



Review Article
Young Latin American Geneticists - Special Issue

Genetic associations with disease in populations with Indigenous American ancestries

Lucas Vicuña¹ 

¹University of Chicago, Department of Medicine, Section of Genetic Medicine, Chicago, USA.

Abstract

The genetic architecture of complex diseases affecting populations with Indigenous American ancestries is poorly understood due to their underrepresentation in genomics studies. While most of the genetic diversity associated with disease trait variation is shared among worldwide populations, a fraction of this component is expected to be unique to each continental group, including Indigenous Americans. Here, I describe the current state of knowledge from genome-wide association studies on Indigenous populations, as well as non-Indigenous populations with partial Indigenous ancestries from the American continent, focusing on disease susceptibility and anthropometric traits. While some studies identified risk alleles unique to Indigenous populations, their effects on trait variation are mostly small. I suggest that the associations rendered by many inter-population studies are probably inflated due to the absence of socio-cultural-economic covariates in the association models. I encourage the inclusion of admixed individuals in future GWAS studies to control for inter-ancestry differences in environmental factors. I suggest that some complex diseases might have arisen as *trade-off* costs of adaptations to past evolutionary selective pressures. Finally, I discuss how expanding panels with Indigenous ancestries in GWAS studies is key to accurately assess genetic risk in populations from the American continent, thus decreasing global health disparities.

Keywords: Indigenous Americans, disease risk, GWAS, genetic ancestry, genetic architecture.

Received: February 06, 2024; Accepted: July 13, 2024.

Preliminary statement

In this review I describe associations between global ancestry proportions and disease traits, as well as genome-wide associations with anthropometric traits and complex diseases of diverse categories. Even though environmental factors contribute an enormous extent to complex diseases, this review focuses on genetic diversity and its association with disease; thus, I do not discuss environmental causes of disease risk in depth. However, I do emphasize socio-economic factors affecting disease risk when applicable. Also, I focus on associations reported to be unique to populations with Indigenous American genetic ancestries, regardless of their effect sizes on disease traits. For simplicity, in most cases I refer to *genetic ancestries* as *ancestries* and to *significant association* as *association*. It is noteworthy that the overrepresentation of certain Indigenous populations in medical genomic studies can create the false impression that these populations are particularly prone to disease.

Populations with Indigenous American ancestries are a broad group that can be roughly divided into two groups: Indigenous and non-Indigenous populations. Indigenous groups self-identify as *Indigenous* and in general derive all or most of their genetic diversity from ancient North, Central, and South American populations. Non-Indigenous populations in general inhabit the same regions do not self-identify as

Indigenous, but derive part of their ancestry from ancient Indigenous populations. For simplicity, I refer to these groups as *Indigenous* and *admixed non-Indigenous* populations, respectively. Although imperfect, these terms are more suitable than US-centric labels such as *Latin Americans* or *Hispanics*, which are used by several studies cited in this review.

Genetic risk for complex diseases

Heritable genetic variation in humans comprises at least 125 million single nucleotide variants and ~173 thousand structural variants across diverse populations around the world (Byrska-Bishop *et al.*, 2022). However, only ~1 million single nucleotide variants and an undefined number of structural variants have predicted functional consequences (Byrska-Bishop *et al.*, 2022), suggesting that they constitute the main substrate for the heritable component of diseases.

In contrast to classic Mendelian *monogenic* diseases, which are caused by a single locus, *complex* diseases are impacted by the sum of many small-effect loci across the genome. The number of these loci, together with their frequencies and effect sizes, constitute the *genetic architecture* of complex diseases (Simons *et al.*, 2018). The genetic architecture of a disease is estimated from its *polygenic score*, a quantity that summarizes the estimated effect of many genetic variants on an individual's trait. These effects are obtained from *genome-wide association studies* (GWAS), namely, statistical associations between genetic variants across the genome and a trait of interest in a given population.

GWAS studies have identified hundreds of thousands of genetic variants involved in the genetic architecture of common complex traits and diseases (Welter *et al.*, 2014). Thanks to GWAS studies, we know that most genetic variation

Send correspondence to Lucas Vicuña. University of Chicago, Department of Medicine, Section of Genetic Medicine, Knapp Center for Biomedical Discovery, 900 E 57th St, 60637, Chicago IL, USA. E-mail: lkvicuna@gmail.com.

associated with disease susceptibility is broadly shared among human populations around the world (Bergström *et al.*, 2020). However, genetic drift (i.e., random changes in allele frequency over time) driven by demographic processes, as well as natural selection and other evolutionary forces, shift allele frequencies in specific populations (Prohaska *et al.*, 2019). Consequently, some risk alleles are common (allele frequency > 0.01) in understudied populations, but rare (allele frequency ≤ 0.01) in global datasets (Auton *et al.*, 2015; Martin *et al.*, 2017). In addition, there are numerous rare risk alleles that are unique to single or related populations (Auton *et al.*, 2015).

Despite individuals of European ancestries constituting only 14% of the world population (United Nations Report, 2022), two-thirds of GWAS studies have focused on populations with European ancestries (Ju *et al.*, 2022), thus limiting the portability of polygenic scores to other continental populations. For example, polygenic risk scores for schizophrenia derived from Eurocentric data were 50% less accurate in predicting risk for African and East Asian populations (Vilhjálmsón *et al.*, 2015). This underrepresentation has led to health disparities (Martin *et al.*, 2019), inaccuracies in genetic risk assessment, and biased medical recommendations for individuals with non-European ancestries (Wojcik *et al.*, 2019).

Populations with Indigenous ancestries are particularly underrepresented in GWAS studies (Ju *et al.*, 2022; Kang and Ruderfer, 2020). As of 2021, the estimated number of Indigenous individuals from countries other than the United States and Canada was 57.7 million, that is, 9.8% of the 589.2 million inhabitants of these countries (CRS Report, 2023). In addition, the estimated number of Indigenous individuals in the United States and Canada in 2021 was 4.3 million (US Census Bureau, 2022) and 1.8 million (Canada Census, 2021), respectively. Indigenous and admixed non-Indigenous people represent roughly 0.81% and 7.5% of the world population (United Nations Report, 2022), respectively. However, only 0.03% and 3.85% of GWAS studies have been performed on these populations, respectively (Ju *et al.*, 2022). Moreover, to my knowledge there are only a few GWAS performed on Indigenous populations from the United States (Malhotra *et al.*, 2011; Hanson *et al.*, 2014; Brown *et al.*, 2017; Peng *et al.*, 2019; Ramirez-Luzuriaga *et al.*, 2024) and none from Canada. In conclusion, the inclusion of Indigenous and admixed non-Indigenous populations in large-scale GWAS holds significant potential to uncover risk variants specific to Indigenous populations that can impact hundreds of millions of people. Figure 1 depicts the geographic location of populations whose Indigenous genetic diversity has been associated with complex diseases. Table 1 summarizes the genetic factors associated with complex diseases and traits.

Regarding the genetic risk of particular ancestry groups, including populations with Indigenous ancestries, our current understanding is derived from GWAS studies following three general cohort designs: (i) within-population analyses on a particular ancestry group; (ii) cross-population analyses

involving different ancestry groups, whereby variant effect sizes are estimated separately in each group; and (iii) analyses on admixed populations with ancestry from two or more groups, where *global* (i.e., individual) and in some cases *local* (i.e., per variant-wise) genetic ancestries are modeled as covariates. In addition, independent association studies based on one or more of the aforementioned designs can be aggregated in large-scale meta-analyses.

Categories of complex diseases

Infectious diseases

On the American continent, infectious diseases such as tuberculosis, influenza and COVID-19 are usually more prevalent in Indigenous groups compared to other ancestries (Rojas, 2007; Tollefson *et al.*, 2013; Soto-Cabezas *et al.*, 2022; Freitas Vaz *et al.*, 2024), largely due to socio-economic disparities (Lee *et al.*, 2023). The role of genetic ancestry in infectious disease risk is unclear, as socio-economic and environmental factors are often unassessed in genetic studies and correlate with genetic ancestry among Indigenous descendants. However, a few studies have found ancestry effects by accounting for socio-demographic factors. For example, among Peruvians from Lima, who in average have a proportion of 80% of Indigenous ancestries, there was a threefold increase in tuberculosis (TB) risk between individuals with the most and least Indigenous ancestries (Asgari *et al.*, 2022). This effect was mostly driven by gene expression regulation, as revealed by ~ 800 Indigenous-specific expression quantitative trait loci (eQTLs) for TB immune response (Supplementary Table 2 from Asgari *et al.*, 2022). Another study from the same group identified a single locus within an enhancer region that possibly alters monocyte function through the action of *ATPIB3* gene (Luo *et al.*, 2019).

If higher proportions of some Indigenous ancestries have an effect on susceptibility for TB, it is possible that genetic diversity at key immune loci could play a role. For example, human leukocyte antigen (*HLA*) and killer-cell immunoglobulin-like receptors (*KIR*) genes have the lowest genetic diversity among Amazonian Indigenous populations compared to many worldwide populations (de Brito Vargas *et al.*, 2022), which could translate in a less diverse immune repertoire. Arguably, this low genetic diversity might have resulted from lower effective population sizes among Indigenous groups compared to the other continental populations (Wang *et al.*, 2007).

Overall, additional studies with proper sample sizes and which account for socio-economic status are needed to attest whether or not genetic ancestry has an effect on susceptibility to some infectious diseases. These studies would also be important to identify more infectious disease-associated loci, as well as the distribution of their effects and phenotypic consequences.



Figure 1 - Complex diseases affecting present-day populations with Indigenous ancestries. Map showing diseases or medical conditions (represented by icons) affecting Indigenous and admixed non-Indigenous populations throughout the American continent.

Table 1 - Genetic factors affecting diseases and traits in populations with Indigenous ancestries. Populations are sorted from north to south along the American continent.

Population	Region	Disease/Trait	Gene(s)/Ancestry	Reference
Indigenous	American continent	Lower LDL cholesterol	Indigenous ancestries	Acuña-Alonzo <i>et al.</i> (2010) Sohail <i>et al.</i> (2023)
Admixed	USA, Mexico, Peru, Chile, Argentina	Systemic lupus erythematosus	<i>IRF5/TNPO3</i>	Alarcón-Riquelme <i>et al.</i> (2016) Langefeld <i>et al.</i> (2017)
Tsimshian	Canada (British Columbia)	Smallpox	<i>HLA-DQA1</i>	Lindo <i>et al.</i> (2016)
Indigenous	USA (Arizona)	Pubertal traits	<i>NUDT3, PACSIN1</i>	Ramirez-Luzuriaga <i>et al.</i> (2024)
Indigenous	USA (California)	Alcohol use disorders	<i>PRKG2, FSTL5</i>	Peng <i>et al.</i> (2019)
Pima	USA	Type 2 diabetes	<i>DNER</i>	Hanson <i>et al.</i> (2014)
Pima	USA	Chronic renal disease	<i>BCL2L11</i>	Brown <i>et al.</i> (2017)
Admixed	USA	Chronic renal disease	<i>BCL2L11</i>	Brown <i>et al.</i> (2017)
Admixed	USA (California)	Acute lymphoblastic leukemia	<i>PDE4B</i> , Indigenous ancestries	Yang <i>et al.</i> (2011)
Admixed USA Mexicans	USA	Short stature	Indigenous ancestries	Spear <i>et al.</i> (2020)
Admixed USA Mexicans	USA	High triglyceride & glucose levels (metabolic disorders)	Indigenous ancestries	Spear <i>et al.</i> (2020) Ko <i>et al.</i> (2014)
Admixed USA Mexicans	USA	Type 2 diabetes in obese women	Indigenous ancestries	Hu <i>et al.</i> (2015)
Indigenous	Mexico	Lung function	Indigenous subancestry (West-East cline)	Moreno-Estrada <i>et al.</i> (2014)
Indigenous	Mexico	Neuromyelitis optica	<i>HLA</i>	Romero-Hidalgo <i>et al.</i> (2020)
Mayan, Tojolabal	Mexico	Shorter stature than Northern Mexico	Indigenous subancestry	Sohail <i>et al.</i> (2023)
Mayan, Tojolabal	Mexico	Higher triglyceride & glucose levels than Northern Mexico	Indigenous subancestry	Sohail <i>et al.</i> (2023)
Mazateco	Mexico	Systemic lupus erythematosus	<i>IRF5</i>	Reddy <i>et al.</i> (2007)
Admixed	Mexico	Lower HDL cholesterol	<i>ABCI*C230</i> , Indigenous ancestries	Spear <i>et al.</i> (2020) Sohail <i>et al.</i> (2023)
Admixed	Mexico, USA	Type 2 diabetes	<i>SLC16A11</i>	Williams <i>et al.</i> (2014) Romero-Hidalgo <i>et al.</i> (2020)
Admixed	Puerto Rico	Lung function	<i>SLIT3</i> , Indigenous ancestries	Lee <i>et al.</i> (2020)
Admixed	Central America (El Salvador, Guatemala, Honduras, Nicaragua)	Type 2 diabetes	<i>EPHB1, KY</i> ; Indigenous ancestries	Horimoto <i>et al.</i> (2022)
Admixed	Peru (Lima)	Pulmonary Tuberculosis	<i>ATP1B3</i> , Indigenous ancestries; eQTLs	Luo <i>et al.</i> (2019) Asgari <i>et al.</i> (2022) Nathan <i>et al.</i> (2022)
Admixed	Peru (Lima)	Short stature	<i>FBN1</i>	Asgari <i>et al.</i> (2020)
Aymara-Quechua-Uro	Peru (Puno)	Preeclampsia	<i>PROZ, F7, F10</i>	Nieves-Colón <i>et al.</i> (2022)
Admixed	Chile	Newborn death due to pregnancy disorders	Indigenous subancestry (Aymara-Quechua)	Koenigstein <i>et al.</i> (2021)
Admixed	Chile	Gallbladder disease	<i>ACBG8</i> , Indigenous subancestry (Mapuche)	Bustos <i>et al.</i> (2019)
Admixed	Chile	Gallbladder, stomach & esophagus cancers	Indigenous subancestry (Mapuche)	Lorenzo-Bermejo <i>et al.</i> (2017)
Admixed	Chile	Early pubertal onset	Indigenous subancestry (Mapuche)	Vicuña <i>et al.</i> (2021) Vicuña <i>et al.</i> (2023)
Admixed	Chile	Pulmonary diseases	Indigenous subancestry (Mapuche)	Lorenzo-Bermejo <i>et al.</i> (2017)
Admixed	Chile, Peru, Brazil, Colombia, Mexico	Nose morphology	<i>EDAR, PAX3, DCHS2, SUPT3H/RUNX2, GLI3, PAX1</i> ; Mapuche subancestry relative to Central Andean subancestry	Adhikari <i>et al.</i> (2016) Bonfante <i>et al.</i> (2021) Chacón-Duque <i>et al.</i> (2018)

Autoimmune diseases

Systemic lupus erythematosus (SLE) is an autoimmune disease with a high heritability (~66%), according to a twin study (Lawrence *et al.*, 1987). Compared with European ancestries, increases in the proportion of Indigenous ancestries correlate with an earlier SLE onset. However, as SLE morbidity has been shown to correlate with lower socio-economic status, it cannot be ruled out that socio-economic factors also affect SLE onset (Sánchez *et al.*, 2012). Differences in genetic susceptibility for SLE between populations with Indigenous and European ancestries might be explained at least in part by differing allele frequencies at genes involved in SLE risk. One example pertains to *IRF5*, a transcription factor with key roles in the innate immune response (Reddy *et al.*, 2007). A case–control association study on the *IRF5* gene found the strongest association with SLE at the risk allele *rs2070197-C*. This variant allele has a substantially higher frequency in the Indigenous Mazateco from Mexico than in admixed Mexicans, and it is also higher in the latter group than in populations with European ancestries (Reddy *et al.*, 2007). Similarly, a GWAS performed on a multi-national cohort with enriched Indigenous ancestries from Argentina, Chile, Peru, Mexico, and the United States captured the strongest hit at a region mapping *IRF5* and its neighboring gene *TNPO3* (Alarcón-Riquelme *et al.*, 2016). Importantly, a large-scale transancestry GWAS replicated the *IRF5/TNPO3* signal as the top hit for admixed non-Indigenous individuals, but not for individuals with European nor African ancestries; in these ancestries the top associations were *HLA* variants (Langefeld *et al.*, 2017). These findings further support the hypothesis that inter-population differences in SLE genetic risk are partially due to differences in allele frequencies.

Another autoimmune disease that has been linked to Indigenous ancestries is neuromyelitis optica, a disease of the central nervous system that affects the optic nerve and spinal cord. A case–control GWAS study on admixed Mexicans showed that affected individuals had a higher proportion of Indigenous ancestries than healthy controls. Specifically, *HLA* variants comprised the strongest associated GWAS loci, all of which had increased means of local Indigenous ancestries (Romero-Hidalgo *et al.*, 2020). These observations contribute to the idea that among Indigenous descendants, genetic diversity of the *HLA* region contributes to the development of the disease.

Metabolic diseases

Increases in proportions of Indigenous ancestries seem to correlate with higher levels of lipid-related traits and diseases, at least in Mexicans (Acuña-Alonzo *et al.*, 2010; Ko *et al.*, 2014; Hu *et al.*, 2015; Huerta-Chagoya *et al.*, 2020; Spear *et al.*, 2020; Sohail *et al.*, 2023). However, as mentioned before, genetic ancestry usually covaries with environmental factors, such as socio-economic status (Sohail *et al.*, 2023), making the genetic component difficult to disentangle—occasionally, genetic and environmental factors act in opposite directions, obscuring disease associations altogether. Among admixed Mexicans, higher proportions of Indigenous ancestries associate with higher waist to hip ratio, triglyceride levels, glucose levels (Spear *et al.*, 2020), and risk for type 2 diabetes (T2D)

among obese women (Hu *et al.*, 2015), but also with lower LDL cholesterol levels (colloquially called *bad cholesterol*) (Sohail *et al.*, 2023). However, a case–control GWAS for T2D performed in Mexicans and in an admixed non-Indigenous cohort from the United States identified an association for the solute carrier gene *SLC16A11*, at a variant that increases intracellular triglyceride levels. Interestingly, the risk allele has a frequency of 0.5 in Indigenous populations, at least five times higher than in other continental ancestries (Williams *et al.*, 2014). Further, higher Indigenous ancestries are associated with lower HDL cholesterol in 80% of Indigenous populations from the American continent (Acuña-Alonzo *et al.*, 2010), thus increasing their collective risk of cardiovascular disease. HDL, usually referred to as *good cholesterol*, is the lipoprotein that removes cholesterol from the blood and transports it back to the liver. This association may be partially explained by a variant in the *ABCA1* transporter gene, which has a key function in the biosynthesis of HDL. The *ABCA1**C230 allele is found exclusively in Indigenous groups and shows the strongest association with low HDL among Pima from North America's Southwest and Mayans from Central America (Acuña-Alonzo *et al.*, 2010; Hünemeier *et al.*, 2012). However, when all these Indigenous groups are considered together and environmental covariates like socio-cultural factors and diet are included, the genetic association with low HDL does not reach the genome-wide association threshold (Sohail *et al.*, 2023). This suggests that the aforementioned environmental factors, and/or possibly an ancient selection event acting at this locus (Hünemeier *et al.*, 2012), are responsible for the lack of significant associations. We see a similar interplay between genetics and environment with BMI among admixed Mexicans. *Runs of homozygosity*, which correlate with lower genetic diversity and are more frequent in Indigenous groups, including those from Mexico, are associated with lower BMI (Sohail *et al.*, 2023). However, urban environments are associated with higher BMI in admixed Mexicans. We see the same correlation for Indigenous people from Mexico, but only when they live in urban environments (Sohail *et al.*, 2023). These findings suggest that if there are genetic risk variants of metabolic diseases among Indigenous groups from Mexico, these variants need to interact with the urban environment (e.g. diet, sedentarism) for obesity to be triggered (Sohail *et al.*, 2023). Taken together, these studies suggest that while genetic factors might contribute to disease risk, environmental factors have a profound effect on cardiovascular disease in Indigenous populations.

In addition to lipid dysregulation, T2D is a significant concern among many Indigenous populations. While modern Western dietary trends likely have a significant influence on the wide-spread development of T2D, recent studies have also uncovered genetic factors unique to Indigenous populations that may contribute to susceptibility. In a cohort of Central American heritage (El Salvador, Guatemala, Honduras, Nicaragua) from the United States, an *admixture mapping* study—associations between a trait and the per SNP-wise mean local ancestry along the genome—identified an haplotype block harboring the multifunctional *KY* and *EPHBI* genes, where Indigenous ancestries conferred increased risk for T2D, after accounting for socio-demographic variables (Horimoto

et al., 2022). The Indigenous Pima have one of the highest incidences of T2D in the world (Schulz and Chaudhari, 2015), and part of this risk might be affected by genetic loci, as suggested by a GWAS study, which identified a *DNER* gene risk variant associated with T2D. *DNER* encodes a growth factor that mediates signaling in the insulin-secreting pancreatic β -cells (Hanson *et al.*, 2014). A related GWAS for BMI on the same population identified several associations, but none reached the genome-wide significance threshold (Malhotra *et al.*, 2011). These two GWAS were not adjusted for socio-economic factors. Thus, while the results suggest that Indigenous populations may be more genetically susceptible to the negative effects of the modern Western diet on metabolic function, this hypothesis needs to be tested by rigorously assessing the contribution of environmental factors.

Cancer

Increased proportions of Indigenous ancestries have been associated with an increased incidence of certain cancers. A genome-wide study performed in a cohort from California, showed that higher individual proportions of Indigenous ancestries, as well as a higher mean of local Indigenous ancestries at *PDE4B* variants, were associated with higher recurrence of acute lymphoblastic leukemia. However, as noted by the authors, the effect of environmental, socio-economic, and dietary factors in these associations cannot be ruled out (Yang *et al.*, 2011).

Indigenous subancestries might have opposite effects on some kinds of cancers. A genetic epidemiology study based on aggregated data, performed associations between regional proportions of Indigenous subancestries in admixed Chileans and their mortality rates due to several cancers, adjusting for individual socio-economic variables (Lorenzo-Bermejo *et al.*, 2017). The most striking finding was the strong positive association between Mapuche subancestry (Northern Patagonia lowlands) and mortality risk for gallbladder cancer; specifically, a 3.7% increased mortality per 1% increase in Mapuche subancestry. Nevertheless, hitherto no variant has achieved a genome-wide association with gallbladder cancer among Mapuche ancestry-bearing populations. However, in admixed Chileans a *ACBG8* gene variant is associated with gallbladder disease, which is the main risk factor for gallbladder cancer (Bustos *et al.*, 2019). The same study from Lorenzo-Bermejo *et al.* (2017) reported that while Mapuche subancestry is associated with increased mortality rates due to esophagus and stomach cancer, Aymara subancestry (Central Andes highlands) has a protective effect for these cancer types. In addition, the mortality rate due to skin, bladder, larynx, bronchus, and lung cancers decreases with augmented Mapuche subancestry proportions. While the findings on subancestry differences in mortality rates are promising for health policies targeting particular ethnic groups, they need to be validated using individual paired genetic-phenotypic data.

Pulmonary diseases

As with cancer, Indigenous subancestries also seem to have contrasting effects on pulmonary disease outcomes. While Mapuche subancestry among admixed Chileans is associated with increased mortality due to asthma, pneumonia,

and chronic lower respiratory diseases, Aymara subancestry appears to have a protective effect (Lorenzo-Bermejo *et al.*, 2017). It is possible that positive selection among the Aymara for physiological traits enabling life at the high altitudes of the Andes Mountains, such as hypoxia and cardiovascular function (Crawford *et al.*, 2017), might have contributed to these differences. Among Mexicans, Indigenous subancestry from the western states of the country (e.g., Sonora) associates with a 7% change in lung function (i.e., severity of obstructive lung diseases) compared with the eastern states (e.g., Yucatán) (Moreno-Estrada *et al.*, 2014). In a cohort of ~5,500 admixed individuals with African, European, and Indigenous ancestries recruited in the United States, Indigenous ancestries associated with lower odds of asthma, after adjusting for early life exposures, air pollution and socioeconomic status (Pino-Yanes *et al.*, 2015). An admixture mapping study performed on admixed Puerto Ricans identified a chromosomal region where each Indigenous allele was associated with an increase in lung function. The lead SNP mapped the tumor suppression *SLIT3* gene (Lee *et al.*, 2020). It remains to be determined which genetic and non-genetic variables contribute the most to such fine-scale differences in lung function among Indigenous subancestries.

Renal diseases

Chronic kidney disease (CKD) is a serious condition characterized by kidney damage and impaired ability to filter waste from the blood. In Central American countries with high Indigenous heritage like Nicaragua, Costa Rica, El Salvador, and Guatemala, CKD is highly prevalent (Horimoto *et al.*, 2022). Also, in the United States the prevalence of CKD among Indigenous populations is double that of other ethnic groups. An admixture mapping study of albuminuria—increased albumin excretion, which is a sign of kidney damage—performed on ~12,000 admixed individuals with sociodemographic assessment, identified two signals enriched in Indigenous ancestries associated with increased urine albumin excretion. One of them harbors a variant mapping the *BCL2L11* gene, which has putative roles in kidney function and disease. Interestingly, the associated allele has a frequency higher than 0.5 among the Indigenous Pima but is non-existent in European and African populations (Brown *et al.*, 2017). A similar admixture mapping study from the same group performed on a cohort of Central American origin identified a haplotype spanning the *RGS6* gene, whereby the allele enriched in Indigenous ancestries has protective effects for albuminuria over the allele enriched in European ancestries (Horimoto *et al.*, 2022). These results suggest that at least part of the genetic diversity of Indigenous populations underlies specific effects on susceptibility or protection to kidney diseases.

Neuropsychiatric disorders

Indigenous ancestries appear to confer a reduced risk of developing Alzheimer's disease, as suggested by genetic studies in admixed Brazilians (Benedet *et al.*, 2012), Colombians (Moreno *et al.*, 2017), and admixed Caribbean individuals (Horimoto *et al.*, 2021). An admixture mapping study in the latter population found that an excess of Indigenous ancestries at a region harboring five protein-coding genes, associated with

a protective effect against Alzheimer's disease risk. Within that region, one variant allele mapping the *DUBR* gene was associated with a reduced risk for Alzheimer's disease in individuals with European ancestries (Horimoto *et al.*, 2021). *DUBR* encodes a long noncoding RNA (lncRNA) implicated in brain development and function (Huang *et al.*, 2022).

Alcohol dependence has a relevant genetic component, with an estimated heritability of 50%, according to twin and adoption studies (Verhulst *et al.*, 2015). Epidemiological data suggests that some Indigenous groups are more prone to alcohol use disorders than other populations (Ehlers and Gizer, 2013). As of 2005, alcohol dependence rates were 5-6 times higher among Indigenous people from the United States compared to other United States ethnic populations (Ehlers and Wilhelmsen 2005). To (partially) explain these observations, Ehlers and Gizer (2013), hypothesized a scenario of gene-by-environment interaction, whereby the disorder's morbidity is increased by the exposure of genetic risk variants to certain environments, such as low socio-economic status and historical trauma, that are more common among Indigenous populations than in other populations.

A GWAS study for alcohol dependence performed on an Indigenous cohort from California, as well as on a cohort of European ancestries revealed a few associations. While common variants in the uncharacterized *FSTL5* gene, associated with alcohol consumption in the population with European ancestries, rare variants in *FSTL5* and *PRKG2* genes associated with alcohol-related life events and with affective symptoms when cutting down alcohol only in the Indigenous group (Peng *et al.*, 2019). Interestingly, *PRKG2* has been associated with obesity-related traits in diverse worldwide populations (Peng *et al.*, 2019), suggesting that variants that favor food intake might favor alcohol consumption as well (Ehlers and Gizer, 2013). While these results are promising, the effect of specific variants need to be formally tested in an admixed cohort, controlling for environmental differences between ancestries.

Pregnancy disorders

Preeclampsia is a medical pregnancy condition that causes 40% of all premature births, and is characterized by high blood pressure as well as signs of liver or kidney damage. The risk for preeclampsia increases in high altitude regions for all ancestries. However, women from Puno—located ~3,800m above sea level in the Peruvian Andes and inhabited by Quechua, Aymara, and Uro Indigenous populations—exhibit one of the highest incidences of preeclampsia in the world, suggesting the susceptibility may have a genetic component in this population. A parent-offspring trio GWAS study among Puno residents found that common genetic variants mapping the clotting factor genes *PROZ*, *F7*, and *F10* in the fetal genome increase susceptibility to preeclampsia among these women (Nieves-Colón *et al.*, 2022). Future studies will help understand whether genetic diversity also underlies risk for preeclampsia in other high-altitude populations worldwide.

Indigenous populations show the lowest genetic diversity worldwide (Wang *et al.*, 2007), and some of them exhibit large runs of homozygosity (Koenigstein *et al.*, 2021), which may result in a higher incidence of birth defects and other diseases.

A genetic epidemiology study performed on admixed Chileans, found that runs of homozygosity and Aymara subancestry are associated with increased risk of child mortality due to natural intracranial hemorrhage of fetuses and newborns as well as to related gestation disorders (Koenigstein *et al.*, 2021), after correcting for several socio-economic/education variables. However, this study did not use paired genetic-phenotypic data. It remains to be determined whether or not these disease outcomes are similar in other worldwide populations with similar inbreeding levels.

Anthropometric traits

Stature is a polygenic trait explained by the contribution of thousands of variants (Yengo *et al.*, 2022), each exerting small effect sizes. In the population of Lima, Peru, which as mentioned before has on average 80% Indigenous ancestry and is among the world's shortest populations, 700 variants explain 7% of the variance in height (Asgari *et al.*, 2020). However, in this population, a single missense variant allele in the *FBNI* gene accounts for 1% of the total height. Hence, one person carrying a single copy of that allele is 2.2 cm shorter on average than people carrying the alternative allele (Asgari *et al.*, 2020). *FBNI* encodes an extracellular matrix protein fibrillin-1, which provides mechanical stability to tissues. While candidate genes were not sought, another study found the composition of Indigenous ancestries may have an effect on height: subancestry from Central America (Oaxaca, Maya) is associated with a shorter stature relative to subancestry from Northern Mexico (Huichol, Tarahumara) (Sohail *et al.*, 2023). This is another case where quantitative genetics analyses can detect subtle effects of Indigenous subancestries over a phenotype.

When comparing the effect of ancestry on facial traits in admixed populations from Brazil, Chile, Colombia, Mexico, and Peru, the strongest correlation (opposite effects in Indigenous vs. European ancestries) was found for a measure of nose position (Bonfante *et al.*, 2021). A related GWAS performed in the same cohort identified six SNPs associated with nose shape, mapping the genes *EDAR*, *PAX3*, *DCHS2*, *SUPT3H/RUNX2*, *GLI3*, and *PAX1* to the trait (Adhikari *et al.*, 2016). In a further study on the same sample it was found that Mapuche subancestry strongly associates with a less protruded nose and broader nose-tip angle when contrasted with Central Andean ancestries (Quechua, Colla, and Aymara populations pooled together). Further, SNP alleles associated with these features show differences in their frequencies between these subancestries (Chacón-Duque *et al.*, 2018).

Abnormal pubertal growth is associated with adult risk for cancer, diabetes and cardiometabolic disorders. Pubertal growth variability has a strong genetic component. In an Indigenous population from Arizona, it was found that the heritability of pubertal traits is between 25% and 71% (Ramirez-Luzuriaga *et al.*, 2024). Also, genetic ancestry underlies variation in pubertal traits. Two studies based on the same admixed longitudinal pediatric cohort, quantified how growth traits differ in predicted individuals with 100% Mapuche versus 100% European ancestries. The first study (Vicuña *et al.*, 2021) analyzed the peak height velocity, which is the period where maximum rate of growth occurs during

puberty. The authors predicted that the age at peak height velocity is 0.7 years earlier in adolescents with exclusively Mapuche ancestry versus exclusively European ancestries. However, the study did not correct for individual socio-economic status. The second study, which did account for maternal education level, a proxy for socio-economic status, analyzed child growth, by focusing on longitudinal BMI (Vicuña *et al.*, 2023). BMI was higher in children with exclusively Mapuche ancestry at all ages older than 6 years old when compared with children with exclusively European ancestries. Further, the age at adiposity rebound (Age-AR)—the age when the minimal BMI is reached during childhood growth—was lower by 1.9 years, and the BMI at Age-AR was higher by 1.2 kg/m² in children with exclusively Mapuche ancestry versus exclusively European ancestries. These results suggest that Mapuche children have an earlier pubertal onset than children with European ancestries, which can be explained in part by differences in genetic diversity. A recent GWAS study performed on an Indigenous population from Arizona, identified six variants associated with peak height velocity and two with the duration of growth spurt (Ramirez-Luzuriaga *et al.*, 2024). The associated genes include *NUDT3* and *PACSINI*, whose role in puberty is unknown. However, since there are only a few longitudinal studies on puberty, it cannot be attested if these associations are specific to Indigenous populations.

Evolutionary causes of genetic risk among Indigenous populations

The first Indigenous Americans entered the North American continent no earlier than 23,000 years ago (Raghavan *et al.*, 2015). Within a few millennia they reached virtually all corners of North, South, and Central America (Figure 2A). During their journey, they experienced complex demographic events including founder events and population admixture. Also, since arriving in the American continent, the genomes of Indigenous populations have been shaped by their local environments. This began in the ancestors of all Indigenous people, leading to shared signatures of selection across the continent. For example, variant alleles of Denisovan origin in the immune-related *MUC19* gene helped ancient Indigenous populations to adapt to local selective pressures (Villanea *et al.*, 2023). Populations then adapted to their local environments, including selection to high levels of inorganic arsenic in water among Andean populations from Argentina (Collas, Calchequies) (Schlebusch *et al.*, 2015); cardiovascular function among high-altitude Aymara from Bolivia (Crawford *et al.*, 2017); TB infection among Ecuadorian highland populations (Joseph *et al.*, 2023); parasitic infections among Amazonian groups from Brazil (Couto-Silva *et al.*, 2023); as well as changing diets among ancient Mesoamericans (Hünemeier *et al.*, 2012). Selection continued in response to changing environmental conditions. For example, the introduction of European pathogens likely imposed strong selection across the American continent, causing genetic adaptation in Indigenous populations (Lindo *et al.*, 2016) (Figure 2A). Moreover, assortative mating conditioned on ancestry-related traits left footprints in the genomes of admixed populations from the American continent (Mas-Sandoval *et al.*, 2023). The cumulative effects of these selective events, genetic drift, as

well as genetic forces such as mutation, recombination, and gene conversion, have led to broad differences in disease susceptibility across the American continent, as well as between Indigenous descendants and individuals from other parts of the world.

The precise evolutionary events leading to genetic risk are unknown for most diseases, including those affecting Indigenous populations. However, there are a few instances where these events have been inferred and contextualized, mostly as *trade-offs* between fitness advantages conferred by positive selection acting over target traits in ancestral groups, and fitness costs associated with increased genetic risk at causal, linked, and *pleiotropic loci* (i.e., loci affecting many traits simultaneously) (Figure 2B).

For example, selection on the *HLA* locus appears to have favored *HLA-DQA1* among the Indigenous Tsimshian from the Canadian Pacific Northwest before the encounter with Europeans. *HLA-DQA1* alleles were fixed in pre-contact Tsimshian as a result of positive selection, most likely to pathogens endemic to that region. However, population genetics analyses showed that after the introduction of smallpox from Europe, some of these alleles drastically decreased in frequency (as low as 0.37) as a result of negative selection (Lindo *et al.*, 2016). Thus, this example highlights how an increased genetic risk for infectious diseases among Indigenous populations might result from prior positive selection acting on immune variants during environmental (i.e., pathogenic) change (Figure 2B).

Trade-off hypotheses posit adaptations to famine at genes involved in lipid metabolism among ancient Indigenous populations at the cost of increased risk for dyslipidemias in modern descendants (Ko *et al.*, 2014) [see Neel's *thrifty gene hypothesis* (Neel 1962)] (Figure 2B). Some Indigenous groups from North America seem to have a higher general risk to obesity and lipid-related disorders (i.e., dyslipidemias) than other populations (Breathett *et al.*, 2020; Huerta-Chagoya *et al.*, 2020). Interestingly, admixed Mexicans show Indigenous-specific positive selection signals at a haplotype region harboring the salt-inducible kinase 3 (*SIK3*) gene (Ko *et al.*, 2014). The selected *SIK3* haplotype is associated with high triglyceride levels in Mexicans, but not in a European population, as suggested by an allele frequency of less than 0.01 in the latter group. Similarly, evidence suggests that the Indigenous-specific *ABCI* C230* allele, which favors the retention of intracellular cholesterol, experienced a selection event in an ancient Indigenous population (Acuña-Alonzo *et al.*, 2010; Hünemeier *et al.*, 2012).

There are other hypotheses of putative *trade-offs* involving food metabolism. A risk variant allele at the gene *SLC16A11* that increases T2D risk in Indigenous populations, introgressed into modern humans via admixture with Neanderthals (SIGMA Type 2 Diabetes Consortium *et al.*, 2014). However, it is unknown whether that risk allele conferred selective advantages to the Neanderthal ancestors or human ancestors of present-day Indigenous populations. Another hypothesis posits a higher susceptibility for eating behavior among Indigenous populations, which would be pleiotropically linked with a higher genetic susceptibility to alcohol abuse disorders (Ehlers and Gizer, 2013). This effect



Figure 2 - Adaptations experienced by ancient Indigenous groups during the peopling of the American continent. **A.** Representative examples of adaptations experienced by some populations. Arrows represent approximate migration routes of ancient Indigenous populations and arrow colors represent different lineages. Adapted from Posth *et al.* (2018). **B.** Hypotheses for *trade-offs* between past adaptations leading to diseases in the present. Scales represent the trade-offs between present day diseases and past adaptations.

would be mediated by variants involved in both traits, such as *PRKG2* variants (see GWAS findings on *PRKG2* mentioned above) (Peng *et al.*, 2019). However, this hypothesis is only supported by indirect evidence.

Trade-offs might also exist between selective advantages provided by early pubertal onset and diseases occurring later in life (Figure 2B). In pygmy groups, early pubertal onset might be favored to achieve higher reproductive rates in conditions of short lifespans and resource availability (Migliano *et al.*, 2007), and positive selection favoring early pubertal onset might play a role here. Interestingly, the Indigenous Hiwi from northern South America are well known for their significantly higher pre-pubertal growth and earlier pubertal onset compared to other groups (Walker *et al.*, 2006). Also, they exhibit short lifespans (e.g., as of 2006, life expectancy at age 15 was 36 years). On the other hand, early puberty can exacerbate disease risk for several diseases in adulthood, including breast and endometrial cancer, T2D, cardiovascular diseases, and psychopathologies (Graber, 2013); (Day *et al.*, 2015; 2017). Even though there is no epidemiological data comparing early puberty-triggered adulthood diseases between the Hiwi and other continental groups, there is evidence for such diseases in other Indigenous groups. For example, among Indigenous males from Arizona, an earlier and accelerated pubertal onset was associated with a higher risk for T2D (Ramirez-Luzuriaga *et al.*, 2023). Also, two genetic studies mentioned earlier showed that Mapuche subancestry leads to an earlier pubertal onset compared to European ancestries (Vicuña *et al.*, 2023, 2021); however, it is not known whether this is the result of positive selection or drift, nor whether this phenomenon has been linked to adult diseases among them. In summary, it is possible that some diseases among Indigenous groups arise as a consequence of early puberty, resulting from an adaptation to short life spans; however, this hypothesis is speculative and would need to be tested.

Future directions

Many of the disease risk variants have been identified via GWAS based on relatively small cohorts genotyped on SNP arrays, which capture a modest fraction of the genetic architecture of complex diseases. Hence, bigger sample sizes and deeper sequence coverage are needed to estimate the effect of common variants, which should be independent of the population (Hou *et al.*, 2023), as well as to capture the effect of rare variants unique to Indigenous populations. Some initiatives are addressing these gaps through the assembly of relatively large cohorts and/or biobanks, such as Maule Cohort in Chile (MAUCO) (Ferreccio *et al.*, 2020), Estudo Longitudinal de Saúde do Adulto (ELSA) in Brazil (Schmidt *et al.*, 2015), and the Mexican Biobank project (Sohail *et al.*, 2023). Association studies in these populations promise to identify variants associated with disease in Indigenous populations.

However, association studies are only one piece to identifying causal genetic variants involved in complex diseases that affect populations with Indigenous ancestries. Mapping of molecular quantitative trait loci (mQTLs) holds enormous potential for disentangling the *functional architecture* of these diseases. Thanks to QTL mapping we know that, in addition to

~10% of GWAS hits mapping variants producing changes in amino acids (Hindorff *et al.*, 2009), ~40% of GWAS hits are associated with the expression (eQTLs) and splicing (sQTLs) levels of particular genes (Mu *et al.*, 2021). An additional ~25% GWAS hits mediate epigenetic regulation through their effects on chromatin accessibility (caQTLs) (Zepeng Mu, unpublished data through personal communication), leaving 25% of GWAS hits uncharacterized.

Fully understanding the causal variants and biological pathways involved in complex diseases affecting Indigenous populations also requires in depth functional characterization using experimental models. For example, expression regulation of the *FTO* gene, a classic gene involved in obesity risk, was thoroughly characterized in adipocytes and brain neurons using functional genomics approaches, leading to key conclusions about pleiotropy in the context of obesity (Sobreira *et al.*, 2021).

More generally, the implementation of high throughput methods, such as single-cell sequencing and genome editing, is key to identifying genetic variants likely to be functionally important, such as those underlying molecular QTLs. This is particularly relevant to capture inter-population differences in disease risk, given that most of these differences are expected to arise as a result of slight differences in gene regulation across multiple tissues (Watanabe *et al.*, 2019).

Finally, and maybe most importantly, genetic research involving Indigenous populations raises significant ethical considerations due to historical injustices, cultural sensitivities, and the potential for exploitation (Claw *et al.*, 2018). Thus, fostering ethical partnerships and engagement with Indigenous communities is key to achieve a higher representation of Indigenous populations in genomic studies, which in turn has the potential to reduce the risk for certain diseases for which they are disproportionately susceptible (Claw *et al.*, 2018).

Conclusions

The genetic architecture of diseases is poorly understood in populations with Indigenous ancestries due to their underrepresentation in medical genomic studies. While most of the genetic architecture of complex diseases is expected to be shared among all human populations, a smaller but relevant fraction is expected to vary among continental populations, including Indigenous populations. Increases in Indigenous genetic ancestries are associated with higher or lower susceptibility for certain diseases. However, in some cases Indigenous subancestries might also have contrasting effects on particular diseases. Indigenous ancestries covary with environmental factors such as socio-economic status, frequently leading to wrong conclusions about genetic risk to complex diseases. The interaction of risk variants with an unfavorable environment (e.g., unhealthy diet) is key for triggering certain diseases (e.g., metabolic disorders) among Indigenous populations. Several findings on higher risk for complex diseases among Indigenous groups are based on cross-population comparisons, mostly with populations with European ancestries. However, it is likely that a substantial fraction of such inter-population risk differences is inflated because of uncontrolled environmental variables and differences in the exact phenotypes assessed across groups

(Hou *et al.*, 2023). Future genetic association studies that compare effects across Indigenous and non-Indigenous local ancestries within admixed individuals are needed to reassess many GWAS findings.

Acknowledgements

I thank Tauras Vilgalys, Raúl Aguirre-Gamboa and Tyler Hansen for their critical comments and feedback, Sarah Sumner for editorial review, Mauricio Álvarez for the illustrations, and Erroll Garner for miscellaneous support. This work was supported by ANID FONDECYT grant 11200324, and the Ancestry Networks for the Human Cell Atlas: “Immune Cell Atlas of Indigenous South American Populations”, funded by the Chan-Zuckerberg Initiative.

Conflict of Interest

The author declares that there is no conflict of interest that could be perceived as prejudicial to the impartiality of the reported research.

Author Contributions

LV conceived the study and wrote the manuscript.

References

- Canada Census - Indigenous People 9 (2021) Government of Alberta March 22, 2023. Treasury Board and Finance, Alberta.
- United Nations Report 2022 United Nations Department of Economic and Social Affairs, Population Division (2022) World Population Prospects 2022: Summary of Results. UN DESA/POP/2022/TR/NO. 3.
- Acuña-Alonzo V, Flores-Dorantes T, Kruit JK, Villarreal-Molina T, Arellano-Campos O, Hünemeier T, Moreno-Estrada A, Ortiz-López MG, Villamil-Ramírez H, León-Mimila P *et al.* (2010) A functional ABCA1 gene variant is associated with low HDL-cholesterol levels and shows evidence of positive selection in Native Americans. *Hum Mol Genet* 19:2877-2885.
- Adhikari K, Fuentes-Guajardo M, Quinto-Sánchez M, Mendoza-Revilla J, Camilo Chacón-Duque J, Acuña-Alonzo V, Jaramillo C, Arias W, Lozano RB, Pérez GM *et al.* (2016) A genome-wide association scan implicates DCHS2, RUNX2, GLI3, PAX1 and EDAR in human facial variation. *Nat Commun* 7:11616.
- Alarcón-Riquelme ME, Ziegler JT, Molineros J, Howard TD, Moreno-Estrada A, Sánchez-Rodríguez E, Ainsworth HC, Ortiz-Tello P, Comeau ME, Rasmussen A *et al.* (2016) Genome-wide association study in an Amerindian ancestry population reveals novel systemic lupus erythematosus risk loci and the role of European admixture. *Arthritis Rheumatol Hoboken NJ* 68:932-943.
- Asgari S, Luo Y, Akbari A, Belbin GM, Li X, Harris DN, Selig M, Bartell E, Calderon R, Slowikowski K *et al.* (2020) A positively selected FBN1 missense variant reduces height in Peruvian individuals. *Nature* 582: 234-239.
- Asgari S, Luo Y, Huang CC, Zhang Z, Calderon R, Jimenez J, Yataco R, Contreras C, Galea JT, Lecca L *et al.* (2022) Higher native Peruvian genetic ancestry proportion is associated with tuberculosis progression risk. *Cell Genom* 2:100151.
- Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, Clark AG, Donnelly P, Eichler EE *et al.* (2015) A global reference for human genetic variation. *Nature* 526:68-74.
- Benedet AL, Moraes CF, Camargos EF, Oliveira LF, Souza VC, Lins TC, Henriques AD, Carmo DGS, Machado-Silva W, Araújo CN *et al.* (2012) Amerindian genetic ancestry protects against Alzheimer's disease. *Dement Geriatr Cogn Disord* 33:311-317.
- Bergström A, McCarthy SA, Hui R, Almarri MA, Ayub Q, Danecek P, Chen Y, Felkel S, Hallast P, Kamm J *et al.* (2020) Insights into human genetic variation and population history from 929 diverse genomes. *Science* 367:eaay5012.
- Bonfante B, Faux P, Navarro N, Mendoza-Revilla J, Dubied M, Montillot C, Wentworth E, Poloni L, Varón-González C, Jones P *et al.* (2021) A GWAS in Latin Americans identifies novel face shape loci, implicating VPS13B and a Denisovan introgressed region in facial variation. *Sci Adv* 7:eabc6160.
- Breathett K, Sims M, Gross M, Jackson EA, Jones EJ, Navas-Acien A, Taylor H, Thomas KL and Howard BV (2020) Cardiovascular health in American Indians and Alaska Natives: A scientific statement from the American Heart Association. *Circulation* 141:e948-e959.
- Brown LA, Sofer T, Stilp AM, Baier LJ, Kramer HJ, Masindova I, Levy D, Hanson RL, Moncrieff AE, Redline S *et al.* (2017) Admixture mapping identifies an Amerindian ancestry locus associated with albuminuria in Hispanics in the United States. *J Am Soc Nephrol* 28:2211-2220.
- Bustos BI, Pérez-Palma E, Buch S, Azócar L, Riveras E, Ugarte GD, Toliat M, Nürnberg P, Lieb W, Franke A *et al.* (2019) Variants in ABCG8 and TRAF3 genes confer risk for gallstone disease in admixed Latinos with Mapuche Native American ancestry. *Sci Rep* 9:772.
- Byrska-Bishop M, Evani US, Zhao X, Basile AO, Abel HJ, Regier AA, Corvelo A, Clarke WE, Musunuri R, Nagulapalli K *et al.* (2022) High-coverage whole-genome sequencing of the expanded 1000 Genomes Project cohort including 602 trios. *Cell* 185:3426-3440.e19.
- Chacón-Duque J-C, Adhikari K, Fuentes-Guajardo M, Mendoza-Revilla J, Acuña-Alonzo V, Barquera R, Quinto-Sánchez M, Gómez-Valdés J, Everardo Martínez P, Villamil-Ramírez H *et al.* (2018) Latin Americans show wide-spread Converso ancestry and imprint of local Native ancestry on physical appearance. *Nat Commun* 9:5388.
- Claw KG, Anderson MZ, Begay RL, Tsosie KS, Fox K and Garrison NA (2018) A framework for enhancing ethical genomic research with Indigenous communities. *Nat Commun* 9:2957.
- Couto-Silva CM, Nunes K, Venturini G, Araújo Castro e Silva M, Pereira LV, Comas D, Pereira A and Hünemeier T (2023) Indigenous people from Amazon show genetic signatures of pathogen-driven selection. *Sci Adv* 9:eabo0234.
- Crawford JE, Amaru R, Song J, Julian CG, Racimo F, Cheng JY, Guo X, Yao J, Ambale-Venkatesh B, Lima JA *et al.* (2017) Natural selection on genes related to cardiovascular health in high-altitude adapted Andeans. *Am J Hum Genet* 101:752-767.
- Day FR, Elks CE, Murray A, Ong KK and Perry JRB (2015) Puberty timing associated with diabetes, cardiovascular disease and also diverse health outcomes in men and women: The UK Biobank study. *Sci Rep* 5:11208.
- Day FR, Thompson DJ, Helgason H, Chasman DI, Finucane H, Sulem P, Ruth KS, Whalen S, Sarkar AK, Albrecht E *et al.* (2017) Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. *Nat Genet* 49:834-841.
- de Brito Vargas L, Beltrame MH, Ho B, Marin WM, Dandekar R, Montero-Martín G, Fernández-Viña MA, Hurtado AM, Hill KR, Tsuneto LT *et al.* (2022) Remarkably low KIR and HLA diversity in Amerindians reveals signatures of strong purifying selection shaping the centromeric KIR region. *Mol Biol Evol* 39:msab298.
- Ehlers CL and Gizer IR (2013) Evidence for a genetic component for substance dependence in Native Americans. *Am J Psychiatry* 170:154-164.

- Ehlers CL and Wilhelmsen KC (2005) Genomic scan for alcohol craving in Mission Indians. *Psychiatr Genet* 15:71-75.
- Ferreccio C, Huidobro A, Cortés S, Bambs C, Toro P, Wyngard VVD, Acevedo J, Paredes F, Venegas P, Verdejo H *et al.* (2020) Cohort Profile: The Maule Cohort (MAUCO). *Int J Epidemiol* 49:760.
- Graber JA (2013) Pubertal timing and the development of psychopathology in adolescence and beyond. *Horm Behav* 64:262-269.
- Hanson RL, Muller YL, Kobes S, Guo T, Bian L, Ossowski V, Wiedrich K, Sutherland J, Wiedrich C, Mahkee D *et al.* (2014) A Genome-wide association study in American Indians implicates DNER as a susceptibility locus for type 2 diabetes. *Diabetes* 63:369-376.
- Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS and Manolio TA (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* 106:9362-9367.
- Horimoto ARVR, Xue D, Thornton TA and Blue EE (2021) Admixture mapping reveals the association between Native American ancestry at 3q13.11 and reduced risk of Alzheimer's disease in Caribbean Hispanics. *Alzheimers Res Ther* 13:122.
- Horimoto ARVR, Sun Q, Lash JP, Daviglius ML, Cai J, Haack K, Cole SA, Thornton TA, Browning SR and Franceschini N (2022) Admixture mapping screening of CKD traits and risk factors in U.S. Hispanic/Latino individuals from Central America country-of-origin. *medRxiv*. DOI: 10.1101/2022.06.17.22276554.
- Hou K, Ding Y, Xu Z, Wu Y, Bhattacharya A, Mester R, Belbin GM, Buyske S, Conti DV, Darst BF *et al.* (2023) Causal effects on complex traits are similar for common variants across segments of different continental ancestries within admixed individuals. *Nat Genet* 55:549-558.
- Hu H, Huff CD, Yamamura Y, Wu X and Strom SS (2015) The relationship between Native American ancestry, body mass index and diabetes risk among Mexican-Americans. *PLoS One* 10:e0141260.
- Huang J, Jiang B, Li G-W, Zheng D, Li M, Xie X, Pan Y, Wei M, Liu X, Jiang X *et al.* (2022) m6A-modified lincRNA DUBR is required for neuronal development by stabilizing YTHDF1/3 and facilitating mRNA translation. *Cell Rep* 41:111693.
- Huerta-Chagoya A, Moreno-Macías H, Sevilla-González M, Rodríguez-Guillén R, Ordóñez-Sánchez ML, Gómez-Velasco D, Muñoz-Hernández L, Segura-Kato Y, Arellano-Campos O, Cruz-Bautista I *et al.* (2020) Contribution of known genetic risk variants to dyslipidemias and type 2 diabetes in Mexico: A population-based nationwide study. *Genes* 11:114.
- Hünemeier T, Amorim CEG, Azevedo S, Contini V, Acuña-Alonzo V, Rothhammer F, Dugoujon J-M, Mazières S, Barrantes R, Villarreal-Molina MT *et al.* (2012) Evolutionary responses to a constructed niche: Ancient Mesoamericans as a model of gene-culture coevolution. *PLoS One* 7:e38862.
- Joseph SK, Migliore NR, Olivieri A, Torroni A, Owings AC, DeGiorgio M, Ordóñez WG, Aguilú JJO, González-Andrade F, Achilli A *et al.* (2023) Genomic evidence for adaptation to tuberculosis in the Andes before European contact. *iScience* 26:106034.
- Ju D, Hui D, Hammond DA, Wonkam A and Tishkoff SA (2022) Importance of including non-European populations in large human genetic studies to enhance precision medicine. *Annu Rev Biomed Data Sci* 5:321-339.
- Kang J and Ruderfer DM (2020) Expanding cultural and ancestral representation in psychiatric genetic studies. *Neuropsychopharmacology* 45:1593-1594.
- Ko A, Cantor RM, Weissglas-Volkov D, Nikkila E, Reddy PMVL, Sinsheimer JS, Pasaniuc B, Brown R, Alvarez M, Rodriguez A *et al.* (2014) Amerindian-specific regions under positive selection harbour new lipid variants in Latinos. *Nat Commun* 5:3983.
- Koenigstein F, Boekstegers F, Wilson JF, Fuentes-Guajardo M, Gonzalez-Jose R, Bedoya G, Bortolini MC, Acuña-Alonzo V, Gallo C, Ruiz Linares A *et al.* (2021) Inbreeding, Native American ancestry and child mortality: Linking human selection and paediatric medicine. *Hum Mol Genet* 31:975-984.
- Langefeld CD, Ainsworth HC, Graham DSC, Kelly JA, Comeau ME, Marion MC, Howard TD, Ramos PS, Croker JA, Morris DL *et al.* (2017) Transancestral mapping and genetic load in systemic lupus erythematosus. *Nat Commun* 8:16021.
- Lawrence JS, Martins CL and Drake GL (1987) A family survey of lupus erythematosus. 1. Heritability. *J Rheumatol* 14:913-921.
- Lee EY, Mak ACY, Hu D, Sajuthi S, White MJ, Keys KL, Eckalbar W, Bonser L, Huntsman S, Urbanek C *et al.* (2020) Whole-genome sequencing identifies novel functional loci associated with lung function in Puerto Rican youth. *Am J Respir Crit Care Med* 202:962-972.
- Lee NR, King A, Vigil D, Mullaney D, Sanderson PR, Ametepee T and Hammitt LL (2023) Infectious diseases in Indigenous populations in North America: Learning from the past to create a more equitable future. *Lancet Infect Dis* 23:e431-e444.
- Lindo J, Huerta-Sánchez E, Nakagome S, Rasmussen M, Petzelt B, Mitchell J, Cybulski JS, Willerslev E, DeGiorgio M and Malhi RS (2016) A time transect of exomes from a Native American population before and after European contact. *Nat Commun* 7:13175.
- Lorenzo-Bermejo J, Boekstegers F, González Silos R, Marcelain K, Baez Benavides P, Barahona Ponce C, Müller B, Ferreccio C, Koshiol J, Fischer C *et al.* (2017) Subtypes of Native American ancestry and leading causes of death: Mapuche ancestry-specific associations with gallbladder cancer risk in Chile. *PLoS Genet* 13:e1006756.
- Luo Y, Suliman S, Asgari S, Amariuta T, Baglaenko Y, Martínez-Bonet M, Ishigaki K, Gutierrez-Arcelus M, Calderon R, Lecca L *et al.* (2019) Early progression to active tuberculosis is a highly heritable trait driven by 3q23 in Peruvians. *Nat Commun* 10:3765.
- Malhotra A, Kobes S, Knowler WC, Baier LJ, Bogardus C and Hanson RL (2011) A genome-wide association study of BMI in American Indians. *Obes Silver Spring Md* 19:2102-2106.
- Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S, Daly MJ, Bustamante CD and Kenny EE (2017) Human demographic history impacts genetic risk prediction across diverse populations. *Am J Hum Genet* 100:635-649.
- Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM and Daly MJ (2019) Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet* 51:584-591.
- Mas-Sandoval A, Mathieson S and Fumagalli M (2023) The genomic footprint of social stratification in admixing American populations. *eLife* 12:e84429.
- Migliano AB, Vinicius L and Lahr MM (2007) Life history trade-offs explain the evolution of human pygmies. *Proc Natl Acad Sci U S A* 104:20216-20219.
- Moreno DJ, Ruiz S, Ríos Á, Lopera F, Ostos H, Via M and Bedoya G (2017) Association of GWAS top genes with late-onset Alzheimer's disease in Colombian population. *Am J Alzheimers Dis Dementias* 32:27-35.
- Moreno-Estrada A, Gignoux CR, Fernández-López JC, Zakharia F, Sikora M, Contreras AV, Acuña-Alonzo V, Sandoval K, Eng C, Romero-Hidalgo S *et al.* (2014) The genetics of Mexico recapitulates Native American substructure and affects biomedical traits. *Science* 344:1280-1285.
- Mu Z, Wei W, Fair B, Miao J, Zhu P and Li YI (2021) The impact of cell type and context-dependent regulatory variants on human immune traits. *Genome Biol* 22:122.

- Nathan A, Asgari S, Ishigaki K, Valencia C, Amariuta T, Luo Y, Beynor JI, Baglaenko Y, Suliman S, Price AL *et al.* (2022) Single-cell eQTL models reveal dynamic T cell state dependence of disease loci. *Nature* 606:120-128.
- Neel JV (1962) Diabetes mellitus: A “thrifty” genotype rendered detrimental by “progress”? *Am J Hum Genet* 14:353-362.
- Nieves-Colón MA, Badillo Rivera KM, Sandoval K, Villanueva Dávalos V, Enriquez Lencinas LE, Mendoza-Revilla J, Adhikari K, González-Buenfil R, Chen JW, Zhang ET *et al.* (2022) Clotting factor genes are associated with preeclampsia in high-altitude pregnant women in the Peruvian Andes. *Am J Hum Genet* 109:1117-1139.
- Peng Q, Bizon C, Gizer IR, Wilhelmsen KC and Ehlers CL (2019) Genetic loci for alcohol-related life events and substance-induced affective symptoms: Indexing the “dark side” of addiction. *Transl Psychiatry* 9:71.
- Pino-Yanes M, Thakur N, Gignoux CR, Galanter JM, Roth LA, Eng C, Nishimura KK, Oh SS, Vora H, Huntsman S *et al.* (2015) Genetic ancestry influences asthma susceptibility and lung function among Latinos. *J Allergy Clin Immunol* 135:228-235.
- Posth C, Nakatsuka N, Lazaridis I, Skoglund P, Mallick S, Lamnidis TC, Rohland N, Nägele K, Adamski N, Bertolini E *et al.* (2018) Reconstructing the deep population history of Central and South America. *Cell* 175:1185-1197.e22.
- Prohaska A, Racimo F, Schork AJ, Sikora M, Stern AJ, Ilardo M, Allentoft ME, Folkersen L, Buil A, Moreno-Mayar JV *et al.* (2019) Human disease variation in the light of population genomics. *Cell* 177:115-131.
- Raghavan M, Steinrücken M, Harris K, Schiffels S, Rasmussen S, DeGiorgio M, Albrechtsen A, Valdiosera C, Ávila-Arcos MC, Malaspina A-S *et al.* (2015) Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science* 349:aab3884.
- Ramirez-Luzuriaga MJ, Kobes S, Sinha M, Knowler WC and Hanson RL (2023) Adolescent growth spurt and type 2 diabetes risk in southwestern American Indians. *Am J Epidemiol* 192:1304-1314.
- Ramirez-Luzuriaga MJ, Kobes S, Hsueh WC, Baier LJ and Hanson RL (2024) Novel signals and polygenic score for height are associated with pubertal growth traits in Southwestern American Indians. *Hum Mol Genet* 33:981-990.
- Reddy MVPL, Velázquez-Cruz R, Baca V, Lima G, Granados J, Orozco L and Alarcón-Riquelme ME (2007) Genetic association of IRF5 with SLE in Mexicans: Higher frequency of the risk haplotype and its homozygosity than Europeans. *Hum Genet* 121:721-727.
- Rojas F (2007) Poverty determinants of acute respiratory infections among Mapuche indigenous peoples in Chile’s Ninth Region of Araucanía, using GIS and spatial statistics to identify health disparities. *Int J Health Geogr* 6:26.
- Romero-Hidalgo S, Flores-Rivera J, Rivas-Alonso V, Barquera R, Villarreal-Molina MT, Antuna-Puente B, Macías-Kauffer LR, Villalobos-Comparán M, Ortiz-Maldonado J, Yu N *et al.* (2020) Native American ancestry significantly contributes to neuromyelitis optica susceptibility in the admixed Mexican population. *Sci Rep* 10:13706.
- Sánchez E, Rasmussen A, Riba L, Acevedo E, Kelly JA, Langefeld CD, García-De La Torre I, Maradiaga-Ceceña MA, Cardiel MH, Esquivel-Valerio JA *et al.* (2012) Impact of genetic ancestry and socio-demographic status on the clinical expression of systemic lupus erythematosus in Amerindian-European populations. *Arthritis Rheum* 64:3687-3694.
- Schlebusch CM, Gattepaille LM, Engström K, Vahter M, Jakobsson M and Broberg K (2015) Human adaptation to arsenic-rich environments. *Mol Biol Evol* 32:1544-1555.
- Schmidt MI, Duncan BB, Mill JG, Lotufo PA, Chor D, Barreto SM, Aquino EML, Passos VMA, Matos SMA, Molina M del CB *et al.* (2015) Cohort profile: Longitudinal study of adult health (ELSA-Brasil). *Int J Epidemiol* 44:68-75.
- Schulz LO and Chaudhari LS (2015) High-risk populations: The Pimas of Arizona and Mexico. *Curr Obes Rep* 4:92-98.
- SIGMA Type 2 Diabetes Consortium, Williams AL, Jacobs SBR, Moreno-Macias H, Huerta-Chagoya A, Churchhouse C, Márquez-Luna C, García-Ortiz H, Gómez-Vázquez MJ, Burt NP *et al.* (2014) Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. *Nature* 506:97-101.
- Simons YB, Bullaughey K, Hudson RR and Sella G (2018) A population genetic interpretation of GWAS findings for human quantitative traits. *PLoS Biol* 16:e2002985.
- Sobreira DR, Joslin AC, Zhang Q, Williamson I, Hansen GT, Farris KM, Sakabe NJ, Sinnott-Armstrong N, Bozek G, Jensen-Cody SO *et al.* (2021) Extensive pleiotropism and allelic heterogeneity mediate metabolic effects of IRX3 and IRX5. *Science* 372:1085-1091.
- Sohail M, Palma-Martínez MJ, Chong AY, Quinto-Cortés CD, Barberena-Jonas C, Medina-Muñoz SG, Ragsdale A, Delgado-Sánchez G, Cruz-Hervert LP, Ferreyra-Reyes L *et al.* (2023) Mexican Biobank advances population and medical genomics of diverse ancestries. *Nature* 622:775-783.
- Soto-Cabezas MG, Reyes MF, Soriano AN, Rodríguez JPV, Ibarguen LO, Martel KS, Jaime NF and Munayco CV (2022) COVID-19 among Amazonian indigenous in Peru: Mortality, incidence, and clinical characteristics. *J Public Health (Oxf)* 44:e359-e365.
- Spear ML, Diaz-Papkovich A, Ziv E, Yracheta JM, Gravel S, Torgerson DG and Hernandez RD (2020) Recent shifts in the genetic ancestry of Mexican Americans may alter the genetic architecture of biomedical traits. *eLife* 9:e56029.
- Tollefson D, Bloss E, Fanning A, Redd JT, Barker K and McCray E (2013) Burden of tuberculosis in indigenous peoples globally: A systematic review. *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis* 17:1139-1150.
- Verhulst B, Neale MC and Kendler KS (2015) The heritability of alcohol use disorders: A meta-analysis of twin and adoption studies. *Psychol Med* 45:1061-1072.
- Freitas Vaz I, Santana Paiva N and de Sousa Viana PV (2024) Spatial-temporal evolution of tuberculosis incidence rates in indigenous and non-indigenous people of Brazil, from 2011 to 2022. *Rev Bras Epidemiol* 26:e230055.
- Vicuña L, Norambuena T, Miranda JP, Pereira A, Mericq V, Ongaro L, Montinaro F, Santos JL and Eyheramendy S (2021) Novel loci and Mapuche genetic ancestry are associated with pubertal growth traits in Chilean boys. *Hum Genet* 140:1651-1661.
- Vicuña L, Barrientos E, Norambuena T, Alvares D, Gana JC, Leiva-Yamaguchi V, Meza C, Santos JL, Mericq V, Pereira A *et al.* (2023) New insights from GWAS on BMI-related growth traits in a longitudinal cohort of admixed children with Native American and European ancestry. *iScience* 26:106091.
- Vilhjálmsón BJ, Yang J, Finucane HK, Gusev A, Lindström S, Ripke S, Genovese G, Loh P-R, Bhatia G, Do R *et al.* (2015) Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *Am J Hum Genet* 97:576-592.
- Villanea FA, Peede D, Kaufman EJ, Añorve-Garibay V, Witt KE, Villalobos V, Zeloni R, Marnetto D, Moorjani P, Jay F *et al.* (2023) The MUC19 gene in Denisovans, Neanderthals, and modern humans: An evolutionary history of recurrent introgression and natural selection. *bioRxiv*. DOI: 10.1101/2023.09.25.559202.
- Walker R, Gurven M, Hill K, Migliano A, Chagnon N, De Souza R, Djurovic G, Hames R, Hurtado AM, Kaplan H *et al.* (2006) Growth rates and life histories in twenty-two small-scale societies. *Am J Hum Biol* 18:295-311.

- Wang S, Lewis CM, Jakobsson M, Ramachandran S, Ray N, Bedoya G, Rojas W, Parra MV, Molina JA, Gallo C *et al.* (2007) Genetic variation and population structure in Native Americans. *PLoS Genet* 3:e185.
- Watanabe K, Stringer S, Frei O, Umićević Mirkov M, de Leeuw C, Polderman TJC, van der Sluis S, Andreassen OA, Neale BM and Posthuma D (2019) A global overview of pleiotropy and genetic architecture in complex traits. *Nat Genet* 51:1339-1348.
- Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Flicek P, Manolio T, Hindorf L *et al.* (2014) The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res* 42:D1001-D1006.
- Williams AL, Jacobs SBR, Moreno-Macías H, Huerta-Chagoya A, Churchhouse C, Márquez-Luna C, García-Ortíz H, José Gómez-Vázquez M, Burt NP, Aguilar-Salinas CA *et al.* (2014) Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. *Nature* 506:97-101.
- Wojcik GL, Graff M, Nishimura KK, Tao R, Haessler J, Gignoux CR, Highland HM, Patel YM, Sorokin EP, Avery CL *et al.* (2019) Genetic analyses of diverse populations improves discovery for complex traits. *Nature* 570:514-518.
- Yang JJ, Cheng C, Devidas M, Cao X, Fan Y, Campana D, Yang W, Neale G, Cox NJ, Scheet P *et al.* (2011) Ancestry and pharmacogenomics of relapse in acute lymphoblastic leukemia. *Nat Genet* 43:237-241.
- Yengo L, Vedantam S, Marouli E, Sidorenko J, Bartell E, Sakaue S, Graff M, Eliassen AU, Jiang Y, Raghavan S *et al.* (2022) A saturated map of common genetic variants associated with human height. *Nature* 610:704-712.

Internet Resources

- CRS Report (2023) Congressional Research Service Report. Indigenous Peoples in Latin America: Statistical Information 2023, <https://crsreports.congress.gov/R4622> (accessed 27 October 2023).
- US Census Bureau (2022) Population Division. Annual Estimates of the Resident Population by Sex, Race and Hispanic Origin for the United States: April 1, 2020 to July 1, 2022 (NC-EST2022-SR11H), (American Indian and Alaska Native) (2022), <https://www.census.gov/> (accessed 16 January 2022).

Associate Editor: Carlos F. M. Menck

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License (type CC-BY), which permits unrestricted use, distribution and reproduction in any medium, provided the original article is properly cited.