired SNP test statistics, respectively, on the LD scores estimated within a matched population reference panel.[26] We used a Bonferroni factor of 2495 for the univariate analyses of all disease pairs, 4088 and 261, respectively, for the bivariate analyses of the GWAS and the ImmunoChip disease pairs to define the statistical significance thresholds. We did not adjust for sample overlap in the LAVA analyses since datasets were near-independent for each disease pair. This was further confirmed in a cross-trait linkage disequilibrium score regression analysis, where the estimated intercept was near-zero.

Pleiotropic enrichment and jointly associated SNPs

We used pleiotropy-informed false discovery rate (pleioFDR), consisting of conditional FDR (condFDR) and conjunctional FDR (conjFDR), to complement the LAVA analyses.[30] CondFDR corrects nominal p values of SNPs of a primary disease by leveraging pleiotropy with a secondary disease where an observed genetic enrichment supports shared genetic susceptibility between the two diseases.[30] To avoid artificial genetic enrichment, we adjusted the nominal p value of each SNP by using intergenic inflation control and excluded SNPs located in the HLA and the chromosome 8 inversion regions from fitting the conditional empirical cumulative distribution function.[30] In the intergenic inflation control, intergenic SNPs, instead of all SNPs, were used to calculate the genomic inflation factor, which was used to adjust the test statistics of SNPs. It has been shown that intergenic SNPs have relative depleted associations with polygenic diseases and thus they better represent true null effects than all SNPs.[31] To visualise genetic enrichment, we constructed fold enrichment plots for each disease in a pair by stratifying the $-\log_{10}$ of the nominal p values of SNPs of the primary disease according to the p values of the corresponding SNPs of the secondary disease into three categories ($-\log_{10}(p) \geq 1$, $-\log_{10}(p) \geq 2$ and $-\log_{10}(p) \geq 3$). The fold enrichment was computed as the ratio between the cumulative distribution of SNPs in a given $-\log_{10}(p)$ stratum and the cumulative distribution of all SNPs.[30] If pleiotropy existed, we expected to observe the strongest fold enrichment in the stratum of the lowest p values of SNPs of a B cell lymphoma or IIM subtype for SNPs with low p values of the paired IIM or B cell lymphoma subtype. The strength of the genetic enrichment was expected to be less in strata with higher p values.

In the pleioFDR analyses, for any disease pair, we computed the condFDR of SNPs of an IIM subtype based on the conditional empirical cumulative distribution function of the p values of the corresponding SNPs of the paired B cell lymphoma subtype. We repeated this process by reversing the disease subtypes to obtain condFDR of SNPs of the paired B cell lymphoma subtype. By maximising the $\text{condFDR}_{\text{IIM|B cell lymphoma}}$ and $\text{condFDR}_{\text{B cell lymphoma|IIM}}$ of a SNP, we obtained a conservative conjFDR of that SNP.[30] We used 0.05 as the significance threshold to determine a SNP jointly associated with both diseases in a pair. We determined the lead jointly associated SNPs/loci ($r^2<0.1$ in LD window 10 000 kb) using the clump function in sumstat.py package in Python (V.2.7.6). Loci closer than 250 kb were merged.

Gene mapping

Positional coding gene mapping using the gnomAD browser V.2.1.1 was performed for non-HLA loci showing significant local genetic correlations to provide insights into functional variants associated with both IIM and B cell lymphomas.[32] We further annotated the mapped protein-coding genes to biological processes defined by Gene Ontology using FUMA GENE2FUN V.1.5.3.[33] Furthermore, the lead SNPs jointly associated with IIM and B cell lymphomas were annotated to the nearest protein-coding genes using FUMA SNP2GENE V.1.5.2.[33]

RESULTS

Local genetic correlations

Table 1 summarises the number of common SNPs and loci tested in the LAVA analyses for each disease pair. In the univariate analyses, some loci were dropped due to the number of common SNPs <2 or negative variance estimates for both diseases in the loci.

The GWAS data

Among the GWAS disease pairs, we identified in total seven loci with moderate-to-strong genetic correlations: [chr17:62 112 374–63 548 724] for DM-FL; [chr6:32 629 240–32 682 213] and [chr9:114 443 476–115 390 112] for DM-CLL; [chr1:238 094 456–238 704 857] for PM-DLBCL; [chr3:158 460 059–159 478 751], [chr7:55 161 395–56 303 513] and [chr12:126 871 453–127 545 377] for PM-MZL (table 2). All detected genetic correlations were positive except in the locus [chr17:62 112 374–63 548 724] (genetic correlation

Table 1 The number of SNPs and loci included in the local analysis of (co)variant association (LAVA) analyses and in the pleiotropy-informed false discovery rate (pleioFDR) analyses for the GWAS and the ImmunoChip disease pairs of IIM and common B cell lymphoma subtypes

		LAVA		pleioFDR
	Number of common SNPs*	Number of total loci for univariate analysis (both:IIM:B cell lymphoma)†	Number of loci for bivariate analysis	Number of common SNPs*
GWAS data				
DM-DLBCL	6 415 387	2471 (1836:1876:2431)	680	6 260 170
DM-FL	6 416 691	2469 (1829:1889:2409)	611	6 259 969
DM-CLL	6 417 297	2474 (1828:1881:2421)	741	6 260 163
DM-MZL	6 384 382	2451 (1786:2337:1900)	813	6 239 939
PM-DLBCL	6 415 351	2466 (1073:1111:2428)	326	6 260 153
PM-FL	6 416 654	2465 (1052:1111:2406)	284	6 259 952
PM-CLL	6 417 260	2472 (1068:1118:2422)	369	6 260 145
PM-MZL	6 384 346	2460 (1025:1147:2338)	264	6 239 921
ImmunoChip data				
DM-DLBCL	882 537	1697 (614:1020:1291)	37	850 547
DM-FL	882 935	1694 (607:1017:1284)	34	850 673
DM-CLL	882 730	1698 (616:1019:1295)	36	850 596
DM-MZL	880 432	1671 (567:1023:1215)	24	849 056
PM-DLBCL	882 537	1740 (627:1078:1289)	35	850 547
PM-FL	882 935	1746 (614:1078:1282)	36	850 673
PM-CLL	882 730	1712 (661:1078:1295)	36	850 596
PM-MZL	880 432	1691 (602:1080:1213)	23	849 056

Common B cell lymphoma subtypes included diffuse large B cell lymphoma (DLBCL), follicular lymphoma (FL), chronic lymphocytic leukaemia (CLL) and marginal zone lymphoma (MZL).
*Non-overlapping SNPs between diseases and SNPs that were not aligned to the reference panel in the LAVA or in the pleioFDR were excluded.
†Some loci were dropped in the univariate analyses due to number of SNPs <2 in the locus or negative variance estimate for both diseases.
DM, dermatomyositis; GWAS, genome-wide association study; IIM, idiopathic inflammatory myopathy; PM, polymyositis; SNP, single-nucleotide polymorphism.

coefficient, r_g=−0.50 95% CI −0.72 to −0.32). We also reported the proportion of the heritability of an IIM subtype that could be explained by that of the paired B cell lymphoma subtype in a given locus (r_g^2). In the locus [chr6:32 629 240–32 682 213] showing a strong genetic correlation between DM and CLL, half of the local heritability of DM could be explained by that of CLL in this locus. For the other detected local genetic correlations, the r_g^2 ranged from 23% to 39%. Notably, the local heritability of the loci showing significant genetic correlations was low for both IIM and B cell lymphoma subtypes.

The ImmunoChip data

We identified more signals of local genetic correlations, particularly within the HLA region, for the ImmunoChip disease pairs despite the GWAS data having a higher overall coverage of the genome (table 2). For the disease pair DM-DLBCL, we observed significant genetic correlations in three HLA loci [chr6:30 715 007–31 106 493], [chr6:31 320 269–31 427 209] and [chr6:31 427 210–32 208 901]. Heritability of DM and FL was also correlated in locus [chr6:31 427 210–32 208 901] and additional significant genetic correlation was detected in the non-HLA locus [chr1:205 917 549–208 162 951] and three HLA loci [chr6:32 208 902–32 454 577], [chr6:32 539 568–32 586 784] and [chr6:32 586 785–32 629 239]. We noted a very strong genetic correlation in the locus [chr6:32 208 902–32 454 577] for disease pair DM-CLL, as well as moderate-to-strong genetic correlations in the loci [chr6:32 586 785–32 629 239] and [chr6:32 629 240–32 682 213]. There were significant local genetic correlations in one non-HLA locus [chr9:138 995 792–140 097 759] and three HLA loci [chr6:30 070 718–30 715 006], [chr6:30 715 007–31 106 493] and [chr6:31 320 269–31 427 209] detected for the disease pair PM-DLBCL where the genetic correlation in [chr9:138 995 792–140 097 759] was negative (r_g=−0.46 95% CI −0.70 to −0.23). The disease pair PM-CLL also showed significant genetic correlation in the same locus [chr6:32 629 240–32 682 213] as DM-CLL did. Among these genetic correlations, the r_g^2 varied from 13% to 70%.

Table 2 The loci with significant local genetic correlations for the GWAS and the ImmunoChip disease pairs of IIM and common B cell lymphoma subtypes

Disease pair	Chr	Start	Stop	HLA region	n.snps	h^2_{IIM}	$h^2_{B\ cell\ lymphoma}$	r_g (95% CI)	r_g^2 (95% CI)	P value
GWAS data										
DM-FL	17	62 112 374	63 548 724		2483	0.02	0.01	−0.50 (−0.72 to −0.32)	0.25 (0.10 to 0.52)	5.29×10^{-7}
DM-CLL	6	32 629 240	32 682 213	Class II	45	0.04	0.01	0.72 (0.53 to 0.89)	0.51 (0.28 to 0.79)	1.38×10^{-8}
DM-CLL	9	114 443 476	115 390 112		2459	0.02	0.02	0.62 (0.54 to 0.70)	0.39 (0.30 to 0.49)	3.39×10^{-38}
PM-DLBCL	1	238 094 456	238 704 857		2081	0.00	0.02	0.48 (0.30 to 0.66)	0.23 (0.09 to 0.44)	7.62×10^{-7}
PM-MZL	3	158 460 059	159 478 751		1961	0.00	0.00	0.54 (0.33 to 0.78)	0.29 (0.11 to 0.61)	3.31×10^{-6}
PM-MZL	7	55 161 395	56 303 513		3322	0.02	0.01	0.53 (0.32 to 0.77)	0.28 (0.10 to 0.59)	3.03×10^{-6}
PM-MZL	12	126 871 453	127 545 377		2021	0.01	0.01	0.57 (0.38 to 0.78)	0.33 (0.15 to 0.61)	6.83×10^{-8}
ImmunoChip data										
DM-DLBCL	6	30 715 007	31 106 493	Class I	544	0.03	0.01	0.77 (0.48 to 1.00)	0.59 (0.23 to 1.00)	1.05×10^{-5}
DM-DLBCL	6	31 320 269	31 427 209	Class I	299	0.04	0.01	0.70 (0.45 to 0.98)	0.49 (0.20 to 0.96)	1.08×10^{-5}
DM-DLBCL	6	31 427 210	32 208 901	Class I and III	400	0.04	0.01	0.58 (0.33 to 0.88)	0.34 (0.11 to 0.77)	5.86×10^{-5}
DM-FL	1	205 917 549	208 162 951		2918	0.01	0.01	0.51 (0.27 to 0.78)	0.26 (0.07 to 0.60)	6.59×10^{-5}
DM-FL	6	31 427 210	32 208 901	Class I and III	400	0.04	0.03	0.35 (0.20 to 0.51)	0.13 (0.04 to 0.26)	2.47×10^{-5}
DM-FL	6	32 208 902	32 454 577	Class II	152	0.03	0.04	0.42 (0.25 to 0.58)	0.18 (0.06 to 0.34)	3.43×10^{-6}
DM-FL	6	32 539 568	32 586 784	Class II	23	0.01	0.04	0.43 (0.22 to 0.63)	0.19 (0.05 to 0.40)	1.65×10^{-4}
DM-FL	6	32 586 785	32 629 239	Class II	75	0.04	0.04	0.50 (0.35 to 0.63)	0.25 (0.12 to 0.40)	3.01×10^{-9}
DM-CLL	6	32 208 902	32 454 577	Class II	151	0.03	0.01	0.84 (0.62 to 1.00)	0.70 (0.39 to 1.00)	2.18×10^{-7}
DM-CLL	6	32 586 785	32 629 239	Class II	75	0.04	0.01	0.77 (0.53 to 1.00)	0.59 (0.28 to 1.00)	1.52×10^{-6}
DM-CLL	6	32 629 240	32 682 213	Class II	45	0.07	0.01	0.54 (0.37 to 0.70)	0.29 (0.14 to 0.49)	1.77×10^{-8}
PM-DLBCL	6	30 070 718	30 715 006	Class I	617	0.03	0.01	0.63 (0.36 to 0.96)	0.40 (0.13 to 0.92)	5.4×10^{-5}
PM-DLBCL	6	30 715 007	31 106 493	Class I	544	0.05	0.01	0.62 (0.35 to 0.94)	0.39 (0.12 to 0.88)	6.54×10^{-5}
PM-DLBCL	6	31 320 269	31 427 209	Class I	299	0.06	0.01	0.68 (0.44 to 0.94)	0.46 (0.19 to 0.88)	3.3×10^{-6}
PM-DLBCL	9	138 995 792	140 097 759		1042	0.02	0.01	−0.46 (−0.70 to −0.23)	0.21 (0.05 to 0.48)	1.69×10^{-4}
PM-CLL	6	32 629 240	32 682 213	Class II	45	0.10	0.01	0.51 (0.33 to 0.67)	0.26 (0.11 to 0.45)	4.48×10^{-6}

Common B cell lymphoma subtypes included diffuse large B cell lymphoma (DLBCL), follicular lymphoma (FL), chronic lymphocytic leukaemia (CLL) and marginal zone lymphoma (MZL).
Chr, chromosome number; DM, dermatomyositis; GWAS, genome-wide association study; $h^2_{B\ cell\ lymphoma}$, local heritability of B cell lymphoma at observed scale; h^2_{IIM}, local heritability of IIM at observed scale; IIM, idiopathic inflammatory myopathy; n.snps, number of single-nucleotide polymorphisms in each locus; PM, polymyositis; r_g, genetic correlation coefficient; r_g^2, square of r_g, the proportion of the heritability of an IIM subtype that could be explained by that of the paired B cell lymphoma subtype in a given locus.

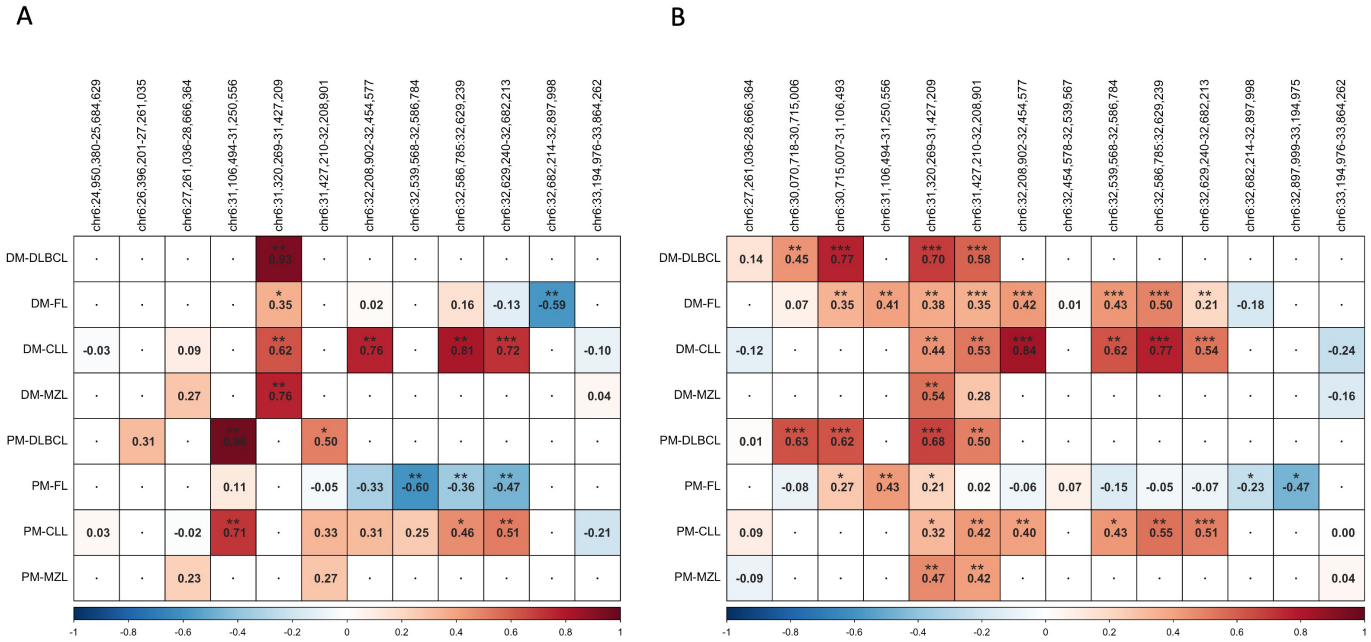

Figure 1 Heatmaps presenting the local genetic correlations (r_g) of all tested human leucocyte antigen loci for the GWAS (A) and the ImmunoChip (B) disease pairs of IIM and common B cell lymphoma subtypes. Common B cell lymphoma subtypes included diffuse large B cell lymphoma (DLBCL), follicular lymphoma (FL), chronic lymphocytic leukaemia (CLL) and marginal zone lymphoma (MZL); ***p value <0.00 (the threshold of significance); **p value <0.01; *p value <0.05. DM, dermatomyositis; GWAS, genome-wide association study; IIM, idiopathic inflammatory myopathy; PM, polymyositis.

Heatmaps of the genetic correlations in the HLA region for the GWAS and the ImmunoChip disease pairs

Figure 1 presents the genetic correlations of all tested HLA loci for the GWAS and the ImmunoChip disease pairs. The SNPs in the HLA loci of the GWAS data were predominantly shared with those of the ImmunoChip data, exhibiting an overlapping rate ranging from 69% to 100% (online supplemental tables 3 and 4). For loci that were tested in both the GWAS and the ImmunoChip disease pairs, their findings are in general consistent for the corresponding disease pairs. We detected no significant local genetic correlation for disease pairs DM-MZL and PM-FL (table 2).

Pleiotropic enrichment

We found no typical pleiotropic enrichment in any of the GWAS disease pairs (online supplemental figure 1) but pleiotropic enrichment was observed for the ImmunoChip data in the disease pairs DM-CLL and PM-CLL (online supplemental figure 2). For SNPs with p values <0.001 in CLL, there were up to 400-fold and 225-fold enrichment of SNPs with $-\log_{10}(p)=7$ for DM and PM compared with all SNPs, respectively (online supplemental figure 2i, k). There were 45-fold and 60-fold enrichment of SNPs with $-\log_{10}(p)=7$ for CLL when conditioning on SNPs with p values <0.001 in DM and PM, respectively (online supplemental figure 2j, l).

Jointly associated SNPs

The GWAS data

Among the GWAS disease pairs, we identified one lead SNP jointly associated with DM-DLBCL, two with DM-FL, one with DM-CLL, one with DM-MZL, three with PM-DLBCL, two with PM-FL, one with PM-CLL and one with PM-MZL (table 3 and online supplemental table 5). All of these jointly associated SNPs except rs1611929 on chromosome 3 and rs117408955 on chromosome 13 for PM-DLBCL resided in the HLA region. The majority of these SNPs affected IIM and B cell lymphoma subtypes in the same direction and had corrected p values below the GWAS-level of significance (p value $<5.00\times10^{-8}$) or suggestive significance (2.25×10^{-5}).

The ImmunoChip data

We detected more jointly associated SNPs among the ImmunoChip disease pairs (table 3 and online supplemental table 6). The two lead jointly associated SNPs rs9270493 and rs2596500 detected, respectively, for the disease pairs PM-FL and PM-CLL of the GWAS data were replicated in the ImmunoChip data. All of these jointly associated SNPs except rs10173316 on chromosome 2 for DM-FL, rs12203592 on chromosome 6 for DM-CLL and PM-CLL and rs2872812 on chromosome 17 for DM-CLL were located in the HLA region. Most of these SNPs were associated with both diseases in the same direction. All of these SNPs had corrected p values below either the GWAS level of significance or suggestive significance level.

Gene mapping

Online supplemental table 7 presents the coding genes located in the significant loci outside the HLA region from the GWAS and the ImmunoChip datasets. None of these genes has been reported to be associated with the corresponding IIM and B cell lymphoma subtypes. In the

Table 3 The lead jointly associated SNPs detected for the GWAS and the ImmunoChip disease pairs of IIM and common B cell lymphoma subtypes

Disease pair	SNP ID	Reference allele	Effect allele	Z_{IIM}	$Z_{B\ cell\ lymphoma}$	Chr	Position	HLA region	P $value_{IIM}$	P $value_{B\ cell\ lymphoma}$	conjFDR
GWAS data											
DM-DLBCL	rs3094005	G	T	7.00	4.32	6	31 465 047	Class I	2.64×10^{-12}	1.56×10^{-5}	8.21×10^{-3}
DM-FL	rs2596462	T	C	4.23	4.80	6	31 414 580	Class I	2.39×10^{-5}	1.56×10^{-6}	9.00×10^{-3}
DM-FL	rs7450278	C	T	6.41	–4.97	6	32 439 048	Class II	32 439 048	6.68×10^{-7}	3.36×10^{-4}
DM-CLL	rs9273325	G	A	8.61	5.86	6	32 623 193	Class II	7.50×10^{-18}	4.54×10^{-9}	2.60×10^{-6}
DM-MZL	rs3130490	G	T	7.25	4.74	6	31 739 120	Class III	4.02×10^{-13}	2.09×10^{-6}	1.04×10^{-3}
PM-DLBCL	rs1611929	G	A	–3.81	–4.00	3	21 457 961		1.41×10^{-4}	6.31×10^{-5}	4.29×10^{-2}
PM-DLBCL	rs3130923	G	A	8.51	4.28	6	31 462 135	Class I	1.69×10^{-17}	1.87×10^{-5}	8.86×10^{-3}
PM-DLBCL	rs117408955	G	A	3.90	–4.28	13	112 538 767		9.45×10^{-5}	1.85×10^{-5}	3.07×10^{-2}
PM-FL	rs130071	G	A	4.24	5.25	6	31 116 210	Class I	2.26×10^{-5}	1.5×10^{-7}	9.43×10^{-3}
PM-FL	rs9270493	T	C	5.20	–5.66	6	32 559 110	Class II	2.01×10^{-7}	1.51×10^{-8}	6.95×10^{-5}
PM-CLL	rs2596500	A	C	8.47	4.04	6	31 321 267	Class I	2.49×10^{-17}	5.42×10^{-5}	4.17×10^{-2}
PM-MZL	rs3130923	G	A	8.52	5.09	6	31 462 135	Class I	1.65×10^{-17}	3.62×10^{-7}	1.89×10^{-4}
ImmunoChip data											
DM-DLBCL	rs3093958	A	G	12.30	4.84	6	31 410 521	Class I	9.56×10^{-35}	1.28×10^{-6}	7.68×10^{-5}
DM-FL	rs10173316	C	T	3.54	3.31	2	152 170 442		4.00×10^{-4}	9.30×10^{-4}	3.78×10^{-2}
DM-FL	rs9273504	T	C	11.18	3.26	6	32 628 407	Class II	5.22×10^{-29}	1.12×10^{-3}	4.40×10^{-2}
DM-CLL	rs12203592	C	T	6.83	7.32	6	396 321		8.78×10^{-12}	2.51×10^{-13}	5.74×10^{-10}
DM-CLL	rs9348747	A	G	–3.41	3.22	6	27 002 406	Class I	6.60×10^{-4}	1.27×10^{-3}	3.21×10^{-2}
DM-CLL	rs3093958	A	G	12.30	3.76	6	31 410 521	Class I	8.73×10^{-35}	1.72×10^{-4}	4.14×10^{-3}
DM-CLL	rs511515	A	G	–3.49	5.03	6	33 541 507	Class II	4.78×10^{-4}	4.99×10^{-7}	1.54×10^{-2}
DM-CLL	rs2872812	G	A	–3.42	–3.36	17	38 758 650		6.22×10^{-4}	7.89×10^{-4}	1.99×10^{-2}
DM-MZL	rs3093958	A	G	12.31	5.13	6	31 410 521	Class I	8.24×10^{-35}	2.86×10^{-7}	1.46×10^{-5}
PM-DLBCL	rs2596500	A	C	15.37	5.06	6	31 321 267	Class I	2.54×10^{-53}	4.18×10^{-7}	2.09×10^{-5}
PM-FL	rs2023472	A	G	–4.39	3.40	6	30 075 864	Class I	1.13×10^{-5}	6.81×10^{-4}	4.55×10^{-2}
PM-FL	rs9270493	T	C	12.22	–5.63	6	32 559 110	Class II	2.36×10^{-34}	1.82×10^{-8}	8.89×10^{-7}
PM-CLL	rs12203592	C	T	5.38	7.32	6	396 321		7.36×10^{-8}	2.51×10^{-13}	4.9×10^{-6}
PM-CLL	rs10946859	A	G	–3.60	–3.38	6	26 851 785	Class I	3.22×10^{-4}	7.18×10^{-4}	2.92×10^{-2}
PM-CLL	rs2523990	A	G	3.84	3.27	6	30 077 229	Class I	1.25×10^{-4}	1.06×10^{-3}	4.33×10^{-2}
PM-CLL	rs2596500	A	C	15.37	3.97	6	31 321 267	Class I	2.80×10^{-53}	7.20×10^{-5}	2.62×10^{-3}
PM-MZL	rs2596500	A	C	15.37	5.31	6	31 321 267	Class I	2.44×10^{-53}	1.07×10^{-7}	4.81×10^{-6}

Common B cell lymphoma subtypes included diffuse large B cell lymphoma (DLBCL), follicular lymphoma (FL), chronic lymphocytic leukaemia (CLL) and marginal zone lymphoma (MZL).
P values were corrected for genomic inflation.
chr, chromosome number; conjFDR, conjunctional false discovery rate; DM, dermatomyositis; GWAS, genome-wide association study; IIM, idiopathic inflammatory myopathy; PM, polymyositis; SNP, single-nucleotide polymorphism; $Z_{B\ cell\ lymphoma}$, Z score of SNP of B cell lymphoma; Z_{IIM}, Z score of SNP of IIM.

GENE2FUN analyses, we only observed significant findings in the locus [chr1:205 917 549–208 162 951], associated with DM and FL. The various gene sets residing in this locus were annotated to 37 Gene Ontology-defined biological processes, including humoral immune response, complement regulation, cytokine production, innate immune response, signal transducer and activation of transcription signalling pathway and stress response (online supplemental figure 3).

Additionally, we mapped the nearest protein-coding genes to each of the lead jointly associated SNPs (online supplemental table 8). Given the extended LD structure of the HLA region, many HLA and immune-related genes were tagged by the lead jointly associated HLA SNPs. For the lead jointly associated non-HLA SNPs, SNP rs1611929 on chromosome 3 for the GWAS disease pair PM-DLBCL was linked to the gene *ZNF385D*; SNP rs10173316 on chromosome 2 for the ImmunoChip disease pair DM-FL was mapped to the genes *RBM43* and *NMI*; SNP rs1220359 on chromosome 6 for the ImmunoChip disease pairs DM-CLL and PM-CLL was mapped to the gene *IRF4*; SNP rs2872812 on chromosome 17 for the ImmunoChip disease pair was mapped to the genes *SMARCE1* and *KRT222*.

DISCUSSION

Evidence supporting the role of genetic susceptibility in the development of IIM and B cell lymphomas is scarce. Previous studies, one of which consisted of a limited number of cases, reported elevated but non-significant familial associations between IIM and B cell lymphomas, with strength of effect estimates ranging from 1.2 to 2.0.[13 14] A familial association suggests involvement of both genetic and environmental factors shared within families. However, the previous studies did not disentangle the genetic from the environmental contribution, leaving the magnitude of genetic contribution to the development of IIM and B cell lymphomas unknown. Our study tried to fill this knowledge gap and found shared genetic susceptibility between IIM and common B cell lymphoma subtypes in several genomic loci, including both the HLA and non-HLA regions.

All of the studied disease pairs exhibited moderate-to-strong genetic correlations and/or jointly associated SNPs located in the HLA region. We also found that there were several HLA loci associated with more than one disease pair. These findings align with the important role of HLA alleles in the genetic susceptibility to both IIM and common B cell lymphoma subtypes. First, we found significant genetic overlap and jointly associated SNPs in loci encoding the HLA class II genes for the disease pairs DM-FL, PM-FL, DM-CLL and PM-CLL. There are associations with haplotype *HLA-DRB1*01:01~DQA1*01:01~DQB1*05:01*, and *HLA-DPB1* reported for FL,[16 21 34–37] while for CLL, associations with haplotype *HLA-DRB4*01:01~DRB1*07:01~DQB1*03:03*, *HLA-DRB5*, *HLA-DQA1* and *BAK1* have been identified.[35 38–40] There are various HLA class II alleles in the HLA 8.1 ancestral haplotype including *HLA-DRB1*03:01*, *HLA-DQA1*05:01*, *HLA-DQB1*02:01* and *HLA-DPB1*01:01* found to confer risk of IIM.[21] For the disease pairs DM-DLBCL, DM-MZL, PM-DLBCL and PM-MZL, we detected shared genetic susceptibility in the HLA region encoding the *HLA-B* gene. A SNP (rs2922994) tagging a variant of this gene has been reported to increase the risk of MZL.[17] Furthermore, *HLA-B*08:01* has been identified as a genetic risk variant for DM, PM and DLBCL in previous studies.[15 19 21] Moreover, we found a significant correlation in the loci encoding HLA class III genes for all DM disease pairs. However, associations with different HLA class III genes have been identified for IIM (*C4A*, *AGER* and *NOTCH4*), DLBCL (*RDBP*), FL (*CFB*, *MSH5* and *TNXB*), CLL and MZL (*RDBP*).[41–44] Given the pleiotropic nature of the HLA region and its established role in autoimmune diseases, these findings likely reflect a broader pattern of shared genetic susceptibility beyond IIM and B cell lymphomas. For example, systemic lupus erythematosus—genetically correlated with IIM in the HLA class I and II regions—also shares genetic susceptibility with DLBCL and MZL in the HLA class I region, and with FL and CLL in the class II region.[41]

Antigen presentation and chronic B cell activation are key to the development of IIM and several B cell lymphoma subtypes.[1 2] Our findings suggest that there are germline variations in the HLA genes, which are important to antigen presentation, lymphocyte maturation and inflammation, making individuals more susceptible to develop IIM and B cell lymphomas. Interestingly, all DM disease pairs exhibited shared genetic susceptibility in the HLA class I and III regions, while the PM disease pairs showed genetic overlap primarily in the HLA class I and II regions. These observations suggest that there might be overlapping activities of inflammation and CD8+ cytotoxic T cells, which are activated by antigen presented by the HLA class I molecules, in the development of DM and common B cell lymphoma subtypes.[45] For PM disease pairs, activities of CD8+ cytotoxic T cells (all pairs) and CD4+ T effector cells (PM-FL and PM-CLL) could be the shared pathological components.[45] Given the extended LD structure in the HLA region, further investigation is needed to elaborate on the overlapping biological mechanism between IIM and common B cell lymphoma subtypes.

There was disease pair-specific shared genetic susceptibility found outside the HLA region for several disease pairs. It remains unclear how these non-HLA loci might contribute to the co-occurrence of IIM and B cell lymphomas. Functional annotation analyses revealed significant findings only for one locus. This locus [chr1:205 917 549–208 162 951], associated with DM and FL, encodes genes linked to primarily immune-related responses. The top five enriched gene sets, with overlapping genes *C4BPA*, *CD55*, *CR2*, *CR1* and *CD46*, are involved in B cell-mediated immunity and complement activation, suggesting shared biological processes between DM and

FL. Furthermore, this finding is in line with the genetic correlation between DM and FL found in the HLA class III region. For the disease pair PM-DLBCL, although gene sets at the locus [chr9:138 995 792–140 097 759] were not significantly enriched in any Gene Ontology-defined biological processes, this locus encodes the *NOTCH1* gene, of which somatic mutation is related to DLBCL.[3]

The findings from the conjFDR analyses provide insights into other non-HLA biological factors contributing to the development of IIM and B cell lymphomas. The risk variant rs10173316:T on chromosome 2 was jointly associated with DM and FL and linked to the *RBM43* and *NMI* genes. The *NMI* gene is a paralog of the *IFI35* gene, which has been found to be associated with IIM.[44] The interferon-induced protein 35 and N-Myc and STAT interactor can interact and form complex to regulate interferon-induced signalling pathways.[46] Moreover, the risk variant rs12203592:T on chromosome 6 jointly associated with DM, PM and CLL was linked to the *IRF4* gene, encoding the key transcriptional factor interferon regulatory factor 4 that is involved in the interferon signalling pathways.[47] This variant has been reported to be associated with IIM and CLL.[24 48–50] These findings suggest that the development of DM and FL, DM and CLL, and PM and CLL might overlap in pathways related to interferon-regulated inflammatory responses. Further investigations are warranted to validate these suggested overlapping biological process and pathways for IIM and B cell lymphomas.

Our findings reveal both overlapping and distinct genetic susceptibility features between DM and PM with common B cell lymphoma subtypes. The pathogenesis of malignancy in IIM appears to differ between DM and PM, as indicated by variations in the temporal relationship with cancer and associated cancer types.[5] However, both DM and PM show an increased risk of haematological malignancies, including NHL, within 3 years of diagnosis.[5 6 8 9] Genetically, our findings suggest different pathological links between DM and PM with B cell lymphomas. For example, pleiotropic HLA loci were found for both DM-DLBCL and PM-DLBCL, while shared genetic susceptibility outside the HLA region was also detected for PM-DLBCL. Further research is needed to determine whether this difference in shared genetic susceptibility translates into biological differences.

There was a notable disparity between the GWAS and the ImmunoChip findings, particularly beyond the HLA region, probably due to variations in genomic coverage and sample size. The GWAS data encompassed the whole genome, while the ImmunoChip data focused on immune-related loci and had a larger sample size, making it better suited for unveiling local genetic correlations within immune-related loci. Nevertheless, the GWAS data might be more effective in identifying genetic correlations in non-immune-related regions, owing to its superior genomic coverage, as evidenced in our study.

The replication of the GWAS findings in the ImmunoChip data could be considered a reinforcement of the reliability of these findings. Similarities were observed for certain HLA loci; for DM-CLL, shared genetic susceptibility was detected in the locus [chr6:32 629 240–32 682 213] in both datasets. Jointly associated SNPs rs9270493 and rs2596500 for PM-FL and PM-CLL, respectively, were also found in both datasets. Furthermore, correlation trends in the loci [chr6:31 320 269–31 427 209] for DM-DLBCL and [chr6:32 208 902–32 454 577] and [chr6:32 586 785–32 629 239] for DM-CLL from the GWAS data were statistically significant in the ImmunoChip data.

Apart from unveiling to what extent genetics might contribute to the co-occurrence of IIM and B cell lymphomas, our findings provide guidance to future genetic research exploring the genetic predisposition to either disease. Missing heritability has been suggested for both IIM and B cell lymphomas.[49 51–54] The loci showing significant genetic correlations between IIM and B cell lymphomas, but containing no SNPs previously reported to be associated with either IIM or B cell lymphomas, might serve as candidate genomic regions for discovering novel genetic associations in future genetic studies of IIM and B cell lymphomas.

The statistical power of LAVA depends on the sample size of the input data. Moreover, compared with continuous traits, LAVA has much lower power to detect local genetic correlations for binary traits. A simulation analysis using UK Biobank data, which included 276 731 individuals, found that for binary traits with a prevalence of 0.2 and without sample overlap between datasets, the power to detect local genetic correlations of 0.3, 0.6 and 0.9 was approximately 17%, 21% and 30%, respectively.[29] Although we used data from the largest GWASs of IIM and common B cell lymphoma subtypes in European ancestry worldwide, the corresponding sample sizes were far smaller than those needed to discover mild to strong local genetic correlations for rare diseases. Therefore, our analyses were potentially underpowered, and the observed local genetic correlations between IIM and B cell lymphomas might represent only a proportion of the total genetic correlations across all genomic loci. Nevertheless, this power issue is not specific to LAVA, which, under the same conditions, has better statistical efficiency compared with other alternative methods estimating local genetic correlations.[29] Furthermore, although LD between loci was minimised in the LAVA analyses, it is still possible that the observed genetic correlation was confounded by adjacent loci.[26] This issue might be more profound within the HLA region. Further investigation through HLA imputation and fine-mapping may help identify causal HLA alleles and amino acid residues jointly associated with both IIM and B cell lymphomas. Regrettably, reliable and robust HLA imputation requires individual-level data, which was absent for our study.[55]

LAVA cannot determine whether the observed correlation was horizontal (common biological pathway) or vertical (causal relationship) pleiotropy,[56] or a

combination of both, which has been suggested for other complex diseases.[57 58] Therefore, we interpreted findings from the LAVA analyses as proxy measures where shared genetic susceptibility between IIM and common B cell lymphoma subtypes might locate in the genome and used pleioFDR as a complementary approach. The causal relationship between IIM and B cell lymphomas could be explored using Mendelian randomisation methods in future studies. However, insufficient statistical power could be an issue.[59]

There are other limitations of the study. Patients with PM included in the GWAS and the ImmunoChip study were clinically defined based on the Bohan and Peter criteria.[19 21] Thus, the PM groups could include other IIM subtypes such as antisynthetase syndrome and IMNM that could be defined as PM using these criteria. It is known that features of clinical presentations, serology and immunohistology differ between these IIM subtypes, as well as associations with cancer.[1] Thus, our findings of the disease pairs of PM with B cell lymphoma subtypes might present an averaged effect over these IIM subtypes. Nevertheless, analyses by myositis-specific autoantibodies associated with an increased risk of cancer (ie, anti-transcriptional intermediary factor 1 γ autoantibodies) were not possible due to the small sample size and lack of imputed data. We also lacked information on cancer occurrence among patients from the GWAS and the ImmunoChip studies. However, patients with a cancer diagnosis within 2 years of the IIM diagnosis were excluded from both studies. Since frequencies of HLA alleles exhibit considerable variation across ethnic groups,[45] our findings may be primarily applicable to Caucasian populations.

Despite these limitations, we present the first attempt to explore the shared genetic susceptibility at a locus-specific level between IIM and common B cell lymphoma subtypes using GWAS summary statistics from the largest corresponding disease consortium in the world. All datasets were imputed to increase genomic coverage and make the screening of shared genetic susceptibility more comprehensive. To mitigate the infeasibility of global genetic correlation measure for diseases with low heritability like IIM, DLBCL, FL and MZL, we applied LAVA to estimate local genetic correlation instead.[51 52 60]

We revealed that IIM and B cell lymphomas shared genetic susceptibility across several genomic loci, primarily within the HLA region and in several non-HLA regions. Our findings suggest potential immune-related biological processes and molecular pathways overlapping between DM and FL, DM and CLL, PM and DLBCL, and PM and CLL. We also demonstrated heterogeneity in shared genetic susceptibility both between and within DM and PM in relation to B cell lymphomas. These findings are important for understanding contributing factors to the development of both IIM and B cell lymphomas.

Author affiliations
[1]Department of Public Health and Medicinal Administration, Faculty of Health Sciences, University of Macau, Taipa, Macao
[2]Clinical Epidemiology Division, Department of Medicine, Solna, Karolinska Institutet, Stockholm, Sweden
[3]Institute for Clinical and Translational Research, Baylor College of Medicine, Houston, Texas, USA
[4]Medical unit Hematology, Theme Cancer, Karolinska University Hospital, Stockholm, Sweden
[5]Rheumatology Division, Department of Medicine, Solna, Karolinska Institutet, Stockholm, Sweden
[6]ME Gastro, Dermatology and Rheumatology, Theme Inflammation and Aging, Karolinska University Hospital, Stockholm, Sweden

X Marie Holmqvist @marie_holmqvist

Acknowledgements This study is partially based on the author's doctoral thesis titled "Idiopathic Inflammatory Myopathies and Cancer: Familial Risk, Genetics and Consequences", which includes preliminary findings on the genetic correlation between idiopathic inflammatory myopathies and B cell lymphomas (see reference 25). The present manuscript has been expanded with new analysis, results and discussions. We would like to thank Lina Marcela Diaz-Gallo for outstanding advice on study design. We also want to acknowledge IMACS, the MYOGEN members, the LM GWAS Coordinating Committee and the InterLymph Data Coordinating Committee for their indispensable support in data application, transfer and critical manuscript review. Special thanks to Arial Harrison, Dakai Zhu and Dennis Robinson for assisting the data transfer. Other IMACS investigators from the MYOGEN contributing to this study and the list of individuals of studies from the InterLymph Consortium are listed in the Collaborators section and the online supplementary appendix, respectively.

Collaborators IMACS Myositis Genetics Scientific Interest Group (MYOGEN): Chris Amos (Baylor College of Medicine, Houston, Texas). Olivier Benveniste (Department of Internal Medicine and Clinical Immunology, Pitié-Salpêtrière Hospital, Paris, France). Hector Chinoy (National Institute for Health Research Manchester Biomedical Research Centre, Manchester University NHS Foundation Trust, The University of Manchester, Manchester, UK, and Department of Rheumatology, Salford Royal Hospital, Northern Care Alliance NHS Foundation Trust, Manchester Academic Health Science Centre, Salford, UK, and Centre for Musculoskeletal Research, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK). Jan L De Bleecker (Department of Neurology, Ghent University, Ghent, Belgium). Boel De Paepe (Department of Neurology, Ghent University, Ghent, Belgium). Andrea Doria (Rheumatology Unit, Department of Medicine, University of Padova, Padova, Italy). Zoltan Griger (University of Debrecen, Faculty of Medicine, Department of Internal Medicine, Division of Clinical Immunology, Debrecen, Hungary). Peter Gregersen (The Robert S. Boas Center for Genomics and Human Genetics, The Feinstein Institute, Manhasset, New York). Michael G Hanna (Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology, and Centre for Rheumatology, UCL Division of Medicine, University College London, London, UK). Vidya Limaye (Rheumatology Unit, Royal Adelaide Hospital and Discipline of Medicine, Adelaide University, Adelaide, Australia). Leonid Padyukov (Division of Rheumatology, Department of Medicine, Solna, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden). Pedro Machado (Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology, and Centre for Rheumatology, UCL Division of Medicine, University College London, London, UK). Pernille Mathiesen (Paediatric Department, Slagelse Hospital and Paediatric Rheumatology Unit, Rigshospitalet, Copenhagen, Denmark). Britta Maurer (Department of Rheumatology and Immunology, University Hospital, Bern, Switzerland). Frederick W Miller (Environmental Autoimmunity Group, National Institute of Environmental Health Sciences, NIH, Bethesda, Maryland). Lisa G Rider (Environmental Autoimmunity Group, National Institute of Environmental Health Sciences, NIH, Bethesda, Maryland). Dr. Miller's and Dr. Lisa Rider's work was supported by the Intramural Research Program of the National Institute of Health (NIH) (Project ZIA ES101074). Øyvind Molberg (Department of Rheumatology, Oslo University Hospital, Oslo, Norway). Lauren Pachman (Ann & Robert H. Lurie Children's Hospital of Chicago, Northwestern University Feinberg School of Medicine, Chicago, Illinois). Timothy Radstake (Department of Rheumatology and Clinical Immunology, University Medical Center, Utrecht, the Netherlands). Ann M Reed (Department of Pediatrics, Duke University, Durham, North Carolina). Simon Rothwell (Centre for Genetics and Genomics Versus Arthritis, Centre for Musculoskeletal Research, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK). Adam Schiffenbauer (Environmental Autoimmunity Group, National Institute of Environmental Health Sciences, NIH, Bethesda, Maryland.) Albert Selva-O'Callaghan (Internal Medicine Department, Vall d'Hebron General Hospital, Universitat Autonoma de Barcelona, Barcelona, Spain). Jiri Vencovksy (Institute of Rheumatology and Department of Rheumatology, First Medical Faculty, Charles University, Prague, Czech Republic). Lucy R Wedderburn (NIHR Biomedical Research Centre at Great Ormond Street Hospital, and Arthritis Research UK Centre for Adolescent Rheumatology, UCL Great Ormond Street Institute of Child Health, University College London, London, UK). International

Lymphoma Epidemiology Consortium (InterLymph):Sonja Berndt. Parveen Bhatti. Lauren Teras. Alain Monnereau. Roel Vermeulen. Delphine Casabonne. Alexandra Nieters. James McKay. Brenda Birmann. Pierluigi Cocco. James Cerhan. Brian Link. Roger Milne. Harindra Jayasekara. Vijai Joseph. Kenneth Offit. Nathaniel Rothman. Claire Vajdic. Alan Arslan. Mark Purdue. Henrik Halgrim. Paige Bracci. Eleanor Kane. Lesley Tinker. Qing Lan. Susan Slager. Immaculata De Vivo. Gilles Salles. Nicola Camp.

Contributors MH is guarantor. All authors conceived and designed the work. CZ imputed the GWAS and ImmunoChip data and performed the association testing. WIC and KP manipulated and standardised the GWAS summary statistics. WIC performed all analyses, interpreted the findings and drafted the manuscript. All authors contributed to revising the manuscript critically and for intellectual content. All authors read and approved the final version of manuscript for submission. Microsoft Copilot was used for language editing assistance, with all authors actively reviewing and approving all textual modifications.

Funding This study is supported by the Swedish Research Council and UM Macao Fellow scheme of the UM Talent Programme of the University of Macau (File no. UMTP2023-MF02). The previous GWAS and ImmunoChip studies of IIM are supported by the Association Francaise Contre Les Myopathies (AFM); European Union Sixth Framework Programme (project AutoCure; LSH-018661); European Science Foundation (ESF) in the framework of the Research Networking Programme European Myositis Network (EUMYONET); European Reference Network for Rare Neuromuscular Diseases (ERN EURO-NMD); The Swedish Research Council and grants provided by the Stockholm County Council (ALF project), the Swedish Rheumatism Association, King Gustaf V 80 year Foundation; the intramural research program of the National Institute of Environmental Health Sciences (NIEHS), the National Institutes of Health (NIH) (Project ZIA ES101074) ; Myositis UK; Arthritis Research UK (awards 18474, 21593, 21552 and 20380); Medical Research Council (MR/N003322/1); The Cure JM Foundation; the Wellcome Trust (awards 085860 and 105610); the Henry Smith Charity UK; Action Medical UK; the Remission Charity; BMA Dorris Hillier Grant 2012 and the Bath Institute of Rheumatic Diseases; Prinses Beatrix Spierfonds (W.OR12-15); the Research Fund of Region Zeeland, Denmark; the Ministry of Health, Czech Republic (Project for Conceptual Development of Research Organization grant 00023728); National Institute for Health Research (NIHR) Biomedical Research Centres (BRC) at Great Ormond Street Hospital (GOSH), Manchester and University College London Hospitals. The contributions of the NIH author(s) were made as part of their official duties as NIH federal employees, are in compliance with agency policy requirements, and are considered Works of the United States Government. However, the findings and conclusions presented in this paper are those of the author(s) and do not necessarily reflect the views of the NIH or the U.S. Department of Health and Human Services. The InterLymph consortium is supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institute of Health. These funding sources were not involved in any parts of the presented research.

Competing interests IL and MH have received grants from Swedish Research Council. MH is a member of the medical advisory board of the Myositis Association. The other authors declare no conflicts of interest. The funding sources for the investigators from the IMACS Myositis Genetics Scientific Interest Group and the individual studies within the InterLymph consortium are listed in this manuscript and the online supplementary appendix, respectively.

Patient and public involvement statement The data sources used in the present study were derived from previous genetic studies of the corresponding diseases where patients and control individuals contributed substantially by providing clinical and biological data. We however did not involve patients in any stage of the present study.

Patient consent for publication Not applicable.

Ethics approval This study was approved by the Swedish Ethical Review Authority (2020-04529 and 2023-01343-02).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information. Not applicable.

ORCID iDs
Weng Ian Che http://orcid.org/0000-0002-8355-2024
Anton Öberg Sysojev http://orcid.org/0000-0003-3221-7979
Ingrid E Lundberg http://orcid.org/0000-0002-6068-9212
Helga Westerlind http://orcid.org/0000-0003-3380-5342
Marie Holmqvist http://orcid.org/0000-0001-8996-5260

REFERENCES

1 Lundberg IE, Fujimoto M, Vencovsky J, *et al*. Idiopathic inflammatory myopathies. *Nat Rev Dis Primers* 2021;7:86.
2 Smedby KE, Ponzoni M. The aetiology of B-cell lymphoid malignancies with a focus on chronic inflammation and infections. *J Intern Med* 2017;282:360–70.
3 Rosenquist R, Beà S, Du M -Q., *et al*. Genetic landscape and deregulated pathways in B-cell lymphoid malignancies. *J Intern Med* 2017;282:371–94.
4 Campo E, Jaffe ES, Cook JR, *et al*. The International Consensus Classification of Mature Lymphoid Neoplasms: a report from the Clinical Advisory Committee. *Blood* 2022;140:1229–53.
5 Dani L, Ian Che W, Lundberg IE, *et al*. Overall and site-specific cancer before and after diagnosis of idiopathic inflammatory myopathies: A nationwide study 2002–2016. *Semin Arthritis Rheum* 2021;51:331–7.
6 Hill CL, Zhang Y, Sigurgeirsson B, *et al*. Frequency of specific cancer types in dermatomyositis and polymyositis: a population-based study. *Lancet* 2001;357:96–100.
7 Stockton D, Doherty VR, Brewster DH. Risk of cancer in patients with dermatomyositis or polymyositis and follow-up implications: a Scottish population-based cohort study. *Br J Cancer* 2001;85:41–5.
8 Kang EH, Lee SJ, Ascherman DP, *et al*. Temporal relationship between cancer and myositis identifies two distinctive subgroups of cancers: impact on cancer risk and survival in patients with myositis. *Rheumatology (Sunnyvale)* 2016;55:1631–41.
9 Hu S, Zhou D, Wu Y, *et al*. Autoimmune disease-associated non-Hodgkin's lymphoma—a large retrospective study from China. *Ann Hematol* 2019;98:445–55.
10 Hemminki K, Huang WQ, Sundquist J, *et al*. Autoimmune diseases and hematological malignancies: Exploring the underlying mechanisms from epidemiological evidence. *Semin Cancer Biol* 2020;64:114–21.
11 Anderson LA, Gadalla S, Morton LM, *et al*. Population-based study of autoimmune conditions and the risk of specific lymphoid malignancies. *Intl Journal of Cancer* 2009;125:398–405.
12 Marie I, Guillevin L, Menard J-F, *et al*. Hematological malignancy associated with polymyositis and dermatomyositis. *Autoimmun Rev* 2012;11:615–20.
13 Hemminki K, Försti A, Sundquist K, *et al*. Familial associations of lymphoma and myeloma with autoimmune diseases. *Blood Cancer J* 2017;7:e515.
14 Che WI, Baecklund F, Hellgren K, *et al*. Familial co-aggregation of idiopathic inflammatory myopathies and cancer - a Swedish population-based study. *Arthritis Rheumatol* 2023.
15 Cerhan JR, Berndt SI, Vijai J, *et al*. Genome-wide association study identifies multiple susceptibility loci for diffuse large B cell lymphoma. *Nat Genet* 2014;46:1233–8.
16 Skibola CF, Berndt SI, Vijai J, *et al*. Genome-wide Association Study Identifies Five Susceptibility Loci for Follicular Lymphoma outside the HLA Region. *Am J Hum Genet* 2014;95:462–71.
17 Vijai J, Wang Z, Berndt SI, *et al*. A genome-wide association study of marginal zone lymphoma shows association to the HLA region. *Nat Commun* 2015;6:5751.
18 Berndt SI, Camp NJ, Skibola CF, *et al*. Meta-analysis of genome-wide association studies discovers multiple loci for chronic lymphocytic leukemia. *Nat Commun* 2016;7:10933.
19 Rothwell S, Cooper RG, Lundberg IE, *et al*. Dense genotyping of immune-related loci in idiopathic inflammatory myopathies confirms HLA alleles as the strongest genetic risk factor and suggests different genetic background for major clinical subgroups. *Ann Rheum Dis* 2016;75:1558–66.
20 Miller FW, Cooper RG, Vencovský J, *et al*. Genome-Wide Association Study of Dermatomyositis Reveals Genetic Overlap With Other Autoimmune Disorders. *Arthritis Rheum* 2013;65:3239–47.

21 Miller FW, Chen W, O'Hanlon TP, *et al*. Genome-wide association study identifies HLA 8.1 ancestral haplotype alleles as major genetic risk factors for myositis phenotypes. *Genes Immun* 2015;16:470–80.
22 Taliun D, Harris DN, Kessler MD, *et al*. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature New Biol* 2021;590:290–9.
23 Zhu C, Han Y, Byun J, *et al*. n.d. Meta-analyses uncover the genetic architecture of idiopathic inflammatory myopathies. In press. *Arthritis Rheumatol*.
24 Zhu C, Han Y, Byun J, *et al*. Genetic Architecture of Idiopathic Inflammatory Myopathies From Meta-Analyses. *Arthritis Rheumatol* 2024.
25 Che WI. Idiopathic inflammatory myopathies and cancer: familial risk, genetics and consequences [Ph.d]. Sweden, Karolinska Institutet (Sweden), 2024
26 Werme J, van der Sluis S, Posthuma D, *et al*. An integrated framework for local genetic correlation analysis. *Nat Genet* 2022;54:274–82.
27 Leeuw Jw. LAVA: local analysis of [co]variant association. R package version 0.1.0. 2023.
28 Team RC. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2021.
29 Zhang C, Zhang Y, Zhang Y, *et al*. Benchmarking of local genetic correlation estimation methods using summary statistics from genome-wide association studies. *Brief Bioinformatics* 2023;24.
30 Andreassen OA, Thompson WK, Schork AJ, *et al*. Improved Detection of Common Variants Associated with Schizophrenia and Bipolar Disorder Using Pleiotropy-Informed Conditional False Discovery Rate. *PLoS Genet* 2013;9:e1003455.
31 Schork AJ, Thompson WK, Pham P, *et al*. All SNPs Are Not Created Equal: Genome-Wide Association Studies Reveal a Consistent Pattern of Enrichment among Functionally Annotated SNPs. *PLoS Genet* 2013;9:e1003449.
32 Karczewski KJ, Francioli LC, Tiao G, *et al*. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature New Biol* 2020;581:434–43.
33 Watanabe K, Taskesen E, van Bochoven A, *et al*. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* 2017;8:1826.
34 Conde L, Halperin E, Akers NK, *et al*. Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. *Nat Genet* 2010;42:661–4.
35 Smedby KE, Foo JN, Skibola CF, *et al*. GWAS of Follicular Lymphoma Reveals Allelic Heterogeneity at 6p21.32 and Suggests Shared Genetic Susceptibility with Diffuse Large B-cell Lymphoma. *PLoS Genet* 2011;7:e1001378.
36 Cerhan JR, Fredericksen ZS, Novak AJ, *et al*. A Two-Stage Evaluation of Genetic Variation in Immune and Inflammation Genes with Risk of Non-Hodgkin Lymphoma Identifies New Susceptibility Locus in 6p21.3 Region. *Cancer Epidemiol Biomarkers Prev* 2012;21:1799–806.
37 Skibola CF, Conde L, Foo J-N, *et al*. A meta-analysis of genome-wide association studies of follicular lymphoma. *BMC Genomics* 2012;13:516.
38 Slager SL, Rabe KG, Achenbach SJ, *et al*. Genome-wide association study identifies a novel susceptibility locus at 6p21.3 among familial CLL. *Blood* 2011;117:1911–6.
39 Slager SL, Skibola CF, Di Bernardo MC, *et al*. Common variation at 6p21.31 (BAK1) influences the risk of chronic lymphocytic leukemia. *Blood* 2012;120:843–6.
40 Gragert L, Fingerson S, Albrecht M, *et al*. Fine-mapping of HLA associations with chronic lymphocytic leukemia in US populations. *Blood* 2014;124:2657–65.
41 Din L, Sheikh M, Kosaraju N, *et al*. Genetic overlap between autoimmune diseases and non-Hodgkin lymphoma subtypes. *Genet Epidemiol* 2019;43:844–63.
42 Zhou D, King EH, Rothwell S, *et al*. Low copy numbers of complement C4 and C4A deficiency are risk factors for myositis, its subgroups and autoantibodies. *Ann Rheum Dis* 2023;82:235–45.
43 Song J, Kim D, Hong J, *et al*. Meta-Analysis of Polymyositis and Dermatomyositis Microarray Data Reveals Novel Genetic Biomarkers. *Genes (Basel)* 2019;10:864.
44 Bianchi M, Kozyrev SV, Notarnicola A, *et al*. Contribution of rare genetic variation to disease susceptibility in a large Scandinavian myositis cohort. *Arthritis Rheumatol* 2022.
45 Trowsdale J, Knight JC. Major Histocompatibility Complex Genomics and Human Disease. *Annu Rev Genom Hum Genet* 2013;14:301–23.
46 De Masi R, Orlando S, Bagordo F, *et al*. IFP35 Is a Relevant Factor in Innate Immunity, Multiple Sclerosis, and Other Chronic Inflammatory Diseases: A Review. *Biology (Basel)* 2021;10:1325.
47 Jefferies CA. Regulating IRFs in IFN Driven Disease. *Front Immunol* 2019;10:325.
48 Di Bernardo MC, Crowther-Swanepoel D, Broderick P, *et al*. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. *Nat Genet* 2008;40:1204–10.
49 Speedy HE, Di Bernardo MC, Sava GP, *et al*. A genome-wide association study identifies multiple susceptibility loci for chronic lymphocytic leukemia. *Nat Genet* 2014;46:56–60.
50 Berndt SI, Skibola CF, Joseph V, *et al*. Genome-wide association study identifies multiple risk loci for chronic lymphocytic leukemia. *Nat Genet* 2013;45:868–76.
51 Rothwell S, Lamb JA, Chinoy H. New developments in genetics of myositis. *Curr Opin Rheumatol* 2016;28:651–6.
52 Che WI, Westerlind H, Lundberg IE, *et al*. Familial aggregation and heritability: a nationwide family-based study of idiopathic inflammatory myopathies. *Ann Rheum Dis* 2021;80:1461–6.
53 Chiara Di Bernardo M, Broderick P, Catovsky D, *et al*. Common genetic variation contributes significantly to the risk of developing chronic lymphocytic leukemia. *Haematologica* 2013;98:e23–4.
54 Law PJ, Berndt SI, Speedy HE, *et al*. Genome-wide association analysis implicates dysregulation of immunity genes in chronic lymphocytic leukaemia. *Nat Commun* 2017;8:14175.
55 Sakaue S, Gurajala S, Curtis M, *et al*. Tutorial: a statistical genetics guide to identifying HLA alleles driving complex disease. *Nat Protoc* 2023;18:2625–41.
56 van Rheenen W, Peyrot WJ, Schork AJ, *et al*. Genetic correlations of polygenic disease traits: from theory to practice. *Nat Rev Genet* 2019;20:567–81.
57 Verbanck M, Chen CY, Neale B, *et al*. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet* 2018;50:693–8.
58 Zhu Z, Zheng Z, Zhang F, *et al*. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat Commun* 2018;9:224.
59 Burgess S, Davey Smith G, Davies NM, *et al*. Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res* 2019;4:186.
60 Berndt SI, Vijai J, Benavente Y, *et al*. Distinct germline genetic susceptibility profiles identified for common non-Hodgkin lymphoma subtypes. *Leukemia* 2022;36:2835–44.