

RESEARCH ARTICLE

<https://doi.org/10.1158/2767-9764.CRC-25-0014>

OPEN ACCESS



Genetic Ancestry, Intrinsic Tumor Subtypes, and Breast Cancer Survival in Latin American Women

Daniela Alves da Quinta^{1,2}, Darío Rocha³, Cristian Yáñez⁴, Renata Binato⁵, Sheila Coelho Soares-Lima⁵, Xiaosong Huang⁶, Daiana Ganiewich¹, Valentina A. Zavala⁷, Monica Sans⁸, Alejandra Lopez-Vazquez⁹, Jael Quintero⁹, Olivia Valenzuela⁹, Antonio Quintero-Ramos¹⁰, Alicia Del Toro-Arreola¹⁰, Mauricio Cerda¹¹, Katherine Marcelain¹², Susanne Crocamo⁵, Maria Aparecida Nagai¹³, Dirce M. Carraro¹⁴, Marcia Maria Chiquitelli Marques¹⁵, Jorge Gómez¹⁶, Nora Artagaveytia⁸, Adrian Daneri-Navarro¹⁰, Bettina G. Müller¹⁷, Javier Retamales¹⁸, Carlos Velazquez⁹, Elmer A. Fernández^{19,20,21}, Osvaldo L. Podhajcer¹, Eliana Abdelhay⁵, Ricardo A. Verdugo^{4,12}; the LACRN Investigators²², Andrea S. Llera¹, and Laura Fejerman^{5,23,24}

ABSTRACT

This study investigates the relationship between genetic ancestry, breast cancer subtypes, and survival outcomes among 951 locally advanced breast cancer cases from Argentina, Brazil, Chile, Mexico, and Uruguay, participating in the Molecular Profile of Breast Cancer Study. Array-based genotyping and ADMIXTURE analysis were used for genetic ancestry evaluation. Breast cancer subtypes were defined by IHC and the gene expression-based PAM50 algorithm. The distribution of genetic ancestry, including European, Indigenous American (IA), African (AFR), and East Asian components, revealed a heterogeneous genetic admixture across countries, with the highest IA ancestry observed in Chile (30.9%) and Mexico (30.8%). Testing the relationship between genetic ancestry and breast cancer subtypes demonstrated that a 10% increase in European ancestry was significantly associated with a 14% decrease in the odds of developing HER2-enriched breast cancer, after adjustment by age, nodal status, and the AFR component (adj. $P = 0.021$, luminal A as reference).

Accordingly, a 10% increase in IA ancestry was associated with a 21% increase in the probability of having HER2-enriched breast cancer (adj. $P = 0.022$). IA ancestry also significantly increased overall survival after adjustment by age, nodal status, and AFR ancestry, although this result is controversial and may be affected by the size and heterogeneity of the Molecular Profile Breast Cancer Study cohort. Our research confirms previous findings of a high prevalence of HER2-dependent breast tumors among Hispanic/Latina women and strengthens the hypotheses of the existence of either population-specific genetic variant(s) or of other ancestry-correlated factors that impact HER2 expression in breast cancer consistently across different Latin American regions.

Significance: The evidence in this work supports the idea that factors linked to genetic ancestry influence the prevalence of breast cancer subtypes in Latin America, potentially affecting treatment needs in the region.

¹Laboratorio de Terapia Molecular y Celular, Fundación Instituto Leloir-CONICET, Ciudad de Buenos Aires, Argentina. ²Universidad Argentina de la Empresa (UADE), Instituto de Tecnología (INTEC), Buenos Aires, Argentina. ³Universidad Nacional de Córdoba, Facultad de Ciencias Exactas, Físicas y Naturales, Córdoba, Argentina. ⁴Programa de Genética Humana, ICBM, Facultad de Medicina, Universidad de Chile, Santiago, Chile. ⁵Instituto Nacional de Câncer, Rio de Janeiro, Brazil. ⁶Genome Center, College of Biological Sciences, University of California Davis, Davis, California. ⁷Department of Public Health Sciences, University of California, Davis, Davis, California. ⁸Hospital de Clínicas Manuel Quintela, Universidad de la República, Montevideo, Uruguay. ⁹Universidad de Sonora, Hermosillo, Mexico. ¹⁰Universidad de Guadalajara, Guadalajara, Mexico. ¹¹Integrative Biology Program, Instituto de Ciencias Biomedicas (ICBM), Centro de Informática Médica y Telemedicina, Facultad de Medicina, Universidad de Chile, Santiago, Chile. ¹²Departamento de Oncología Básico Clínico and Centro para la Prevención y el Control del Cáncer (CeCAN), Facultad de Medicina, Universidad de Chile, Santiago, Chile. ¹³Instituto de Cancer de São Paulo, São Paulo, Brazil. ¹⁴AC Camargo Cancer Center, São Paulo, Brazil. ¹⁵Hospital de Cancer de Barretos, Barretos, Brazil. ¹⁶Health Sciences Center, Texas A&M University, Bryan, Texas. ¹⁷Instituto Nacional del Cáncer, Santiago de Chile, Chile. ¹⁸Grupo Oncológico Cooperativo Chileno de Investigación, Santiago de Chile, Chile. ¹⁹Fundación para el Progreso de la Medicina, ScireLab, Córdoba, Argentina. ²⁰CONICET, Córdoba, Argentina. ²¹CFEfyN, Depto. de Computación, Escuela de

Ingeniería Biomédica, Universidad Nacional de Córdoba, Córdoba, Argentina. ²²The list of LACRN Investigators with their affiliations is included at the end of the Supplementary Information. ²³Department of Public Health Sciences, School of Medicine, University of California Davis, Davis, California. ²⁴University of California Davis Comprehensive Cancer Center, School of Medicine, University of California Davis, Davis, California.

D. Alves da Quinta and D. Rocha contributed equally to this article.

A.S. Llera and L. Fejerman are sharing senior authorship.

Valentina A. Zavala, currently employed by Regeneron Pharmaceuticals, Inc., contributed to this article as an employee of University of California Davis and the views expressed do not necessarily represent the views of Regeneron Pharmaceuticals, Inc.

Corresponding Author: Andrea S. Llera, Fundación Instituto Leloir, Av. Patricias Argentinas 435, Ciudad de Buenos Aires 1405BWE, Argentina. E-mail: allera@leloir.org.ar

doi: 10.1158/2767-9764.CRC-25-0014

This open access article is distributed under the Creative Commons Attribution 4.0 International (CC BY 4.0) license.

©2025 The Authors; Published by the American Association for Cancer Research

Introduction

Breast cancer is the leading cause of cancer death in Latin American women (1). In this region, breast tumors are frequently diagnosed at stages II or higher, mainly because of relatively low rates of mammography screening (2, 3) and delays in access to care for diagnostic studies (4). Breast cancer is also a heterogeneous disease including tumors with different responses to treatment and survival, and more aggressive subtypes are often associated with more advanced stages at diagnosis (5). Although some studies have shown a higher proportion of aggressive breast cancer subtypes in the region (6), in general, Latin American studies are not population-based and tend to overrepresent aggressive disease. However, registry-based studies in the United States have reported higher proportions of hormone receptor–negative (HR–) disease and HER2-positive (HER2+) disease in Hispanic/Latina women, suggesting that observations in Latin America might reflect a real difference in subtype distribution (7). An analysis conducted on patients with breast cancer from Peru showed an association between Indigenous American (IA) ancestry and HER2+ disease, which was replicated in an independent study including Colombian and Mexican women (8). Genetic ancestry is correlated with both genetic and nongenetic factors (9, 10), therefore, additional testing of the observed association, including better characterization of tumor heterogeneity in women from multiple regions, could give important insights leading to the refinement of our understanding of relevant factors.

IHC markers (HR and HER2) are routinely tested in Latin American patients; however, data defining intrinsic molecular subtypes have been sparse. The Molecular Profile of Breast Cancer Study (MPBCS) of the Latin America Cancer Research Network (LACRN) collected molecular, clinical, and epidemiologic data from an observational cohort of more than 1,200 Latin American patients with breast cancer. Previous publications of the MPBCS have described the distribution of PAM50 subtypes from transcriptomic data and their impact on survival (11, 12), showing a worse prognosis for basal-like tumors [followed by HER2-enriched (HER2E) and luminal B tumors] than for luminal A (LumA) cases, in concordance with other advanced breast cancer cohorts. To test whether the degree of genetic admixture presented in patients of this study might correlate with breast cancer subtype distribution and survival, we evaluated the association between genetic ancestry, tumor subtype, recurrence, and survival in a sample of 951 women from public-health institutions in Argentina, Brazil, Chile, Mexico, and Uruguay participating in the MPBCS. We particularly focused on the intrinsic subtypes as they recapitulate the actual driver pathways of breast cancer more closely and estimate survival more precisely than IHC-based classifications (13).

Materials and Methods

Study participants

The MPBCS detailed eligibility criteria have been described previously (11). Briefly, women with clinical stage II or III (American Joint Committee on Cancer 7) breast adenocarcinoma were deemed eligible for this study. Patients with bilateral or inflammatory breast cancer or metastatic disease were excluded. As recruitment occurred prior to the collection of pathologic diagnosis samples to facilitate the acquisition of additional tissue for molecular studies, eligibility was confirmed retrospectively. Patients who were subsequently reassessed as stage I remained eligible for participation. The protocol received approval from the NCI Ethics Committee and local institutional review boards in each country. The MPBCS was registered

at [ClinicalTrials.gov](https://clinicaltrials.gov) (identifier: NCT02326857) and adhered to the principles of the Declaration of Helsinki and local regulations. Before the study procedures, all participants signed the study-specific written informed consent form. Participants were monitored for 5 years to track survival and recurrence. Electronic case report forms were utilized to capture clinical data, with local data managers ensuring accuracy.

Tumor and blood samples were collected at the time of diagnosis, prior to the initiation of any treatment. From a total of 1,278 patients, DNAs extracted from blood samples from 1,001 eligible patients were successfully genotyped, and 951 constituted the final dataset. From those, 827 also had transcriptomic information available from treatment-naïve tumor samples and constituted the dataset used for association and survival analyses.

Demographic data

After signing the consent form, trained study personnel applied a questionnaire containing questions about socioeconomic and demographic characteristics and lifestyle factors, including age, education, alcohol and tobacco use, access to healthcare, familial cancer, hormonal and reproductive history, and physical activity (14).

Clinical data

Clinical stages for each patient were established according to the American Joint Committee on Cancer staging manual seventh edition. Given that some patients underwent chemotherapy before surgery, the lymph node status at diagnosis was defined clinically as negative (i.e., no clinical evidence of involved nodes) or positive (when at least one node was detected at physical examination). Body mass index was calculated using height and weight reported in medical charts. IHC for HR [i.e., estrogen receptor (ER) and progesterone receptor (PR)], HER2, and Ki67 were determined locally following standard operating procedures (11). All local pathology departments were accredited by the College of American Pathologists. Patients were classified according to IHC subtype based on ER/PR and HER2 status. We used a cutoff of 1% to define ER/PR positivity. HER2 positivity was defined 3+ staining by IHC or 2+ with positive gene amplification by FISH or chromogenic in situ hybridization testing. Detailed information regarding molecular subtype determination (PAM50 subtypes: LumA, luminal B, HER2E, and basal-like) by microarray-based transcriptomic assay has been described previously (11). Quality control measures, including principal component analysis, were implemented to avoid bias. Patients classified as normal by PAM50 were not considered in the analyses.

Genetic ancestry estimation

DNA was extracted from the 1,001 available whole blood samples following standard protocols. DNA samples were genotyped with Infinium Multi-Ethnic Global-8 (MEGA) array version 1.0 (Illumina WG-316-1004), comprising 1,748,250 markers (single-nucleotide variants and insertions/deletions). Genotypes were obtained in GenoType Compressor format using the Illumina Array Analysis Platform Genotyping version 1.1.0 GenCall program. The genotypes and probes of each marker were aligned per sample against the hg19 reference genome with bcftools (<https://samtools.github.io/bcftools/>) and gtc2vcf (<https://github.com/freeseek/gtc2vcf>), obtaining a binary variant call format file (BCL). Quality control was performed using PLINK 1.9 (RRID: SSR_001757; ref. 15). All variants that were not present in at least 10%

TABLE 1 By-country epidemiologic and clinical characteristics of the 951 patients of the LACRN-MPBCS included in this study

Parameter	Country						Univariate analysis ^a
	Total	Argentina	Brazil	Chile	Mexico	Uruguay	
Number of patients included in this study	951	207	207	140	321	76	
Demographic/anthropometric							
Age at diagnosis–mean (SD)	951	55.6 (12.0)	52.5 (11.5)	56.2 (12.7)	52.1 (12.2)	58.4 (12.3)	<i>P</i> < 0.001^b
Years of education– <i>n</i> (%)							<i>P</i> = 0.016 <i>V</i> = 0.13
Up to 8 years	357	88 (42.5)	90 (43.5)	45 (32.1)	101 (31.5)	33 (43.4)	
9 years or +	338	81 (39.1)	73 (35.3)	70 (50.0)	97 (30.2)	17 (22.4)	
Unknown/missing	256	38 (18.3)	44 (21.2)	25 (17.8)	123 (38.3)	26 (34.2)	
BMI– <i>n</i> (%)							<i>P</i> = 0.532 <i>V</i> = 0.06
<25.0 kg/m ²	231	57 (27.5)	56 (27.0)	32 (22.8)	64 (19.9)	22 (28.9)	
25.0–29.99 kg/m ²	294	65 (31.4)	72 (34.8)	49 (35.0)	94 (29.3)	14 (18.4)	
>30.0 kg/m ²	294	62 (29.9)	71 (34.3)	52 (37.1)	87 (27.1)	22 (28.9)	
Unknown/missing	132	23 (11.1)	8 (3.86)	7 (5.00)	76 (23.7)	18 (23.7)	
Ancestry–median (IQR)							
EUR	951	72.2 (59.2–87.4)	79.9 (63.5–90.7)	65.0 (62.1–68.4)	58.5 (52.6–66.5)	83.9 (76.0–91.0)	<i>P</i> < 0.001^b
IA	951	22.9 (9.7–33.6)	5.3 (2.9–8.4)	30.9 (28.4–33.1)	30.8 (24.8–36.6)	10.4 (6.2–17.0)	<i>P</i> < 0.001^b
AFR	951	1.2 (0.001–3.3)	11.9 (2.9–25.2)	0.1 (0.001–1.5)	3.8 (2.3–5.6)	3.2 (0.7–6.3)	<i>P</i> < 0.001^b
EAS	951	2.8 (1.1–4.0)	0.03 (0.001–0.9)	3.2 (2.3–4.1)	5.6 (4.2–6.9)	1.0 (0.001–2.4)	<i>P</i> < 0.001^b
Clinical							
Clinical stage– <i>n</i> (%)							<i>P</i> < 0.001 <i>V</i> = 0.19
Early (IIA–IIB)	347	92 (44.4)	89 (42.9)	41 (29.3)	84 (26.2)	41 (53.9)	
Locally advanced (IIIA–IIIB)	583	114 (55.1)	117 (56.5)	98 (70.0)	223 (69.5)	31 (40.8)	
Missing/other	21	1 (0.5)	1 (0.6)	1 (0.7)	14 (4.3)	4 (5.3)	
Lymph node status– <i>n</i> (%)							<i>P</i> < 0.001 <i>V</i> = 0.22
Negative	406	115 (55.5)	103 (49.7)	44 (31.4)	102 (31.8)	42 (55.3)	
Positive	526	92 (44.4)	103 (49.7)	95 (67.8)	206 (64.2)	30 (39.5)	
Missing	19	—	1 (0.6)	1 (0.8)	13 (4.0)	4 (5.2)	
HER2 status– <i>n</i> (%)							<i>P</i> < 0.001 <i>V</i> = 0.12
Negative	730	171 (82.6)	153 (73.9)	107 (76.4)	236 (73.5)	63 (82.9)	
Positive	189	36 (17.4)	52 (25.1)	23 (16.4)	69 (21.5)	9 (11.8)	
Missing/equivocal	32	—	2 (0.97)	10 (7.14)	16 (4.98)	4 (5.26)	
IHC subtype– <i>n</i> (%)							<i>P</i> < 0.001 <i>V</i> = 0.11
HR(+) HER2(–)	568	134 (64.7)	122 (58.9)	85 (60.7)	175 (54.5)	52 (68.4)	
HR(+) HER2(+)	108	19 (9.2)	36 (17.4)	13 (9.3)	35 (10.9)	5 (6.6)	
HR(–) HER2(+)	81	17 (8.2)	16 (7.7)	10 (7.1)	34 (10.6)	4 (5.3)	
HR(–) HER2(–)	154	37 (17.9)	31 (15.0)	21 (15.0)	57 (17.8)	8 (10.5)	
Missing	40	—	2 (1.0)	11 (7.9)	20 (6.2)	7 (9.2)	
PAM50 subtype– <i>n</i> (%)							<i>P</i> = 0.007 <i>V</i> = 0.10
LumA	376	104 (50.2)	76 (36.7)	54 (38.6)	101 (31.5)	41 (53.9)	
LumB	189	32 (15.5)	44 (21.3)	37 (26.4)	55 (17.1)	21 (27.6)	
HER2E	112	20 (9.7)	24 (11.6)	17 (12.1)	44 (13.7)	7 (9.2)	
Basal-like	150	35 (16.9)	28 (13.5)	23 (16.4)	59 (18.4)	5 (6.7)	
Normal	46	16 (7.7)	6 (2.9)	6 (4.3)	16 (5.0)	2 (2.6)	
Missing	78	—	29 (14.0)	3 (2.3)	46 (14.3)	—	

(Continued on the following page)

TABLE 1 By-country epidemiologic and clinical characteristics of the 951 patients of the LACRN-MPBCS included in this study (Cont'd)

Parameter	Country						Univariate analysis ^a
	Total	Argentina	Brazil	Chile	Mexico	Uruguay	
5-year survival- <i>n</i> (%)							<i>P</i> = 0.015 <i>V</i> = 0.10
Alive	748	166 (80.2)	157 (75.8)	112 (80.0)	254 (79.1)	59 (77.6)	
Dead	152	33 (15.9)	42 (20.3)	26 (18.6)	39 (12.1)	12 (15.8)	
Unknown/missing	51	8 (3.9)	8 (3.9)	2 (1.4)	28 (8.8)	5 (6.6)	

Abbreviations: AFR, African; BMI, body mass index; EAS, East Asian; EUR, European; HER2E, HER2-enriched; IA, Indigenous American; IHC, immunohistochemical; IQR, inter-quartile range; LumA, luminal A; LumB, luminal B.

Percentages (%) are defined as the proportion of individuals of a country that showed the variable level respect to the total individuals of such country (columns).

^a χ^2 *P* value and Cramer's *V* for categorical variables. Bold numbers highlight significant associations between the variable distribution and the country. A Cramer's *V* value of 0.20 or less indicates a weak association, between 0.20 and 0.30 a moderate association, and higher than 0.30 a strong association.

^bKruskal–Wallis test for non-normally distributed continuous variables.

of the samples were removed. Variants were filtered by a Hardy–Weinberg equilibrium test, discarding 9,248 variants with a *P* value lower than 1E-7. Fifty samples with a genotyping call rate of less than 90% were eliminated. From the remaining samples, 391,714 monomorphic variants were eliminated using a minor frequency allele filter >0 (0%). After filtering by 1% minor frequency allele, 873,197 variants were available in the final set of 951 samples.

The following two genome datasets were used to generate the IA ancestry reference group: (i) 17 genomes of Patagonian origin (3 from Laitec Island of putative Chono ancestry, 4 Kaweskar, 3 Huilliche, 3 Pehuenche, and 4 Yamana) and (ii) 21 unrelated individuals of IA origin from Mexico, Brazil, Argentina, Peru, and Colombia [from the Simons Genome Diversity Project (16)], including 1 Chane, 3 Karitiana, 2 Surui, 2 Piapoco, 2 Mayan, 2 Mixe, 2 Mixtec, 2 Pima, 2 Zapotec, and 3 Quechua. All individuals used as IA reference showed >90% IA ancestry in a previous ADMIXTURE analysis (see below). As we had a limited number (*n* = 38) of IA reference genomes, we decided to match as closely as possible the number of reference individuals in the other ancestries to have a similar level of information for each ancestry. Thus, 40 representative samples from Southeast Asia (40 Han Chinese, Europe (20 Iberian and 20 Northern Europeans from Utah), and Africa (40 Yoruba) were randomly selected among the 1000 Genomes phase three individuals with >90% specific ancestry (17). All three sets were converted to PLINK format independently. Hardy–Weinberg equilibrium filtering discarded those variants with a *P* value lower than 1E-7 in the three sets. Triallelic variants were also removed.

A set of reference samples plus those from the MPBCS were created with the intersection of 165,820 variants. After filtering by linkage disequilibrium with PLINK 1.9 (window size 50, number of variants 5, and variance inflation factor threshold 1.2), 41,054 variants were available.

Genetic ancestry estimation, defined as the estimated genetic similarity to reference populations, was performed using the unsupervised ADMIXTURE version 1.23 program (RRID: SCR_001263; ref. 18) with four populations to capture IA, European (EUR), African (AFR), and East Asian (EAS) ancestry based on the known major continental influences to the population of Latin America. As a result, individual fractions of the EUR, IA, AFR, and EAS ancestral components were obtained for the 951 samples. To visualize the

ancestral structure of the MPBCS participants we conducted principal component analysis using the program PLINK 1.9.

Statistical analyses

To evaluate the significance of differences in the distribution of variables among countries, χ^2 tests (stats, RRID: SCR_025968) were used for categorical variables with Cramer's *V* [*rcompanion* R package (19)] as a measure of the strength of association. Age was tested as a continuous variable using a Kruskal–Wallis test (stats, RRID: SCR_025968). The univariate association between EUR ancestry and breast cancer subtype was tested using a Kruskal–Wallis test with a *post hoc* Dunn test and Benjamini–Hochberg *P* value correction.

Multinomial logistic regression models were applied to study the association of the scaled EUR ancestry (i.e., 1 unit difference in the ancestry coefficient is equivalent to a change in 10% of the ancestry component) with breast cancer subtypes [HER2 status (negative or positive), IHC-based subtypes, or PAM50 subtypes] using the *nnet* R package (20). HER2–, HR+ HER2–, and LumA patients were defined as reference groups for each model, respectively. For studying the association between ancestry and survival, Cox proportional hazard regression models with the *survival* R package were performed considering ancestry as a continuous scaled variable (1 U equivalent to a change in 10% of the ancestry component). To select the most important potential confounders in the logistic and Cox models, we analyzed previous evidence in the literature and our own univariate and collinearity analysis (see Extended Figs. E1, E2, and E3 in Supplementary Information for a detailed analysis). Clinical nodal and tumor statuses were strongly correlated with clinical stage (coefficients of 0.78 and 0.57, respectively, Extended Fig. E3 in Supplementary Information), as they are parameters used to calculate clinical stage. For this reason, we chose clinical lymph nodal status (i.e., negative vs. positive) as the simplest and more complete (i.e., more subjects had this variable with data) confounder representative of stage. We also selected age at diagnosis as a continuous variable (correlation coefficients of 0.11 with EUR ancestry and –0.09 with IA ancestry), and AFR ancestry (coefficients of –0.35 with EUR ancestry and –0.01 with IA ancestry) and country (coefficients of 0.21 with EUR ancestry and 0.48 with IA

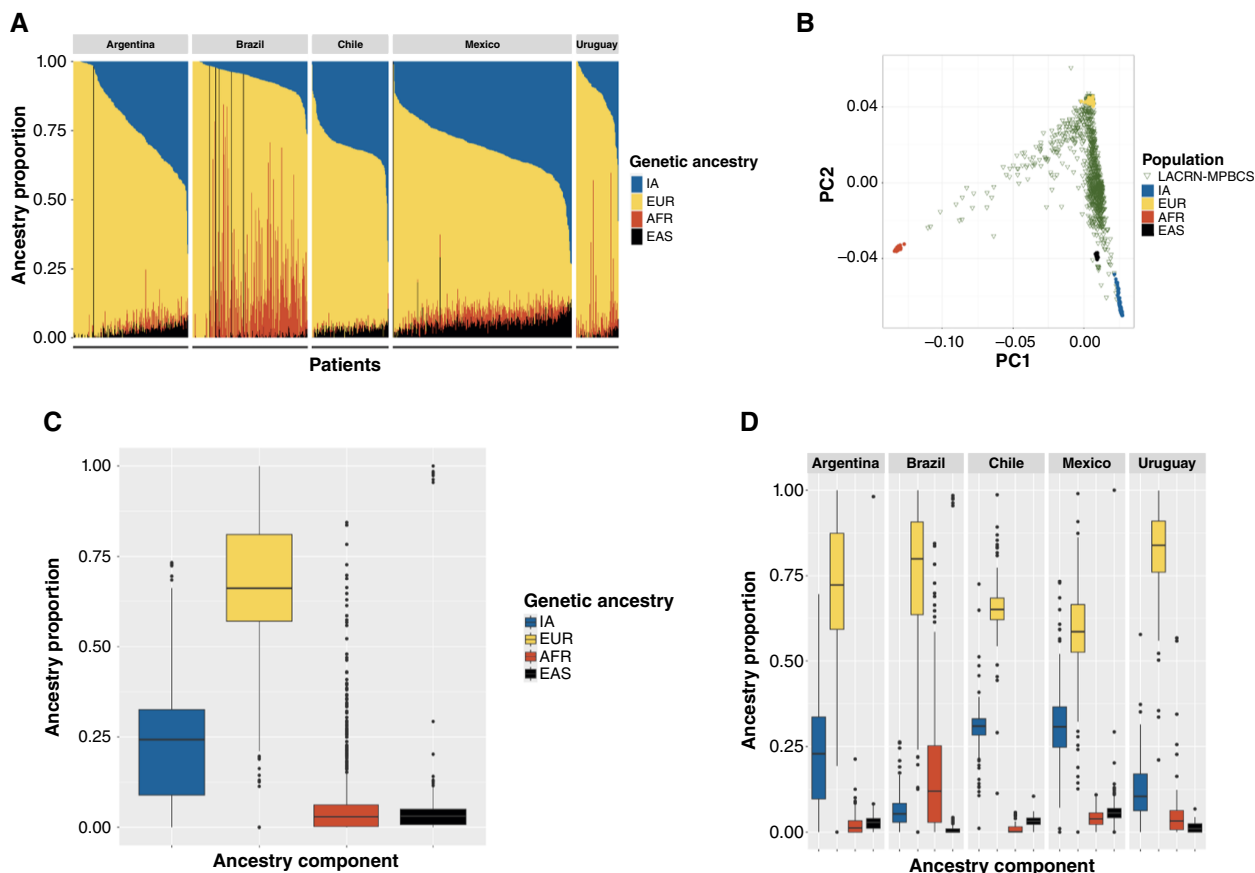


FIGURE 1 Population structure of the 951 patients with breast cancer of the LACRN-MPBCS cohort included in this study. **A**, ADMIXTURE ancestry estimations obtained assuming four ancestral components for each country of origin (**B**) Principal component analysis (PCA) of data from all LACRN-MPBCS patients (triangles) and 158 reference subjects (dots). The two principal components of variation are shown as PC1 and PC2. **C** and **D**, Boxplots showing the distribution of each ancestry component for the whole cohort and by country, respectively. Median ancestry and IQRs are depicted; whiskers extend up to 1.5 IQR. Values outside the inter-quartile range are shown as solid dots. AFR, African; EAS, East Asian; EUR, European; IA, Indigenous American.

ancestry) were included as potential confounders (Extended Fig. E3 in Supplementary Information).

All analyses were performed in R version 4.2.2. P values ≤ 0.050 were considered significant.

Data availability

The processed data and scripts used for this study are available at <https://github.com/danielaalvesdq/LACRN-MPBCS>. Raw genotyping array data generated in this study are currently protected by data policies of the LAGENO-BC and Confluence Consortia and will be available from the corresponding author in the future upon reasonable request.

Results

The distribution of relevant clinical and epidemiologic data of this hospital-based cohort is summarized in Table 1 and reflects the previously published description of larger versions of the cohort (11, 14). In this dataset of 951 patients, heterogeneity between countries was evident for age at diagnosis, education level, genetic ancestry, lymph node status, IHC-based and

PAM50 subtypes, and survival but not for body mass index (Table 1). According to the Cramer's V , the magnitude of the association between variables and country was weak (Cramer's $V \leq 0.20$).

Genetic ancestry distribution in the LACRN-MPBCS

The distribution of the ancestry components of this multicountry cohort showed important differences between individuals and countries (Fig. 1A–D). Study sites in Argentina, Uruguay, and Brazil showed the highest medians for EUR ancestry (72.2%, 83.9%, and 79.9%, respectively), whereas those in Chile and Mexico have the lowest EUR medians (65.0% and 58.5%, respectively). The AFR component is well-represented in Brazil (a median of 11.9%) and to a lesser extent in Mexico (3.8%) and Uruguay (3.2%), whereas it is minimal in Argentina and Chile. The IA median proportion was lowest (5.3%) in Brazil and highest in Chile and Mexico (30.9% and 30.8%, respectively; Table 1; Fig. 1D). A complementarity between the EUR and IA ancestries was evident for most individuals of the cohort (Fig. 1A), with the exception of Brazil in which the AFR component was most relevant (i.e., higher than 15% of the ancestry) in 40% (83/207) of patients. The EAS component was relatively low; only seven individuals were with more than

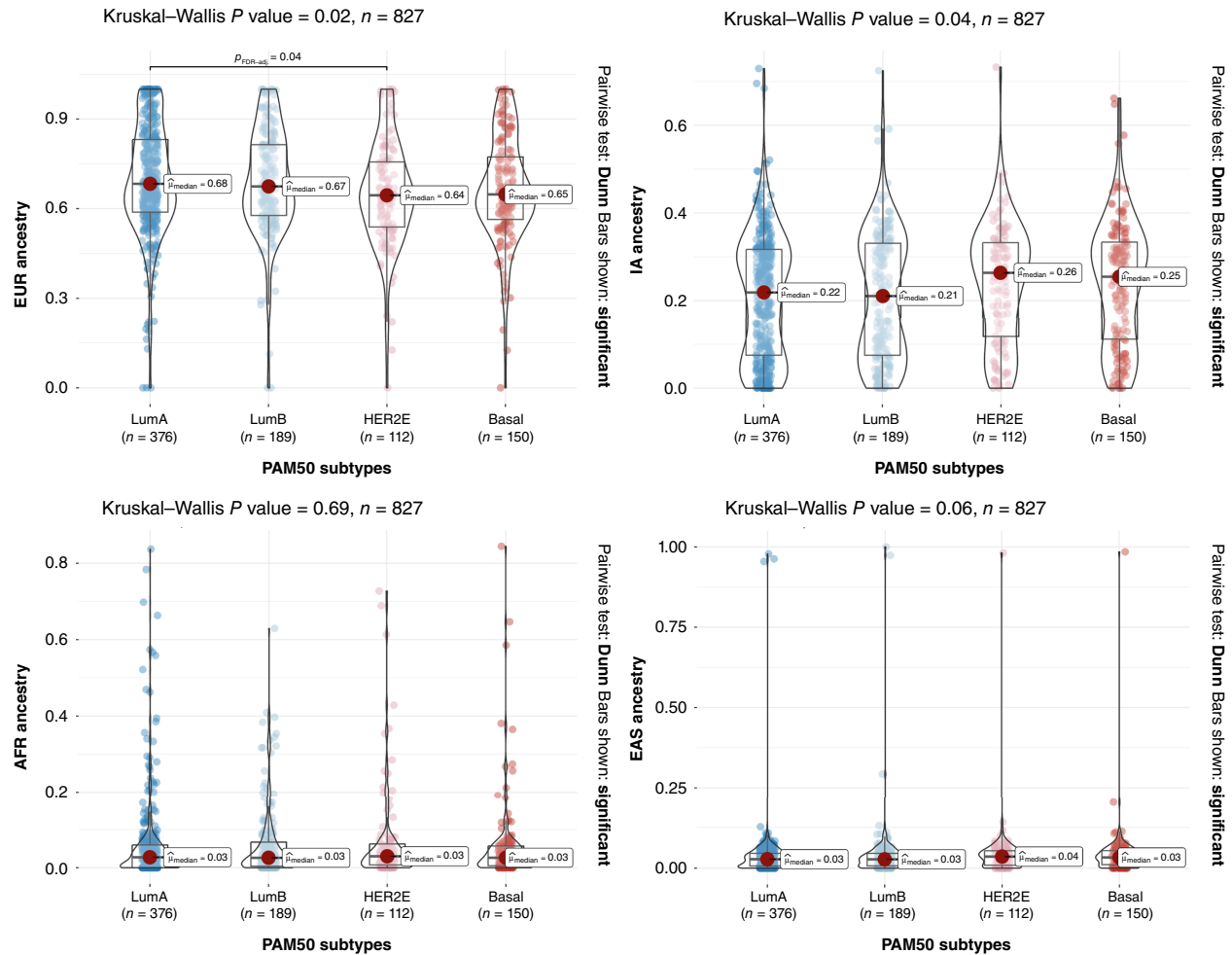


FIGURE 2 Distribution of PAM50 breast cancer subtypes according to the four most prevalent ancestral components in Latin America: European (EUR), Indigenous American (IA), African (AFR), and East Asian (EAS). Bars shown correspond to *post hoc* Dunn test pairwise comparisons <0.05 . HER2E, HER2-enriched; LumA, luminal A; LumB, luminal B.

95% EAS ancestry, corresponding to patients identified as members of the Asian immigrant communities within each country.

At the whole-cohort level, the EUR (median of 66.1%) and IA (median of 24.2%) components were the most represented, followed by AFR (2.9%) and EAS (3.1%; Fig. 1A and C). When differences in the EUR median proportion among countries were tested, all intercountry comparisons were statistically significantly different ($P < 0.050$) except for Argentina versus Brazil and Uruguay versus Brazil, in which no differences in the representation of EUR ancestry were observed ($P = 0.701$ and $P = 1.000$, respectively). We also identified a difference in the EUR and IA ancestry variances between countries (Table 1; Fig. 1D). IQRs for Argentina (59.2–87.4), Brazil (63.5–90.7), and Uruguay (76.0–91.0) showed a larger variance in EUR coefficients than that of Chile (IQR = 62.1–68.4) and Mexico (IQR = 52.6–66.5). A similar, complementary picture was seen for the IA component (Table 1).

To estimate biases in ancestry representation among institutions, we explored differences in genetic ancestry estimates between institutions within countries (Supplementary Fig. S1). For most countries, no significant

differences were observed except for Brazil ($P = 0.003$ for EUR, $P < 0.001$ for AFR, and $P > 0.050$ for IA and EAS components).

Association between breast cancer subtypes and genetic ancestry

To study whether there was any association between ancestry and tumor subtypes, the median proportion of the different genetic ancestry components was compared among breast cancer subtypes defined both by IHC and PAM50. An analysis of differences of medians showed that EUR and IA proportions (but not AFR and EAS) were significantly different among PAM50 subtypes (Fig. 2). In the case of EUR ancestry, we could further demonstrate that the statistical significance was driven by the LumA–HER2E contrast and explained by the lower EUR ancestry associated with the HER2E subtype (LumA vs. HER2E adjusted $P = 0.038$ for EUR ancestry, Fig. 2). We could not see statistically significant differences in ancestry distribution for IHC-based subtypes (Supplementary Fig. S2).

Given the heterogeneity of the distribution of IA and AFR ancestries among countries, the limited number of patients, and the strong correlation between

TABLE 2 Association between HER2 status, IHC and PAM50 subtypes, and EUR genetic ancestry (for every 10% increase in EUR or IA ancestry component)

			EUR ancestry		IA ancestry	
	Subtype	<i>n</i>	OR (CI)	<i>P</i> value	OR (CI)	<i>P</i> value
By HER status						
Univariate	HER2−	919 ^a	Ref		Ref	
	HER2+		0.94 (0.86–1.03)	0.190	1.02 (0.91–1.14)	0.720
Nodal status	HER2−	903	Ref		Ref	
	HER2+		0.95 (0.87–1.04)	0.250	1.00 (0.89–1.12)	0.954
Nodal status + age at Dx	HER2−	903	Ref		Ref	
	HER2+		0.96 (0.87–1.05)	0.334	0.99 (0.88–1.12)	0.928
Nodal status + age at Dx + AFR ancestry	HER2−	903	Ref		Ref	
	HER2+		0.97 (0.88–1.07)	0.535	1.01 (0.90–1.14)	0.814
Nodal status + age at Dx + AFR ancestry + country	HER2−	903	Ref		Ref	
	HER2+		0.94 (0.84–1.06)	0.331	1.08 (0.91–1.27)	0.377
By IHC subtypes						
Univariate	HR+/HER2−	911 ^b	Ref		Ref	
	HR+/HER2+		0.98 (0.88–1.10)	0.755	0.98 (0.85–1.13)	0.759
	HR−/HER2+		0.89 (0.78–0.99)	0.040	1.14 (0.97–1.34)	0.109
	HR−/HER2−		0.95 (0.86–1.05)	0.308	1.12 (0.99–1.27)	0.071
Nodal status	HR+/HER2−	895	Ref		Ref	
	HR+/HER2+		0.99 (0.88–1.12)	0.906	0.95 (0.82–1.11)	0.554
	HR−/HER2+		<i>0.88 (0.78–1.00)</i>	<i>0.059</i>	1.11 (0.94–1.31)	0.205
	HR−/HER2−		0.97 (0.87–1.07)	0.556	1.08 (0.95–1.23)	0.226
Nodal status + age at Dx	HR+/HER2−	895	Ref		Ref	
	HR+/HER2+		1.01 (0.88–1.13)	0.844	0.94 (0.81–1.09)	0.406
	HR−/HER2+		0.87 (0.78–1.01)	0.065	1.11 (0.94–1.31)	0.217
	HR−/HER2−		0.98 (0.88–1.09)	0.716	1.07 (0.94–1.21)	0.307
Nodal status + age at Dx + AFR ancestry	HR+/HER2−	895	Ref		Ref	
	HR+/HER2+		0.99 (0.88–1.12)	0.906	0.95 (0.81–1.11)	0.550
	HR−/HER2+		<i>0.88 (0.78–1.00)</i>	<i>0.059</i>	1.14 (0.96–1.36)	0.143
	HR−/HER2−		0.97 (0.87–1.07)	0.556	1.07 (0.94–1.23)	0.287
Nodal status + age at Dx + AFR ancestry + country	HR+/HER2−	895	Ref		Ref	
	HR+/HER2+		0.97 (0.84–1.13)	0.763	1.08 (0.87–1.34)	0.505
	HR−/HER2+		0.91 (0.77–1.06)	0.226	1.12 (0.89–1.41)	0.330
	HR−/HER2−		0.99 (0.87–1.14)	0.937	1.10 (0.92–1.31)	0.303
By PAM50						
Univariate	LumA	827 ^c	Ref		Ref	
	LumB		0.97 (0.88–1.07)	0.572	1.01 (0.89–1.14)	0.848
	HER2E		0.86 (0.77–0.97)	0.011	1.16 (1.00–1.35)	0.044
	Basal		0.90 (0.81–0.99)	0.047	1.17 (1.02–1.33)	0.021
Nodal status	LumA	810				
	LumB		0.97 (0.88–1.08)	0.629	1.01 (0.89–1.15)	0.871
	HER2E		0.86 (0.76–0.96)	0.009	1.16 (1.00–1.36)	0.050
	Basal		0.92 (0.82–1.02)	0.120	1.13 (0.98–1.29)	0.095
Nodal status + age at Dx	LumA	810				
	LumB		0.97 (0.88–1.08)	0.631	1.01 (0.89–1.15)	0.875
	HER2E		0.86 (0.76–0.97)	0.011	<i>1.16 (0.99–1.35)</i>	<i>0.056</i>
	Basal		0.93 (0.83–1.04)	0.182	1.11 (0.97–1.28)	0.130

(Continued on the following page)

TABLE 2 Association between HER2 status, IHC and PAM50 subtypes, and EUR genetic ancestry (for every 10% increase in EUR or IA ancestry component) (Cont'd)

	Subtype	n	EUR ancestry		IA ancestry	
			OR (CI)	P value	OR (CI)	P value
Nodal status + age at Dx + AFR ancestry	LumA	810				
	LumB		0.97 (0.87–1.08)	0.635	1.01 (0.89–1.15)	0.847
	HER2E		0.86 (0.76–0.98)	0.021	1.21 (1.03–1.42)	0.022
	Basal		0.92 (0.81–1.03)	0.145	1.12 (0.97–1.30)	0.126
Nodal status + age at Dx + AFR ancestry + country	LumA	810				
	LumB		0.97 (0.85–1.10)	0.672	1.02 (0.85–1.22)	0.857
	HER2E		0.90 (0.78–1.04)	0.151	1.18 (0.95–1.45)	0.124
	Basal		0.96 (0.83–1.10)	0.547	1.07 (0.89–1.30)	0.461

Ancestry was modeled as a continuous variable and coefficients were scaled to reflect a 10% increase in ancestry. The number of individuals (*n*) in each analysis depends on the completeness of the variables used for adjustment. Bold numbers denote statistically significant differences, and italic numbers denote marginally nonsignificant values.

Abbreviations: CI, confidence interval; Dx, diagnosis; HER–, HER2 nonamplified; HER+, HER2 amplified; HER2E, HER2-enriched; LumA, luminal A; LumB, luminal B; OR: odds ratio.

^aFrom the total of 951 genotyped patients, 32 had a HER2-missing status (Table 1), and from those, 16 lacked the nodal status.

^bFrom the total of 951 genotyped patients, 40 had missing status of any of the HR or HER2 markers (Table 1), and from those, 16 lacked the nodal status.

^cFrom the total of 951 genotyped patients, 124 had either missing data or belonged to the normal PAM50 subtype, which was not considered in this study (Table 1); from those, 17 also lacked the nodal status.

EUR and IA coefficients (–0.78, see Extended Fig. E3 in Supplementary Information), we first decided to perform subsequent analyses using the EUR ancestry proportion as a proxy of admixture. To evaluate whether there was an association between EUR ancestry and HER2 status, IHC or PAM50 subtypes in the context of a multivariable model, we conducted a multinomial logistic regression model with the subtype as the dependent variable and scaled EUR ancestry as the main predictor. The results from the univariate analysis showed that the HR– HER2+ subtype, but not HER2 status alone, was associated with lower EUR ancestry (Table 2, left). In the HR– HER2+ subtype, a 10% increase in EUR ancestry was significantly associated with an 11% decrease in the OR of presenting with this tumor subtype. The addition of the selected covariates kept the direction of the OR but affected the significance (Table 2). In the model with PAM50-intrinsic subtypes as the outcome, we also observed an inverse association between HER2E and EUR ancestry. In this case, a 10% increase in EUR ancestry was significantly associated with a 14% decrease in the odds of presenting HER2E breast cancer. The OR for the basal subtype was also significantly decreased (10%) with a 10% increase in EUR ancestry. The inverse association between the HER2E subtype and EUR ancestry maintained statistical significance with the addition of age, lymph node status, and AFR ancestry as covariates (Table 2). The incorporation of the variable country as an additional covariate did not significantly affect the OR but it did result in an increase in the *P* value of the model, rendering it statistically nonsignificant.

We further evaluated the association between the scaled IA ancestry and the breast cancer subtypes. This model did not reach significance for the IHC-based subtypes but showed a 16% increase in the odds of presenting HER2E breast cancer for every 10% additional IA ancestry (*P* = 0.044, Table 2, right). The incorporation of nodal status, age, and AFR ancestry covariables to the model rendered higher odds (21%) of presenting HER2E breast cancer

for every 10% additional IA ancestry (*P* = 0.021). The inclusion of the “country” covariable abrogated statistical significance, although the direction and size of the OR was consistent with the association. In addition, in the univariate model, we could detect a 17% increase in the odds of presenting basal-like breast cancer for every 10% additional IA ancestry, but this association lost significance with the inclusion of covariables to the model (Table 2, right).

Genetic ancestry, tumor subtype, and overall survival

We then evaluated the association between genetic ancestry and survival in univariate and adjusted Cox models, including the same covariates as in the previous analysis (age, lymph node status, AFR ancestry, and country) and adding the PAM50 subtypes (see Supplementary Information for a detailed description of the selection of covariables). Neither EUR nor IA ancestry was significantly associated with overall survival in univariate analysis (Table 3). Adjustment by confounders such as PAM50 subtype, age, AFR ancestry, and lymph node status resulted in an apparent increase in the hazard ratio with increasing EUR ancestry that was reverted by the addition of country as an additional confounder (Table 3).

Discussion

There have been a limited number of studies conducted in diverse cohorts of Latin American patients with breast cancer that explored the association of genetic ancestry and tumor molecular characteristics (7, 8, 21, 22). These studies included women from Perú, Mexico, and Colombia. In this work, we further tested the association between genetic ancestry and breast cancer subtypes, defined both by IHC markers and PAM50, in the MPBCS cohort, which includes patients from Argentina, Brazil, Chile, Mexico, and Uruguay.

TABLE 3 Univariate and multivariate Cox proportional hazard models for overall survival for every 10% increase of EUR or IA ancestry

	<i>n</i>	Hazard ratio (CI)	<i>P</i> value
EUR ancestry ^a			
Univariate	793	1.07 (0.97-1.18)	0.189
PAM50 subtypes	793	1.11 (1.00-1.23)	0.043
PAM50 subtypes + nodal status	780	1.14 (1.03-1.27)	0.014
PAM50 subtypes + nodal status + age at Dx	780	1.14 (1.02-1.27)	0.015
PAM50 subtypes + nodal status + age at Dx + AFR ancestry	780	1.15 (1.03-1.29)	0.012
PAM50 subtypes + nodal status + age at Dx +AFR ancestry + country	780	1.05 (0.92-1.21)	0.449
IA ancestry ^a			
Univariate	793	0.94 (0.83-1.06)	0.337
PAM50 subtypes	793	0.90 (0.79-1.02)	0.093
PAM50 subtypes + nodal status	780	0.86 (0.76-0.98)	0.021
PAM50 subtypes + nodal status + age at Dx	780	0.86 (0.75-0.98)	0.020
PAM50 subtypes + nodal status + age at Dx + AFR ancestry	780	0.85 (0.74-0.97)	0.013
PAM50 subtypes + nodal status + age at Dx + AFR ancestry + country	780	0.94 (0.79-1.11)	0.466

Abbreviations: CI, confidence interval; Dx, diagnosis.

^aAncestry was modeled as a continuous variable, and coefficients were scaled to reflect a 10% increase in the ancestry proportion. The number of individuals (*n*) in each analysis depends on the completeness of the variables used for adjustment.

The genetic ancestry distributions of these countries are heterogenous and close to those previously described (23–26), with our data showing that EUR ancestry is predominantly represented across all study sites in the different countries. In Chile and Mexico, the contribution of IA ancestry is higher compared with other countries. Additionally, Brazil shows an important proportion of AFR ancestry.

Our findings support previous observations of a higher frequency of HER2-dependent tumors in patients with increased IA and decreased EUR ancestry (8). In the MPBCS cohort, the association is seen only for HR– tumors, an observation already suggested by the Peruvian and Colombian studies (7). Moreover, even when the MPBCS cohort included both IHC and gene expression-based subtypes, the association seems to be more specific to the PAM50 HER2E subtype as the ORs were higher and the statistical significance was stronger for this intrinsic subtype than for its IHC counterpart. Interestingly, the HER2E subtype includes those tumors in which the HER2 pathway is active, regardless of the amplification status of the *ERBB2* gene. It is our hypothesis (to be explored) that the effect of IA ancestry on the HER2 pathway may not only be related to the amplification of *ERBB2* but to the activation of the HER2 pathway by various mechanisms.

Of note, the significance of the ORs was affected by the addition of the “country” variable, likely because of the power limitations when considering effects within each country. Sequentially adjusted models maintained the magnitude of the ORs and *P* values, suggesting that there are not strong mediators or confounders in the association between ancestry and subtype, except for the model including country.

In univariate Cox proportional hazard regression models, we showed that EUR and IA ancestries were not significantly associated with overall survival. However, adjustment for PAM50-intrinsic subtypes, age, AFR ancestry, and

lymph node status rendered the model significant, showing an increase in mortality with higher EUR ancestry concomitant with a decrease in mortality with higher IA ancestry. The effect of the addition of country as a confounding variable in the model also abrogated these effects. Previous reports showed contradictory evidence of the effect of ancestry in breast cancer survival. On one hand, a lack of association between genetic ancestry and overall or cancer-specific survival was shown in a Californian Hispanic/Latina breast cancer cohort with homogeneous access to care (27). On the other hand, a more heterogeneous Hispanic/Latina cohort showed a twofold increase in mortality in women with more than 50% IA ancestry compared with women with 50% or less IA ancestry (28). Evidently, the complex and context-specific interplay between biological and nonbiological determinants of survival in admixed populations should be clarified with larger, comprehensive datasets from admixed cohorts (7).

This study has some limitations. First, the LACRN-MPBCS cohort is hospital-based and may not be representative of Latin American breast cancer in terms of clinical and/or pathologic characteristics. In addition, on average, participants from Mexico had lower IA ancestry than expected based on previous literature. Guadalajara and Sonora, known to have a more important Spanish contribution than other regions in Mexico (25, 29), was the source of patients in the Mexican LACRN-MPBCS cohort. It is also possible that patients recruited from Mexican sites had higher EUR ancestry than the general Mexican population. Alternatively, a higher proportion of EUR ancestry among LACRN-MPBCS participants could be explained by the previously described positive association between EUR genetic ancestry and breast cancer risk in Hispanic/Latina and Latin American women (10, 30–32). The skewed EUR ancestry proportion for the Mexican site and the limited proportion of IA ancestry in Brazil and Uruguay had an impact on the representation of the IA ancestry in the cohort, thus limiting the power of the

analysis to evaluate the influence of IA in subtype distribution. Another limitation was the number of subjects for whom all data was available, which suggests that small but significant effects may have been missed in multivariable analyses because of missingness for some of the covariates.

In summary, the admixed LACRN-MPBCS cohort, with representation of Latin American countries that were not present in other studies, supports an association between IA ancestry and the HER2E breast cancer subtype. These results strengthen the hypotheses of the existence of either population-specific genetic variant(s) or of other ancestry-linked or correlated factors that affect HER2 expression in breast cancer in a consistent manner across different Latin American regions. We have already shown that SNPs specific to IA ancestry can affect cancer incidence in a subtype-specific manner (33–35). We can speculate that not yet discovered, ancestry-specific expression quantitative trait locus may be either affecting HER2 expression or signaling pathways relevant to HER2 expression (36). Other possible explanations may involve the existence of ancestry-specific splice variants (37) or genetic variants in other genes that affect the probability of HER2 pathway activation in tumor cells (38–40). On the other hand, non-genetic factors other than the ones included in our models may be acting as confounders on the association between genetic ancestry and the HER2E subtype (41, 42). This is especially relevant given that the association seen in this study is abrogated by the inclusion of “country” as a confounding variable. For example, Hispanic/Latino ancestry has been associated with lower socioeconomic status in the United States (9, 43). Individuals from lower socioeconomic backgrounds tend to seek medical attention at more advanced stages of breast cancer, often presenting with more aggressive tumor subtypes (44). This evidence may result from disparities in access to health services as a consequence of living in remote places and/or from a lack of awareness of slower-growing tumors (9, 45). These factors may induce a bias in the proportion of subtypes that reach medical care. We are actively pursuing studies that might shed light on the biological explanation for this observation.

Ethics statement

The MPBCS was registered at [ClinicalTrials.gov](https://clinicaltrials.gov) (identifier: NCT02326857) and adhered to the principles of the Declaration of Helsinki and local regulations. The study protocol was approved by the NCI Ethics Committee and local institutional review boards in each country. Before the study procedures, all participants signed study-specific written informed consent forms.

Authors' Disclosures

D. Ganiewich reports personal fees from Fundación Leo Messi during the conduct of the study. V.A. Zavala is currently employed at Regeneron Pharmaceuticals; however, he contributed to this study as a postdoctoral student at the University of California, Davis under Laura Fejerman's mentorship. The views expressed do not necessarily represent the views of Regeneron Pharmaceuticals Inc. A. Daneri-Navarro reports other support from the Center for Global Health at the U.S. NCI/NIH during the conduct of the study. J. Retamales reports grants from Pfizer Competitive Grant Program outside the submitted work. A.S. Llera reports personal fees from CONICET and grants from the Center for Global Health, Fogarty International Center, NIH, Department of Health and Human Services (HHS), Susan G. Komen for the Cure, Instituto Nacional del Cáncer Argentina, Fundación Argentina de Nanotecnología, and Agencia

Nacional de Promoción Científica y Tecnológica during the conduct of the study. No disclosures were reported by the other authors.

Authors' Contributions

D. Alves da Quinta: Formal analysis, visualization, methodology, writing—original draft. **D. Rocha:** Data curation, formal analysis, methodology, writing—original draft. **C. Yáñez:** Formal analysis, methodology, writing—review and editing. **R. Binato:** Data curation, formal analysis, methodology, writing—review and editing. **S.C. Soares-Lima:** Methodology. **X. Huang:** Visualization, writing—review and editing. **D. Ganiewich:** Data curation, writing—review and editing. **V.A. Zavala:** Visualization, writing—review and editing. **M. Sans:** Conceptualization, writing—review and editing. **A. Lopez-Vazquez:** Data curation, writing—review and editing. **J. Quintero:** Data curation, writing—review and editing. **O. Valenzuela:** Data curation, writing—review and editing. **A. Quintero-Ramos:** Data curation, writing—review and editing. **A. Del Toro-Arreola:** Data curation, writing—review and editing. **M. Cerda:** Data curation, writing—review and editing. **K. Marcelain:** Conceptualization, writing—review and editing. **S. Crocamo:** Data curation, writing—review and editing. **M.A. Nagai:** Data curation, writing—review and editing. **D.M. Carraro:** Data curation, writing—review and editing. **M.M.C. Marques:** Data curation, writing—review and editing. **J. Gómez:** Funding acquisition, writing—review and editing. **N. Artagaveytia:** Conceptualization, writing—review and editing. **A. Daneri-Navarro:** Conceptualization, writing—review and editing. **B.G. Müller:** Conceptualization, writing—review and editing. **J. Retamales:** Data curation, writing—review and editing. **C. Velazquez:** Conceptualization, writing—review and editing. **E.A. Fernández:** Conceptualization, writing—review and editing. **O.L. Podhajcer:** Conceptualization, writing—review and editing. **E. Abdelhay:** Conceptualization, supervision, writing—review and editing. **R.A. Verdugo:** Conceptualization, data curation, supervision, methodology, writing—review and editing. **A.S. Llera:** Conceptualization, supervision, investigation, methodology, writing—original draft. **L. Fejerman:** Conceptualization, supervision, investigation, methodology, writing—original draft.

Acknowledgments

Past and current members of the LACRN are acknowledged for their essential work on data collection and curation. Their names and affiliations are listed at the end of Supplementary Information. This work was supported by the Center for Global Health at the U.S. NCI/NIH (contract award No. HHSN2612010000871/NO2-PC-2010-00087), Fogarty International Center, NIH, HHS, and Susan G. Komen for the Cure; Argentina: Instituto Nacional del Cáncer (Ministry of Health), Fundación Argentina de Nanotecnología, Agencia Nacional de Promoción Científica y Tecnológica, and CONICET (Ministry of Science, Technology, and Productive Innovation); Brazil: Ministério da Saúde (Ministry of Health); Chile: Instituto de Salud Pública (Public Health Institute), Ministerio de Salud (Ministry of Health), and Agencia Nacional de Investigación y Desarrollo (ACT210079, FONDEF ID21I10355, FONDAP 152220002); and Mexico: Consejo Estatal de Ciencia y Tecnología de Jalisco (COECYTJAL) and Universidad de Sonora (University of Sonora). Financial support for Fejerman L. comes from the Placer Breast Cancer Endowed Chair, University of California, Davis (L. Fejerman) and the NCI (R01-CA286650, R01-CA273313, and R01-CA204797). LACRN investigators: Juan Abarca, Eliana Abdelhay, Pamela Acevedo,

Graciela Acosta, Gissel Acosta, Ana Acosta, Gabriela Acosta Haab, Keyla Teresa Acosta-Torres, Marta Aghazarian, Carola Aguayo, Bernardo Aizen, Gustavo Alarcon-Lopez, Elsa Alcoba, Liz Almeida, Isabel Alonso, Ana Alvarez, Viviane Andrade, Wenceslao Angeles-Bueno, Roberto Arai, Priscila Elvira Arambula-Barreras, Ma. Isabel Arámburo-Rubio, Estrellita Araus, Gonzalo Ardao, Lilia A. Arellano-Jimenez, Felipe Argandoña, Claudia Arias, Ricardo Armisen, Nora Artagaveytia, Mauricio Aspee, Rodrigo Assar, Itzel René Astiazarán-Rascón, Sebastian Astorga, Maxwell Avilés-Rodríguez, Antônio Bailão Junior, Adolfo E. Barragan-Curiel, Adelfo Barragan-Ruiz, Fernanda Bermudez, Julia Bernachin, Wilfrido Bernal-Herrera, Renata Binato, Mara Bonet, Alicia I. Bravo, Sarah Brnich, Claudio Bustamante, Miguel Angel Bustamante, Julio Bustos-Gomez, Felipe de J. Bustos-Rodriguez, Janett Caballero-Jasso, Angie Calfuman, Natalia Camejo, Antonio Hugo José Froes Marques Campos, Mónica Campos, Soledad Cano, Juan C. Canton-Romero, Ricardo Cappelletta, Paulina Carmona, Dirce Maria Carraro, Fernando Carrizo, André Lopes Carvalho, Erika Carvalho, Julio Carzogio, Monica Casalnuovo, Benedicta Caserta, Alvaro Castillo, César Castillo, Mónica Castro, Juan M. Castro-Cervantes, Sandra Cataldi, Alfonso Cayota, Mauricio Cerda, Yascara Cerda, Roger Chammas, Mario Alberto Chavez-Zamudio, Loreto Chia, Elisa Chiarello, Daniela Chirico, Esther Cisneros-Quirarter, Alicia Colombo, Minor Raul Cordero-Bautista, Valeria Cornejo, Baldemar Corral-Villegas, Andrés Cortés, Sandra Cortés, Laura Cortes-Sanabria, German Salvador Cortez-Zamorano, Alejandro Corvalan, Susanne Crocamo, Adolfo Cruz, Alba d'Aurora, Adrian Daneri-Navarro, Sandra De la Fuente, Soledad De la Peña, Roberto de Leon-Caballero, Mirian de Souza, César Del Castillo, Alicia Del Toro-Arreola, Azucena Del-Toro-Valero, Raul Delgadillo-Cristerna, Lucía Delgado, Mirtha Di Pretoro, Andrea Digonzelli, Marisa Dreyer Breitenbach, Jose El Ters, Paula Escobar, Marcela Estolaza, Adriane Feijo Evangelista, Marcelo Fanelli, Paulo Farias, Graciela Fernandez, Elmer Fernández, Jorge Fernández, Wanda Fernández, Natalia Filgueiras, Diego Flaks, Edgar G. Flores-Ayala, Maria R. Flores-Marquez, David Franco-Hughes, Ramon A. Franco-Topete, Karina Franco-Topete, Jimena Freire, Cristobal Fresno, Carolina Gabay, Romina Gabrielli, Fancy Gaete, Mario Gallegos, Jorge Gamboa, Daiana Ganievich, Carlos Garbovesky, Ricardo Garcia-Gaeta, Alma C. Garcia-Martinez, Rubén Alejandro García-Munguía, Adriana Garibay-Escobar, Liliana Gimenez, Hector Gómez Silveira, Mariana M. Gomez-Del Toro, Marcela Gonzalez, Alicia Gonzalez, Germán González, César Osbaldo González-Mondaca, Leivy P. Gonzalez-Ramirez, Beatriz Gonzalez-Ulloa, Susana Gorostidy, Mariela Grass, Gonzalo Greif, Marisol Guerrero, Alfonso G. Guevara-Torres, Lorena Gutierrez, Susan A. Gutierrez-Rubio, Adrián Hannois, Andrew Hart, Steffen Härtel, Marcos Henriquez, Miriam E. Hernandez-Franco, Rafael Hernandez-Guevara, Manuel I. Herrera-Miramontes, Graciela Horton, Gladys Ibañez, Martín Ipiña, Beatriz Jalfin, Lilian Jara, Raul Jara, Maria Luisa Jaramillo, Maria Eugenia Jimenez, Victor M. Jimenez-Moreno, Hugo Ju, Nazareth Juárez Rusjan, Karen Juneman, Ligia Maria Kerr, Alejandra Krupelis, Flor Esmeralda Larios-Jimenez, Jose Domingo Latorre, Guillermo Laviña, Fernando Lavista, Irma Leticia León-Duarte, Alberto Lescano, Verónica Lezano, Andrea S. Llera, Rossana Mendoza Lopez, Jose Guillermo López-Cervantes, Miguel Enrique Lopez-Muñoz,

Alejandra Lopez-Vazquez, Dora Loria, Alejandra Luque, Alejandro Maass, Maria do Socorro Maciel, Silvina Maldonado, Flavia Rotea Mangone, Jorge Mansilla, Katherine Marcelain, Carolina Mariani, Marcia Maria Chiquit Marques, Reyna J. Martinez-Arriaga, Hector R. Martinez-Ramirez, Marcela Martins, Alma G. Maya-Gonzalez, Brenda Mazzafferri, Mariana Menini, Silvia Míguez, Soledad Milans, Soledad Montes, Ana Verónica Morales-Hernández, Andres de J. Moran-Mendoza, Giberto Morga-Villela, Carla Morong, Bettina Müller, Homero Muñoz, Ignacio Miguel Muse, Carina Mussetti, Eduardo Mussetti, Maria Aparecida Nagai, Luis J. Najar-Acosta, Elisa Napolitano e Ferreira, Nancy E. Navarro-Ruiz, Cristina Noblía, João Soares Nunes, Daniela Núñez, Fabiola Núñez, Antonio Ocegüera-Villanueva, Nilton Onari, Emma M. Oropeza-De Anda, David Ortega-Tirado, Miguel Angel Ortiz-Martinez, Cynthia Aparecida Bueno de Toledo Osório, Carlos Eduardo Paiva, Paulina Peñaloza, Miguel Peredo-Navarro, David Pereira, Laura Perez-Michel, Francisca Pino, Tania Pino, Natalia Pinto, Jessica Pizarro, Osvaldo L. Podhajcer, Carlos Pressa, Jael Quintero, Antonio Quintero-Ramos, Enrique Ramirez, Gladys E. Ramirez-Rosales, Claudia Ramis, Maritza Ramos-Ramirez, Adela Rascon-Alcantar, Silvana Ravaglio, Rui M. Reis, Javier Retamales, Francois Richard, Omar Rios-Méndez, Ernesto Rivera-Claiss, Ramón E. Robles-Zepeda, Iara Santana Rocha, Natalia Rodriguez, Vilma Rodriguez, Maria Teresa Rodriguez, Robinson Rodriguez, Diego Rodriguez-Gonzalez, Rosemeire A. Roela, Ana M. Romero-Gomez, Cristina Rosales, Ana M. Rosales-Sandoval, Lidia A. Rubio-Chavez, Omar V. Rubio-Plascencia, Florencia Russo, Gaciela Sabini, Isabel Saffie, Efrain Salas-Gonzalez, Brenda Samaniego, Julio San Martino, Benito Sanchez-Llamas, Verónica Sanchotena, Daniel Sat-Muñoz, Mariana Savignano, Cristovam Scapulatempo Neto, Laura Segovia, Juan M. Sendoya, Max Mano Senna, Carolina Silva, Aida A. Silva-Garcia, Jaime Silvera, Isabele Small, Fernando Soares, Iberê Soares, Silvana Soares dos Santos, Evandro Sobrosa de Mello, José Antonio Sola, Irene Sorín, Anabella Sosa, Alejandra Sosa, Claudio Sosa, Sandra Soto, Cristiano de Pádua Souza, Lucía Spangenberg, Gustavo Steffanof, Florencia Straminsky, Mónica Tapia, Raziel O. Tapia-Llanos, Geronimo M. Tavares-Macias, Guillermo Temperley, Verónica Terzieff, Vicente Teti, Javier Tognarelli, Verónica Toledo, Paulina Toro, Roberto Torres, Mariana Torres-Palomares, Alejandra Trincherro, Rogelio Troyo-San Roman, Hernan Urbano, Nicolas Vacca, Daniel Vaimberg, María Lourdes Valencia-Peña, Olivia Valenzuela, Maria Lujan Vaselevich, Jaime Vazquez-Nares, Carlos Velazquez, Ezequiel Velez-Gomez, Laura N. Venegas-Godinez, Patricia Vercelli, Ricardo Verdugo, René Aloisio da Costa Vieira, Marta Vilensky, María José Villarubias, Manuel Isaac Villegas-Gómez, Stella Viña, Silvia Vornetti, Anapaula Hidemi Uema Watanabe, Livia Zagame, Carlos Zamorano, Luis Zapata, and Zdenka Zlatar.

Note

Supplementary data for this article are available at Cancer Research Communications Online (<https://aacrjournals.org/cancerrescommun/>).

Received January 11, 2025; revised April 03, 2025; accepted June 03, 2025; posted first June 06, 2025.

References

- Piñeros M, Laversanne M, Barrios E, Cancela MdC, de Vries E, Pardo C, et al. An updated profile of the cancer burden, patterns and trends in Latin America and the Caribbean. *Lancet Reg Health Am* 2022;13:100294.
- Ayala N, Barchuk S, Inurrigarro G, Celano C, Soriano-García JL, Bolaños P, et al. Status of breast cancer in Latin American: results of the breast cancer revealed initiative. *Crit Rev Oncol Hematol* 2023;181:103890.

3. Justo N, Wilking N, Jönsson B, Luciani S, Cazap E. A review of breast cancer care and outcomes in Latin America. *Oncologist* 2013;18:248–56.
4. Barrios CH, Werutsky G, Mohar A, Ferrigno AS, Müller BG, Bychkovsky BL, et al. Cancer control in Latin America and the Caribbean: recent advances and opportunities to move forward. *Lancet Oncol* 2021;22:e474–87.
5. Loibl S, Poortmans P, Morrow M, Denkert C, Curigliano G. Breast cancer. *Lancet* 2021;397:1750–69.
6. Serrano-Gómez SJ, Fejerman L, Zabaleta J. Breast cancer in Latinas: a focus on intrinsic subtypes distribution. *Cancer Epidemiol Biomarkers Prev* 2018;27:3–10.
7. Tamayo LI, Day-Friedland E, Zavala VA, Marker KM, Fejerman L. Genetic ancestry and breast cancer subtypes in Hispanic/Latina women. In: *Advancing the science in Latinos*. Cham (Switzerland): Springer; 2023.
8. Marker KM, Zavala VA, Vidaurre T, Lott PC, Vázquez JN, Casavila-Zambrano S, et al. Human epidermal growth factor receptor 2-positive breast cancer is associated with indigenous American ancestry in Latin American women. *Cancer Res* 2020;80:1893–901.
9. Zavala VA, Bracci PM, Carethers JM, Carvajal-Carmona L, Coggins NB, Cruz-Correa MR, et al. Cancer health disparities in racial/ethnic minorities in the United States. *Br J Cancer* 2021;124:315–32.
10. Fejerman L, Ramirez AG, Nápoles AM, Gomez SL, Stern MC. Cancer epidemiology in hispanic populations: what have we learned and where do we need to make progress? *Cancer Epidemiol Biomarkers Prev* 2022;31:932–41.
11. Llera AS, Abdelhay ESFW, Arttagaveytia N, Daneri-Navarro A, Müller B, Velazquez C, et al. The transcriptomic portrait of locally advanced breast cancer and its prognostic value in a multi-country cohort of Latin American patients. *Front Oncol* 2022;12:835626.
12. Investigators of the US–Latin America Cancer Research Network. Translational cancer research comes of age in Latin America. *Sci Transl Med* 2015;7:319fs50.
13. Prat A, Pineda E, Adamo B, Galván P, Fernández A, Gaba L, et al. Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast* 2015;24:S26–35.
14. de Almeida LM, Cortés S, Vilensky M, Valenzuela O, Cortes-Sanabria L, de Souza M, et al. Socioeconomic, clinical, and molecular features of breast cancer influence overall survival of Latin American women. *Front Oncol* 2022;12:845527.
15. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 2015;4:7.
16. Mallick S, Li H, Lipson M, Mathieson I, Gymrek M, Racimo F, et al. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* 2016;538:201–6.
17. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, et al. A global reference for human genetic variation. *Nature* 2015;526:68–74.
18. Alexander DH, Lange K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* 2011;12:246.
19. Mangiafico SS. Rcompanion: functions to support extension education program Evaluation. Rutgers cooperative extension, New Brunswick, New Jersey. Version 2.4.36 [Internet]. 2024 [cited 2025 Jan 11]. Available from: <https://CRAN.R-project.org/package=rcompanion/>.
20. Venables WN, Ripley BD. Modern applied statistics with S, 4th ed [Internet]. 2002 [cited 2025 Jan 11]. Available from: <https://www.stats.ox.ac.uk/pub/MASS4/>.
21. Rey-Vargas L, Bejarano-Rivera LM, Mejia-Henao JC, Sua LF, Bastidas-Andrade JF, Ossa CA, et al. Association of genetic ancestry with HER2, GRB7 AND estrogen receptor expression among Colombian women with breast cancer. *Front Oncol* 2022;12:899761.
22. Serrano-Gómez SJ, Sanabria-Salas MC, Garay J, Baddoo MC, Hernández-Suarez G, Mejia JC, et al. Ancestry as a potential modifier of gene expression in breast tumors from Colombian women. *PLoS One* 2017;12:e0183179.
23. Homburger JR, Moreno-Estrada A, Gignoux CR, Nelson D, Sanchez E, Ortiz-Tello P, et al. Genomic insights into the ancestry and demographic history of South America. *PLoS Genet* 2015;11:e1005602.
24. Wang S, Ray N, Rojas W, Parra MV, Bedoya G, Gallo C, et al. Geographic patterns of genome admixture in Latin American Mestizos. *PLoS Genet* 2008;4:e1000037.
25. Sohail M, Palma-Martínez MJ, Chong AY, Quinto-Cortés CD, Barberena-Jonas C, Medina-Muñoz SG, et al. Mexican Biobank advances population and medical genomics of diverse ancestries. *Nature* 2023; 622:775–83.
26. Bonilla C, Bertoni B, Hidalgo PC, Arttagaveytia N, Ackermann E, Barreto I, et al. Breast cancer risk and genetic ancestry: a case-control study in Uruguay. *BMC Womens Health* 2015;15:11.
27. Engmann NJ, Ergas IJ, Yao S, Kwan ML, Roh JM, Ambrosone CB, et al. Genetic ancestry is not associated with breast cancer recurrence or survival in U.S. Latina women enrolled in the Kaiser permanente pathways study. *Cancer Epidemiol Biomarkers Prev* 2017;26:1466–9.
28. Fejerman L, Hu D, Huntsman S, John EM, Stern MC, Haiman CA, et al. Genetic ancestry and risk of mortality among U.S. Latinas with breast cancer. *Cancer Res* 2013;73:7243–53.
29. Martínez-Cortés G, Salazar-Flores J, Fernández-Rodríguez LG, Rubi-Castellanos R, Rodríguez-Loya C, Velarde-Félix JS, et al. Admixture and population structure in Mexican-Mestizos based on paternal lineages. *J Hum Genet* 2012;57:568–74.
30. Zollner L, Torres D, Briceno I, Gilbert M, Torres-Mejía G, Dennis J, et al. Native American ancestry and breast cancer risk in Colombian and Mexican women: ruling out potential confounding through ancestry-informative markers. *Breast Cancer Res* 2023;25:111.
31. Fejerman L, Romieu I, John EM, Lazcano-Ponce E, Huntsman S, Beckman KB, et al. European ancestry is positively associated with breast cancer risk in Mexican women. *Cancer Epidemiol Biomarkers Prev* 2010;19:1074–82.
32. Fejerman L, John EM, Huntsman S, Beckman K, Choudhry S, Perez-Stable E, et al. Genetic ancestry and risk of breast cancer among U.S. Latinas. *Cancer Res* 2008;68:9723–8.
33. Huang X, Lott PC, Hu D, Zavala VA, Jamal ZN, Vidaurre T, et al. Evaluation of multiple breast cancer polygenic risk score panels in women of Latin American Heritage. *Cancer Epidemiol Biomarkers Prev* 2025;34:234–45.
34. Fejerman L, Ahmadiyeh N, Hu D, Huntsman S, Beckman KB, Caswell JL, et al. Genome-wide association study of breast cancer in Latinas identifies novel protective variants on 6q25. *Nat Commun* 2014;5:5260.
35. Fejerman L, Chen GK, Eng C, Huntsman S, Hu D, Williams A, et al. Admixture mapping identifies a locus on 6q25 associated with breast cancer risk in US Latinas. *Hum Mol Genet* 2012;21:1907–17.
36. Su Y, Jiang Y, Sun S, Yin H, Shan M, Tao W, et al. Effects of HER2 genetic polymorphisms on its protein expression in breast cancer. *Cancer Epidemiol* 2015;39:1123–7.
37. Caswell JL, Camarda R, Zhou AY, Huntsman S, Hu D, Brenner SE, et al. Multiple breast cancer risk variants are associated with differential transcript isoform expression in tumors. *Hum Mol Genet* 2015;24:7421–31.
38. LaFramboise T, Weir BA, Zhao X, Beroukhir R, Li C, Harrington D, et al. Allele-specific amplification in cancer revealed by SNP array analysis. *PLoS Comput Biol* 2005;1:e65.
39. Martínez-Nava GA, Urbina-Jara LK, Lira-Albarrán S, Gómez HL, Ruiz-García E, Nieto-Coronel MT, et al. Somatic mutations in Latin American breast cancer patients: a systematic review and meta-analysis. *Diagnostics (Basel)* 2024;14:287.
40. Bhat-Nakshatri P, Gao H, Khatpe AS, Adebayo AK, McGuire PC, Erdogan C, et al. Single-nucleus chromatin accessibility and transcriptomic map of breast tissues of women of diverse genetic ancestry. *Nat Med* 2024;30:3482–94.
41. Gaudet MM, Gierach GL, Carter BD, Luo J, Milne RL, Weiderpass E, et al. Pooled analysis of nine cohorts reveals breast cancer risk factors by tumor molecular subtype. *Cancer Res* 2018;78:6011–21.
42. Holm J, Eriksson L, Ploner A, Eriksson M, Rantalainen M, Li J, et al. Assessment of breast cancer risk factors reveals subtype heterogeneity. *Cancer Res* 2017;77:3708–17.
43. Williams DR, Priest N, Anderson NB. Understanding associations among race, socioeconomic status, and health: patterns and prospects. *Health Psychol* 2016;35:407–11.
44. Werutsky G, Villarreal-Garza C, Gómez H, Donaire JM, Bines J, Fein L, et al. Abstract PO2-10-05: the impact of socioeconomic factors on breast cancer diagnosis in Latin America: the LATINA study (LACOG 0615/MO39485). *Cancer Res* 2024;84:PO2-10-05.
45. Sarma EA, Rendle KA, Kobrin SC. Cancer symptom awareness in the US: Sociodemographic differences in a population-based survey of adults. *Prev Med* 2020;132:106005.