

Mendelian Randomization Analysis of the Relationship Between Native American Ancestry and Gallbladder Cancer Risk

Linda Zollner, MSc¹, Felix Boekstegers, PhD¹, Carol Barahona Ponce, PhD¹, Dominique Scherer, PhD¹, Katherine Marcelain, PhD², Valentina Gárate-Calderón, MSc^{1,2}, Melanie Waldenberger, PhD³, Erik Morales, MD^{4,5}, Armando Rojas, PhD⁵, César Munoz, MD^{4,5}, Bettina Müller, MD⁶, Javier Retamales, MD⁶, Gonzalo de Toro, MD^{7,8}, Allan Vera Kortmann⁷, Olga Barajas, MD^{2,9}, María Teresa Rivera, MD¹⁰, Analía Cortés, MD¹⁰, Denisse Loader, MD¹¹, Javiera Saavedra, MD¹¹, Lorena Gutiérrez, MD¹², Alejandro Ortega, MD¹³, Maria Enriqueta Bertrán, MD¹⁴, Leonardo Bartolotti, MD¹⁵, Fernando Gabler, MD¹⁶, Mónica Campos, MD¹⁶, Juan Alvarado, MD¹⁷, Fabricio Moisés, MD¹⁷, Loreto Spencer, MD¹⁷, Bruno Nervi, MD¹⁸, Daniel Carvajal, MD¹⁹, Héctor Losada, MD²⁰, Mauricio Almau, MD²¹, Plinio Fernández, MD²¹, Jordi Olloquequi, PhD^{22,23}, Alice R. Carter²⁴, Juan Francisco Miquel Poblete, MD²⁵, Bernabe Ignacio Bustos, PhD²⁶, Macarena Fuentes Guajardo, PhD²⁷, Rolando Gonzalez-Jose, PhD²⁸, Maria Cátira Bortolini, PhD²⁹, Victor Acuña-Alonzo, PhD³⁰, Carla Gallo, PhD³¹, Andres Ruiz Linares, PhD^{32,33,34}, Francisco Rothhammer, PhD³⁵, Justo Lorenzo Bermejo, PhD^{1,*}

¹ Statistical Genetics Research Group, Institute of Medical Biometry, Heidelberg University, Germany

² Department of Basic and Clinical Oncology, Medical Faculty, University of Chile, Chile

³ Research Unit Molecular Epidemiology, Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

⁴ Hospital Regional de Talca, Talca, Chile

⁵ Facultad de Medicina, Universidad Católica del Maule, Talca, Chile

⁶ Instituto Nacional del Cáncer, Santiago, Chile

⁷ Hospital de Puerto Montt, Puerto Montt, Chile

⁸ Escuela de Tecnología Médica, Universidad Austral de Chile sede Puerto Montt, Puerto Montt, Chile

⁹ Hospital Clínico Universidad de Chile, Santiago, Chile

¹⁰ Hospital del Salvador, Santiago, Chile

¹¹ Hospital Padre Hurtado, Santiago, Chile

¹² Hospital San Juan de Dios, Santiago, Chile

¹³ Hospital Regional, Arica, Chile

¹⁴ Unidad Registro Hospitalario de Cáncer, Hospital Base de Valdivia, Valdivia, Chile

¹⁵ Hospital Base de Valdivia, Chile

¹⁶ Hospital San Borja Arriarán, Santiago, Chile

¹⁷ Hospital Regional Guillermo Grant Benavente, Concepción, Chile

¹⁸ Departamento de Hematología y Oncología, Escuela de Medicina Pontificia Universidad Católica de Chile, Santiago, Chile

¹⁹ Facultad de Medicina, Clínica Alemana Universidad del Desarrollo, Santiago, Chile

²⁰ Departamento de Cirugía, Universidad de la Frontera, Temuco, Chile

²¹ Hospital de Rancagua, Rancagua, Chile

²² Department of Biochemistry and Physiology, Faculty of Pharmacy and Food Sciences, University of Barcelona, Barcelona, Spain

²³ Facultad de Ciencias de la Salud, Universidad Autónoma de Chile, Talca, Chile

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

²⁴ MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

²⁵ Departamento de Gastroenterología, Escuela de Medicina, Pontificia Universidad Católica de Chile, Santiago, Chile

²⁶ Ken and Ruth Davee Department of Neurology and Simpson Querrey Center for Neurogenetics, Northwestern University, Feinberg School of Medicine, Chicago, IL, USA

²⁷ Departamento de Tecnología Médica, Facultad de Ciencias de la Salud, Tarapacá University, Arica, Chile

²⁸ Instituto Patagónico de Ciencias Sociales y Humanas, Centro Nacional Patagónico, CONICET, Puerto Madryn, Argentina

²⁹ Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

³⁰ National Institute of Anthropology and History, Mexico City, Mexico

³¹ Laboratorios de Investigación y Desarrollo, Facultad de Ciencias y Filosofía, Universidad Peruana Cayetano Heredia, Lima, Perú

³² Ministry of Education Key Laboratory of Contemporary Anthropology and Collaborative Innovation Center of Genetics and Development, School of Life Sciences and Human Phenome Institute, Fudan University, Shanghai, China

³³ ADES (Anthropologie Bio-Culturelle, Droit, Éthique et Santé), UFR de Médecine, Aix-Marseille University, Marseille, France

³⁴ Department of Genetics, Evolution and Environment, and UCL Genetics Institute, University College London, London, UK

³⁵ Instituto de Alta Investigación, Tarapacá University, Arica, Chile

17

18 * Corresponding Author

19 Email: lorenzo@imbi.uni-heidelberg.de

20 Full postal address: Statistical Genetics Research Group, Institute of Medical Biometry, INF

21 130.3, 69120 Heidelberg, Germany

22

23 **Short title:** Mapuche Ancestry and Gallbladder Cancer

24 **Keywords:** Genetic admixture, Native American Mapuche ancestry, Gallbladder cancer,

25 Gallstone disease, Mendelian randomization

26

Abstract

Background: A strong association between the proportion of Native American ancestry and the risk of gallbladder cancer (GBC) has been reported in observational studies. Chileans show the highest incidence of GBC worldwide, and the Mapuche are the largest Native American people in Chile. We set out to investigate the causal association between Native American Mapuche ancestry and GBC risk, and the possible mediating effects of gallstone disease and body mass index (BMI) on this association.

Methods: Markers of Mapuche ancestry were selected based on the informativeness for assignment measure and then used as instrumental variables in two-sample mendelian randomization (MR) analyses and complementary sensitivity analyses.

Results: We found evidence of a causal effect of Mapuche ancestry on GBC risk (inverse variance-weighted (IVW) risk increase of 0.8% for every 1% increase in Mapuche ancestry proportion, 95% CI 0.4% to 1.2%, $p = 6.6 \times 10^{-5}$). Mapuche ancestry was also causally linked to gallstone disease (IVW risk increase of 3.6% per 1% increase in Mapuche proportion, 95% CI 3.1% to 4.0%, $p = 1.0 \times 10^{-59}$), suggesting a mediating effect of gallstones in the relationship between Mapuche ancestry and GBC. In contrast, the proportion of Mapuche ancestry showed a negative causal effect on BMI (IVW estimate -0.006 kg/m^2 per 1% increase in Mapuche proportion, 95% CI -0.009 to -0.003 , $p = 4.4 \times 10^{-5}$).

Conclusions: The results presented here may have significant implications for GBC prevention and are important for future admixture mapping studies. Given that the association between Mapuche ancestry and GBC risk previously noted in observational studies appears to be causal, primary and secondary prevention strategies that take into account the individual proportion of Mapuche ancestry could be particularly efficient.

Introduction

Every year, around 116,000 people are diagnosed with gallbladder cancer (GBC) and 85,000 die due to this aggressive disease worldwide (The Global Cancer Observatory, 2020). The malignancy of the biliary tract primarily affects women in low- and middle-income countries, and relatively little effort has been invested in research on GBC (Mehrotra et al., 2018; Stinton & Shaffer, 2012).

The development of GBC is probably driven by a combination of environmental exposures and genetic predisposition (Hundal & Shaffer, 2014). Symptoms are often absent or unspecific until the disease has progressed to a non-curative stage, leaving patients with few treatment options (Hundal & Shaffer, 2014; Roa & de Aretxabala, 2015; Stinton & Shaffer, 2012). As GBC is mostly diagnosed at an advanced stage, the 5-year survival rate is low, reported at between 5% and 30% depending on the country of origin of the study population (Bertran et al., 2010; Hundal & Shaffer, 2014; Kanthan et al., 2015; Stinton & Shaffer, 2012; Zhu et al., 2020). Gallbladder removal (cholecystectomy) is a valuable tool for primary and secondary prevention of GBC, but little progress has been made in individualized risk prediction and early diagnosis (Roa & de Aretxabala, 2015).

The incidence of GBC shows wide geographical and ethnic variation (Mehrotra et al., 2018). High-income regions such as Western Europe, the United States and Australia have 1–2 cases per 100,000 person-years. In contrast, the largest Native American people in Chile – the Mapuche – shows the highest incidence of GBC in the world, with more than 20 cases per 100,000 person-years (Campbell et al., 2017). Observational studies have found a strong association between the individual proportion of Native American Mapuche ancestry and GBC risk: each 1% increase in the proportion of Mapuche ancestry translates into a 3.7% increase in GBC mortality (Lorenzo Bermejo et al., 2017). However, the observed association may arise from other established GBC risk factors, especially gallstones and elevated body mass index (BMI) (Hundal & Shaffer, 2014; Roa & de Aretxabala, 2015; Stinton & Shaffer, 2012). The Mapuche have a high prevalence of gallstone disease, a relative GBC risk of 4.9 has been observed in individuals with a history of gallstones, and recent studies have found evidence of a causal effect of gallstones on GBC risk for genetically admixed Chileans (odds ratio [OR] = 1.97) (Barahona Ponce et al., 2021;

Mehrotra et al., 2018; Stinton & Shaffer, 2012). Native American ancestry has also been associated with increased BMI: the World Cancer Research Fund considers that higher body fatness marked by BMI probably causes GBC, and relative GBC risks of 1.59 for women and 1.09 for men per five-point increase in BMI have been reported (Campbell et al., 2017; Hundal & Shaffer, 2014; Roa & de Aretxabala, 2015; World Cancer Research Fund/American Institute for Cancer Research, 2018). Furthermore, multiparity in women, low socio-economic status, chronic inflammation and lifestyle choices such as cigarette smoking and alcohol consumption could also confound the association between ethnicity and GBC, justifying the use of MR to assess the causal relationship between Mapuche ancestry and GBC risk (Lugo et al., 2020; Roa & de Aretxabala, 2015; Stinton & Shaffer, 2012; Wistuba & Gazdar, 2004).

In response to the high GBC mortality rates and lack of treatment options, but based on weak scientific evidence, the Chilean government is implementing a GBC prevention program that provides financial support to gallstone patients for prophylactic cholecystectomy (Roa & de Aretxabala, 2015). Prioritizing prophylactic cholecystectomy for those at high risk of GBC, taking into account causal risk factors rather than observational associations potentially attributable to confounding, would optimize the efficiency of current GBC prevention measures. Mapuche ancestry per se neither causes nor prevents disease, but statistical inference on the causal effect of the individual's proportion of Mapuche ancestry on GBC risk allows for a distinction between correlation and confounding-free causation, allowing for prioritization of causal risk factors for prevention and a better understanding of the underlying etiological mechanisms leading to GBC development.

With this goal in mind, we preselected markers of Mapuche ancestry in a panel of Native American, European and African individuals, relying on the informativeness for assignment measure (Rosenberg et al., 2003). We then used a subset of ancestry markers as instrumental variables for the individual proportion of Mapuche ancestry in two-sample MR to assess the causal effect of Mapuche ancestry on GBC risk. It should be noted that population stratification, typically a limitation in MR studies, was the exposure of interest in the present study.

Results

The flowchart in **Figure 1** describes the performed analyses with some intermediate results. Preselection of markers of Mapuche ancestry in a reference panel composed of Native American Mapuche (n = 28), Native American Aymara (n = 63), European (CEU and IBS, n = 206) and African (YRI, n = 108) individuals relying on the informativeness for assignment measure yielded 21,854 candidate markers. Of these, 985 genetic variants were robustly associated ($p < 5 \times 10^{-8}$) with the individual proportion of Mapuche ancestry, which was previously estimated in a dataset with genome-wide genotype data from 1,861 genetically admixed Chileans. Exclusion of markers associated with established GBC risk factors (PheWAS $p < 5 \times 10^{-8}$) complemented with LD pruning ($r^2 < 0.01$) resulted in 430 instrumental variables preliminarily retained for the subsequent MR analyses. Radial MR based on summary statistics of the association between the instrumental variables and the proportion of Mapuche ancestry (sample I: 1,861 admixed Chileans) and that between the instrumental variables and GBC status (sample II: 412 Chilean GBC patient–control pairs), both summary statistics adjusted for age and sex, detected 33 outlying instruments. The subsequent MR analysis of the association Mapuche ancestry $\rightarrow$ GBC was therefore based on 397 instrumental variables, which explained 13.2% of the variance in the proportion of Mapuche ancestry.

Figure 2 depicts the genetic principal components, and the estimated proportions of Mapuche, Aymara and European ancestry in the two samples used for the main MR analysis. Individuals with a proportion of Mapuche ancestry above the 95th percentile in the sample are represented in orange; blue and green highlight individuals with the largest proportions of Aymara and European ancestry, respectively. The first principal component in admixed Chileans distinguished between European and Native American ancestry, and the second principal component separated the two main types of Native American ancestry in Chile: Mapuche and Aymara. A strong correlation between the second principal component and the proportion of Mapuche ancestry was found in the two samples used for MR analysis (panels C and F in **Figure 2**). In the three plots depicting sample II (panels D, E and F in **Figure 2**), crosses represent GBC patients and circles represent population-based controls. The concentration of GBC patients among individuals with the highest proportion of Native American ancestry, especially Mapuche ancestry, was striking.

MR analysis of the association Mapuche ancestry → GBC risk revealed no heterogeneity among instrumental variables as a proxy for pleiotropy (inverse variance-weighted (IVW) Cochran's Q statistic $p = 0.99$) and no horizontal pleiotropy (MR-Egger intercept $p = 0.87$; **Table 1**). Neither outliers nor weak instrument biases were apparent in the scatter and funnel plots (panels A and B in **Figure 3**). We found evidence of a causal effect of Mapuche ancestry on GBC risk (IVW OR = 1.008, which translates into a GBC risk increase of 0.8% for every 1% increase in the proportion of Mapuche ancestry, 95% confidence interval [CI] 0.4% to 1.2%, $p = 6.6 \times 10^{-5}$). MR-Egger (OR = 1.009, 95% CI 1.004 to 1.015, $p = 6.7 \times 10^{-4}$) and weighted median estimates (OR = 1.009, 95% CI 0.991 to 1.028, $p = 0.33$) were consistent with IVW estimates. As described above, the second principal component of genetic variability in admixed Chileans reflects the individual proportion of Mapuche ancestry, and, as expected, the causal effect vanished after including the second principal component in the calculation of summary statistics ($p = 0.62$). In contrast, consideration of the first and third to tenth principal components translated into stronger estimates of the causal effect (1.4% risk increase, 95% CI 0.8% to 1.9%, $p = 5.9 \times 10^{-7}$).

Table 1 MR results on the causal relationship between Native American Mapuche ancestry as exposure and GBC, gallstone disease and BMI as outcomes of interest. Cochran's Q statistic p-values higher than 0.05 are suggestive of no instrument heterogeneity as a proxy for pleiotropy, and MR-Egger intercept p-values higher than 0.05 are consistent with no horizontal pleiotropy. Results from sensitivity analyses on the association Mapuche ancestry → GBC risk after including the first to the tenth principal component, both with and without the second principal component, which reflects the individual proportion of Mapuche ancestry in admixed Chileans, are also shown.

Outcome	IV	Inverse variance weighted					MR-Egger
		β /OR	95% CI		β p-value	Q p-value	Intercept p-value
GBC	397	1.008	1.004	1.012	6.6×10^{-5}	0.99	0.87
Including PC1-PC10	397	1.007	0.979	1.037	0.62	0.99	0.17
Including PC1 & PC3-PC10	397	1.014	1.008	1.019	5.9×10^{-7}	0.99	0.36
Gallstone disease	388	1.036	1.031	1.040	1.0×10^{-59}	0.95	0.63
BMI	391	-0.006	-0.009	-0.003	4.4×10^{-5}	0.96	0.90

IV: Number of instrumental variables, β /OR: estimated causal effect size for BMI/estimated causal odds ratio for GBC and gallstone disease per 1% increase in the proportion of Mapuche ancestry, CI: confidence interval, Q: Cochran's Q statistic, PC: principal components estimated based on individual genotypes. Bold represents probability values smaller than 0.05.

Further sensitivity analysis considered two-step MR and multivariable MR (MVMR) to investigate the potential mediating effects of gallstone disease and BMI on the link between Mapuche ancestry and GBC. For two-step MR mediation analysis, we first assessed the causal effect of Mapuche ancestry on gallstone disease. Neither heterogeneity among instrumental variables (Q statistic $p = 0.95$) nor horizontal pleiotropy (MR-Egger intercept $p = 0.63$) was noted. We found evidence of a causal effect of Mapuche ancestry on the risk of gallstone disease (IVW risk increase of 3.6% for every 1% increase in Mapuche proportion, $p = 1.0 \times 10^{-59}$; MR-Egger OR = 1.032, 95% CI 1.026 to 1.039, $p = 5.2 \times 10^{-24}$; weighted median OR = 1.031, 95% CI 1.010 to 1.051, $p = 0.003$). These results are visually represented in **Figure 3** (panels C and D).

Since we found a causal effect of Mapuche ancestry on gallstones in the present study, and a causal effect of gallstones on GBC risk has already been demonstrated for Chileans, the two-step MR analysis suggests that gallstones mediate the association between Mapuche ancestry and GBC risk. To validate this hypothesis, we also performed MVMR. The instrumental variables used for MVMR were weak (Mapuche ancestry F statistics = 3.49; gallstones F statistics = 1.37), but horizontal pleiotropy was not found (Q statistic of 425 on 432 degrees of freedom, $p = 0.59$). In agreement with the results of two-step MR, MVMR revealed a causal effect of gallstones on GBC risk (OR = 1.263, $p = 2.0 \times 10^{-9}$) but no concurrent causal effect of Mapuche ancestry on GBC risk (OR=1.0001, $p = 0.95$).

Neither heterogeneity among instrumental variables (Q statistic $p = 0.96$) nor horizontal pleiotropy (MR-Egger intercept $p = 0.90$) was detected in MR analysis of the association Mapuche ancestry $\rightarrow$ BMI. We found evidence of a negative causal effect of Mapuche ancestry on BMI (IVW $\beta = -0.006$ kg/m² per 1% increase in Mapuche proportion, 95% CI -0.009 to -0.003, $p = 4.4 \times 10^{-5}$). The corresponding scatter and funnel plots are shown in panels E and F of **Figure 3**. MR-Egger ($\beta = -0.007$ kg/m² per 1% Mapuche ancestry proportion, 95% CI -0.020 to 0.006, $p = 0.31$) and weighted median estimates ($\beta = -0.006$ kg/m² per 1% Mapuche ancestry proportion, 95% CI -0.011 to -0.002, $p = 3.8 \times 10^{-3}$) showed good agreement with the negative causal effect estimated by IVW.

Discussion

Utilizing ancestry informative markers as instrumental variables for the individual proportion of Mapuche ancestry in admixed Chileans, we applied two-sample MR to assess the causal relationship between Mapuche ancestry and GBC risk as the primary outcome, as well as two-step and MVMR to examine the mediating effects of gallstones and BMI on the relationship Mapuche ancestry $\rightarrow$ GBC risk. According to the MR results, the proportion of Mapuche ancestry is causally associated with GBC development and gallstones mediate this causal association. The practical implication of the identified causal relationships is that intensified surveillance and prophylactic gallbladder removal in gallstone carriers with a high percentage of Mapuche ancestry may improve the efficiency of current GBC prevention programs. The extent to which gallstones mediate the underlying mechanisms of the relationship between Mapuche ancestry and GBC needs to be quantified in future larger studies. Given the causal association between increasing proportions of Mapuche ancestry and decreasing BMI, and the causal association between increasing BMI and increasing GBC risk, it seems rather unlikely that BMI positively mediates between Mapuche ancestry and GBC risk.

Gallstones are highly prevalent in Chileans and one of the most important risk factors for GBC: observational studies report a relative GBC risk of 9.2–10.1 for individuals with gallstones larger than 3 cm (Miquel et al., 1998; Stinton & Shaffer, 2012). We therefore examined the possible mediating effect of gallstone disease on the association between Mapuche ancestry and GBC risk. The evidence for a causal effect of Mapuche ancestry on gallstone disease was quite robust: the estimated causal OR was 1.036 per 1% Mapuche ancestry proportion. This finding, together with the previously reported causal effect of gallstones on GBC risk in Chileans, support a mediating effect of gallstones based on two-step MR (Barahona Ponce et al., 2021). The MVMR results, based on a relatively small sample size and weak instrumental variables, showed good agreement with two-step MR and also supported a mediating effect of gallstones between Mapuche ancestry and GBC risk; however, large prospective Chilean datasets with complete information on GBC, gallstones and individual genotype data are urgently needed to conduct formal mediation analyses and accurately quantify the magnitude of the direct and indirect effects of gallstones, which are highly relevant for more precise GBC prevention.

The International Agency for Research on Cancer, the American Institute for Cancer Research and the World Cancer Research Fund all consider obesity to be a likely cause of GBC, and high BMI is another important GBC risk factor, with a direct effect on GBC risk reported for Chileans and an indirect effect mediated by gallstones for Europeans (Barahona Ponce et al., 2021; Lauby-Secretan et al., 2016; World Cancer Research Fund/American Institute for Cancer Research, 2018). However, studies of the association between Native American ancestry and BMI have yielded contradictory results. BMI decreased by 0.13 m/kg² for each 10% increase in the proportion of Native American ancestry in Mexican–American women (Hu et al., 2015), and the association was also negative in a Hispanic–Mexican study (Tang et al., 2006). In contrast, BMI increased by 0.56 m/kg² per 10% increase in the proportion of Native American ancestry in Native Americans (Norden-Krichmar et al., 2014). Native American ancestry was positively correlated with obesity in Mexico and Peru, whereas no association was found in Brazil, Chile, and Colombia, and higher Native American ancestry was associated with overweight and obesity, but only among foreign-born Latina women (Ziv et al., 2006). In this study we found clear evidence of a negative causal effect of Mapuche ancestry on BMI, probably ruling out a positive mediating effect of BMI on the association between Mapuche ancestry and GBC risk.

The relatively small number of GBC patients investigated was a limitation of this study, especially considering the large sample sizes normally required for MR. However, the variance in the proportion of Mapuche ancestry explained by the instrumental variables was high (13.2%), and observational studies have reported a strong association between Mapuche ancestry and GBC risk. Besides low statistical power, another common limitation of MR studies is pleiotropy. First-order inverse variance weights keep the type I error rate under the causal null, and we calculated Cochran’s Q statistic using first-order weights to detect heterogeneity, which often reflects pleiotropy. We also visually inspected scatter and funnel plots, performed MR-Egger regression to detect potential bias attributable to horizontal pleiotropy, used radial MR to detect outlying instruments, and excluded instrumental variables associated with GBC risk factors not investigated in our study.

Previous studies have demonstrated the importance of subdividing Native American ancestry into its main subcomponents –in the case of admixed Chileans, the two major Native American subcomponents are Mapuche and Aymara ancestry (Lorenzo Bermejo et al., 2017). While combined Native American ancestry showed no association with GBC mortality and the proportion of Aymara ancestry showed a negative association with GBC mortality, each 1% increase in the Mapuche proportion translated into a 3.7% increase in the risk of death due to GBC (95% CI 3.1 to 4.3%). To assess the contribution of health system access to the identified association between Mapuche ancestry and GBC risk, hospitalization rates due to gallbladder removal (cholecystectomy) were included as an additional explanatory variable in the fitted regression models; this showed a minor role of health system access in the estimated increase in GBC mortality risk. The MR-based OR estimated in the present study confirmed a causal effect of Mapuche ancestry on GBC risk, while reducing the effect size to 0.8% (95% CI 0.4% to 1.2%). This result suggests that GBC prevention strategies that consider the individual proportion of Mapuche ancestry may be particularly efficient, and that the difference between the observational (3.7%) and causal (0.8%) increase in GBC risk is due to confounders of the association between Mapuche ancestry and GBC risk, which may include modifiable risk factors. Further research is warranted.

Beyond the practical relevance for GBC prevention, the results presented here may also be useful for designing future studies. On the one hand, the evidence of a causal effect of Mapuche ancestry on GBC risk underpins the potential of admixture mapping to identify novel GBC susceptibility variants, possibly in combination with subsequent association testing (note that the genome-wide significance level is much higher for admixture mapping than for association mapping) (Horimoto et al., 2021). For example, assuming that the average proportion of Mapuche ancestry in Chileans is about 40%, approximately 1100 GBC patients are needed to detect a Mapuche haplotype with a risk ratio of 1.8, consistent with the estimated causal OR of 1.008 per 1% proportion of Mapuche ancestry in the whole Chilean genome (McKeigue, 2005). On the other hand, the 3.7% increased GBC mortality risk per 1% Mapuche ancestry proportion previously reported in observational studies probably overestimates the contribution of Mapuche ancestry to GBC risk and would lead to

underpowered admixture mapping studies. From an implementation perspective, recruiting study participants from the southern regions of Chile would increase the average proportion of Mapuche ancestry and thus the statistical power of the study.

In summary, we assessed the relationship between the individual proportion of Native American Mapuche ancestry and GBC risk using mendelian randomization. To our knowledge, this is the first MR study that considers genetic ancestry as the exposure of interest – it is important to keep in mind that the study was based on genetically admixed Chileans, who show continuous gradients of ancestry. Since statistical evidence of a causal effect of ancestry on disease development is generally more relevant than an observational association, which could also be due to confounding, the present findings provide refined information on the potential of accounting for ethnic differences (in this case Mapuche ancestry) in prevention of disease (in this case gallbladder cancer). While ancestry proportions are not modifiable exposures, personalized preventive measures could prioritize gallstone screening and prophylactic cholecystectomy for individuals with a high percentage of Mapuche ancestry. Future admixture mapping studies could also benefit from the methodology applied in the present investigation to test and quantify the causal effects of genetic ancestry on disease outcomes. From a more applied point of view, we found strong evidence for a causal effect of Mapuche ancestry proportion on GBC risk, most likely mediated by gallstones, with direct implications for the development of more efficient GBC prevention strategies.

Materials and Methods

Ethics approval

The study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki and was approved by the ethics committees of Servicio de Salud Metropolitano Oriente, Santiago de Chile (#06.10.2015, #08.03.2016 and #12.11.2019), Servicio de Salud Metropolitano Sur Oriente, Santiago de Chile (#15.10.2015 and #05.04.2018), Servicio de Salud Metropolitano Central, Santiago de Chile (#1188-2015), Servicio de Salud Coquimbo, Coquimbo, Chile (#01.04.2016), Servicio de Salud Maule, Talca, Chile (#05.11.2015), Universidad Católica del Maule, Talca, Chile (#102-2020), Servicio de Salud Concepción, Concepción, Chile (ID: 16-11-97 and ID:19-12-111), Servicio de Salud Araucanía Sur, Temuco, Chile (#10.02.2020), Servicio de Salud Valdivia, Valdivia, Chile (ID:438), Centro de Bioética, Universidad del Desarrollo, Clínica Alemana de Santiago, Santiago de Chile (#2018-97, ID 678) and Unidad de Investigación Hospital San Juan de Dios, Santiago de Chile (#6182), the Medical Faculties of Universidad de Chile (approval #123-2012 and #11.10.2012) and Pontificia Universidad Católica de Chile (#11-159), and Universidad de Tarapacá and University College London as described in (Ruiz-Linares et al., 2014). All participants provided written informed consent before enrolment.

Study participants

The Chilean genome is a mixture of chromosomal segments from two major Native American peoples, the Aymara in the north and the Mapuche in the south of the country; Europeans; and, to a lesser extent, Africans (Eyheramendy et al., 2015; Hundal & Shaffer, 2014; Ruiz-Linares et al., 2014). The reference panel for preselection of ancestry-informative markers and estimation of individual ancestry proportions in genetically admixed Chileans therefore included 63 Aymara (Reich et al., 2012) and 28 Mapuche (Lindo et al., 2018; Reich et al., 2012), as well as 206 Europeans (99 Utah residents with Northern and Western European ancestry [CEU] and 107 Iberians from Spain [IBS]), and 108 African Yorubans from Ibadan, Nigeria (YRI) from the 1000 Genome Project (The 1000 Genomes Project Consortium, 2015).

Sample I for two-sample MR analyses of the associations Mapuche ancestry → GBC, Mapuche ancestry → gallstone disease and Mapuche ancestry → BMI consisted of 1,861 Chileans from the Consortium for the Analysis of the Diversity and Evolution of Latin America (CANDELA) (Ruiz-Linares et al., 2014). Sample II for MR analysis of the association Mapuche ancestry → GBC risk was composed of 412 Chilean patients diagnosed with GBC and 412 population-based controls matched by age and sex with the GBC patients. Sample II for MR analysis of the association Mapuche ancestry → gallstone disease was composed of 351 Chilean gallstone patients and 351 controls matched by age and sex. Among GBC patients, 77% were diagnosed after surgical removal of the gallbladder (cholecystectomy), and gallstones were found in about 86% of the GBC patients investigated. Gallstone patients were patients who underwent cholecystectomy due to symptomatic gallstones. Population-based controls were selected from the Chilean subgroup of CANDELA and from Chilean studies on COPD and Chagas disease, with GBC and gallstone incidences representative of the general Chilean population (Blandino et al., 2022). Sample I and sample II for MR analyses on Mapuche ancestry → GBC risk and Mapuche ancestry → gallstone disease partially overlapped: 84 of the 412 controls and 91 of the 351 controls in the respective sample II also belonged to sample I. Sample II for MR analysis of the association Mapuche ancestry → BMI was based on 12,216 individuals from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL, dbGaP accession number phs000810.v1.p1).

Preselection of markers of Mapuche ancestry

To preselect genetic variants robustly associated with our exposure of interest, the individual proportion of Mapuche ancestry, we first chose markers of Mapuche ancestry based on the informativeness for assignment measure I_n (Rosenberg et al., 2003). For each genetic variant with $j = 1, \dots, N$ possible alleles in $i = 1, \dots, K$ subpopulations we calculated:

$$I_n(Q, J) = \sum_{j=1}^N \left(-p_j \log(p_j) + \sum_{i=1}^K \frac{p_{ij}}{K} \log(p_{ij}) \right),$$

where p_j denotes the average frequency of allele j in all K subpopulations, p_{ij} the average frequency of allele j in subpopulation i , Q the (random) assignment of an individual to a subpopulation and J the (random) genotype of one of the two alleles of an individual. I_n conditioning on Q and J is a general approach to quantify the amount

of information that multi-allelic markers provide about individual ancestry composition. We used the reference panel described above to estimate I_n for each of the 43,625 genetic variants available for all individuals in this study considering two subpopulations at a time (Mapuche – European, Mapuche – Aymara and Mapuche – African). We selected the 10,000 genetic variants with the highest I_n for each comparison and retained the genetic variants that were present in at least one comparison for the subsequent analyses.

Selection of instrumental variables

In order to ensure that preselected ancestry informative markers fulfilled the first MR assumption (relevance), we used the software ADMIXTURE (version 1.3) for supervised estimation of the individual proportions of Mapuche ancestry in sample I, relying on the reference panel of Aymara, Mapuche, European and African individuals (Alexander et al., 2009), and retained as instrumental variables (IV) the genetic variants that showed a robust association ($p < 5 \times 10^{-8}$) with the estimated proportion of Mapuche ancestry, adjusted for age and sex. The second MR assumption (independence) was assessed by a phenome-wide association study (PheWAS), excluding genetic variants associated ($p < 5 \times 10^{-8}$) with potential confounders (menopause, educational level, diabetes, body circumference, smoking, alcohol consumption or gallstones) in MR-Base (Hemani et al., 2018). We then used the summary statistics of the association between the genetic variants and the estimated proportion of Mapuche ancestry (β) and the minor allele frequency (MAF) to calculate the variance in Mapuche ancestry proportion explained by each variant:

$$\text{Explained variance} = \beta^2 \times 2 \times \text{MAF} \times (1 - \text{MAF}),$$

and excluded variants in linkage disequilibrium (LD) ($r^2 > 0.01$) with other variants that explained a larger variance.

To try to fulfil the third MR assumption (exclusion restriction), radial MR was applied separately for each investigated outcome to detect and subsequently exclude outlying and influential genetic variants in combination with further techniques described in the next section.

Two-sample MR, sensitivity and mediation analyses (two-step and multivariable MR)

Using the previously selected IV and the respective sample I and sample II as described above, we performed MR on the association Mapuche ancestry → GBC risk as the primary analysis. In addition to radial MR, we (1) visually inspected the funnel and scatter plots of the IV outcome (GBC) against IV exposure (proportion of Mapuche ancestry) association statistics, both adjusted for age and sex, to detect weak instrument bias; (2) calculated Cochran's Q statistic using first-order inverse variance weights to detect heterogeneity, which indicates a possible violation of the instrumental variable or modelling assumptions, of which pleiotropy is a likely major cause; and (3) used the p-value for non-zero MR-Egger intercept to assess horizontal pleiotropy. Our primary objective was to identify statistical causal effects, which requires weaker MR assumptions than estimation of their magnitude, and we tested causality based on random-effect inverse variance-weighted (IVW) p-values. As a secondary objective we estimated the causal effect sizes and assessed their robustness by comparing IVW, weighted median and MR-Egger regression parameter estimates (ORs for GBC and gallstone disease, and beta values for BMI per 1% increase in the proportion of Mapuche ancestry).

Population stratification is particularly relevant in MR studies of genetically admixed individuals. To deal with potential stratification, association statistics are typically adjusted for the main principal components of genetic variability. However, the individual proportion of Mapuche ancestry was the exposure of interest in this study of Chileans, who show continuous gradients of ancestry. To check the sensitivity of MR results to population stratification, we estimated the genetic principal components in sample I and sample II using the *eigenstrat* function (Patterson et al., 2006); examined the correlation between the estimated principal components and the proportion of Mapuche ancestry; adjusted the association statistics for (1) all of the first ten principal components or (2) the first ten principal components with the exception of the principal component correlated with the proportion of Mapuche ancestry; and repeated the MR analyses.

Since a causal effect of gallstones and BMI on GBC risk has recently been reported (Barahona Ponce et al., 2021), we investigated the mediating effects of gallstone

disease and BMI as a surrogate marker for obesity, as potential mechanisms explaining the relationship between Mapuche ancestry and GBC risk. The unavailability of large prospective Chilean datasets with complete information on GBC, gallstone, BMI and individual genotypes precluded the implementation of formal mediation analyses, and we decided first to apply two-step MR to assess mediation (Zheng et al., 2017). In the first MR step, IV for the proportion of Mapuche ancestry were used to assess the causal effect of the Mapuche ancestry proportion on the potential mediator (gallstone disease or BMI) by MR as described above. The second MR step was based on published findings, also based on genetically admixed Chileans, on the causal effect of the mediators on GBC risk. Evidence of association in the two MR steps (Mapuche ancestry $\rightarrow$ gallstone disease and gallstone disease $\rightarrow$ GBC) would imply some degree of mediation between Mapuche ancestry and GBC risk on the part of the intermediate trait (gallstones). We also performed MVMR to simultaneously assess the causal effects of Mapuche ancestry and gallstones on GBC risk. MVMR allows the estimation of mediation effects when considering potentially correlated risk factors utilizing genetic instruments within the MR framework. We used the R package 'MVMR', following the workflow suggested by the software developers (Sanderson et al., 2021). We calculated F statistics to monitor instrument strength, Q statistics to assess instrument validity including horizontal pleiotropy, and finally estimated the direct effects of the considered exposures (Mapuche ancestry and gallstone disease) simultaneously on the outcome of interest (GBC risk). The IV used included the previously selected IV for Mapuche ancestry along with five genetic variants (one palindromic variant was removed) robustly associated with gallstone disease in Chileans (Joshi et al., 2016). The sample I and sample II used to assess the association between IV and Mapuche ancestry, and between IV and gallstone disease, did not overlap. They consisted of 1,703 Chileans (1,861 minus those used as controls in sample II) and 351 gallstone patients and matched controls, respectively. The sample used to assess the direct effect between IV and gallstone disease on GBC risk was the same as the sample II utilized for the MR analysis of Mapuche ancestry $\rightarrow$ GBC risk (412 Chilean GBC cases and 412 matched controls).

MR analyses were conducted using the R version of MR-Base, which provides convenient tools for the harmonization of the association statistics, including

standardization of the effect alleles and removal of problematic palindromic genetic variants (Hemani et al., 2018). The R package ‘RadialMR’ was used for radial MR (Bowden et al., 2018). For general data processing we used Plink version 1.9 (Purcell et al., 2007), and the R software environment for statistical computing and graphics (version 3.6.2). The source code to reproduce all the results described is provided as supplementary material, and the necessary input files are available at www.biometrie.uni-heidelberg.de/StatisticalGenetics/Software_and_Data.

References

- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Barahona Ponce, C., Scherer, D., Brinster, R., Bookstegers, F., Marcelain, K., Gárate-Calderón, V., Müller, B., de Toro, G., Retamales, J., Barajas, O., Ahumada, M., Morales, E., Rojas, A., Sanhueza, V., Loader, D., Rivera, M. T., Gutiérrez, L., Bernal, G., Ortega, A., ... Lorenzo Bermejo, J. (2021). Gallstones, Body Mass Index, C-Reactive Protein, and Gallbladder Cancer: Mendelian Randomization Analysis of Chilean and European Genotype Data. *Hepatology*, 73(5), 1783–1796. <https://doi.org/10.1002/hep.31537>
- Bertran, E., Heise, K., Andia, M. E., & Ferreccio, C. (2010). Gallbladder cancer: Incidence and survival in a high-risk area of Chile. *International Journal of Cancer*, 127(10), 2446–2454. <https://doi.org/10.1002/ijc.25421>
- Blandino, A., Scherer, D., Rounge, T. B., Umu, S. U., Bookstegers, F., Barahona Ponce, C., Marcelain, K., Gárate-Calderón, V., Waldenberger, M., Morales, E., Rojas, A., Munoz, C., Retamales, J., de Toro, G., Barajas, O., Rivera, M. T., Cortés, A., Loader, D., Saavedra, J., ... Bermejo, J. L. (2022). Identification of Circulating lncRNAs Associated with Gallbladder Cancer Risk by Tissue-Based Preselection, Cis-eQTL Validation, and Analysis of Association with Genotype-Based Expression. *Cancers*, 14(3). <https://doi.org/10.3390/cancers14030634>
- Bowden, J., Spiller, W., Del Greco, F. M., Sheehan, N., Thompson, J., Minelli, C., & Smith, G. D. (2018). Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression. *International Journal of Epidemiology*, 47(4), 1264–1278. <https://doi.org/10.1093/ije/dyy101>
- Campbell, P. T., Newton, C. C., Kitahara, C. M., Patel, A. V., Hartge, P., Koshiol, J., McGlynn, K. A., Adami, H.-O., Berrington De Gonzalez, A., Beane Freeman, L. E., Bernstein, L., Buring, J. E., Freedman, N. D., Gao, Y. T., Giles, G. G., Gunter, M. J., Jenab, M., Liao, L. M., Milne, R. L., ... Gapstur, S. M. (2017). Body Size Indicators and Risk of Gallbladder Cancer: Pooled Analysis of Individual-Level Data from 19 Prospective Cohort Studies. *Cancer Epidemiology Biomarkers and Prevention*, 26(4), 597–606. <https://doi.org/10.1158/1055-9965.EPI-16-0796>
- Eyheramendy, S., Martinez, F. I., Manevy, F., Vial, C., & Repetto, G. M. (2015). Genetic structure characterization of Chileans reflects historical immigration patterns. *Nature Communications*, 6, 6472. <https://doi.org/10.1038/ncomms7472>
- Hemani, G., Zheng, J., Elsworth, B., Wade, K. H., Haberland, V., Baird, D., Laurin, C., Burgess, S., Bowden, J., Langdon, R., Tan, V. Y., Yarmolinsky, J., Shihab, H. A., Timpson, N. J., Evans, D. M., Relton, C., Martin, R. M., Davey Smith, G., Gaunt, T. R., & Haycock, P. C. (2018). The MR-Base platform supports systematic causal inference across the human phenome. *ELife*, 7, e34408. <https://doi.org/10.7554/eLife.34408>
- Horimoto, A. R. V. R., Xue, D., Thornton, T. A., & Blue, E. E. (2021). Admixture mapping reveals the association between Native American ancestry at 3q13.11 and reduced risk of Alzheimer's disease in Caribbean Hispanics. *Alzheimer's Research and Therapy*, 13. <https://doi.org/10.1186/s13195-021-00866-9>
- Hu, H., Huff, C. D., Yamamura, Y., Wu, X., & Strom, S. S. (2015). The Relationship between Native American Ancestry, Body Mass Index and Diabetes Risk among Mexican-Americans. *PLoS ONE*, 10(10), e0141260.

- 571 <https://doi.org/10.1371/journal.pone.0141260>
- 572 Hundal, R., & Shaffer, E. A. (2014). Gallbladder cancer: epidemiology and outcome.
- 573 *Clinical Epidemiology*, 6, 99–109. <https://doi.org/10.2147/CLEP.S37357>
- 574 Joshi, A. D., Andersson, C., Buch, S., Stender, S., Noordam, R., Weng, L. C.,
- 575 Weeke, P. E., Auer, P. L., Boehm, B., Chen, C., Choi, H., Curhan, G., Denny, J.
- 576 C., De Vivo, I., Eicher, J. D., Ellinghaus, D., Folsom, A. R., Fuchs, C., Gala, M.,
- 577 ... Johnson, A. D. (2016). Four Susceptibility Loci for Gallstone Disease
- 578 Identified in a Meta-analysis of Genome-Wide Association Studies.
- 579 *Gastroenterology*, 151(2), 351-363.e28.
- 580 <https://doi.org/10.1053/j.gastro.2016.04.007>
- 581 Kanthan, R., Senger, J.-L., Ahmed, S., & Kanthan, S. C. (2015). Gallbladder Cancer
- 582 in the 21st Century. *Journal of Oncology*, 2015, 967472.
- 583 <https://doi.org/10.1155/2015/967472>
- 584 Lauby-Secretan, B., Scoccianti, C., Loomis, D., Grosse, Y., Bianchini, F., & Straif, K.
- 585 (2016). Body Fatness and Cancer — Viewpoint of the IARC Working Group. In
- 586 *New England Journal of Medicine* (Vol. 375).
- 587 <https://doi.org/10.1056/NEJMSr1606602>
- 588 Lindo, J., Haas, R., Hofman, C., Apata, M., Moraga, M., Verdugo, R. A., Watson, J.
- 589 T., Llave, C. V., Witonsky, D., Beall, C., Warinner, C., Novembre, J.,
- 590 Aldenderfer, M., & Di Rienzo, A. (2018). The genetic prehistory of the Andean
- 591 highlands 7000 years BP though European contact. *Science Advances*, 4(11),
- 592 eaau4921. <https://doi.org/10.1126/sciadv.aau4921>
- 593 Lorenzo Bermejo, J., Boekstegers, F., González Silos, R., Marcelain, K., Baez
- 594 Benavides, P., Barahona Ponce, C., Müller, B., Ferreccio, C., Koshiol, J.,
- 595 Fischer, C., Peil, B., Sinsheimer, J., Fuentes Guajardo, M., Barajas, O.,
- 596 Gonzalez-Jose, R., Bedoya, G., Cátira Bortolini, M., Canizales-Quinteros, S.,
- 597 Gallo, C., ... Rothhammer, F. (2017). Subtypes of Native American ancestry and
- 598 leading causes of death: Mapuche ancestry-specific associations with
- 599 gallbladder cancer risk in Chile. *PLOS Genetics*, 13(5), e1006756.
- 600 <https://doi.org/10.1371/journal.pgen.1006756>
- 601 Lugo, A., Peveri, G., & Gallus, S. (2020). Should we consider gallbladder cancer a
- 602 new smoking-related cancer? A comprehensive meta-analysis focused on dose–
- 603 response relationships. *International Journal of Cancer*, 146(12), 3304–3311.
- 604 <https://doi.org/10.1002/ijc.32681>
- 605 McKeigue, P. M. (2005). Prospects for admixture mapping of complex traits.
- 606 *American Journal of Human Genetics*, 76(1), 1–7.
- 607 <https://doi.org/10.1086/426949>
- 608 Mehrotra, R., Tulsyan, S., Hussain, S., Mittal, B., Singh Saluja, S., Singh, S., Tanwar,
- 609 P., Khan, A., Javle, M., Hassan, M. M., Pant, S., De Aretxabala, X., Sirohi, B.,
- 610 Rajaraman, P., Kaur, T., & Rath, G. K. (2018). Genetic landscape of gallbladder
- 611 cancer: Global overview. *Mutation Research - Reviews in Mutation Research*,
- 612 778, 61–71. <https://doi.org/10.1016/j.mrrev.2018.08.003>
- 613 Miquel, J. F., Covarrubias, C., Villaroel, L., Mingrone, G., Greco, A. V., Puglielli, L.,
- 614 Carvalho, P., Marshall, G., Del Pino, G., & Nervi, F. (1998). Genetic epidemiology
- 615 of cholesterol cholelithiasis among Chilean Hispanics, Amerindians, and Maoris.
- 616 *Gastroenterology*, 115(4), 937–946. [https://doi.org/10.1016/S0016-](https://doi.org/10.1016/S0016-5085(98)70266-5)
- 617 [5085\(98\)70266-5](https://doi.org/10.1016/S0016-5085(98)70266-5)
- 618 Norden-Krichmar, T. M., Gizer, I. R., Libiger, O., Wilhelmsen, K. C., Ehlers, C. L., &
- 619 Schork, N. J. (2014). Correlation analysis of genetic admixture and social
- 620 identification with body mass index in a Native American community. *Human*
- 621 *Biology*, 26(3), 347–360. <https://doi.org/doi:10.1002/ajhb.22521>

- Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genetics*, 2(12), 2074–2093. <https://doi.org/10.1371/journal.pgen.0020190>
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *American Journal of Human Genetics*, 81(3), 559–575. <https://doi.org/10.1086/519795>
- Reich, D., Patterson, N., Campbell, D., Tandon, A., Mazieres, S., Ray, N., Parra, M. V., Rojas, W., Duque, C., Mesa, N., García, L. F., Triana, O., Blair, S., Maestre, A., Dib, J. C., Bravi, C. M., Bailliet, G., Corach, D., Hünemeier, T., ... Ruiz-Linares, A. (2012). Reconstructing Native American population history. *Nature*, 488(7411), 370–374. <https://doi.org/10.1038/nature11258>
- Roa, I., & de Aretxabala, X. (2015). Gallbladder cancer in Chile: what have we learned? *Current Opinion in Gastroenterology*, 31(3), 269–275. <https://doi.org/10.1097/MOG.0000000000000164>
- Rosenberg, N. A., Li, L. M., Ward, R., & Pritchard, J. K. (2003). Informativeness of Genetic Markers for Inference of Ancestry. *American Journal of Human Genetics*, 73(6), 1402–1422. <https://doi.org/10.1086/380416>
- Ruiz-Linares, A., Adhikari, K., Acuña-Alonzo, V., Quinto-Sanchez, M., Jaramillo, C., Arias, W., Fuentes, M., Pizarro, M., Everardo, P., de Avila, F., Gómez-Valdés, J., León-Mimila, P., Hunemeier, T., Ramallo, V., Silva de Cerqueira, C. C., Burley, M. W., Konca, E., de Oliveira, M. Z., Veronez, M. R., ... Gonzalez-José, R. (2014). Admixture in Latin America: Geographic Structure, Phenotypic Diversity and Self-Perception of Ancestry Based on 7,342 Individuals. *PLoS Genetics*, 10(9). <https://doi.org/10.1371/journal.pgen.1004572>
- Sanderson, E., Spiller, W., & Bowden, J. (2021). Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization. *Statistics in Medicine*, 40(25), 5434–5452. <https://doi.org/10.1002/sim.9133>
- Stinton, L. M., & Shaffer, E. A. (2012). Epidemiology of Gallbladder Disease: Cholelithiasis and Cancer. *Gut and Liver*, 6(2), 172–187. <https://doi.org/10.5009/gnl.2012.6.2.172>
- Tang, H., Jorgenson, E., Gadde, M., Kardia, S. L. R., Rao, D. C., Zhu, X., Schork, N. J., Hanis, C. L., & Risch, N. (2006). Racial admixture and its impact on BMI and blood pressure in African and Mexican Americans. *Human Genetics*, 119(6), 624–633. <https://doi.org/10.1007/s00439-006-0175-4>
- The 1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature*, 526(7571), 68–74. <https://doi.org/10.1038/nature15393>
- The Global Cancer Observatory. (2020). *Global Cancer Observatory: Cancer Today - Gallbladder cancer fact sheet*. Gco.iarc.Fr. <https://gco.iarc.fr/today/data/factsheets/cancers/12-Gallbladder-fact-sheet.pdf>
- Wistuba, I. I., & Gazdar, A. F. (2004). Gallbladder cancer: lessons from a rare tumour. *Nature Reviews Cancer*, 4(9), 695–706. <https://doi.org/10.1038/nrc1429>
- World Cancer Research Fund/American Institute for Cancer Research. (2018). *Diet, Nutrition, Physical Activity and Cancer: A Global Perspective. Continuous Update Project Expert Report 2018*. Dietandcancerreport.Org. <http://dietandcancerreport.org>
- Zheng, J., Baird, D., Borges, M.-C., Bowden, J., Hemani, G., Haycock, P., Evans, D. M., & Davey Smith, G. (2017). Recent Developments in Mendelian Randomization Studies. *Current Epidemiology Reports*, 4(4), 330–345. <https://doi.org/10.1007/s40471-017-0128-6>

- Zhu, X., Zhang, X., Hu, X., Ren, H., Wu, S., Wu, J., Wu, G., Si, X., & Wang, B. (2020). Survival analysis of patients with primary gallbladder cancer from 2010 to 2015: A retrospective study based on SEER data. *Medicine*, 99(40), e22292. <https://doi.org/10.1097/MD.00000000000022292>
- Ziv, E., John, E. M., Choudhry, S., Kho, J., Lorizio, W., Perez-Stable, E. J., & Burchard, E. G. (2006). Genetic ancestry and risk factors for breast cancer among Latinas in the San Francisco Bay Area. *Cancer Epidemiology Biomarkers and Prevention*, 15(10), 1878–1885. <https://doi.org/10.1158/1055-9965.EPI-06-0092>

Figure legends

Figure 1 Flowchart describing the main analyses, from the selection of markers of Mapuche ancestry in a reference panel composed of Mapuche, Aymara, European and African individuals, to the two-sample MR based on 397 selected instrumental variables after exclusion of outlying instruments based on radial MR

Figure 1 Scatter plots of first and second genetic principal components (PC) and the individual proportion of Mapuche ancestry. Individuals with proportions of Mapuche, Aymara and European ancestry above the 95th percentile are shown in orange, blue and green, respectively. Panels A, B and C refer to sample I; panels D, E and F refer to sample II used in the main MR analysis, respectively. In sample II, GBC patients are represented by crosses and population-based controls by circles.

Figure 3 Scatter and funnel plots from MR analyses after exclusion of outlying instruments according to radial MR on the associations between Mapuche ancestry and GBC risk (panels A and B), gallstone disease (panels C and D) and BMI (panels E and F)

Funding

This study was supported by the European Union's Horizon 2020 research and innovation program (grant 825741); the Deutsche Forschungsgemeinschaft (DFG; grant LO 1928/11-1, project number 424112940); and the Biobank of the University of Chile (BTUCH). For the publication fee we acknowledge financial support from the DFG within the funding programme „Open Access Publikationskosten“ and from Heidelberg University. Part of the research was conducted using the Hispanic Community Health Study/Study of Latinos (HCHS/SOL, dbGaP accession number phs000810.v1.p1). The HCHS/SOL study was supported by the National Heart, Lung, and Blood Institute (NHLBI) at the University of North Carolina (N01-HC65233), University of Miami (N01-HC65234), Albert Einstein College of Medicine (N01-HC65235), Northwestern University (N01-HC65236) and San Diego State University (N01-HC65237). The funders had no role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; the preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication.

Conflict of Interest Statement

The authors declare that they have no conflicts of interest.

Acknowledgements

The authors gratefully acknowledge the data storage service SDS@hd supported by the Ministry of Science, Research, and the Arts Baden-Württemberg (MWK) and the German Research Foundation (DFG) through grants INST 35/1314-1 FUGG and INST 35/1503-1 FUGG, and the access to computing facilities provided by Steven J. Lubbe at Northwestern University - Feinberg School of Medicine.

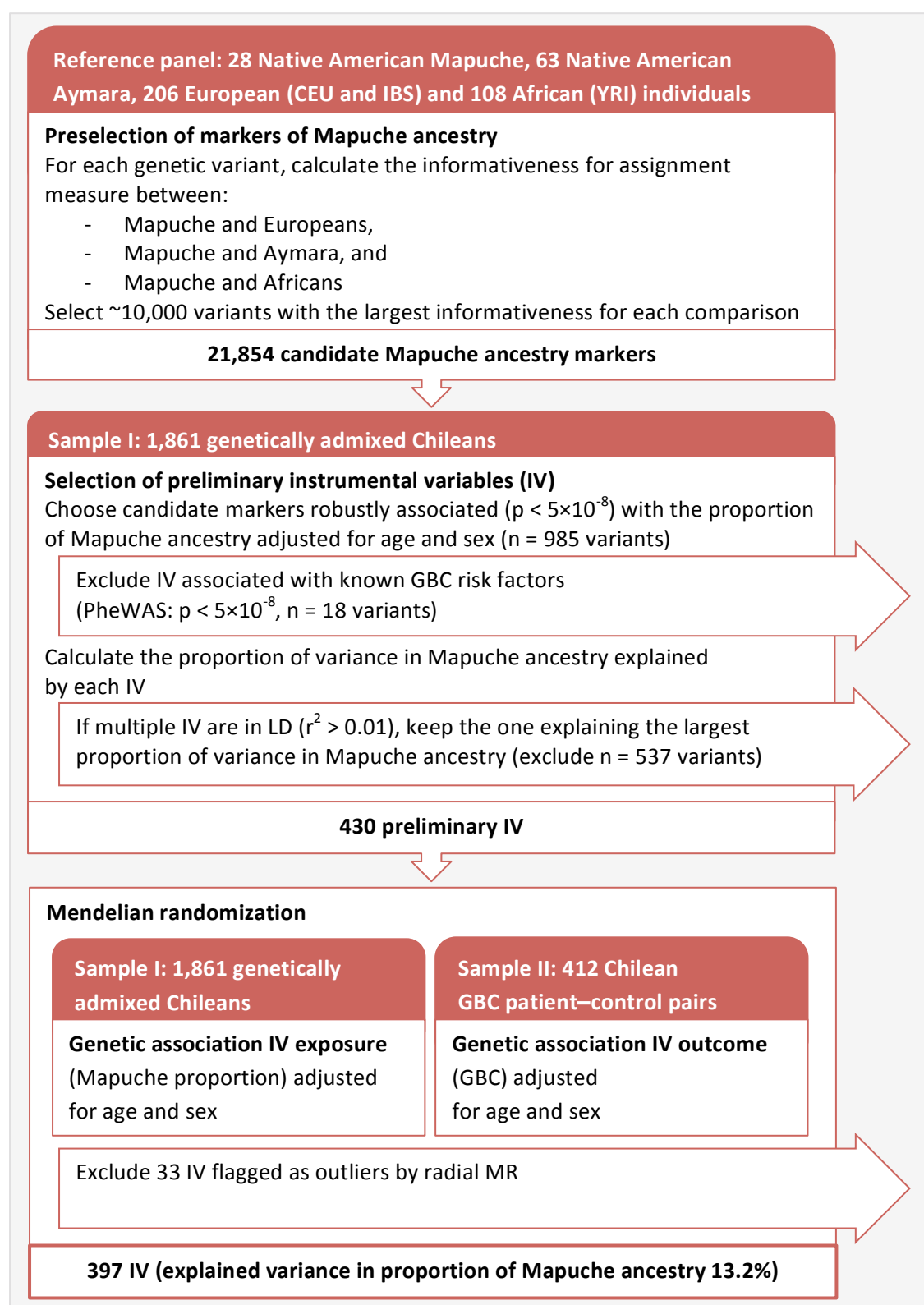


Figure 1 Flowchart describing the main analyses, from the selection of markers of Mapuche ancestry in a reference panel composed of Mapuche, Aymara, European and African individuals, to the two-sample MR based on 397 selected instrumental variables after exclusion of outlying instruments based on radial MR

Figure 2 Scatter plots of first and second genetic principal components (PC), and the individual proportion of Mapuche ancestry. Individuals with proportions of Mapuche, Aymara and European ancestry above the 95th percentile are shown in orange, blue and green, respectively. Panels A, B and C refer to sample I; panels D, E and F refer to sample II used in the main MR analysis, respectively. In sample II, GBC patients are represented by crosses and population-based controls by circles.

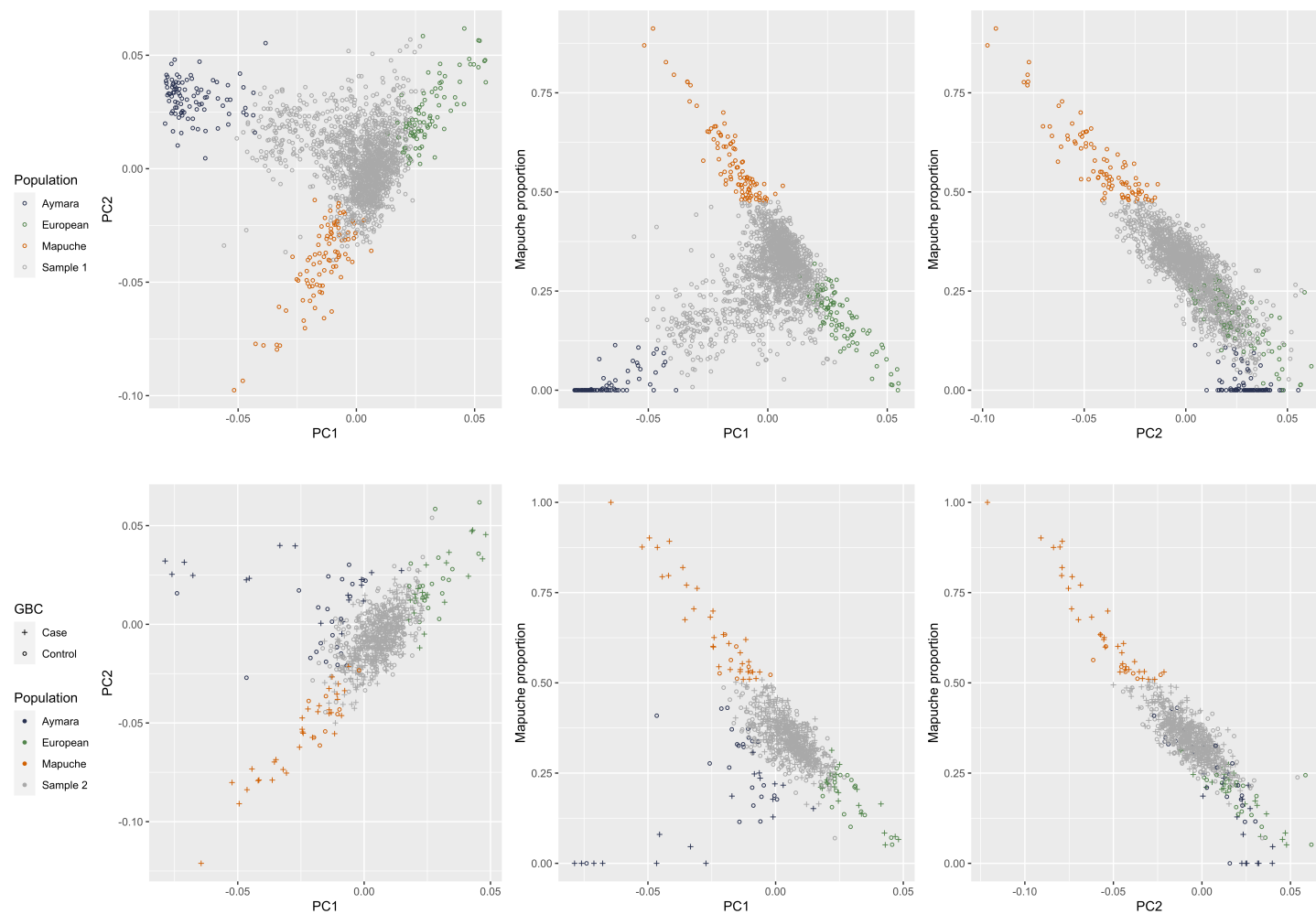


Figure 3 Scatter and funnel plots from MR analyses after exclusion of outlying instruments according to radial MR on the association between Mapuche ancestry and GBC risk (panels A and B), gallstone disease (panels C and D) and BMI (panels E and F)

