Relationship between Genetic Risk and Age of Diagnosis of Systemic Lupus Erythematosus

Daniela Dominguez MSc, Division of Rheumatology, Hospital for Sick Children, Hospital for Sick Children, Research Institute, University of Toronto, Toronto, Ontario

Sylvia Kamphuis, MD PhD, Division of Rheumatology Department of Pediatrics, Sophia Children's Hospital, Erasmus Medical Center, Rotterdam, Netherlands

Joseph Beyene, PhD, Department of Clinical Epidemiology & Biostatistics, McMaster University, Hamilton, Ontario

Joan Wither, MD PhD, Division of Genetics and Development, Krembil Research Institute, Arthritis Centre of Excellence, Division of Rheumatology, Toronto Western Hospital, University Health Network, Departments of Medicine and Immunology, University of Toronto, Toronto, Ontario.

John B. Harley, MD PhD Center for Autoimmune Genomics and Etiology(CAGE), Cincinnati Children's Hospital Medical Center, Cincinnati, OH

Irene Blanco MD, Albert Einstein College of Medicine, Division of Rheumatology Bronx, NY

Catarina Vila-Inda MD, Albert Einstein College of Medicine, Division of Rheumatology Bronx, NY

Hermine Brunner, MD MSc, Division of Rheumatology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH,

Marissa Klein-Gitleman MD, Ann & Robert H. Lurie Children's Hospital of Chicago Northwestern University Feinberg School of Medicine, Chicago, IL

Deborah McCurdy MD Division of Pediaitric Rheumatology, University of California Los Angeles, Los Angeles, California

Dawn M. Wahezi, MD Children's Hospital at Montefiore, Division of Pediatric Rheumatology, Albert Einstein College of Medicine, Bronx, New York

Thomas Lehman MD, Division of Pediatric Rheumatology, Hospital for Special Surgery New York, New York

Marija Jelusic MD, Department of Pediatric Rheumatology, University of Zagreb School of Medicine, Croatia

Canadian Network for Improved Outcomes in SLE (CaNIOS)

Christine A. Peschken, MD MSc, Departments of Medicine and Community Health Sciences, University of Manitoba, Winnipeg, Manitoba;

Janet E. Pope, MD, MPH, Professor of Medicine, Department of Medicine, University of Western, Ontario, London;

Dafna D. Gladman, MD, Department of Medicine, University of Toronto, Toronto, Ontario;

John G Hanly MD, Division of Rheumatology, Department of Medicine and Department of Pathology, Queen Elizabeth II Health Sciences Center and Dalhousie University, Halifax, Nova Scotia

Ann E. Clarke, MD, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada

Sasha Bernatsky, MD, PhD, Department of Medicine, McGill University, Montreal, Quebec,

Christian Pineau, MD, Department of Medicine, McGill University Hospital, Montreal, Quebec;

C. Douglas. Smith, MD, Department of Medicine, Faculty of Health Sciences, University of Ottawa, Ottawa, Ontario;

Susan Barr MD Division of Rheumatology, Department of Medicine, University of Calgary, Calgary, Alberta,

Gilles Boire MD Division of Rheumatology, Department of Medicine, Faculty of Medicine and Health Sciences, Universite´ de Sherbrooke, Sherbrooke, Quebec

Eric Rich, Division of Rheumatology, Centre Hospitalier de l'Universite´ de Montreal, Department of Medicine, University of Montreal School of Medicine, Montreal, Quebec.

Earl Silverman, Division of Rheumatology, Department of Pediatrics, SickKids Hospital, SickKids Hospital Research Institute, University of Toronto, Toronto, Ontario.

APPLE Investigators

Corresponding author: Earl Silverman MD FRCPC
Professor Emeritus of Pediatrics
Division of Rheumatology
Hospital for Sick Children
555 University Ave.
Toronto, Ontario, M5N 1R2
Phone: 416-813-6249
FAX: 416-813-4989
Email:earl.silverman@sickkids.ca

Competing Interests: There are no competing interest for the PI or any co-investigators

Contributions of co-authors: EDS, DD, SK, JB and JW were instrumental in the design of the study and writing the paper. All authors contributed patient data and material and read and edited the paper.

Disclosure: There were no financial or other commercial support for any of the work associated with this manuscript in anyway.

Ethics Approval and Patient Involvement: Coordinating center ethics approval-Hospital for Sick Children REB number 100007761. In addition, ethics was obtained at all centers as were data and material transfer agreements. All patients gave informed consent for genetic and clinical data for use in this study

Data sharing: Anonymized data will be available upon request to DD or EDS.

Key words- Genetic risk score, SLE, SNPs, age of onset

Abstract

Objective

Specific risk alleles for childhood-onset SLE (cSLE) versus adult-onset SLE (aSLE) patients have not been identified. The aims of this study were to determine if: 1) There is an association between non-HLA-related genetic risk score (GRS) and age of SLE diagnosis; and if 2) There is an association between HLA-related genetic risk score and age of SLE diagnosis.

Methods

Genomic DNA was obtained from 2,001 multi-ethnic patients and genotyped using the Immunochip. Following quality control, genetic risk counting (GRCS), weighted (GRWS) and standardized counting (GRSCS) and standardized weighted (GRSWS) scores were calculated based on independent SNPs from validated SLE-loci. Scores were analyzed in a regression model and adjusted by sex and ancestral population.

Results

The analysed cohort consisted of 1,540 patients: 1,351 females and 189 males (675 cSLE and 865 aSLE). There were significant negative associations with age of SLE diagnosis $p=0.011$ and $r^2=0.175$ for GRWS, $p=0.008$ and $r^2=0.178$ for GRSCS, $p=0.002$ and $r2=0.176$ for GRSWS for all non-HLA genetic risk scores (higher GRS the lower the age of diagnosis.) All HLA genetic risk scores showed significant positive associations with age of diagnosis $p=0.049$ and $r2=0.176$ for GRCS, $p=0.022$ and $r^2=0.176$ for GRWS, $p=0.022$ and $r^2=0.176$ for GRSCS, $p=0.011$ and $r^2=0.177$ for GRSWS: higher genetic scores correlated with higher age of diagnosis.

Conclusion

Our data suggested that there is a linear relationship between genetic risk and age of SLE diagnosis and that HLA and non-HLA genetic risk scores are associated with age of diagnosis in opposite directions.

Introduction

Systemic Lupus Erythematosus (SLE) is a chronic multi-system autoimmune disease with a peak incidence in females during childbearing years. SLE tends to be more severe in: men, patients with younger age of onset, and specific genetic ancestries [1-4]. Multiple genes have been implicated in its pathogenesis with >90 genes/loci associated with predisposition to SLE [5]. Although the majority of SLE-susceptibility genes across ancestral populations are in the same gene, the associated single nucleotide polymorphisms (SNPs) may differ or convey different risks for the development of SLE [6-9]. SLE-associated genetic risk variants are located in both HLA and non-HLA regions with alleles within the HLA region showing the strongest disease association [8,10-13].

For polygenic diseases such as SLE, it is accepted that a genetic risk score provides better information on the genetic contribution to the development of autoimmune diseases than investigating a single SNP [14-16]. There are two popular models for the calculation of polygenic risk scores: 1) A counting score that is a simple sum of risk and protective alleles present in an individual and 2) A weighted score that takes into account the effect size of the SNP. Although there have been previous publications of polygenic risk scores in SLE, few examined if a risk score can be used as a predictor of age of disease onset and only one examined a non-HLA genetic risk score over a large multi-ethnic population [17-24]. This is the first study to examine the influence of HLA and non-HLA genes separately in a large combined pediatric and adult across a multi-ethnic population.

The aims of this study were to determine if:

1) There is an association of genetic risk score and age of SLE diagnosis for HLA genes.

2) There is an association of genetic risk score and age of diagnosis of SLE for non-HLA genes.

Methods

SLE Cohort

This is a multi-centre, multi-national genetic study of patients with both childhood-onset SLE (cSLE) and adult-onset SLE (aSLE). cSLE was defined as SLE diagnosed at age <18 years and aSLE as age of diagnosis of $\geq$ 18 years. Genomic DNA was collected from 2,001 patients from 17 centres within North America and one from Europe (Supplementary Table 1) who fulfilled $\geq$ 4 of 11 ACR classification criteria for SLE [25], with an age of disease diagnosis range of 1-82 years. The following clinical variables were available in 1,979 of the 2001 patients: Date of diagnosis, date of birth, age at diagnosis, and sex. The study was approved by the Research Ethics Board at each centre and all participants signed an informed consent form prior to obtaining DNA and clinical information.

Genotyping

All of the 1979 DNA samples with clinical information were sent for genotyping using the Immunochip Illumina Infinium genotyping chip [Illumina, Inc San Diego CA, USA], at one of two centres: HudsonAlpha Genomics Services Laboratory and Cincinnati Children's Hospital Medical Center (CCHMC)-Harley Laboratory. Quality control (QC) of the genotype data was conducted using SNP & Variation Suite v8 software [Golden Helix, Inc., Bozeman, MT, www.goldenhelix.com] for each of the 4 different genotyping runs. Poor quality samples with low call rates (<95%), sample contamination, mixed samples or duplication and close relatedness to another sample in the study were excluded. Relatedness between subjects was estimated by identity by descent analysis (IBDA). The total number of samples after QC was 1773/1979 samples genotyped.

SNP Selection

By literature review of genome wide association studies (GWAS), meta-GWAS [9,26,27] and candidate gene and replication studies [28-31], we found 106 SLE-associated SNPs of which 93 SNPs were present on the Immunochip (Supplementary Table 2). These 93 SNPs represented 39 different SLE-associated loci. All SNPs with a call rate <0.99, minor allele frequency (MAF) <0.05 and Hardy-Weinberg Equilibrium p-value <0.001 were excluded. After QC there were 68 SNPs in 33 different loci; 58 in non-HLA areas and 10 in the HLA complex (Supplemental Table 3). Of the 13 SNPs that were not present on the Immunochip we identified 4 proxies (all in HLA region) using SNP Annotation and Proxy Search (SNAP) online tool version 2.2 from the Broad Institute (Harvard University which result in a total of 72 SNPs (58 non-HLA and 14 HLA). SNPs that were located in the same gene were tested for pairwise high linkage disequilibrium (LD) using the SNAP online tool. In the case of pairs of SNPs that met the threshold criteria ($r^2 \geq 0.8$)

in a specific gene for LD, the one with the lower odds ratio (OR) reported in the literature was excluded. A further 14 SNPS were eliminated leaving 58 independent SNPs (48 non-HLA and 10 HLA) from 33 SLE-susceptibility loci (Supplemental Table 3).

Genetic Score Risk Calculation

Odds Ratios (OR) for all SNPs were classified as protective (OR <1) or risk (OR >1) by ancestral population (Supplementary Table 3).

Four types of genetic risk scores were calculated for the statistical analysis: 1) genetic risk counting scores (GRCS), 2) genetic risk weighted scores (GRWS) and 3) genetic risk standardized counting score (GRSCS) and 4) genetic risk standardized weighted score (GRSWS). All scores were calculated separately for HLA and non-HLA SNPs and analyzed in the total population.

1) The GRCS was an additive genetic model based on the presence of the risk or protective allele and was determined as the sum of those alleles present in each individual (Supplementary Table 3): simple count of the risk alleles minus the protective alleles.

2) The GRWS accounts for the relative effect of each risk/protective allele by using the OR. The GRWS was calculated as the sum of the natural logarithm of the OR of each risk/protective allele present. The OR of a risk allele was positive and OR of a protective allele was negative (Supplementary Table 3).

3) Considering the large variation in the number of HLA and non-HLA SNPs identified in each individual ancestral population, we standardized the maximum GRCS and GRWS to 10 for each population to produce standardized GRCS (GRSCS) and GRWS (GRSWS) using only the SNPs available for the individual ancestral population. This allowed for comparisons across ancestral populations, to weigh each population equally (Figure 1).

Determination of ancestral population

Principal component analysis (PCA) was used to determine the ancestral identity of each patient. We first ran the analysis in the whole population comparing the first two principal components against reference samples of known ethnicities from HapMap3. Samples that were outliers from the calculated clusters were dropped from the study. Multidimensional Outlier

Detection (MOD) analysis was performed for each ancestral population individually until we did not detect any outliers. PCA/MOD analysis eliminated 131 patients and a further 102 patients were eliminated as they were identified as South Asian ancestry (an ancestry without any applicable gene studies). Therefore, the population analyzed was 1,540.

Statistical Analysis

Since sex and ancestral population are suggested to influence genetic susceptibility to SLE, we first determined the distribution of both across all age groups. We also determined the association of these factors with genetic risk scores and age of SLE diagnosis in each model. We used linear regression analysis to determine if genetic risk scores varied by age of SLE diagnosis for the whole cohort. As the age of diagnosis distribution did not follow a normal distribution, the natural logarithm (ln) of the age of diagnosis was used for statistical analysis as it followed a normal distribution.

Both ln of age of SLE diagnosis (dependent variable) and genetic risk score (predictor variable) were analyzed as continuous variables. A p-value of ≤ 0.05 was considered statistically significant. The percent variation in the dependent variable explained by the predictor is quantified using the adjusted R squared statistic in each model. The effects of sex and ancestral population were tested for interactions in the final model. SNP & Variation Suite v8 software [Golden Helix, Inc., Bozeman, MT, www.goldenhelix.com], R version 3.1.2 statistical package [R Foundation for Statistical Computing, Vienna, Austria, www.R-project.org] and StatPlus:Mac [AnalystSoft, USA, www.analystsoft.com] were used for the statistical analyses.

Results

Descriptive statistics

The analysed cohort consisted of 1,540 patients: 1,351 females and 189 males; 1,094 were Caucasian (71.0%), 196 Black (12.7%), 129 Hispanic (8.4%) and 121 of Asian (7.9%) ancestry. Mean age of diagnosis in the total population of 25.3 years (SD=14.2) and median age of 21.0 years (Table 1). The mean age at diagnosis of females at 25.8 (SD=14.3) years was statistically significantly higher than that of males at 21.5 (SD=13.1) (p=4.16x10^{-5}). The percentage of cSLE patients in the total male population (59.3%) was higher than the percentage of cSLE patients in the total female population (41.7%) with lower absolute number (112 males vs. 563 females).

Non-HLA genetic risk scores

We initially determined the association of sex and ancestral population with genetic risk scores and age of SLE diagnosis. We found that for all non-HLA genetic risk scores there were statistically significant associations between sex and age of SLE diagnosis. As a result, sex was included in our statistical model. There was a statistically significant association between sex and GRSWS (p=0.013, regression coefficient= 0.180), but not between sex and the other non-HLA genetic risk scores. Significant associations were seen with age of diagnosis and specific ancestries for GRCS (p<2x10^{-16} for Caucasian and p=0.033 for Hispanic ancestry) and GRSCS (p<2x10^{-16} for Caucasian and p=0.001 for Black ancestry) while for GRWS and GRSWS the only statistically significant association with age of diagnosis was in Caucasian ancestry (p<2x10^{-16} for GRWS and GRSWS). Therefore, the final statistical model for GRCS and GRSCS included sex and all 4 ancestral populations while the final model for GRWS and GRSWS included sex and Caucasian vs. non-Caucasian (C/NC) ancestry as covariates.

The final linear regression models for all non-HLA genetic risk scores, except the GRCS, showed a statistically significant negative association with age (Table 2). These models explained similar percentages of the variance of the genetic contribution (17.8%, 17.5% and 17.6% respectively) (Supplementary Figures 1-3).

HLA genetic risk scores

There were statistically significant associations between sex and age of SLE diagnosis for all 4 HLA genetic risk scores. Therefore, sex was included in our statistical model. However, there were no statistically significant associations between sex and any of the 4 HLA genetic risk

scores. All the ancestral populations showed statistically significant associations with age of SLE diagnosis for all 4 risk scores. However, in contrast to the results for non-HLA genetic risk scores, the final linear regression models of all 4 HLA genetic risk scores with age of SLE diagnosis, which included sex and all 4 ancestral populations as covariates, showed statistically significant positive associations with age (Table 3). All 4 models explained almost identical percentages of the variance of the genetic contribution (17.6% or 17.7%) (Supplementary Figures 4-7).

Discussion

It has been suggested that the genetic contribution to the development of SLE likely differs between cSLE and aSLE. However, candidate gene studies have not identified any genes that were specific or unique to cSLE, no GWAS has been performed in cSLE although a study has suggested that unique SNPs were found in a Korean cSLE population [21,24,32]. For these reasons, it was necessary to apply a different genetic approach to better understand the genetics of SLE across all ages. In this study we used polygenic risk scores to better understand the genetic association with age of SLE diagnosis. Our data has suggested that there is a linear relationship between genetic risk and age of SLE diagnosis and that HLA and non-HLA genetic risk scores influence age of diagnosis of SLE differently.

HLA and non-HLA genes could play different roles in disease susceptibility due to their different degrees of relative risk for the development of SLE [10,12] and therefore could have different effects in predicting the age of onset of SLE. We found that for non-HLA SNPs there was a negative association of genetic risk score with age of SLE diagnosis, i.e. the higher genetic risk score the younger the patient and 18% of the variation in age of onset of SLE was explained by our model. However, when HLA genetic risk scores were determined, there was a positive association of HLA SNPs with age of SLE diagnosis, i.e. the higher the risk score the older the patient. Therefore, we have for the first time shown that HLA and non-HLA SNPs may contribute differently to the age of SLE diagnosis. This may explain why our findings appear to be different from previous studies that combined both HLA and non-HLA SNPs in determining the association of genetic risk score and age of onset of SLE [17,18,21,24]. We suggest that the contribution of non-HLA SNPs maybe more important in the development of SLE in the younger patient while the contribution of HLA SNPs maybe more important in the development of the disease in later years.

Previous publications have used different genetic risk scores (counting, weighted or both) to predict the risk of multiple autoimmune diseases [33-35]. In SLE, there have been 7 publications that used polygenic risk scores to determine risk of SLE but only 4 examined them as a predictor of age of disease onset [17-22,24,36]. These 4 studies used either a counting score, and/or weighted score [17,18,21,24]. We therefore took the approach to examine multiple different genetic risk scores in order to determine which gave the most robust results. We found that genetic risk weighted scores were optimal. These scores were more robust than additive scores as they are not affected by sample size, strength of marker interactions and account for differences in effect size [15].

Previous studies in SLE have shown conflicting results regarding sex and genetic risk scores. An initial study using a weighted score showed that men had a higher genetic risk than women (largely the result of HLA SNPs) [19]. However, a second study, using a weighted score but smaller cohort and different HLA SNPs (only 13 SNPs were shared), did not replicate the finding [22]. Since the main difference between the 2 previous studies was the different HLA SNPs included, we covered all of the HLA SNPs used in both investigations (either with the same SNP or with a SNP in high LD). A small study of only 75 cSLE patients, using only 7 SNPs, none in the HLA region, did not find a significant difference in the genetic risk counting score between the sexes [21]. We found a significant association between sex and age of SLE diagnosis in the analysis of all the genetic risk scores (HLA and non-HLA) for all ages; therefore, we used sex as a covariant as it was a potential confounder. When we controlled for the variation in sex between ages, only the non-HLA GRSWS showed a statistically significant association with sex: male patients showed an increase of their non-HLA GRSWS compared to female patients. Although this difference was low (regression coefficient= 0.180), it was statistically significant and replicated what was found in the largest study[19]. More investigations are needed to confirm these conclusions.

It is well-recognized that SNPs associated with susceptibility differ across ancestral populations[6,7,37]. In both cSLE and aSLE populations SLE is more prevalent in non-Caucasian populations (Hispanic, Black, Native Americans and Asians)[36]. Thus, in the calculation of the weighted score, each SNP was weighted according to its effect in the population studied. Moreover, when we analyzed the associations between age of SLE diagnosis and ancestral populations we found statistically significant associations for all the genetic risk scores. Therefore, in the final linear regression models of all genetic risk scores with age of SLE

diagnosis, we included ancestral populations as covariates. This is the first time that effects of genetic ancestry on genetic risk scores have been addressed.

One limitation of our study maybe that all genotyping was performed on the Immunochip platform. This platform was designed for use in European populations and therefore is less informative in other ethnic populations with poorer coverage of SNPs associated with the development of SLE in non-Caucasian populations. In addition, there is increasing evidence that rare variants may be important in the development of SLE[38,39], but only a few of these rare variants are present on the Immunochip. Although, we were able to examine only 10/21 (47.6%) HLA SNPs validated in SLE GWAS, this is much greater coverage than in previous studies that used polygenic risk scores. Regarding the SNP coverage in each ancestral population, most of the non-HLA and HLA SNPs validated by GWAS and meta-GWAS for Caucasian and Asian populations were analyzed [7,12,37,40-47]. However, for Black and Hispanic populations there were no published GWAS and therefore candidate gene studies were used[7,12,28,29,37,40-47]. The resulting differences in the number of SNPs covered in each population were overcome by the process of genetic risk score standardization but it is still likely that some relevant SNPs in Black and Hispanic groups were missed. However, by standardizing the genetic score, the impact of any single locus will differ between ethnicities. Finally, there were no publications in South Asian populations and therefore we could not include this population. However, our coverage of both HLA and non-HLA SNPs is the largest published to date in each ancestral population. Another limitation to our study is the possibility of unmeasured confounders that can affect genetic scores and age of diagnosis of SLE. This could be the case in our models because our dependent variable is "age of diagnosis" and not "age of onset" of SLE, there may be differential bias (e.g., time to diagnose in cSLE vs aSLE patients, males vs females; access to care, socioeconomical factors among others).
However, we strongly believe that our results are still valid as start point for futures investigations.

The present study is the first to show that there are different effects of non-HLA and HLA genetic risk scores on age of SLE diagnosis in a multiethnic population. Specifically, non-HLA genetic risk scores showed that the higher the number of SLE associated non-HLA SNPs, the younger age of SLE diagnosis. Conversely, the higher the HLA genetic risk score the older the age of SLE diagnosis. These results were consistent across all methods of estimating their effect. We suggest that the non-HLA GRSWS may be the most robust score to use as it has showed the highest degree of statistical significance. However, for HLA risk, all the risk scores

performed well. Overall genetic risk scores explained 18% of the variance of age of onset of SLE. We suggest that future studies use standardized genetic risk scores which account for the variability in the distribution of the scores across populations, examine the effect of sex on risk scores and determine the effect HLA and non-HLA risk scores separately. These findings emphasize the complexity of the influence of genetic risk on the age of onset of SLE.

References

1. Ambrose N, Morgan TA, Galloway J, Ionnoau Y, Beresford MW, Isenberg DA. Differences in disease phenotype and severity in SLE across age groups. Lupus. 2016.

2. Cui Y, Sheng Y, Zhang X. Genetic susceptibility to SLE: recent progress from GWAS. J Autoimmun. 2013;41:25-33.

3. Hiraki LT, Benseler SM, Tyrrell PN, Harvey E, Hebert D, Silverman ED. Ethnic differences in pediatric systemic lupus erythematosus. J Rheumatol. 2009;36:2539-2546.

4. Levy DM, Peschken CA, Tucker LB, et al. Influence of ethnicity on childhood-onset systemic lupus erythematosus: Results from a multiethnic multicenter Canadian cohort. Arthritis Care Res (Hoboken). 2013;65:152-160.

5. Deng Y, Tsao BP. Updates in Lupus Genetics. Curr Rheumatol Rep. 2017;19:68.

6. Alarcon-Riquelme ME, Ziegler JT, Molineros J, et al. Genome-Wide Association Study in an Amerindian Ancestry Population Reveals Novel Systemic Lupus Erythematosus Risk Loci and the Role of European Admixture. Arthritis Rheumatol. 2016;68:932-943.

7. Gateva V, Sandling JK, Hom G, et al. A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. Nat Genet. 2009;41:1228-1233.

8. Harley JB, Alarcon-Riquelme ME, Criswell LA, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat Genet. 2008;40:204-210.

9. Morris DL, Sheng Y, Zhang Y, et al. Genome-wide association meta-analysis in Chinese and European individuals identifies ten new loci associated with systemic lupus erythematosus. Nat Genet. 2016;48:940-946.

10. Armstrong DL, Reiff A, Myones BL, et al. Identification of new SLE-associated genes with a twostep Bayesian study design. Genes Immun. 2009;10:446-456.

11. Cunninghame Graham DS, Morris DL, Bhangale TR, et al. Association of NCF2, IKZF1, IRF8, IFIH1, and TYK2 with Systemic Lupus Erythematosus. PLoS Genet. 2011;7:e1002341.

12. Morris DL, Taylor KE, Fernando MM, et al. Unraveling multiple MHC gene associations with systemic lupus erythematosus: model choice indicates a role for HLA alleles and non-HLA genes in Europeans. Am J Hum Genet. 2012;91:778-793.

13. Ruiz-Narvaez EA, Fraser PA, Palmer JR, et al. MHC region and risk of systemic lupus erythematosus in African American women. Hum Genet. 2011;130:807-815.

14. Yarwood A, Han B, Raychaudhuri S, et al. A weighted genetic risk score using all known

susceptibility variants to estimate rheumatoid arthritis risk. Ann Rheum Dis. 2015;74:170-176.

15. Che R, Motsinger-Reif AA. Evaluation of genetic risk score models in the presence of interaction and linkage disequilibrium. Front Genet. 2013;4:138.

16. Dudbridge F. Power and predictive accuracy of polygenic risk scores. PLoS Genet. 2013;9:e1003348.

17. Taylor KE, Chung SA, Graham RR, et al. Risk alleles for systemic lupus erythematosus in a large case-control collection and associations with clinical subphenotypes. PLoS Genet. 2011;7:e1001311.

18. Webb R, Kelly JA, Somers EC, et al. Early disease onset is predicted by a higher genetic risk for lupus and is associated with a more severe phenotype in lupus patients. Ann Rheum Dis. 2011;70:151-156.

19. Hughes T, Adler A, Merrill JT, et al. Analysis of autosomal genes reveals gene–sex interactions and higher total genetic risk in men with systemic lupus erythematosus. Ann Rheum Dis. 2012;71:694-699.

20. Sawalha AH, Wang L, Nadig A, et al. Sex-specific differences in the relationship between genetic susceptibility, T cell DNA demethylation and lupus flare severity. J Autoimmun. 2012;38:J216-222.

21. Kadota K, Mori M, Yanagimachi M, et al. Analysis of gender differences in genetic risk: association of TNFAIP3 polymorphism with male childhood-onset systemic lupus erythematosus in the Japanese population. PLoS One. 2013;8:e72551.

22. Alonso-Perez E, Suarez-Gestal M, Calaza M, et al. Lack of replication of higher genetic risk load in men than in women with systemic lupus erythematosus. Arthritis Res Ther. 2014;16:R128.

23. Langefeld CD, Ainsworth HC, Cunninghame Graham DS, et al. Transancestral mapping and genetic load in systemic lupus erythematosus. Nat Commun. 2017;8:16021.

24. Joo YB, Lim J, Tsao BP, Nath SK, Kim K, Bae SC. Genetic variants in systemic lupus erythematosus susceptibility loci, XKR6 and GLT1D1 are associated with childhood-onset SLE in a Korean cohort. Sci Rep. 2018;8:9962.

25. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum. 1997;40:1725.

26. Bentham J, Morris DL, Cunninghame Graham DS, et al. Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. Nat Genet. 2015;47:1457-1464.

27. Zhang J, Zhang L, Zhang Y, et al. Gene-based meta-analysis of genome-wide association study data identifies independent single-nucleotide polymorphisms in ANXA6 as being associated with systemic lupus erythematosus in Asian populations. Arthritis Rheumatol. 2015;67:2966-2977.

28. Sanchez E, Webb RD, Rasmussen A, et al. Genetically determined Amerindian ancestry correlates with increased frequency of risk alleles for systemic lupus erythematosus. Arthritis Rheum. 2010;62:3722-3729.

29. Sanchez E, Comeau ME, Freedman BI, et al. Identification of novel genetic susceptibility loci in African American lupus patients in a candidate gene association study. Arthritis Rheum. 2011;63:3493-3501.

30. Manku H, Langefeld CD, Guerra SG, et al. Trans-ancestral studies fine map the SLE-susceptibility locus TNFSF4. PLoS Genet. 2013;9:e1003554.

31. Han S, Kim-Howard X, Deshmukh H, et al. Evaluation of imputation-based association in and around the integrin-alpha-M (ITGAM) gene and replication of robust association between a nonsynonymous functional variant within ITGAM and systemic lupus erythematosus (SLE). Hum Mol Genet. 2009;18:1171-1180.

32. Jacob CO, Reiff A, Armstrong DL, et al. Identification of novel susceptibility genes in childhoodonset systemic lupus erythematosus using a uniquely designed candidate gene pathway platform. Arthritis Rheum. 2007;56:4164-4173.

33. Meigs JB, Shrader P, Sullivan LM, et al. Genotype score in addition to common risk factors for prediction of type 2 diabetes. N Engl J Med. 2008;359:2208-2219.

34. Prahalad S, Conneely KN, Jiang Y, et al. Susceptibility to childhood-onset rheumatoid arthritis:investigation of a weighted genetic risk score that integrates cumulative effects of variants at five genetic loci. Arthritis Rheum. 2013;65:1663-1667.

35. Hilven K, Patsopoulos NA, Dubois B, Goris A. Burden of risk variants correlates with phenotype of multiple sclerosis. Mult Scler. 2015;21:1670-1680.

36. Tucker LB, Uribe AG, Fernandez M, et al. Adolescent onset of lupus results in more aggressive disease and worse outcomes: results of a nested matched case-control study within LUMINA, a multiethnic US cohort (LUMINA LVII). Lupus. 2008;17:314-322.

37. Han JW, Zheng HF, Cui Y, et al. Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. Nat Genet. 2009;41:1234-1237.

38. Roberts AL, Thomas ER, Bhosle S, et al. Resequencing the susceptibility gene, ITGAM, identifies two functionally deleterious rare variants in systemic lupus erythematosus cases. Arthritis Res Ther. 2014;16:R114.

39. Zhang YM, Cheng FJ, Zhou XJ, Qi YY, Zhao MH, Zhang H. Rare Variants of ATG5 Are Likely to Be Associated With Chinese Patients With Systemic Lupus Erythematosus. Medicine (Baltimore). 2015;94:e939.

40. Graham RR, Cotsapas C, Davies L, et al. Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. Nat Genet. 2008;40:1059-1061.

41. Kamatani Y, Matsuda K, Ohishi T, et al. Identification of a significant association of a single nucleotide polymorphism in TNXB with systemic lupus erythematosus in a Japanese population. J Hum Genet. 2007;53:64-73.

42. Kozyrev SV, Abelson AK, Wojcik J, et al. Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. Nat Genet. 2008;40:211-216.

43. Lea WW, Lee YH. The association between the PTPN22 C1858T polymorphism and systemic lupus erythematosus: a meta-analysis update. Lupus. 2011;20:51-57.

44. Lee YH, Bae SC, Choi SJ, Ji JD, Song GG. Genome-wide pathway analysis of genome-wide association studies on systemic lupus erythematosus and rheumatoid arthritis. Mol Biol Rep. 2012.

45. Okada Y, Shimane K, Kochi Y, et al. A Genome-Wide Association Study Identified AFF1 as a Susceptibility Locus for Systemic Lupus Eyrthematosus in Japanese. PLoS Genet. 2012;8:e1002455.

46. Yang W, Shen N, Ye DQ, et al. Genome-wide association study in Asian populations identifies variants in ETS1 and WDFY4 associated with systemic lupus erythematosus. PLoS Genet. 2010;6:e1000841.

47. Yang W, Tang H, Zhang Y, et al. Meta-analysis followed by replication identifies loci in or near CDKN1B, TET3, CD80, DRAM1, and ARID5B as associated with systemic lupus erythematosus in Asians. Am J Hum Genet. 2013;92:41-51.

Table 1. Demographic Features of Patients

	Female	Male	Caucasian	Black	Hispanic	Asian	Total Population
Total	1,351	189	**1094**	**196**	**129**	**121**	**1,540**
Mean of Age of Diagnosis	25.81 (SD=14.3)	21.48 (SD=13.1)	28.65 (SD=14.360)	17.58 (SD=9.84)	18.48 (SD=11.025)	14.50 (SD=7.49)	25.28 (SD=14.2)
Median of Age of Diagnosis	21.81	16	26.46	14.58	14.90	13.23	20.95
Minimum of Age of Diagnosis	1	5	2	4	1	4.3	1
Maximum of Age of Diagnosis	82.14	64.55	82.14	54.98	56.76	54.40	82.14
Number of cSLE <18 years old)	563 (41.67%)	112 (59.25%)	329	144	90	112	675 (43.83%)
Number of aSLE (≥18 years old)	788 (58.32%)	77 (40.74%)	765	52	39	9	865 (56.16%)

Table 2. Linear regression analyses of the non-HLA genetic risk scores

	Genetic risk counting scores		Genetic risk weighted scores	
	GRCS	**GRSCS**	**GRWS**	**GRSWS**
Final model p-value*	0.103	**0.007**	**0.011**	**0.001**
Slope&	-0.066	-0.044	-0.023	-0.048
Adjusted r^2	0.175	0.178	0.175	0.176

GRCS= Genetic risk counting score, GRSCS= Genetic risk standardized counting score, GRWS= Genetic risk weighted score, GRSWS= Genetic risk standardized weighted score. *Final statistical model of the genetic risk weighted scores included sex and Caucasian vs. non-Caucasian (C/NC) ancestry as covariates, final model of the genetic risk counting scores, included sex and all 4 ancestral populations as covariates. &: Slope represents the rate of change in y (age of diagnosis) as x (non-HLA genetic risk scores) changes.

Table 3. Linear regression analyses of the HLA genetic risk scores

	Genetic risk counting scores		Genetic risk weighted scores	
	GRCS	**GRSCS**	**GRWS**	**GRSWS**
Final model p-value*	**0.049**	**0.022**	**0.022**	**0.011**
Slope&	0.014	0.013	0.026	0.014
Adjusted r^2	0.176	0.176	0.176	0.177

GRCS= Genetic risk counting score, GRSCS= Genetic risk standardized counting score, GRWS= Genetic risk weighted score, GRSWS= Genetic risk standardized weighted score. *Final model of all HLA genetic risk scores included sex and all 4 ancestral populations as covariates. &Slope represents the rate of change in y (age of diagnosis) as x (HLA genetic risk scores) changes.

Figure 1. Genetic Standardized Score Calculation

Step 1

GRCS = $\sum$ risk alleles per patient

Step 2

GRSCS = GRCS x 10/ allele number*

Step 3

Risk OR = 1/protective OR

Step 4

MaxGRWS = ln (O.R SNP$_1$) x 2 + ln (O.R SNP$_2$) x 2 + nSNPs......

Step 5

GRSWS = GRWS/MaxGRWS* x 10