BIODEMOGRAPHY AND SOCIAL BIOLOGY
2021, VOL. 66, NO. 2, 156–171
https://doi.org/10.1080/19485565.2021.1879626

The Chilean socio-ethno-genomic cline

E. Barozet[a], C.Y. Valenzuela[b], L. Cifuentes[b], R.A. Verdugo[b], L. Herrera[b], M. Acuña[b], E. Llop[b], M. Moraga[b], S. Berríos[b], A. Di Genova[c], D. Digman[b], A. Symon[b], S. Asenjo[b], P. López[b], M. L. Bustamante[b,d], P. Pezo-Valderrama[b], J. Suazo[e], F. Caba[f], M. Villalón[g], S. Alvarado[g], D. Cáceres[g], K. Salgado[f], P. Portales[h], N. Loira[c], and A. Maas[c,i]

[a]Departamento de Sociología, Facultad de Ciencias Sociales, Universidad de Chile, Centro de Estudios de Conflicto y Cohesión, Social, Santiago, Chile; [b]Programa de Genética Humana del ICBM, Facultad de Medicina, Universidad de Chile, Santiago, Chile; [c]Mathomics, Centro de Modelamiento Matemático y Centro Para la Regulación del Genoma, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile; [d]Departamento de Psiquiatría y Salud Mental Norte, Facultad de Medicina, Universidad de Chile, Santiago, Chile; [e]Instituto de Investigación en Ciencias Odontológicas, Facultad de Odontología, Universidad de Chile, Santiago, Chile; [f]Escuela de Puericultura, Facultad de Ciencias Médicas, Universidad Bernardo O'Higgins, Chile; [g]Instituto de Salud Poblacional "Escuela de Salud Pública", Universidad de Chile, Santiago, Chile; [h]Corporación Municipal de Desarrollo Social, Iquique, Chile; [i]Departamento de Ingeniería Matemática, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile

ABSTRACT

Studies of the current Chilean population performed using classical genetic markers have established that the Chilean population originated primarily from the admixture of European people, particularly Spaniards, and Amerindians. A socioeconomic-ethno-genetic cline was established soon after the conquest. Spaniards born in Spain or Chile occupied the highest Socioeconomic Strata, while Amerindians belonged to the lowest. The intermediate strata consisted of people with different degrees of ethnic admixture; the larger the European admixture, the higher the Socioeconomic Level. The present study of molecular genomic markers sought to calculate the percentage of Amerindian admixture and revealed a finer distribution of this cline, as well as differences between two Amerindian groups: Aymara and Mapuche. The use of two socioeconomic classifications – Class and Socioeconomic Level – reveals important differences. Furthermore, Self-reported Ethnicity (self-assignment to an ethnic group) and Self-reported Ancestry (self-recognition of Amerindian ancestors) show variations and differing relationships between socioeconomic classifications and genomic Amerindian Admixture. These data constitute a valuable input for the formulation of public healthcare policy and show that the notions of Ethnicity, Socioeconomic Strata and Class should always be a consideration in policy development.

Introduction

The present study uses genomic tools to estimate the autosomal Amerindian-European admixture within the current Chilean population, and its relationship to Self-reported Ethnicity and socioeconomic variables. Chile spans a north-to-south extension of 4,270 km between parallels 17°29′57" S and 56°32′ S. It has an estimated population of

CONTACT E. Barozet ebarozet@uchile.cl Departamento de Sociología, Facultad de Ciencias Sociales, Universidad de Chile, Centro de Estudios de Conflicto y Cohesión, Social, Avda. Capitán Ignacio Carrera Pinto 1045, Ñuñoa, Región Metropolitana, Santiago, Chile.

17.5 million people, almost all of admixed genetic origin. The contemporary Chilean population is the result of the union of two main ethnic components: Spaniards and Amerindians (Encina 1993). Before the arrival of the Spaniards five centuries ago, several Amerindian groups were present in what is now Chile (Thayer-Ojeda 1919). The two main Amerindian ethnic groups today are the Aymara in the north of the country and the Mapuche in the center and south.

The ancestors of the Amerindians migrated from Asia around 35,000 years ago (Raghavan et al. 2015), but the present Chilean population resulted primarily from the admixture of European men and Amerindian women during the Spanish conquest (Valenzuela and Harb 1977). A socio-ethnic-genetic cline was thus established early in the colonization process (Araya and Valenzuela 2010). In the highest socioeconomic strata were Spaniards born in Spain or Chile to a Spanish father and mother. The middle strata comprised several classes with varying proportions of Amerindian and African admixture, with stronger European admixture denoting a higher social stratum. Finally, the lowest socioeconomic strata included Amerindians and the African slaves that were brought to Chile in smaller numbers than to other South American countries (Cussen 2006; Valenzuela, Acuña, and Harb 1987). Later, during the nineteenth and twentieth centuries, European migrants were offered economic and social incentives from the Chilean State (Estrada Turra 2018; Norambuena 1990) and their arrival served to "whiten" the Chilean population, thus reinforcing native Chilean servitude (Bengoa 1996). Finally, prior to the recent surge in immigration from other Caribbean and Latin American countries, African admixture represented only 3% of the global ancestry estimates for the Chilean population. This was due to the presence of Afro-Chileans, especially in the country's northern cities, albeit in smaller proportions than other ethnic groups (Eyheramendy et al. 2015). From a historical point of view, contemporary socioeconomic-ethno-genetic inequality in Chile is a product of both European colonization and the actions of the Chilean State to the present day. In terms of genetics, this socioeconomic-ethno-genetic cline is well documented in Chile (Cifuentes et al. 2004; Pinto-Cisternas et al. 1971; Valenzuela 2011; Valenzuela, Acuña, and Harb 1987) and in other countries (Avena et al. 2012; Beardmore and Karimi-Booshehri 1983).

Using classical genetic markers, the situation can be approximately described as follows. 1) The highest Socioeconomic Stratum comprises 5% of the population with less than 5% autosomal Amerindian admixture; almost 100% European chromosome Y and with a higher proportion of Amerindian mtDNA that we have not yet estimated confidently. 2) The middle Socioeconomic Strata comprise around 20% of the population with around 20–25% autosomal Amerindian admixture, a high proportion of European chromosome Y, and a high proportion of Amerindian mtDNA (Cifuentes et al. 2004; Valenzuela 2011), 3) The lower Socioeconomic Strata comprise 75% of the population with 35–50% autosomal Amerindian admixture, a higher proportion of European chromosome Y (Avena et al. 2012; Vieira-Machado et al. 2016) and Amerindian mtDNA (Rocco et al. 2002). One of the aims of the present study is to update autosomal Amerindian-European admixture estimates using genomic tools.

This socio-genetic cline, as the result of genomic, historical and institutional environmental conditions, has meant that different ethnic groups have different structures of morbidity and mortality, and it would be interesting to establish whether these differences influence these structures within the mixed Chilean population. As we know, each

characteristic, phenotype, behavior or disease, as well as the genome involved in diseases (patho-genome), is determined by interaction between the genome and the environment, social circumstances, and history. Thus, the structure of morbidity and mortality within the population is also determined by this genome-environment interaction. A correct approach to the patho-genome requires first a precise analysis of the distribution of the "non-pathological" genome across socioeconomic strata and ethnic groups, and an accurate estimation of the Amerindian admixture of these different groups and classes. This is the second aim of our research. As an example, we found major differences in skin, hair, and eye color between the higher and lower Socioeconomic Strata of the Chilean population, differences which affect the prevalence of certain skin diseases, particularly skin cancer. The higher strata have lighter skin than the lower strata (Zemelman et al. 2002), and the anatomical distribution of melanoma between the higher and lower strata resembles that found in European and Asian populations, respectively (Zemelman et al. 2006, 2014).

For the most part, socioeconomic stratification has been studied in terms of either Socioeconomic Strata or Classes. In the present study, we take both concepts and analyze their relationships with ethnicity ascertained by self-assignment, and with Amerindian Admixture measured using genomic markers.

Methods

Study Subjects

The present study involved 2,830 unrelated Chilean individuals (46.6% males) recruited for ChileGenomico (Cifuentes 2015; Verdugo et al., 2020), a project that sought to uncover the ethnic ancestry of the Chilean population. Sampling was performed in 2012 and 2013, mainly at public hospital blood banks in cities around the country. In Santiago, the capital, sampling was also conducted at private clinics and during blood drives (Supplementary Table 1). Written informed consent was obtained from each participant. The study's protocols were approved by the Ethics Committees of the University of Chile's Faculty of Medicine and of each hospital or district. Participants donated 5 ml of blood or saliva for DNA extraction – the DNA obtained is the same using both sampling methods and there was no locational or socioeconomic bias in terms of the method used – and completed a socio-demographic questionnaire that included questions about their ancestry and identity. Ethnicity corresponds to self-assignment to a specific ethnic group; Ancestry is based on the declaration of at least one Amerindian ancestor. Respondents also answered questions about their Socioeconomic Strata and Class. Four groups were defined in terms of Ethnicity or Ancestry: Non-Amerindian, Aymara, Mapuche, and Not Answered (NA). It should be noted that despite the presence of other ethnic groups in certain Chilean cities, their small number (less than 0.5% of the sample) would make evaluation impossible using our sampling methods, for which reason they were omitted from the study.

Socioeconomic Level (Stratum) and Class

We used two different variables to measure Socioeconomic Level: Socioeconomic Stratum, widely used in marketing and economic studies; and Class, generally used in sociology.

Table 1. Description of socioeconomic strata and classes.

Socioeconomic level*	Description	Proportion of the population
ABC1	Wealthiest socioeconomic level (A being the elite, B the wealthy, and C1 the upper middle class. A and B are generally not statistically representative as their members tend not to respond to surveys, so they are generally added to C1)	10%
C2	Middle class	20%
C3	Lower middle class	25%
D (35%)	Working class	35%
E (10%)	The poor	10%

* The categorization criteria have been set by marketing companies according to fixed quantiles such that the socioeconomic strata are always of the same size from one measurement to another. It is their characteristics that change according to the evolution of the distribution of the variables within the population.

Socioeconomic Stratum is a five-category variable that ranges from ABC1 (social groups with the highest levels of consumption and education) to E (the poor), as described in Table 1. The variable functions as an index, based on a selection of household goods available – ranging from *common* (cheap and readily-available) to *scarce* (expensive and hard-to-find) – and the educational level of the head of the household (Asociación Internacional de Marketing 2008). This index, which was still in use among marketing firms and researchers when we applied our survey in 2011, was replaced in 2015 by a new instrument based on household income and size. One of the main advantages of the old index is the fact that any member of the household should be able to answer the questions.

The other variable we used as a proxy for social status is Social Class, as defined in the Erikson, Goldthorpe and Portocarrero (EGP) class scheme (Erikson and Goldthorpe 1992). This is a more complex measure involving three variables: ISCO-88, size of the firm or institution that employs the person, and whether the person is a supervisor or supervisee. ISCO-88 is the International Standard Classification of Occupations as defined by the International Labor Organization, which classifies 390 jobs based on a description of the formal training and skills required. We used the version employed for the generation of Chilean official statistics. The job held by the head of the household is classified according to a combination of the other two variables into either eleven or seven classes, depending on the research goals. We used the seven-category EGP scheme that has been adapted to Chile and used by a number of studies, as described in Table 2 (Espinoza, Barozet, and Méndez 2013; Espinoza and Núñez 2014; Torche and Wormald 2004).

In the recent debate concerning social stratification, both instruments are widely used to define Socioeconomic Level, although they measure two different dimensions – consumption

Table 2. Description of socioeconomic classes (classes).

Number	Description	Proportion of the Chilean population (ENES, 2009)
1	Service class: professionals, administrators, managers, and higher-professional or technical occupations	25.8%
2	Non-manual routine employees	10.9%
3	Petty bourgeoisie: small proprietors, self-employed and small employers	17.1%
4	Farmers and smallholders	6.3%
5	Lower grade technical occupations, supervisors of manual workers and skilled workers	14.2%
6	Semi-skilled and unskilled workers	19.7%
7	Agricultural workers	6.1%

capacity (which defines lifestyles) for Socioeconomic Strata and social status associated with work for Class respectively – and concern the head of the household of the person surveyed in this study.

Amerindian-European Genomic Admixture

All DNA samples were genotyped for 147 ancestry-informative autosomal SNPs (Supplementary Table 2) selected for their ability to estimate continental African, Amerindian, and European proportions (Project FONDEF 2018). Most of these ancestry-informative SNPs (141 of them) were chosen because they exhibited large differences in allele frequency among reference populations (30 European and 30 African individuals from the 1000 Genomes Project; 17 Chilean with high Aymara ancestry and 31 Chilean with high Mapuche ancestry), as well as maximum genome coverage. To this set we added 6 SNPs found in the 48 Chilean individuals with high Amerindian ancestry and whose frequency was not known in African and European reference populations at the time of the analysis. Genotyping was performed by LGC Genomics Ltd. (Middlesex, UK). The ADMIXTURE software (Alexander, Novembre, and Lange 2009) was used to infer the percentage of Amerindian, European, and African ancestry of each individual by comparing their genotypes with the genomes of reference populations: 30 European and 30 African individuals from the 1000 Genomes Project, 30 individuals self-assigned to the Aymara community from Puno in Perú (DNA supplied by Andrés Moreno of Stanford University), and 30 individuals with Mapuche ancestry from the Pehuenche and Huilliche communities in southern Chile (De la Fuente et al. 2018; Llop 1996).

Statistics

ANOVAs were performed to compare Genomic Amerindian Admixture (AM-AD) among different categories of Class, Socioeconomic Strata, Self-reported Ethnicity, and Self-reported Ancestry. Following these analyses, two-way ANOVAs were used to study the significance of associations with three variables (Supplementary File). More complex analyses were considered unnecessary. The Chi Square test was used to explore associations between categorical variables. Z- and t-tests were used to compare two means. Statistical analyses were conducted using STATA 13.0.

Results

Classes and Socioeconomic Strata

The distribution of the seven Classes by Socioeconomic Stratum is presented in Table 3; as expected, the positive association was highly significant ($P < .001$). It is remarkable that the distribution of *Farmers and Smallholders* (class 4) in the Socioeconomic Strata is more similar to that of *Petty Bourgeoisie* (class 3) than to *Non-Manual Routine* (class 2). According to Socioeconomic Strata and considering the extreme levels ABC1 and D + E, the EGP Classes are clustered in five distinct socioeconomic groups: I) Class 1 with 64% in ABC1 and 0.8% in D + E; II) Classes 2 + 3 + 4 with approximately 18% in ABC1 and approximately 12% in D + E; III) Class 5 with approximately 6% in ABC1 and 12% in D + E

Table 3. Contingency table for socioeconomic stratum and class.

	Class	Socioeconomic Stratum												
		ABC1		C2		C3		D		E		NA		
		N	%	N	%	N	%	N	%	N	%	N	%	Total
1:	Service Class	646	64.3	305	30.4	41	4.1	8	0.8	0	0.0	4	0.4	1004
2:	Non-Manual Routine	67	18.5	150	41.3	115	31.7	25	6.9	3	0.8	3	0.8	363
3:	Petty Bourgeoisie	49	15.2	117	36.2	102	31.6	45	13.9	2	0.6	8	2.5	323
4:	Farmers and Smallholders	7	21.9	12	37.5	7	21.9	5	15.6	0	0.0	1	3.1	32
5:	Lower Grade Technical and Skilled Workers	21	6.3	154	46.0	111	33.1	36	10.7	5	1.5	8	2.4	335
6:	Semi-Skilled and Unskilled Workers	21	4.2	176	35.4	167	33.6	119	23.9	12	2.4	2	0.4	497
7:	Agricultural Workers	0	0.0	8	5.8	47	33.8	74	53.2	9	6.5	1	0.7	139
	NA	16	11.7	48	35.0	34	24.8	25	18.2	9	6.6	5	3.6	137
	Total	827		970		624		337		40		32		2830

(concentrated in C2 and C3, 79%); IV) Class 6 with 4% in ABC1 and approximately 26% in D + E; and V) Class 7 with 0% in ABC1 and approximately 60% in D + E.

Ethnicity (Self-assignment to an Amerindian Group), Ancestry (Self-recognition of Amerindian Ancestors) and Genomic Amerindian Admixture (AM-AD)

The distribution of the sample according to the two meaningful ethnic groups of Self-reported Ethnicity and Self-reported Ancestry is presented in Table 4. As expected, the distribution was far from random ($P < .001$): even though the classifications differ, there is a strong degree of coincidence, particularly among those people classified as Amerindian or non-Amerindian by both criteria. It is remarkable that there was not a total coincidence in the Amerindian category as was expected (less than 67%; 35.6% of those who declare themselves to be a member of an Amerindian ethnic group do not report having an Amerindian ancestor). Overall genomic ancestry percentages were 44.04 for Amerindian, 50.61 for European, and 4.39 for African, with standard errors of 0.27%, 0.26% and 0.07%, respectively. The distribution of the Mean Genomic Amerindian Admixture (AM-AD), Self-reported Ancestry, and Self-reported Ethnicity is shown in Figures 1 and Figures 2, and in Table 5. Figures 1 and Figures 2 show a cline in AM-AD: Aymara > Mapuche > NA > Non-Amerindian (Self-reported Ancestry), and Aymara > Mapuche > Non-Amerindian ≥ NA (Self-reported Ethnicity). The differences are highly significant, as indicated by the standard errors and ANOVA ($P < .001$) (Supplementary File). The large percentage of AM-AD (approximately 40%) of the Non-Amerindian group is remarkable. Table 5 shows the mean AM-AD found in groups classified according to Self-reported Ancestry and Self-reported Ethnicity; numbers are smaller than those of Table 4 because a few individuals could not be typed using genomic techniques. Individuals self-classified as having Amerindian Ancestry and Amerindian Ethnicity had 63.5% AM-AD; those self-classified as having Amerindian Ancestry but Non-Amerindian Ethnicity had a lower AM-AD of 50.5% (z-test $P < 10^{-6}$). Individuals classified as Non-Amerindian by Self-reported Ancestry, but Amerindian by Self-reported Ethnicity had 42.0% AM-AD, while those classified as Non-Amerindian by Self-reported Ethnicity had 40.8% AD-AM – a non-significant difference ($P = .153$). However, individuals classified by Self-reported Ethnicity as either Amerindian or Non-Amerindian had a much (and significantly) lower percentage of AM-AD when classified as Non-Amerindian according to Self-reported Ancestry. Concordantly, the ANOVA for percentage of AM-AD yielded highly significant values for Self-reported Ancestry ($P < 10^{-5}$), Self-reported Ethnicity ($P < 10^{-6}$), and interaction between Self-reported Ancestry and Self-reported Ethnicity ($P < 10^{-9}$) (Supplementary File). The high percentage of AM-AD in the Non-Amerindian group self-classified according to Ethnicity and Ancestry (40.8%) is noteworthy.

Table 4. Contingency table for self-reported ethnicity and ancestry.

	Self-reported Ethnicity				
	Amerindian		Non-Amerindian		Total
Self-reported Ancestry	N	%	N	%	N
Amerindian	316	66.8	157	33.2	473
Non-Amerindian	175	8.2	1955	91.8	2130
Total	491	18.9	2112	81.1	2603

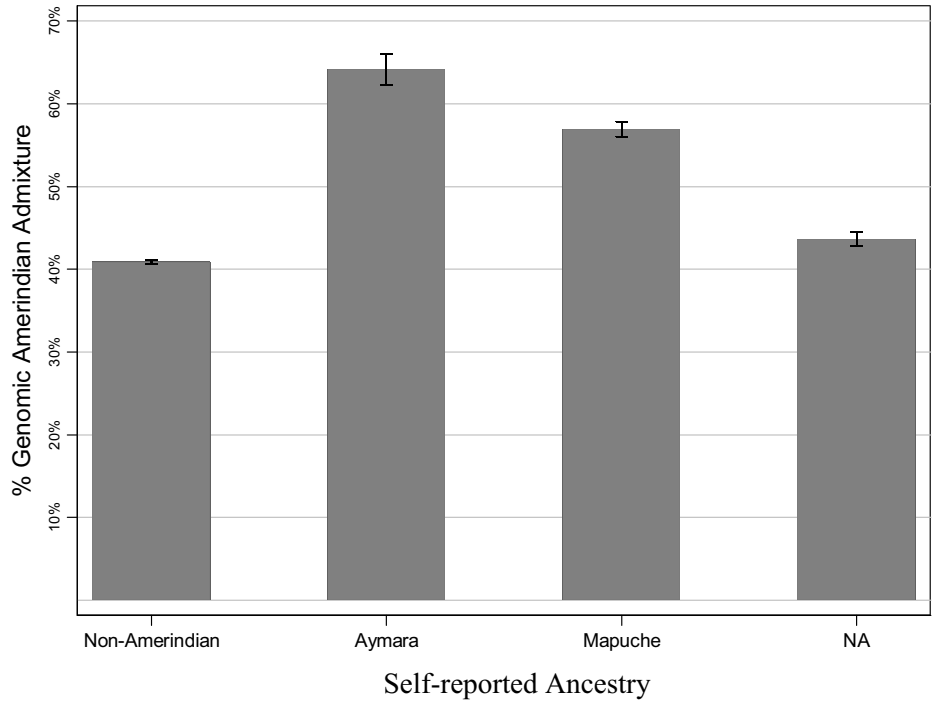

Figure 1. Mean Genomic Amerindian Admixture (AM-AD) with Standard Errors by Self-reported Ancestry nc.docx.

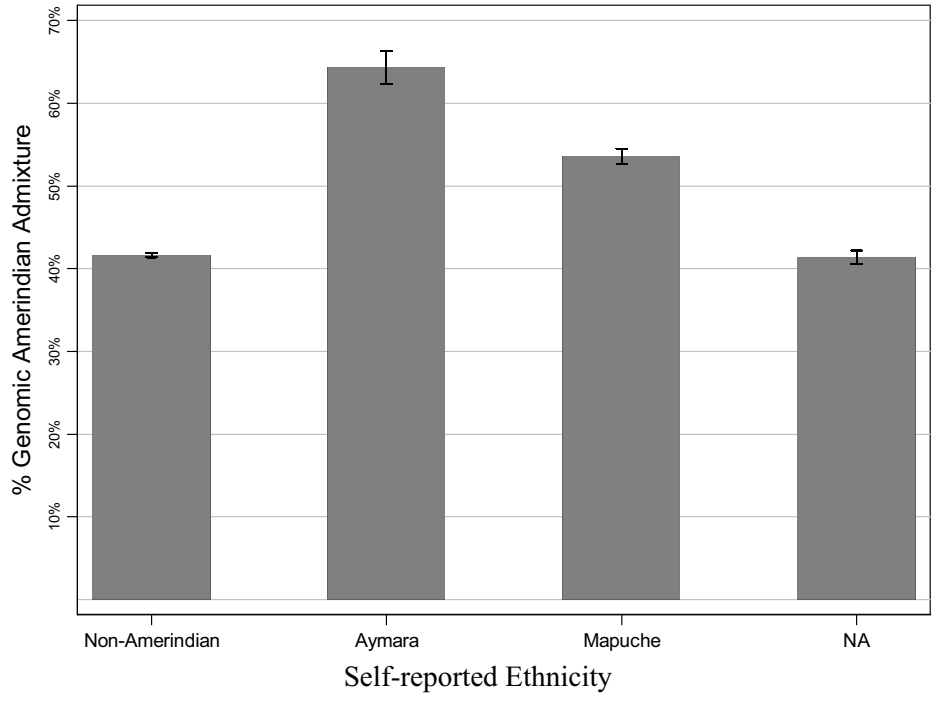

Figure 2. Mean Genomic Amerindian Admixture (AM-AD) with Standard Errors by Self-reported Ethnicity nc.docx.

Table 5. Mean genomic amerindian admixture (AM-AD) by self-reported ancestry and ethnicity.

	Self-reported Ethnicity		
Self-reported Ancestry	Amerindian	Non-Amerindian	Total
Amerindian			
N	305	156	461
Mean	63.5	50.5	57.1
SD	17.8	14.9	18.0
Non-Amerindian			
N	173	1955	2130
Mean	42.0	40.8	40.9
SD	10.5	11.3	11.2
Total			
N	478	2112	2589
Mean	55.5	41.5	44.1
SD	18.7	11.9	14.5

N = number of individuals; Mean = mean % of AM-AD; SD = standard deviation of AM-AD.

Table 6. Contingency table for self-reported ancestry and socioeconomic stratum.

	Self-reported Ancestry								
	Aymara		Mapuche		Non-Amerindian		NA		Total
Socioeconomic Stratum	N	%	N	%	N	%	N	%	N
ABC1	23	2.8	61	7.4	678	82.6	59	7.2	821
C2	47	4.9	108	11.2	756	78.6	51	5.3	962
C3	26	4.2	94	15.3	468	76.0	28	4.5	616
D	12	3.6	72	21.6	235	70.6	14	4.2	333
E	4	10.0	7	17.5	27	67.5	2	5.0	40
NA	4	14.3	4	14.3	17	60.7	3	10.7	28
Total	116		346		2181		157		2800

Percentages represent row proportions.

Table 7. Contingency table for self-reported ancestry and class.

		Self-reported Ancestry								
		Aymara		Mapuche		Non-Amerindian		NA		Total
Class		N	%	N	%	N	%	N	%	N
1:	Service Class	34	3.4	91	9.2	798	80.4	70	7.0	993
2:	Non-Manual Routine	23	6.4	38	10.7	285	79.8	11	3.1	357
3:	Petty Bourgeoisie	15	4.7	43	13.3	245	76.1	19	5.9	322
4:	Farmers and Smallholders	1	3.3	1	3.3	24	80.0	4	13.4	30
5:	Lower Grade Technical and Skilled Workers	12	3.6	46	13.9	254	77.0	18	5.5	330
6:	Semi-Skilled and Non-Skilled Workers	15	3.0	76	15.4	383	77.4	21	4.2	495
7:	Agricultural Workers	4	2.9	34	24.5	94	67.6	7	5.0	139
	NA	12	9.0	17	12.7	98	73.1	7	5.2	134
	Total	116		346		2181		157		2800

Percentages represent row proportions.

Table 8. Mean genomic amerindian admixture (AM-AD) with standard errors in 7 levels of class.

Class	Mean %	Standard Error %
Service Class	39.84	0.44
Non-Manual Routine	46.14	0.76
Petty Bourgeoise	45.10	0.75
Farmers and Smallholders	41.56	2.29
Lower Grade Technicians and Skilled Workers	46.46	0.72
Semi-Skilled and Non-Skilled Workers	46.54	0.60
Agricultural Workers	50.08	1.31
NA	46.19	1.23

Table 9. Mean genomic amerindian admixture (AM-AD) with standard errors in 5 levels of socioeconomic stratum.

Socioeconomic Stratum	Mean %	Standard Error %
ABC1	38.70	0.48
C2	44.81	0.42
C3	46.15	0.54
D	49.43	0.84
E	51.61	2.60
NA	51.08	2.90

Self-reported Ancestry, Socioeconomic Stratum, and Class

Tables 6 and Tables 7 describe the highly significant associations ($P < .001$) between Self-reported Ancestry and Socioeconomic Stratum and Class, respectively. In terms of Socioeconomic Strata, individuals with Mapuche and Aymara Self-reported Ancestry are less strongly represented in ABC1 than Chileans with no AM-AD. However, those with Aymara Self-reported Ancestry have a stronger presence in C2, while those with Mapuche Self-reported Ancestry belong predominantly to C3 and D. In spite of the NA (not answered) groups in both tables, there was a tendency for Self-reported Amerindian Ancestry to increase and Self-reported non-Amerindian Ancestry to decrease as the Socioeconomic Stratum or Class decreased.

Genomic Amerindian Admixture (AM-AD), Class, and Socioeconomic Stratum

This relationship is presented in Table 8 (Class) and 9 (Social Stratum). There is a general tendency for mean AM-AD to increase as Class decreases, except in the cases of Class 4 and NA. Mean AM-AD increases steadily as Socioeconomic Stratum decreases. In each table, statistical significance is illustrated by the standard error.

A two-way ANOVA (Supplementary File) was conducted in order to study this association (see also previous Figures and Tables). It showed that the association of AM-AD with Socioeconomic Stratum was highly significant ($P < .0001$, Table 9); however, the association of AM-AD with Class and the interaction between these social factors were not significant ($P = .207$ and $P = .223$, respectively; compare Tables 8 and Tables 9). The lack of significance of the association between Class and AM-AD is mostly due to the low AM-AD of Class 4 (*Farmers and Smallholders*), the intermediate value of Class 3 (*Petty Bourgeoisie*) and NA

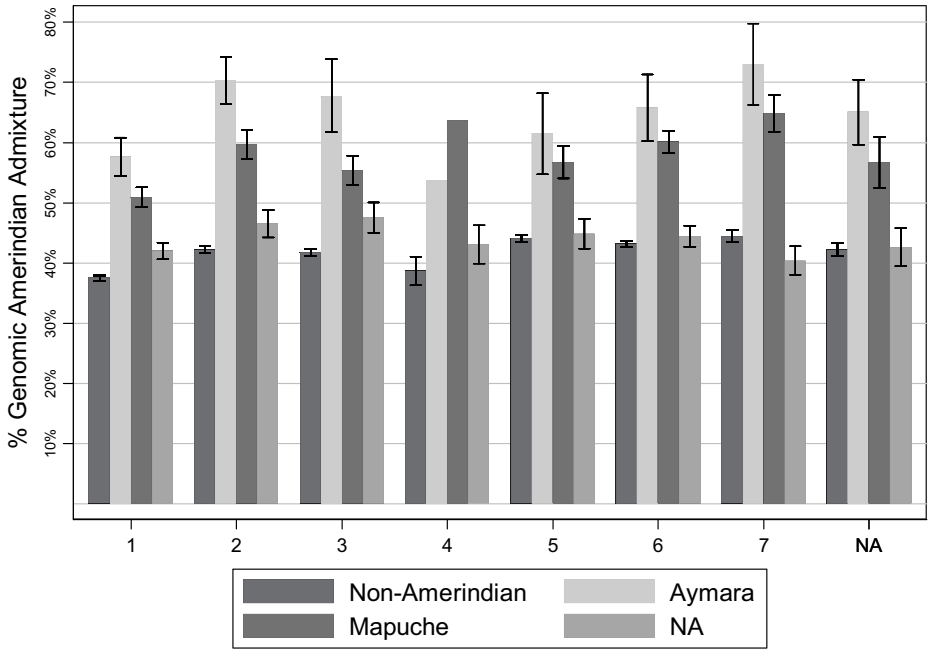

Figure 3. Mean Genomic Amerindian Admixture (AM-AD) with Standard Errors in 7 Classes by Self-reported Ancestry nc.docx

(not answered), and because the relationship between Class and Socioeconomic Stratum is not consistently linear (Table 3 and its analysis), even though they are positively associated as a whole ($P < .001$). Nevertheless, the percentage of AM-AD increases steadily as the order of the Socioeconomic Strata decreases, thus the association is between higher strata or classes and lower AM-AD values.

Self-reported Ethnicity, Self-reported Ancestry, Socioeconomic Strata, Class, and Genomic Amerindian Admixture (AM-AD): Significance Analyses

The effect of the interaction between Socioeconomic Stratum and Self-reported Ethnicity or Self-reported Ancestry on AM-AD was studied using two-factor ANOVA (Supplementary File). The model indicated a significant effect of Ethnicity and Strata ($P < .001$ in both cases), and a significant effect of the interaction ($P < .001$). The same analysis was done for Class (7 classes) and Self-reported Ethnicity; the interactive model shows strong significances: $P < .001$ for Class, $P < .001$ for Ethnicity, and $P = .0018$ for the Ethnicity-Class interaction.

The Aymara and Mapuche ethnic groups were sufficiently numerous to enable analysis of AM-AD according to Self-reported Ancestry and Class. This analysis is presented in Figure 3, where within each Class, the percentage of AM-AD is given for the four Self-reported Ancestry groups. Analysis found a higher AM-AD in the Aymara ethnic group than in the Mapuche group in all Classes except Class 4 (*Farmers and Smallholders*). An important proportion of AM-AD in the Non-Amerindian group (Chileans without known Amerindian descent) is also evident in all Classes. It is remarkable that the proportion of

AM-AD in the Non-Amerindian and NA (not answered according to Amerindian component) groups is distributed differently in two sections of the class scale. Between *Service class* and *Farmers and Smallholders* (classes 1 to 4, the highest classes), the AM-AD was higher in the NA group than in the Non-Amerindian group. By contrast, between *Lower Grade Technical and Skilled Workers* (class 5) and NA (not answered according to Class) (the lowest classes), the AM-AD in NA was equal to or less than in the Non-Amerindian group.

Discussion and Conclusions

The genomic ancestry (AM-AD) estimates produced by the present study are in line with those obtained in other recent studies. Estimations made by Eyheramendy et al. (2015) are 44.7%, 52.3%, and 3.0% for Amerindian, European, and African genomic admixture, respectively. Those obtained by Bermejo et al. (2017) are 40%, 49%, and 3%, respectively. The present study is also consistent with previous research (Cifuentes 2015; Beardmore and Karimi-Booshehri 1983; Bermejo et al. 2017; Eyheramendy et al. 2015; Raghavan et al. 2015; Rocco et al. 2002; Valenzuela 2011; Valenzuela and Harb 1977; Vieira-Machado et al. 2016) regarding a number of traits of the Chilean population, and offers clearer findings on some points. I) The general Chilean population has strong Amerindian ancestry of almost 40%. II) The higher the Socioeconomic Stratum or Class, the lower its Amerindian Admixture. III) Self-reported Ancestry and Self-reported Ethnicity are good predictors of Genomic Amerindian Admixture, although Self-reported Amerindian ancestry is a better marker of Amerindian Admixture than Self-reported Ethnicity is. According to these results, Genomic Amerindian Admixture correlates better with Socioeconomic Stratum than with Class. All inter-factor interactions were highly significant.

The association we found between Socioeconomic Stratum and Amerindian Admixture is not as extreme as that found by Valenzuela, Acuña, and Harb (1987), who used another social classification in which the highest stratum included categories 1 to 3 of a 13-category socioeconomic classification, or categories 1 to 2 of a 5-category classification. In our sample, few or no Amerindian individuals belong to the highest strata. Despite this difference, a strong correlation between Class or Socioeconomic Stratum and AM-AD was found in line with previous and historical studies (Eyheramendy et al. 2015; Raghavan et al. 2015; Thayer-Ojeda 1919; Valenzuela, Acuña, and Harb 1987).

The correlation between Socioeconomic Stratum and the percentage of Genomic Amerindian Admixture has been reported in other Latin-American countries, such as Argentina (Avena et al. 2012), Colombia (Campbell et al. 2012), and Uruguay (Bonilla et al. 2015). Spaniards from Spain or "Criollos" born in Chile to Spanish parents occupied the upper Socioeconomic Strata or intellectual occupations, leaving ethnically mixed people or Amerindians to manual work or the low Socioeconomic Strata, and Africans to slavery.

The relationship between socioeconomic stratum, ancestry and genome, and their incidence in public health policies can be seen in a simple example. In the case of COVID-19, individuals of blood type A are more susceptible to infection, while individuals of blood type O are more resistant and suffer milder symptoms (Wu et al. 2020). However, due to socioeconomic, cultural, and ethnic discrimination, the prevalence of blood type A in the upper Socioeconomic Strata has remained close to that of the Spaniards, while prevalence in the *mestizo* population is approximately that expected as a result of the half admixture of Amerindians, whose blood type was exclusively O. As such, although people of

lower socioeconomic status have a lower prevalence of blood type A, meaning that they could be at less risk of contracting COVID-19, their poorer living conditions place them at greater risk. Future studies should address the different levels of susceptibility determined by both blood type and socioeconomic conditions. Specifically, these relationships must be taken into consideration in the design of health programs, as there is a considerable difference between severe cases of the disease resulting from poor economic conditions and those that are due to a particular ABO phenotype. Since several genetic systems may be associated with different degrees of susceptibility to disease but differently associated with socioeconomic conditions, it is not possible at present to propose guidance regarding health policies. In-depth research is required on this subject.

These correlations have been described in many contemporary populations in the context of etiological research and management of health problems, classifying individuals based on their Self-reported Ethnicity (see Almeida-Filho et al. 2004 for the case of Brazil). As use of EGP and Socioeconomic Strata is relatively recent in Latin America, the present research could constitute the basis of an important debate regarding socio-ethnic-economic class inequality. Furthermore, it may contribute to the broader literature concerning ethnic-racial disparities and public health in Latin America, comparing countries with different admixture, socioeconomic conditions, and health systems.

Nevertheless, we found that, at least in the Chilean population, having Amerindian ancestors (Self-reported Ancestry) is a more representative criterion than Self-reported Ethnicity, given its stronger correlation with the Genomic Amerindian Admixture. It should be remembered that in Chile, socioeconomic variables are always related to Amerindian admixture, and Amerindians are heterogeneous socio-cultural-genomic ethnic groups.

Regarding Self-reported Ancestry and Self-reported Ethnicity, in the higher classes, the AM-AD was higher in NA than in the Non-Amerindian group, while in the lower classes the percentage in NA was equal to or less than in the Non-Amerindian group. The fact that some individuals with even recent Amerindian ancestors do not consider themselves as Amerindian may be unexpected; however, this is a constant and well-known feature of Chilean culture. On the other hand, a person without recent Amerindian ancestors may assume Amerindian ethnicity based on the mixed composition of the Chilean population as a whole, or as a result of ideological notions associated with a new movement of indigenous "pride". This proposal is limited but may explain some such cases. In Chile, the word "Indian" has negative cultural connotations, and the fact we were unable to identify this group may be due to the small proportion of the population that conserve the ancestral memory of belonging to Amerindian groups.

Additionally, Self-reported Ancestry and Self-reported Ethnicity are good predictors of Amerindian Admixture, although the former is a more accurate indicator. However, Self-reported Ethnicity may provide additional information regarding genomic ancestry, which may in turn be useful in the study of disease. It is also likely that Self-reported Ethnicity provides a more accurate reflection of how people have been treated throughout their lives, as well as how they feel and the nature of their cultural heritage. By contrast, Self-reported Ancestry provides a better indication of genomic admixture. The two indicators are closely related but reflect different perspectives of an individual's current situation. They may also have contrasting implications regarding healthcare, especially if Self-reported Ethnicity is conditioned by cultural and/or discriminatory experiences.

A number of limitations to the study must be acknowledged. Firstly, the sample is not representative of the Chilean population as a whole and does not apply to all of the country's regions. However, this does not invalidate the association between variables. Secondly, our study was restricted to nuclear markers, and we know that mtDNA and markers on the Y chromosome would yield different results, as noted in the introduction. Thirdly, when comparing Self-reported Ancestry and Self-reported Ethnicity, only two ethnic groups (Mapuche and Aymara, the most prominent in the general genetic profile of the Chilean population) were considered in the analysis. This was because other indigenous subgroups were barely represented in our sample. Finally, extreme social strata and classes, especially the higher ones, were not included, as members seldom respond to surveys.

Acknowledgments

We are grateful to Lafayette Eaton and Paul Salter who improved both the English style and genetic conceptualizations. We are also grateful to Cristian Garay. Emmanuelle Barozet acknowledges the support of the INCASI network (Horizon 2020, Research and innovation programme, Marie Skłodowska-Curie GA No 691004 coordinated by Dr. Pedro López-Roldán).

This work was supported by project CONICYT/FONDEF under Grant D10I1007 and project ANID/FONDAP under Grant 15130009.

ORCID

F. Caba http://orcid.org/0000-0001-6852-3527

References

Alexander, D. H., J. Novembre, and K. Lange. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* 19 (9):1655–64. doi:10.1101/gr.094052.109.

Almeida-Filho, N., I. Lessa, L. Magalhaes, M. J. Araujo, E. Aquino, J. A. Sherman, and I. Kawachi. 2004. Social inequality and depressive disorders in Bahia, Brazil: Interactions of gender, ethnicity, and social class. *Social Science & Medicine* 59:1339–53. doi:10.1016/j.socscimed.2003.11.037.

Anillo SOC-12. 2009. Encuesta Nacional de Estratificación y Movilidad Social. CONICYT: Santiago de Chile.

Araya, A., and J. Valenzuela, eds. 2010. *América Colonial. Denominaciones, clasificaciones e identidades*. Santiago, Chile: RIL.

Asociación Internacional de Marketing. 2008. Grupos socioeconómicos en Chile. Accessed March 20, 2018. http://www.aimchile.cl/wp-content/uploads/2011/12/Grupos_Socioeconomicos_AIM-2008.pdf

Avena, S., M. Via, E. Ziv, E. J. Pérez-Stable, C. R. Gignoux, C. Dejean, S. Huntsman, Torres-Mejía G, Dutil J, Matta JL, et al. 2012. Heterogeneity in genetic admixture across different regions of Argentina. *PLoS One*. 7(4):e34695. doi:10.1371/journal.pone.0034695.

Beardmore, J. A., and F. Karimi-Booshehri. 1983. ABO genes are differentially distributed in socio-economic groups in England. *Nature* 303 (5917):522–24. doi:10.1038/303522a0.

Bengoa, J. 1996. *Historia del Pueblo Mapuche*. Santiago, Chile: SUR ediciones.

Bermejo, J.L., F. Boekstegers, R. González Silos, K. Marcelain, P. Baez Benavides, C. Barahona Ponce, B. Müller, C. Ferreccio, J. Koshiol, C. Fischer, et al. May 25 2017. Subtypes of Native American ancestry and leading causes of death: Mapuche ancestry-specific associations with gallbladder cancer risk in Chile. *PLoS Genetics* 13 (5).

Bonilla, C., B. Bertoni, P. C. Hidalgo, N. Artagaveytia, E. Ackermann, I. Barreto, P. Cancela, M. Cappetta, A. Egaña, G. Figueiro, et al. 2015. Breast cancer risk and genetic ancestry: A case-control study in Uruguay. *BMC Women's Health* 15:11. doi:10.1186/s12905-015-0171-8.

Campbell, D. D., M. V. Parra, C. Duque, N. Gallego, L. Franco, A. Tandon, T. Hünemeier, A. Villegas, G. Bedoya, M. I. McCarthy, et al. 2012. Amerind ancestry, socioeconomic status and the genetics of type 2 diabetes in a Colombian population. *PLoS One* 7 (4). doi:10.1371/journal.pone.0033570.

Cifuentes, L., R. Morales, D. Sepúlveda, H. Jorquera, and M. Acuña. 2004. DYS19 and DYS199 loci in a Chilean population of mixed ancestry. *American Journal of Physical Anthropology* 125 (1):85–89. doi:10.1002/ajpa.10380.

Cifuentes, L. 2015. Proyecto ChileGenómico: tras la diversidad genética de los chilenos. Bol Acad Chil Med. LII:13–36.

Cussen, C. 2006. El paso de los negros por la historia de Chile. *Cuadernos de Historia, Universidad de Chile* 25:45–58.

De la Fuente, C., M. C. Ávila-Arcos, J. Galimany, M. L. Carpenter, J. R. Homburger, A. Blanco, P. Contreras, D. Cruz Dávalos, O. Reyes, M. San Roman, et al. April, 2018. Genomic insights into the origin and diversification of late maritime hunter-gatherers from the Chilean Patagonia. *Proceedings of the National Academy of Sciences of the United States of America* 24 (115):17.

Encina, F. 1993. *Historia de Chile. Chapter III*. Santiago de Chile: Sociedad Editora Revista Ercilla.

Erikson, R., and J. H. Goldthorpe. 1992. *The constant flux: A study of class mobility in industrial societies*. USA: Oxford University Press.

Espinoza, V., E. Barozet, and M. L. Méndez. 2013. Estratificación y movilidad social bajo un modelo neoliberal: El caso de Chile. *Revista Lavboratorio* 25 (14):169–91.

Espinoza, V., and J. Núñez. 2014. Movilidad ocupacional en Chile 2001-2009. ¿Desigualdad de ingresos con igualdad de oportunidades? *Revista Internacional de Sociolgía* 72 (1):57–82. doi:10.3989/ris.2011.11.08.

Estrada Turra, B. 2018. *La azarosa y difícil aventura de migrar: Españoles en Chile (1880-1950)*. Santiago de Chile: RIL.

Eyheramendy, S., F. I. Martinez, F. Manevy, C. Vial, and G. M. Repetto. 2015. Genetic structure characterization of Chileans reflects historical immigration patterns. *Nature Communications* 6:6472. doi:10.1038/ncomms7472.

Llop, E. 1996. Genetic composition of Chilean aboriginal populations: HLA and other genetic marker variation. *American Journal of Physical Anthropology* 101 (3):325–32. doi:10.1002/(SICI)1096-8644(199611)101:3<325::AID-AJPA2>3.0.CO;2-Y.

Norambuena, C. 1990. *Política y legislación inmigratoria en Chile: 1830-1930*. Santiago, Chile: Universidad de Santiago de Chile.

Pinto-Cisternas, J., H. Figueroa, B. Lazo, C. Salinas, and C. Campusano. 1971. Genetic structure of the population of Valparaiso. V. ABO blood groups, color vision deficiency and their relationship to other variables. *Human Heredity* 21 (5):431–39. doi:10.1159/000152436.

Project FONDEF 2018. Genómica de la Población Chilena: Obtención de Perfiles Genéticos necesarios en Investigación Clínica, Salud Pública y Medicina Forense. www.chilegenomico.cl (accesed March 20 2018).

Raghavan, M., M. Steinrücken, K. Harris, S. Schiffels, S. Rasmussen, M. DeGiorgio, A. Albrechtsen, C. Valdiosera, M. C. Avila-Arcos, A.-S. Malaspinas, et al. 2015. Population Genetics. Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science*. 349 (6250):aab3884. doi:10.1126/science.aab3884.

Rocco, P., C. Morales, M. Moraga, J. F. Miquel, F. Nervi, E. Llop, P. Carvallo, and F. Rothhammer. 2002. Genetic composition of the Chilean population. Analysis of mitochondrial DNA polymorphism. *Revista Médica de Chile* 130 (2):125–31.

Thayer-Ojeda, L. 1919. *Elementos étnicos que han intervenido en la población de Chile*. Santiago, Chile: Imprenta La Ilustración.

Torche, F., and G. Wormald. 2004. *Estratificación y movilidad social en Chile: Entre la adscripción y el logro*. Santiago de Chile: United Nations Publications.

Valenzuela, C. Y. 2011. Human sociogenetics. *Biological Research* 44 (4):393–404. doi:10.4067/S0716-97602011000400012.

Valenzuela, C. Y., M. P. Acuña, and Z. Harb. 1987. Sociogenetic gradient in the Chilean population. *Revista Médica de Chile* 115 (4):295–99.

Valenzuela, C. Y., and Z. Harb. 1977. Socioeconomic assortative mating in Santiago, Chile: A demonstration using stochastic matrices of mother-child relationships applied to ABO blood groups. *Social Biology* 24 (3):225–33. doi:10.1080/19485565.1977.9988285.

Verdugo, R., A. Di Genova, L. Herrera, M. Moraga, M. Acuña, S. Berríos, E. Llop, C. Valenzuela, L. Bustamante, D. Digman, et al.2020. Development of a small panel of SNPs to infer ancestry in Chileans that distinguishes Aymara and Mapuche components. *Biological Research*. 53:15.

Vieira-Machado, C. D., M. Tostes, G. Alves, J. Nazer, L. Martinez, E. Wettig, O. Pizarro, M. Diaz Caamaño, J. Larenas Ascui, P. Pavez, et al. 2016. Uniparental ancestry markers in Chilean populations. *Genetics and Molecular Biology*. 39(4):573–79. doi:10.1590/1678-4685-gmb-2015-0273.

Wu, Y., Z. Feng, P. Li, and Q. Yu. Oct 2020. Relationship between ABO blood group distribution and clinical characteristics in patients with COVID-19. *Clinica Chimica Acta* 509:220–23. doi: 10.1016/j.cca.2020.06.026.

Zemelman, V., P. von Beck, O. Alvarado, and C. Y. Valenzuela. 2002. Sexual dimorphism in skin, eye and hair color and the presence of freckles in Chilean teenagers from two socioeconomic strata. *Revista Médica Chilena* 130 (8):879–84.

Zemelman, V. B., C. Y. Valenzuela, I. Sazunic, and I. Araya. 2014. Malignant melanoma in Chile: Different site distribution between private and state patients. *Biological Research* 47:34. doi:10.1186/0717-6287-47-34.

Zemelman, V. B., J. Roa, S. R. Tagle, and C. Y. Valenzuela. 2006. Malignant melanoma in Chile: An unusual distribution of primary sites in men from low socioeconomic strata. *Clinical and Experimental Dermatology* 31 (3):335–38. doi:10.1111/j.1365-2230.2005.02038.x.