ACR Open Rheumatology
Vol. 7, No. 5, May 2025, e70053
DOI 10.1002/acr2.70053

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Uncovering Genetic Variation in Systemic Lupus Erythematosus Risk Variants in Indigenous Peruvians

Cristina M. Lanata,[1] Richard F. Oppong,[1] Mary K. Horton,[1] Victor Borda,[2] Manuel F. Ugarte-Gil,[3] Joanne Nititham,[1] Eduardo Tarazona-Santos,[4] Timothy D. O'Connor,[2] Heinner Guio,[5] and Lindsey A. Criswell[1]

Objective. Systemic lupus erythematosus (SLE) results in worse clinical outcomes among individuals of Amerindian descent. The genetic basis for this is uncertain, and there is a significant lack of genetic research focused on Amerindian ancestry populations. This study aims to compare the frequencies of SLE risk variants and polygenic risk scores between Indigenous Peruvians and global populations with diverse ancestral backgrounds.

Methods. We studied 670 individuals from the Peruvian Genome Project, 2,068 individuals from the 1000 Genomes Project Phase 3 release, and 47 patients with SLE from Lima, Peru. Ancestry was inferred using admixture and RFMix. Data were imputed with the TOPMed Imputation server and annotated to hg38. We compared the frequencies of 199 SLE-associated risk variants among study participants. We also calculated SLE genetic risk scores and fixation index (FST) statistics.

Results. All 199 SLE risk single-nucleotide polymorphisms had highly significant differences in frequencies across Peruvian and other continental populations (*P* values <0.001). Indigenous Peruvian patients have higher polygenic risk for SLE compared to European, African, South Asian, and East Asian patients. FST analysis of SLE risk variants revealed the largest FST between Peruvian patients and African patients (mean FST 0.12), and the smallest between Peruvian patients and East Asian patients (mean FST 0.09).

Conclusion. SLE-associated variants are common among Indigenous Peruvian patients, with varying frequencies across subpopulations. This underscores the need for ongoing genetic studies in Indigenous populations, potentially explaining SLE heterogeneity.

INTRODUCTION

Autoimmune diseases, particularly systemic lupus erythematosus (SLE), present at a younger age and with worse clinical outcomes among people who are not of European descent.[1] Multiancestry population studies of SLE outcomes have demonstrated significant differences in lupus nephritis risk by ancestry.[2,3] Studies have shown that patients from self-reported Amerindian background experience more frequent renal involvement and flares and worse SLE disease outcomes than other self-reported White Latin American participants with SLE.[4,5] Another study reported an increased frequency of SLE risk alleles with more Amerindian ancestry in a South American SLE population.[6] However, there is a severe underrepresentation of genetics and genomics studies within the Indigenous populations of North and South America.[7]

Peru has a diverse and challenging topography, resulting in Indigenous populations being relatively isolated until the past 200 years.[8] Nonindigenous admixture is lower than other South American countries, with Peruvian people having, on average, 70% Indigenous American continental ancestry.[9] This makes

Supported by the Intramural Research Program of the National Human Genome Research Institute, NIH. Minas Gerais Foundation for Research. Grant Numbers: FAPEMIG, APQ-04228-24, RED-00089-23. Brazilian National Council of Scientific and Technological Development. Grant Numbers: CNPq, 406913/2022-6.

[1]Cristina M. Lanata, MD, Richard F. Oppong, PhD, Mary K. Horton, MPH, PhD, Joanne Nititham, MPH, Lindsey A. Criswell, MD, MPH, DSc: Genomics of Autoimmune Rheumatic Disease Section, National Human Genome Research Institute, NIH, Bethesda, Maryland; [2]Victor Borda, PhD, Timothy D. O'Connor, PhD: School of Medicine, University of Maryland, Baltimore; [3]Manuel F. Ugarte-Gil, MD, MSc: Grupo Peruano de Estudio de Enfermedades Autoinmunes Sistémicas, Universidad Científica del Sur and Rheumatology Department, Hospital Guillermo Almenara Irigoyen, EsSalud, Lima, Peru; [4]Eduardo Tarazona-Santos, PhD: Department of Genetics, Ecology and Evolution, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil; [5]Heinner Guio, MD, PhD: Instituto Nacional de Salud and INBIOMEDIC, Lima, Peru, and Universidad de Huánuco, Huánuco, Peru.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr2.70053).

Address correspondence via email to Cristina M. Lanata, MD, at cristina.lanata@nih.gov.

Submitted for publication February 21, 2025; accepted in revised form April 8, 2025.

Peruvian populations a highly informative group for studying genetic variation of continental South Amerindian populations.

The Peruvian Genome Project (PGP) is a Peruvian-funded initiative that started in 2011 and collected 1,149 samples from 17 traditional indigenous and 13 Mestizo (mixed of Indigenous, African, and European ancestry) communities.[8] The objective of our study was to compare the frequencies of established SLE risk variants in the PGP to other global populations of different ancestral backgrounds.

PATIENTS AND METHODS

Study participants. We analyzed data from 670 individuals from the PGP with available genotyping information.[8,10] Native populations from the coast, Andes, and Amazon and coastal regions of South America possess distinct ancestral divisions, largely imposed by the Andes Mountains.[10] The PGP included Indigenous participants from the following three regions of Peru: the coast (Tallanes and Moche), the Andes (Jaqarus, Chopccas, Quechuass, Qeros, Puno, and Uros), and the Amazon (Ashaninkas, Matsiguenkas, Shipibo, Nahua, and Matses). We also included 2,077 individuals from the 1000 Genomes Project (1KGP) Phase 3 release[11] (representing European,[8] East Asian, South Asian, and African populations). Participants with SLE from Lima were previously consented for genetic research.

Genotype data. The PGP participants were genotyped using the Illumina HumanOmni Array, whereas the Lima participants with SLE were genotyped using the Affymetrix Lat1 World Array. Genotype data from 1KGP participants were generated using a combination of low-coverage whole-genome sequencing, targeted exome sequencing, and high-density single-nucleotide polymorphism (SNP) microarrays. Genotype data from PGP participants and participants with SLE were lifted over to the hg38 reference genome, imputed on the TOPMed Imputation server, and annotated to hg38 reference genome. We applied stringent quality control measures, including the removal of variants with imputation quality score (r^2) below 0.3, the exclusion of second-degree relatives, the removal of multiallelic variants, exclusion of variants with call rates below 90%, and the exclusion of variants failing Hardy-Weinberg equilibrium thresholds ($P < 1 \times 10^{-5}$).

Ancestry calculations. We used a two-step approach to estimate genome-wide ancestry proportions. First, we used admixture[12] using K values ranging from four to eight in a subset of individuals comprising data from the full PGP, a subgroup of high-coverage populations from the 1KGP with unrelated individuals (TSI, IBS, CEU, PEL, CLM, MXL, PUR, LWK, MSL, YRI, GWD, JPT, CHB, and CHS), and Native Americans from the Human Genome Diversity Project (HGDP), including Colombian, Pima, and Maya populations. Genotypes were filtered to unlinked single-nucleotide variants with a minor allele frequency >5%. We performed 10 iterations for each K value and selected the run with the highest log-likelihood. We then performed local ancestry by running RFMix[13] version 2 with four continental references based on admixture results. We selected European (n = 404 individuals; populations: CEU, GBR, IBS, and TSI), African (n = 405 individuals; populations: YRI, ESN, GWD, and LWK), and East Asian (n = 411 individuals; populations: CHB, CHS, JPT, and KHV) individuals from 1KGP populations. The Native American reference included 187 individuals from PGP with more than 99% Native American ancestry based on admixture K = 4 results. We ran RFMix using two expectation maximization steps. Results were used to categorize nonadmixed Peruvian individuals as individuals with greater than 95% Indigenous ancestry.

Analysis. We analyzed the frequencies of 199 SLE-associated variants across all participants (Supplementary Dataset S1). These variants were selected based on recent literature reviews and studies included studies involving individuals with Native American ancestry. The 199 variants are those that passed quality control thresholds in our data set. We calculated the fixation index (FST), defined by Weir and Cockerham (1984), which measures genetic differentiation between populations, for each SLE-associated SNP for each pairwise reference population. FST values range from 0 to 1, with 0 reflecting minimal and 1 reflecting complete differentiation between populations. We used logistic regression to estimate the odds of being a homozygous carrier of the risk alleles among nonadmixed Peruvians individuals compared to individuals of European ancestry (1KGP European). Additionally, we calculated an unweighted SLE genetic risk score (GRS) and a weighted GRS based on the 199 SNPs, weighing each SNP by its reported effects size in SLE genome-wide association study (GWAS) from the populations they were derived (Supplementary Table S1). These GRS values were standardized to a mean of 0 and SD of 1. To assess population differences in the GRSs distribution, we compared the mean GRS values of nonadmixed Peruvians and 1KGP populations using a *t*-test, with Peruvian SLE samples serving as the reference group. For comparison, we performed random sampling without replacement of 199 variants across the genome and calculated corresponding unweighted GRS. This random sampling procedure was repeated 10,000 times to generate an empirical null distribution of genetic scores.

RESULTS

Principal component analysis plots of study participants. Ancestry analysis revealed low levels of non-Indigenous admixture with other continental ancestry groups among participants in the PGP. We classified 282 participants as nonadmixed Peruvian participants, those having greater than 95% Indigenous ancestry. Principal component analysis plots showed that nonadmixed Peruvian participants formed a cluster

clearly separated from the other continental ancestral populations from the 1KGP (Figure 1). The Peruvian cluster is further illustrated to show Native populations from the coast, Andes, and Amazon regions of Peru, showing much less differentiation in comparison to other continental ancestries. This genetic pattern closely resembled that observed in Native Brazilian participants and was distinct from the genetic profiles of Native Mexican participants analyzed in the HGDP (Supplementary Figure S1).

SLE risk variants having different frequencies across ancestral populations. We analyzed allele frequency distributions for 199 SLE-associated variants, all of which showed significant differences across populations ($P < 0.001$, Supplementary Table S2). Among Peruvian participants, we observed no significant differences in allele frequency distributions across Native populations from the coast, Andes, and Amazon regions. To investigate whether specific SLE risk variants were more prevalent in Peruvian participants compared to European participants, the most extensively studied group for SLE genetics, we performed a logistic regression to test the odds of being a homozygous risk variant carrier in nonadmixed Peruvian participants versus European participants (Table 1). The results revealed striking differences in the proportions of homozygous risk variant carriers. For instance, the rs3821236 variant in the *STAT4* gene exhibited a risk homozygosity rate of 75.9% among nonadmixed Peruvian participants, compared to 3.2% in European participants, 8.3% in African participants, 17.3% in East Asian participants, and 4.7% in South Asian participants (odds ratio 93.9, confidence interval 55.3–168.8, Peruvian participants vs European participants). This SNP also had the highest FST value (mean 0.57) between European participants and nonadmixed Peruvian participants, reflecting that **57% of the genetic variation** at this variant is due to differences between European participants and nonadmixed Peruvian participants. Additionally, two variants in the *UBE2LE*, and *C7orf72* genes showed similarly high FST values between nonadmixed Peruvian participants versus European participants, alongside striking differences in homozygosity rate (Table 1). For other risk variants, such as rs7902146 in *ARID5B*, rs7815944 in *LINC00824*, and rs1780813 in *SMYD3*, the risk alleles were present at lower proportions in nonadmixed Peruvian participants compared to European participants, illustrating the heterogeneity in the genetic risk of SLE variants across populations.

Further comparisons of FST values for all studied variants across populations are shown in Supplementary Table S3. Among nonadmixed Peruvian participants, the mean FST values for SLE risk variants were consistently low (mean FST <0.04 for these comparisons) between populations from the Andes, Amazon, and coast. However, the mean FST between Peruvian participants and European participants was 0.10, indicating moderate differentiation overall. Comparisons with other global populations showed the highest mean FST of 0.12 between Peruvian participants and African participants, whereas the lowest FST of 0.09 was observed between Peruvian participants and East Asian participants, consistent with the global pattern of human genetic diversity (may include 1KGP reference).

SLE polygenic GRS higher in nonadmixed Peruvian participants compared to other populations. We calculated both weighted and unweighted SLE GRS for each participant (Figure 2). Mean GRS values for each continental ancestry group (including nonadmixed Peruvian) were compared to Peruvian patients with SLE from Lima as the reference population. Analysis of variance revealed highly significant differences, with P values $<2 \times 10^{-16}$ for both weighted and unweighted GRS (Figure 1). Nonadmixed Peruvian participants exhibited the highest mean GRS (weighted = 1.09; unweighted = 1.14), followed by admixed Peruvian participants, African participants, South Asian participants, East Asian participants, and European participants. As expected,

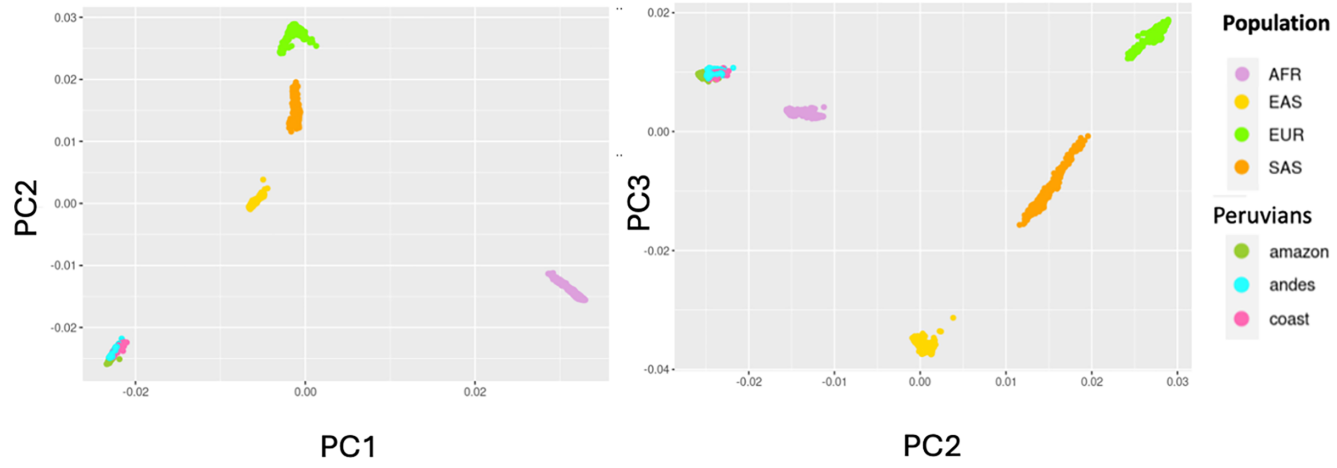

Figure 1. Principal component analysis plot of nonadmixed Peruvian participants and other continental populations from the 1000 Genomes Project. PC, principal component.

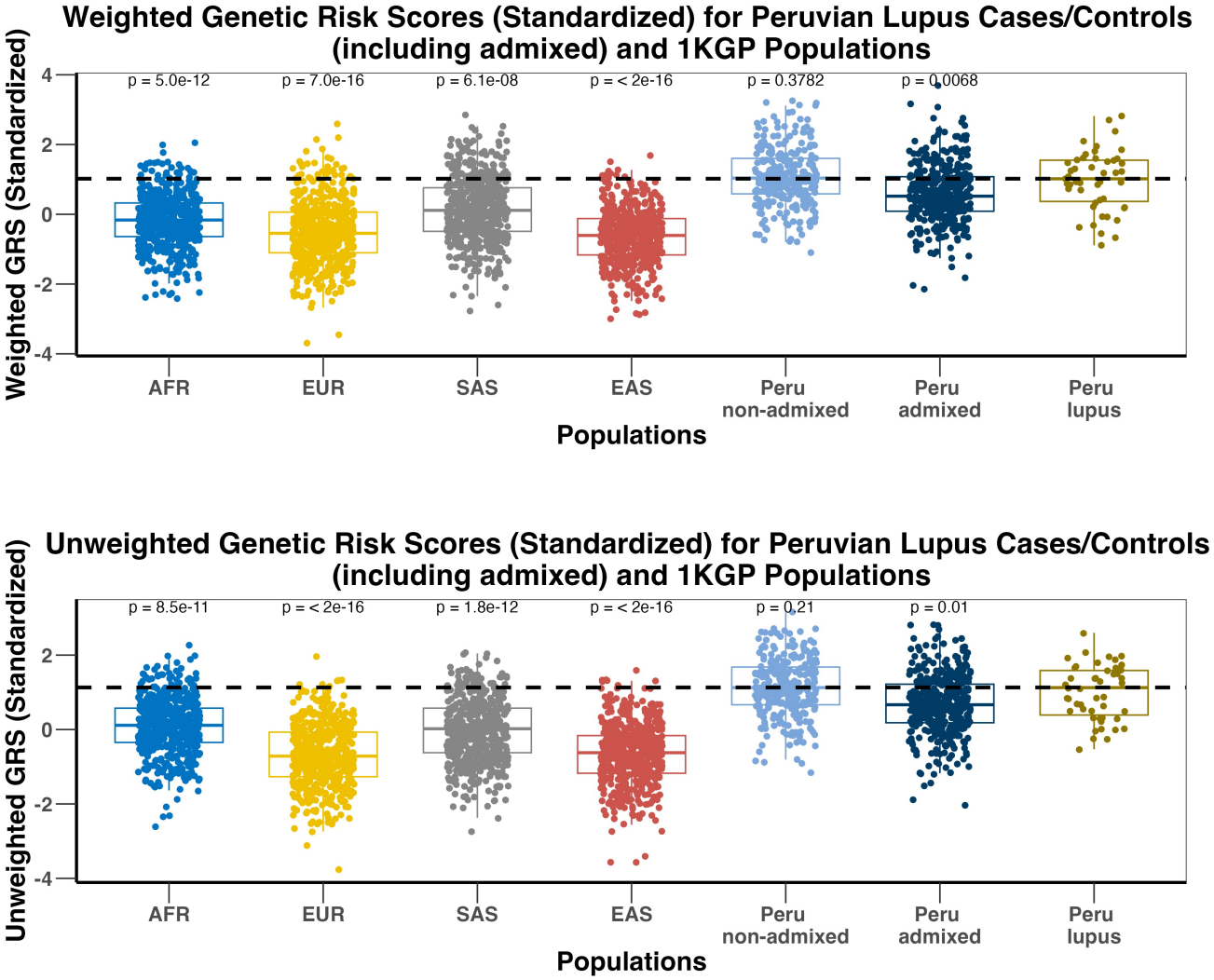

Figure 2. Standardized Systemic lupus erythematosus (SLE) GRS across populations. Box plots display the standardized (top) weighted and (bottom) unweighted SLE GRS for different population groups, including African, European, South Asian, and East Asian from 1KGP, nonadmixed Peruvian participants and admixed Peruvian participants from the Peruvian Genome Project, and Peruvian patients with SLE. *P* values are derived from *t*-tests comparing the mean GRS between each population groups and the Peruvian patients with SLE. The dashed black line represents the median GRS for the Peruvian patients with SLE. The plot shows that nonadmixed Peruvian participants exhibited similarly elevated GRS values as participants with lupus compared to other groups. 1KGP, 1000 Genomes Project; GRS, genetic risk scores.

Peruvian participants with SLE exhibited the highest polygenic risk scores (PRS) compared to healthy individuals from other populations—except for Indigenous Peruvian participants. Notably, this underscores the elevated SLE PRS among Indigenous Peruvian participants relative to other groups. Randomly sampling (n = 199) variants across the genome and creating GRS (10,000 times) did not result in similarly high scores among Indigenous Peruvian participants in comparison to other populations (Supplementary Figure S2).

DISCUSSION

In this study, we leveraged the unique data from the PGP to investigate the frequency of SLE-associated risk variants across commonly used European, Asian, and African reference populations. Our findings reveal a substantial genetic risk signal among Indigenous Peruvian populations. This elevated genetic risk highlights the potential for meaningful insights into disease susceptibility and progression in this population. Indigenous Peruvian is a population with a unique genetic architecture, shaped by centuries of adaptation to diverse environments, including the high-altitude conditions of the Andes and the infectious disease pressures of the Amazon region and by prolonged geographic isolation.[8,10] Our FST analyses highlight the differences in frequencies of SLE risk variants across populations, demonstrating that SLE-related genetic variants have undergone distinct selective pressures across populations.

Table 1. SLE risk variants exhibiting the highest differences in homozygous carrier proportions between non-admixed Peruvians and Europeans. Proportion of homozygous risk variant carriers among African (AFR), European (EUR), South Asian (SAS), East Asian (EAS), non-admixed Peruvians (PER), admixed Peruvians (PER ad), and Peruvian lupus cases (Lupus).

			Proportion of homozygous risk variant carriers (%)							Logistic regression of homozygous risk variant carrier in PER vs EUR			
SNP	CHR	Gene	AFR *n* = 532	EUR *n* = 524	EAS *n* = 509	SAS *n* = 512	PER *n* = 282	PER ad *n* = 352	Lupus *n* = 47	OR	CI	FDR p	Fst*
rs3821236	2	STAT4	8.3	3.2	17.3	4.7	75.9	59.4	53.2	93.9	55.3–168.8	6.51E−56	0.57
rs5998672	22	UBE2L3	26.9	3.8	26.1	17.2	62.1	40.3	44.7	41.2	25.4–70.3	2.14E−45	0.50
rs3747093	22	UBE2L3	26.9	3.4	25.9	17.6	62.1	40.6	44.7	46	27.8–80.4	1.71E−44	0.49
rs4598207	7	C7orf72	10.2	10.5	7.3	6.2	70.2	49.4	27.7	20.1	13.9–29.6	8.53E−53	0.38
rs4690055	4	TNIP2	13.5	18.7	16.3	20.1	91.1	72.7	57.4	44.7	28.6–72.7	9.14E−56	0.38
rs2009453	11	PCNXL3	43.6	18.9	28.7	24.4	87.9	74.4	63.8	31.3	20.8–48.3	6.51E−56	0.37
rs2238577	19	ARID3A	41.9	75.8	28.7	55.1	15.6	29	36.2	0.1	0–0.1	5.16E−47	0.36
rs8016947	14	NFKBIA	42.9	26.3	35	34.2	97.9	83.8	93.6	128.7	61–331.7	3.10E−29	0.35
rs405858	19	ANKRD27	34	12.2	47.2	21.5	66	60.2	51.1	13.9	9.8–20.1	2.14E−45	0.33
rs10892301	11	CXCR5	40.4	30.2	23.2	16.2	96.1	78.4	80.9	57.1	31.8–113.6	6.08E−35	0.32
rs17603856	6	ATXN1	0.2	10.5	2.4	4.7	67	35.8	23.4	17.3	12–25.4	5.25E−49	0.31
rs7902146	10	ARID5B	34.4	51.1	51.7	46.9	7.4	16.8	21.3	0.1	0–0.1	5.17E−25	0.28
rs11059928	12	SLC15A4	100	79.4	65.4	64.5	31.2	45.5	27.7	0.1	0.1–0.2	6.68E−36	0.23
rs10239340	7	TNPO3	19	14.1	31.2	15	58.5	50.9	48.9	8.6	6.1–12.1	9.30E−34	0.23
rs7170151	15	RASGRP1	48.3	6.5	31	11.9	35.8	31.8	31.9	8	5.3–12.4	4.91E−21	0.22
rs10912578	1	LOC100506023	20.7	10.5	17.7	8.2	45	34.1	46.8	7	4.9–10.1	1.57E−24	0.22
rs4760589	12	SLC15A4	49.2	16.8	19.4	22.3	61.3	44.9	46.8	7.9	5.7–11	5.13E−33	0.20
rs12370194	12	SLC15A4	44.7	17.9	19.3	21.9	61.3	44.6	46.8	7.3	5.3–10.1	2.51E−31	0.19
rs704840	1	TNFSF4	6.4	7.8	13.8	5.7	37.2	26.7	34	7	4.7–10.5	1.24E−20	0.18
rs7815944	8	LINC00824	51.1	92	50.7	66.4	61.7	60.5	63.8	0.1	0.1–0.2	2.43E−21	0.18
rs1308020	11	RNASEH2C	67.5	43.7	57.2	60.4	90.4	82.4	72.3	12.2	8–19.1	1.25E−28	0.17
rs1780813	1	SMYD3	85.5	96.6	85.9	90.8	64.9	66.2	74.5	0.1	0–0.1	7.49E−23	0.17
rs10931481	2	STAT4	13.2	9.4	23.8	13.9	40.1	33	40.4	6.5	4.5–9.5	3.10E−21	0.17
rs564976	3	IL12A-AS1	75.9	40.1	74.3	66.8	84	80.4	76.6	7.9	5.5–11.4	1.04E−27	0.15
rs17266594	4	BANK1	56.8	49.2	66.2	66.4	87.9	82.4	91.5	7.5	5.1–11.3	2.22E−22	0.12

*Fst statistic calculated between non-admixed Peruvians and EUR 1000 genomes.

GWAS results have notable limitations in their interpretation.[14] For instance, the causal variants underlying many signals remain unconfirmed, and the tag SNPs identified in one population may not capture the causal variants in others. Moreover, genotyping arrays are biased toward capturing genetic variation predominantly in European populations. Additionally, GRS often perform inconsistently across diverse populations.[15] Despite these challenges, our findings indicate a higher genetic risk for SLE in Indigenous Peruvian individuals. In the post-GWAS era, sequencing studies focused on SLE in Indigenous populations hold promise for uncovering the genetic drivers of autoimmunity in Indigenous populations, offering a more comprehensive understanding of disease mechanisms.

The current paucity of genetic studies in Indigenous populations, including Peruvian individuals, presents both an opportunity and a challenge for advancing personalized medicine. These populations harbor rich untapped genetic diversity, with the potential to uncover novel insights into disease mechanisms and therapeutic targets. However, this paucity also reflects systemic inequalities in research representation, as Indigenous populations have historically been excluded from scientific studies. Efforts to expand genetic studies in these populations must be approached within a framework that prioritizes community engagement, cultural sensitivity, and benefit-sharing. Without such considerations, these efforts risk perpetuating existing disparities. The insights gained from studying the genetic risk in Indigenous Peruvian individuals must ultimately translate into tangible benefits for these communities, such as improved health outcomes and access to tailored medical interventions. Achieving this goal requires a paradigm shift in the conduct of genetic research, moving beyond purely academic pursuits toward initiatives that directly address the health priorities and needs of these populations.

The main strength of this study is the use of Indigenous Peruvian genetic data, which represents an underexplored population. By using GWAS-significant SNPs identified in previous studies, our study focused on well-established SLE-associated variants, enabling a targeted approach to assess risk loci enriched in Indigenous Peruvian individuals. Although our approach effectively highlights population-level differences in disease risk, it is limited by its reliance on previously identified variants, which may overlook novel or population-specific risk loci unique to Indigenous Peruvian individuals. Additionally, the relatively small sample size of the Indigenous Peruvian participants may limit the statistical power to detect associations for less common variants. Future studies incorporating larger Indigenous sample sizes and whole-genome sequencing will allow for discovery of novel loci and provide greater resolution in understanding genetic risks in SLE.

In conclusion, our findings demonstrate the promise of genetic studies in uncovering disease risk factors among Indigenous populations. However, the paucity of such studies emphasizes a significant gap in the field of genomics, particularly within the context of personalized medicine. Addressing this gap will require sustained commitment to collaboration and inclusivity in genetic research.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Lanata confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Parodis I, Lanata C, Nikolopoulos D, et al. Reframing health disparities in SLE: a critical reassessment of racial and ethnic differences in lupus disease outcomes. Best Pract Res Clin Rheumatol 2023;37(4):101894.
2. Lanata CM, Nititham J, Taylor KE, et al. Genetic contributions to lupus nephritis in a multi-ethnic cohort of systemic lupus erythematous patients. PLoS One 2018;13(6):e0199003.
3. Solomon O, Lanata CM, Adams C, et al. Local ancestry at the major histocompatibility complex region is not a major contributor to disease heterogeneity in a multiethnic lupus cohort. Arthritis Rheumatol 2024;76(4):614–619.
4. Reátegui-Sokolova C, Ugarte-Gil MF, Harvey GB, et al; GLADEL. Predictors of renal damage in systemic lupus erythematous patients: data from a multiethnic, multinational Latin American lupus cohort (GLADEL). RMD Open 2020;6(3):e001299.
5. Pons-Estel GJ, Catoggio LJ, Cardiel MH, et al; GLADEL. Lupus in Latin-American patients: lessons from the GLADEL cohort. Lupus 2015;24(6):536–545.
6. 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(12):3722–3729.
7. West KM, Blacksher E, Burke W. Genomics, health disparities, and missed opportunities for the nation's research agenda. JAMA 2017;317(18):1831–1832.
8. Harris DN, Song W, Shetty AC, et al. Evolutionary genomic dynamics of Peruvians before, during, and after the Inca Empire. Proc Natl Acad Sci USA 2018;115(28):E6526–E6535.
9. Sandoval JR, Salazar-Granara A, Acosta O, et al. Tracing the genomic ancestry of Peruvians reveals a major legacy of pre-Columbian ancestors. J Hum Genet 2013;58(9):627–634.
10. Borda V, Alvim I, Mendes M, et al. The genetic structure and adaptation of Andean highlanders and Amazonians are influenced by the interplay between geography and culture. Proc Natl Acad Sci USA 2020;117(51):32557–32565.
11. Auton A, Brooks LD, Durbin RM, et al; Genomes Project Consortium. A global reference for human genetic variation. Nature 2015;526(7571):68–74.

12. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res 2009;19(9):1655–1664.
13. Maples BK, Gravel S, Kenny EE, et al. RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. Am J Hum Genet 2013;93(2):278–288.
14. Tam V, Patel N, Turcotte M, et al. Benefits and limitations of genome-wide association studies. Nat Rev Genet 2019;20(8):467–484.
15. Martin AR, Kanai M, Kamatani Y, et al. Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet 2019;51(4):584–591.